Dissertation
submitted to the
Combined Faculties for the Natural Sciences and for Mathematics
of the Ruperto-Carola University of Heidelberg, Germany
for the degree of
Doctor of Natural Sciences

presented by
Diplom-Physicist Michael Klar,
born in Heilbronn.

Date of oral examination: 30.11.2005.

Design of an endoscopic
3-D Particle-Tracking Velocimetry system
and its application in flow measurements
within a gravel layer

Referees: Prof. Dr. Bernd Jähne

Prof. Dr. Kurt Roth

Zusammenfassung

In der vorliegenden Arbeit wird eine neuartige Technik zur Messung von dreidimensionalen Strömungen in einem porösen Medium vorgestellt, die erstmals den experimentellen Zugang zur Porenströmung in einer Kiessohle erlaubt. Zwei faseroptische Endoskope werden in stereoskopischer Anordnung eingesetzt, um Bildsequenzen des Strömungsfeldes innerhalb einer einzelnen Kiespore zu gewinnen. Zur Auswertung der Bilddaten wird die 3-D Particle-Tracking Velocimetry (3-D PTV) verwendet. Diese ermöglicht die dreidimensionale Rekonstruktion Lagrange'scher Partikeltrajektorien. Die zugrundeliegenden Bildverarbeitungsverfahren werden entscheidend weiterentwickelt und an die speziellen Verhältnisse endoskopischer Bildgewinnung angepasst. Dies beinhaltet Methoden zur Bildvorverarbeitung, zur robusten Kamerakalibrierung, zur Bildsegmentierung sowie zur Partikelverfolgung. Nach einer Leistungs- und Genauigkeitsanalyse wird das Messverfahren in umfangreichen systematischen Untersuchungen der Strömung durch eine Kiessohle in einer Versuchsrinne der Bundesanstalt für Wasserbau in Karlsruhe eingesetzt. Ein erweiterter experimenteller Aufbau ermöglicht neben der Messung der Porenströmung in drei Poren die simultane Erfassung des sohlnahen 3-D Strömungsfelds in der turbulenten Kanalströmung oberhalb der Kiessohle sowie von Kornbewegungen in einer Sandschicht unterhalb der Kiessohle. Somit kann erstmals die Interaktion der Oberflächenströmung mit der Strömung im Porenraum zeitlich und räumlich hoch aufgelöst untersucht werden. Die experimentellen Untersuchungen sind Teil eines internationalen Forschungsprojekts des Filter And Erosion Research Clubs (FERC). Das langfristige Ziel dieses Projekts ist es, den Einfluss turbulenter Geschwindigkeits- und Druckschwankungen auf die Sohlstabilität von Wasserstrassen zu quantifizieren. Die gewonnenen Messdaten ermöglichen neue Einblicke in das Dämpfungsverhalten einer Kiessohle und können zukünftig zum Vergleich mit numerischen, analytischen und phänomenologischen Modellen herangezogen werden.

Abstract

In this thesis a novel method for 3-D flow measurements within a permeable gravel layer is developed. Two fiberoptic endoscopes are used in a stereoscopic arrangement to acquire image sequences of the flow field within a single gravel pore. The images are processed by a 3-D Particle-Tracking Velocimetry (3-D PTV) algorithm, which yields the three-dimensional reconstruction of Lagrangian particle trajectories. The underlying image processing algorithms are significantly enhanced and adapted to the special conditions of endoscopic imagery. This includes methods for image preprocessing, robust camera calibration, image segmentation and particle-tracking. After a performance and accuracy analysis, the measurement technique is applied in extensive systematic investigations of the flow within a gravel layer in an experimental flume at the Federal Waterways Engineering and Research Institute in Karlsruhe. In addition to measurements of the pore flow within three gravel pores, an extended experimental setup enables the simultaneous observation of the near-bed 3-D flow field in the turbulent open-channel flow above the gravel layer and of grain motions in a sand layer beneath the gravel layer. The interaction of the free surface flow and the pore flow can be analyzed for the first time with a high temporal and spatial resolution. The experiments are part of a research project initiated by an international cooperation called Filter and Erosion Research Club (FERC). The long-term goal of this project is to quantify the influence of turbulent velocity and pressure fluctuations on the bed stability of waterways. The obtained experimental data provide new insight into the damping behaviour of a gravel bed and can be used for comparison with numerical, analytical and phenomenological models.

(...) We had made many similar journeys together, but the Danube, more than any other river I knew, impressed us from the very beginning with its aliveness. From its tiny bubbling entry into the world among the pinewood gardens of Donaueschingen, until this moment when it began to play the great river-game of losing itself among the deserted swamps, unobserved, unrestrained, it had seemed to us like following the grown of some living creature. Sleepy at first, but later developing violent desires as it became conscious of its deep soul, it rolled, like some huge fluid being, through all the countries we had passed, holding our little craft on its mighty shoulders, playing roughly with us sometimes, yet always friendly and well-meaning, till at length we had come inevitably to regard it as a Great Personage. How, indeed, could it be otherwise, since it told us so much of its secret life? At night we heard it singing to the moon as we lay in our tent, uttering that odd sibilant note peculiar to itself and said to be caused by the rapid tearing of the pebbles along its bed, so great is its hurrying speed. We knew, too, the voice of its gurgling whirlpools, suddenly bubbling up on a surface previously quite calm; the roar of its shallows and swift rapids; its constant steady thundering below all mere surface sounds; and that ceaseless tearing of its icy waters at the banks. How it stood up and shouted when the rains fell flat upon its face! And how its laughter roared out when the wind blew upstream and tried to stop its growing speed! (...)

ALGERNON BLACKWOOD, THE WILLOWS (1907)

The subject of the flow of fluids, and particularly of water, fascinates everybody. We can all remember, as children, playing in the bathtub or in mud puddles with the strange stuff. As we get older, we watch streams, waterfalls, and whirlpools, and we are fascinated by this substance which seems almost alive relative to solids. The behaviour of fluids is in many ways very unexpected and interesting (...). The efforts of a child trying to dam a small stream flowing in the street and his surprise at the strange way the water works its way out has its analog in our attempts over the years to understand the flow of fluids. We have tried to dam the water up - in our understanding - by getting the laws and the equations that describe the flow. We will describe these attempts in this chapter. In the next chapter, we will describe the unique way in which water has broken through the dam and escaped our attempts to understand it.

RICHARD P. FEYNMAN, THE FEYNMAN LECTURES ON PHYSICS (1964)

Contents

Nomenclature

The notation used in this thesis is summarized in table 1 (related to the image processing in part II of the thesis) and table 2 (related to the hydromechanic analysis in parts I and III of the thesis).

Following Hartley and Zisserman [2000] and Heuel [2002] we do not distinguish between projective points and their coordinate vectors. Homogeneous vectors $\mathbf{P} \in \mathbb{P}^3 \simeq \mathbb{R}^4$ are denoted by capital upright bold letters, e.g. $\mathbf{P}_w = [X_w, Y_w, Z_w, 1]^T$, Euclidean vectors $\boldsymbol{P} \in \mathbb{R}^3$ are denoted by capital italic bold letters, e.g. $\boldsymbol{P}_w = [X_w, Y_w, Z_w]^T$. Euclidean vectors $\boldsymbol{p} \in \mathbb{R}^2$ and homogeneous vectors $\mathbf{p} \in \mathbb{R}^3$ are denoted in lowercase letters. Homogeneous matrices are denoted in capital sans serif letters, e.g. a homography H. Euclidean matrices, e.g. a 3-D rotation matrix $\boldsymbol{R}$, are denoted in capital italic sans serif letters.

In cases where pixel-wise operations shall be highlighted, images are denoted by $g(i, j)$, where i is the row index and j is the column index.

A spatial average of a quantity q is denoted by $\langle q \rangle$, a temporal average by $\widehat{q}$ or $\bar{q}$.

The experimental setup developed and applied in this work comprises eleven cameras, which are all operated simultaneously to gather data from seven different positions of the flow in an experimental test flume. To ease the assignment of particular results to their experimental setup resp. measurement position, the seven positions are referred to by the name of the workstation that was used to acquire the data. For the definition of the coordinate system, see section 9.1.2. In particular, we have the following seven subsystems:

1. **JAGST**: endoscopic stereo system, upstream artificial gravel pore at $x = -18\ cm$

2. **KOCHER**: endoscopic stereo system, middle artificial gravel pore at $x = 1.5\ cm$

3. **HÖLLBACH**: endoscopic stereo system, downstream artificial gravel pore at $x = 26\ cm$

4. **NECKAR**: C-mount lens stereo system, near-wall free surface flow at $x = 46\ cm$ resp. $x = 26\ cm$

5. **ELBE(U)**: single camera at gravel-sand-interface, upstream at $x = 80\ cm$

6. **ELBE(M)**: single camera at gravel-sand-interface, middle at $x = 130\ cm$

7. **ELBE(D)**: single camera at gravel-sand-interface, downstream at $x = 180\ cm$

Table 1: *Notation related to the image processing. Homogeneous quantities are defined only up to a scale factor α.*

variable	unit	denoting
$\mathbf{P}_w = \alpha\left[X_w, Y_w, Z_w, 1\right]^T$	mm	homogeneous world coordinates of 3-D point
$\mathbf{p}_n = \alpha\left[x_n, y_n, 1\right]^T$	mm	homogeneous normalized image coordinates of ideal, undistorted 2-D image point
$\mathbf{p}_u = \alpha\left[x_u, y_u, 1\right]^T$	mm	homogeneous image coordinates of ideal, undistorted 2-D image point
$\mathbf{p}_d = \alpha\left[x_d, y_d, 1\right]^T$	mm	homogeneous image coordinates of observed, distorted 2-D image point
$\mathbf{p}_p = \alpha\left[x_p, y_p, 1\right]^T$	pix	homogeneous pixel coordinates of 2-D image point
$\mathbf{l} = (l_1, l_2, l_3)^T$	pix	homogeneous representation of an image line
P		homogeneous (3×4)-projection matrix
H		homogeneous (3×3)-projection matrix (homography)
$\boldsymbol{R}$		3-D Euclidean rotation matrix
$\boldsymbol{E}$		essential matrix
$\boldsymbol{F}$		fundamental matrix
$\boldsymbol{K}$		camera calibration matrix
$\boldsymbol{C} = [C_x, C_y, C_z]^T$	mm	Euclidean world coordinates of camera center
f_x, f_y	pix	effective focal lengths
$\boldsymbol{c} = [c_x, c_y]^T$	mm	2-D Euclidean image coordinates of principal point
$\boldsymbol{T} = [T_x, T_y, T_z]^T$	mm	3-D Euclidean translation vector
ω, ϕ, κ	$^\circ$	Euler angles
s		camera skew factor
$\boldsymbol{\theta}_d = [k_1, k_2, t_1, t_2]^T$		vector of lens distortion parameters
$\boldsymbol{p} = [x_p, y_p]^T$	pix	pixel coordinates of a particle image

Table 2: _Notation related to the hydromechanic analysis._

variable	unit	denoting
x	cm or mm	spatial coordinate in streamwise direction
y	cm or mm	spatial coordinate in vertical direction
z	cm or mm	spatial coordinate in spanwise direction
v_x	cm/s or mm/s	instantaneous velocity in streamwise direction
v_y	cm/s or mm/s	instantaneous velocity in vertical direction
v_z	cm/s or mm/s	instantaneous velocity in spanwise direction
B	m	(spanwise) flume width
L	m	(streamwise) flume length
H_S	m	height of sand layer
H_P	m	height of gravel layer
d_{mD}	mm	mean grain diameter of gravel
d_{50}	mm	grain diameter at 50% sieve fraction
C_c	-	$= d_{60}/d_{10}$, degree of non-uniformity
d_S	mm	mean grain diameter of sand
$\tau_{0,c}$	Pa	critical shear stress of gravel
k	m	geometric roughness height
k_s	m	equivalent sand roughness
Q	m^3/s	flow rate
h	m	water level
U	m/s	mean bulk velocity of free surface flow
Re	-	Reynolds number
u_*	m/s	shear velocity
Re_*	-	roughness Reynolds number
Re_p	-	pore Reynolds number
Re_κ	-	permeability Reynolds number
μ	$kg/m/s$	dynamic viscosity
ν	m^2/s	kinematic viscosity
v_f	m/s	superficial velocity
v_i	m/s	intrinsic velocity
ϕ	-	porosity
κ	m^2	permeability

x

Chapter 1

Introduction

This thesis presents a new experimental approach to three-dimensional measurements of the pore flow within a permeable and rough wall, driven by a turbulent open-channel flow on top. These measurements have been made possible by a new experimental setup based on **endoscopic imaging**. Stereoscopic image sequences of tracer particles visualizing the pore flow are acquired and analyzed by digital image processing techniques, namely an algorithm for **3-D particle-tracking velocimetry** (3-D PTV). Section 1.1 gives a motivation for this research, showing that this particular kind of flow configuration is relevant for a number of applications. Consequently, a large body of research has already been published on the subject. However, due to the great complexity of experimental investigations, most studies are theoretical or numerical, and thus limited to small Reynolds numbers. Many authors agree that there is a definite lack of experimental data, especially for unsteady and turbulent flows. This thesis presents a step forward to fill this gap. A short summary of related research is given in Section 1.2, pointing out the new contributions and objectives of this work. The structure of the thesis is outlined in Section 1.3.

1.1   Motivation

1.1.1   Flow over and through permeable walls

The experimental study of flow over and through a permeable and rough wall is an interesting and challenging subject of basic hydromechanic research. In addition, the topic is of great importance for a number of applications, with spatial scales ranging from nanometers to kilometers. On the micro-scale, permeable walls are encountered in biochemical and medical applications. For example, blood vessels in the human body are lined with a deformable porous wall layer affecting the gliding motion of blood cells (which interestingly finds a striking analogon in a human skier or snowboarder skiing on compressed powder) [Feng and Weinbaum, 2000]. On the macro-scale, high-density urban areas or areas with dense vegetation like forests also represent permeable wall layers with respect to air flow in the lower atmosphere. Topics of interest in this area are e.g. the exchange of carbon dioxide between forests and the atmosphere or the dispersion of pollutants [Finnigan, 2000]. Other areas of application include the prediction of groundwater flow and dissolved contaminants in hydrology, oil recovery in petroleum engineering or flow through packed-bed reactors in chemical engineering.

All these quite different applications have one thing in common. They all seek to find models for flow

and transport within and/or in the immediate vicinity of porous media. Having established such models, they shall then be used to predict the transport of mass, momentum or heat through the porous medium and the influence of the porous medium on the flow surrounding it. Obviously, a precise description of the flow is essential in understanding the associated transport phenomena. In the laminar flow regime (Stokes flow), analytical approaches to flow modeling are possible, and many theories have been developed, e.g. [Cushman, 1990; Genabeek, 1992; Feng and Weinbaum, 2000; James and Davis, 2001]. For turbulent flows above permeable walls, some studies have been performed by Direct Numerical Simulation (DNS) [Jimenez et al., 2001; Hahn et al., 2002; Breugem, 2004] or by combined analytical and numerical approaches using Lattice-Boltzmann methods [Chen and Doolen, 1998; Martys and Hagedorn, 2002; Davis et al., 2003, 2004]. Koenders et al. [2000] present an analytical study of turbulent flow in a permeable filter layer.

All approaches to investigate flows in and above permeable walls, regardless if they are analytical, numerical or experimental, share the great complexity of the task, which is basically the geometric complexity of the underlying porous structure. It is highly random and disordered, making it infeasible to capture the exact small-scale pore structure (e.g. as boundary conditions in numerical simulations). Turbulence adds even more complexity to the problem, since it introduces dynamic processes acting on many different scales: from the effective pore diameter resp. the Kolmogorov scales to the dimension of the porous layer resp. dimension of macroscopic turbulent flow structures, which in turn scale with the dimensions of the channel flow, e.g. the water level. In such a situation, progress can only be made by introducing simplifications regarding the small-scale structures. Towards this end, the analysis of the physical system is focused on appropriate **length and time scales**. For example, in Large-Eddy-Simulations (LES) of turbulent flows, only the large-scale motions are resolved while the properties of the small-scale eddies are described by a *turbulence model* [Rodi, 1993]. Turbulence models assume some kind of relation between small-scale phenomena (e.g. the Reynolds stresses) and macroscopic quantities, e.g. the mean flow profile or its gradient. Similarly, macroscopic bulk properties of porous media, e.g. permeability or porosity, always represent an average over some suitably selected volume of the medium. Hence, any macroscopic description has a characteristic length scale associated with it. An important line of research on porous media is to find the relation between the properties of the medium on different scales, e.g. the relation of macroscopic permeability to the microscopic pore structure [Sahimi, 1995]. Comparing turbulent channel flow over smooth and impermeable walls with turbulent flow within and/or above a porous medium, additional characteristic length scales are present in the latter case (grain diameter, pore diameter, dimensions of the porous medium). A fundamental problem is to find out the relevant length and time scales governing the flow [Shimizu et al., 1990; Koenders et al., 2000].

To summarize, the analysis of complex disordered and random physical systems like turbulent flows and flows in porous media requires the introduction of physical models describing aspects of the system on particular length and time scales. These models are defined in terms of a set of scale-dependent intrinsic parameters which contain information about the structural and material properties of the system on the chosen scale. The models may be derived empirically or by theoretical approaches. In either case, they can only be verified on the basis of comparison with experimental data. The main goal of this thesis is the development of an appropriate experimental setup to obtain this data for the particular application outlined in Section 1.1.2.

Table 1.1: *Freight traffic in Germany in millions of tons per year (ADACmotorwelt, January 2003, p. 24).*

	rivers	rail	road
freight per year $[10^9 \ kg]$	242	294	982

1.1.2 River bed stability

The particular application motivating the research presented in this thesis is the study of river bed stability in hydraulic and geotechnical engineering. Both natural river beds and man-made bed protection structures (so-called *filter layers* or *geofilters*) typically consist of one ore more layers of porous material, such as gravel. Depending on the properties of the gravel (grain size distribution, porosity, permeability), a river bed represents a porous medium with more or less interconnected pores, through which fluid may flow. Thus, the flow in gravel-bed rivers and canals represents an important example of turbulent open-channel flow that is bounded by a permeable and rough wall below.

In flows over non-cohesive bed material like sand or gravel, the flow-induced forces acting on the upper grain layer may approach a certain stability threshold and initiate grain motion. Above the threshold, *erosion* resp. *sediment transport* occurs [Yalin, 1977; Thorne et al., 1987; Yalin, 1992; Clifford et al., 1993]. The purpose of coarse gravel filter layers is to protect the fine sediment layers below from erosion. The hydrodynamic load is reduced because of the drag introduced by the porous medium [Köhler, 2001]. Obviously, protection is only provided if the filter layer itself is stable.

Erosion of bed material may cause serious degradation of both natural environments and man-made hydraulic structures. The consequences are manifold, reaching from environmental and agricultural damages emerging gradually over long time periods to the complete failure of hydraulic structures in extreme situations like floods. An obvious example are dams protecting populated coastal areas. But also the enduring usage of natural rivers as waterways requires sustained maintenance work and protection measures for river beds, embankments and hydraulic constructions. For example, $170.000 \ m^3$ of gravel feeding per year is needed for the Iffezheim barrage at the river Rhein near Karlsruhe to avoid erosion in the downstream river bed. The costs amount to 5 mio. EUR per year[1].

The German waterway system has a total length of about 7500 km and is used by more than 2500 cargo-ships. An average cargo-ship is 80 m long, 10 m wide and carries about 1500 tons of freight (mainly building materials like stones, gravel or sand, salt, metals, coal or oil), which is equivalent to the freight of about 50 large lorries[2]. To give a concrete example, the river Neckar (which is one of the smaller rivers in Germany) is used as a waterway on a length of $\approx 200 \ km$ between the cities of Stuttgart and Mannheim. In the year 2002, about 9 mio. tons of goods have been transported on the Neckar[3], which is equivalent to the freight of about 500.000 large trucks. Table 1.1 compares the freight traffic in Germany on road, rail and rivers. Only a fraction of the total potential of freight transport on waterways is currently used. This fraction may be increased in the future, both for ecological and financial reasons (transport on waterways is cheaper than on rail or road and creates the least environmental pollution[4]). Hence, erosion protection is important for safety, economic and environmental reasons, and there is a growing need for reliable sediment

---

[1]Wasser- und Schifffahrtsamt Freiburg, 2004
[2]ADACmotorwelt, January 2003, p. 24
[3]Heilbronner Stimme, February 27th, 2003, p.22
[4]ADACmotorwelt, January 2003, p. 24

transport models which accurately predict the morphodynamic development of river beds under different loading scenarios [Jirka et al., 2001].

Since the beginning of the 20th century, research has been carried out to investigate the stability of river beds. Shields [1936] developed the concept of a **critical shear stress** τ_{oc} to describe the transition from a stable to a moving bed in the so-called *Shields-diagram*. This diagram basically relates the mean critical shear stress to the mean grain diameter of the granular material. Until the present day, the majority of design rules for filter layers are based on this 'classical work' or empirical modifications thereof. However, since the Shields approach is based on mean quantities (mean flow velocity and mean grain diameter), it is only valid for stationary flow conditions and uniform bed material. The highly time-dependent dynamical processes taking place in the turbulent boundary layer above actual river beds are not taken into account explicitly. As a result, there is an appreciable amount of scatter in experimental measurements of the Shields curve.

To enable improved, reliable predictions of destabilisation, the stochastic nature of both the loading forces and the geometric structure of a river bed has to be considered. A thorough understanding of the physical mechanisms causing incipient motion of sediment material is necessary. In particular, the influence of the turbulent flow field on the beginning of grain movement is a central question. A satisfactory, physically-founded description of river bed stability has to take into account the flow-induced fluctuating forces (lift and drag) acting on the grains [Dittrich et al., 1996; Dittrich, 1997; Dey et al., 1999]. To derive the probability distributions of the magnitude and frequency of these forces, information on

- the structure and dynamics of the **turbulent free surface flow** above the bed, including the influence of the bed roughness and permeability,

- the structure and dynamics of the **pore flow** within the bed,

- and the **interaction** of the free surface flow and the subsurface flow

is needed. This information can only be provided by experimental measurements. In spite of the fact that the morphodynamical evolution of a river takes place on rather large scales (kilometers resp. years), the stability problem requires a spatially and temporally highly resolved approach, because incipient grain motion is connected to the (turbulent) flow conditions both within and above the filter layer [Garcia et al., 1996; Sechet and Guennec, 1999; Zanke, 2001; Koll, 2002; Hofland, 2004; Vollmer, 2005].

Dancey et al. [2002] use the probability of individual grain motion as a criterion for the experimental characterization of the threshold of sediment motion. This threshold is specified by a fixed value of the probability. The criterion accounts for the statistical nature of sediment movement in turbulent flow, which is related to the typical time scale of the flow fluctuations. However, the latter is only based on a rough estimation (flow depth divided by average flow velocity). The authors point out the need for further experimental investigations of the relevant time scales for sediment entrainment at near threshold conditions on rough surfaces. For example, precise information about the time scales resp. frequency of occurence of burst events is needed.

Other particular open questions in this context are the following:

- In the near-wall boundary layer of the open-channel flow, the dominating pressure fluctuations are related to the square of the instantaneous near-bed velocity [Detert et al., 2004a; Hofland, 2005]. Is it possible to derive empirical relations between fluctuating velocities and pressure fluctuations within the filter layer?

- What is the influence of extra turbulence (generated by an obstacle or by surface waves) on the flow characteristics in the filter?

- It is obvious that the nature of fluctuations is different for an open-channel flow and for the confined flow that takes place within the gravel pores. In the pore flow, mixing mechanisms are partly intrinsically flow-related and partly the result of the erratic geometry. Can the latter effects be included in standard turbulence models, e.g. described by an eddy viscosity? What are the relevant scales?

1.2 Related work and own contribution

1.2.1 Work within the *Filter and Erosion Research Club*

The work presented in this thesis has been carried out within an international research cooperation called *Filter and Erosion Research Club* (FERC). FERC brings together different researchers (from universities, geotechnical and hydraulic engineering consultancy industry and governmental institutions) interested in physical mechanisms in the water-driven motion of granular matter in erosion processes[5]. Various disciplines are present in FERC: theoretical mechanics, hydromechanics, geotechnical and hydraulic engineering, experimental physics and image processing. The long-term goal of the activities within FERC is to find general criterions for the beginning of sediment motion, which would enable the specification of more focused filter design rules for hydraulic engineering and geotechnical applications.

In cooperation with the other FERC members, the Federal Waterways Engineering and Research Institute (Bundesanstalt für Wasserbau, BAW) in Karlsruhe has initiated two research projects:

1. FuE-Proj. Nr. 8123: Geotechnische Filter unter hydraulischer Belastung (Geotechnical filters under hydraulic load)

2. FuE-Proj. Nr. 8140: Stabilität der Sohle von Wasserstraßen (Bed stability of waterways)

In accordance with the studies pursued by FERC, these projects have the following four goals:

- improve the understanding of the interaction between turbulent free surface flow and subsurface flow,

- explain the physical mechanisms that are responsible for bed instability resp. the damping effect of a filter layer,

- develop and validate numerical models for flow over and through rough and permeable walls,

- in the long run: obtain more reliable sediment transport models and design criteria for hydraulic engineering and geotechnical applications, including the influence of turbulence.

To achieve these goals, an interdisciplinary approach is taken, which is based on theoretical, numerical and experimental research. Theoretical subsurface flow models are developed at Kingston University, e.g. [Koenders et al., 2000; Roussell et al., 2000]. These models are also combined with Lattice-Boltzmann

---

[5]The cooperating partners are Kingston University, UK, BAW Karlsruhe and University of Karlsruhe, Germany, Rijkswaterstaat DWW and Delft Geotechnics, The Netherlands.

simulations of the flow over and through a filter layer [Davis et al., 2003, 2004]. LES modeling of the flow over rough walls is carried out at Karlsruhe University [Stoesser et al., 2003, 2004].

The experimental part of the research consists of flow and pressure measurements in an experimental flume at the BAW Karlsruhe. The experiments are carried out in close cooperation with the BAW Karlsruhe and the Institute of Hydromechanics (IfH) of the University of Karlsruhe. A new technique to measure pressure fluctuations has been developed at the IfH [Detert et al., 2004a]. This thesis contributes the experimental technique to measure the flow within and above a gravel filter layer.

1.2.2 Related experimental approaches

Flow measurements in porous media.

Existing experimental approaches to measure flow in porous media are mainly based on x-ray tomography [Bayer, 2005], nuclear magnetic resonance tomography [Nakagawa et al., 1993; Ogawa et al., 2001], or refractive index matching techniques [Yarlagadda and Yoganathan, 1989; Saleh et al., 1992; Cui and Adrian, 1997; Stöhr, 2003]. Some applications of these techniques are briefly reviewed in section 2.8. While being able to reconstruct three-dimensional velocity fields, these techniques are limited to small-scale (sub-millimeter) flows with very low Reynolds numbers (Stokes flow). They are not feasible for large-scale flume experiments under transitional and turbulent flow conditions. However, in the experiments aspired here, large-scale tests are necessary to ensure both Reynolds- and Froude-similarity (especially for surface waves) and to enable a study of macro-turbulent flow structures. The latter scale with the flume dimensions and are considered to play an important role in the structure and dynamics of turbulence in gravel-bed rivers [Roy et al., 2004].

White et al. [2001] show a geotechnical application of a 2-D PIV system to measure soil deformations. This approach is similar to the method presented in section 9.2. However, only preliminary validation experiments are reported. The PIV images are obtained using a digital still camera without the use of endoscopes. Hofland [2004] uses a 2-D PIV setup to measure the flow structures in the near-wall boundary layer around a large stone embedded in a gravel layer.

Endoscopic flow measurements.

Endoscopic 2-D PIV is mainly applied to observe the flow fields in internal combustion engines [Gindele and Spicher, 1998; Dierksheide et al., 2001, 2002; Geis et al., 2002]. These applications require viewing through a small window and at the same time covering a large spatial region of the flow. Hence, rigid endoscopes with wide-angle lenses are used, similar to those described in section 9.2.1. Dierksheide et al. [2002] and Tani et al. [2002] point out the two main problems in quantitative endoscopic imaging:

1. **Transmission of the image through the endoscope.** The image transmission through the fiber bundle of a flexible endoscope resp. through the rod lens system of a rigid endoscope causes a **lower image resolution** and a **lower image contrast** due to a degraded modular transfer function (MTF) as compared to standard lenses.

2. **Imaging by wide-angle lenses with small focal lengths.** Rigid endoscopes usually have short focal lengths and a large field of view. **Radial lens distortions** may be so large that the usual paraxial

approximation of geometrical optics is not valid. Hence, images of rigid endoscopes can be similar to those produced by fish-eye lenses. The image of a planar object projected by such lenses appears on a curved surface. The whole image area cannot be focused equally well onto the planar sensor of the camera, and parts of the image are **out-of-focus**. This effect can also be seen in the calibration images of figure 5.6.

All the endoscopic applications that have been reported in the literature so far are restricted to the standard 2-D/2-C PIV configuration, i.e. they measure 2-component vectors within a plane in space. Reeves and Lawson [2004] propose the use of stereo endoscopic PIV to reduce errors related to perspective effects that are particulary strong in endoscopic images due to the short focal lengths. However, they only show a test implementation using a single camera equipped with an $f = 8\ mm$ micro video imaging lens. No real endoscopes are used, and the stereo setup is simulated by shifting the object plane on a translation stage. To the best of the author's knowledge, the only application of quantitative endoscopic 3-D imaging is presented by Mühlmann [2002]. He presents a method to reconstruct static 3-D surfaces from images obtained by a stereo endoscope, which is used for medical applications.

3-D Particle-Tracking Velocimetry.

While it still cannot be considered a standard method like PIV, many implementations of 3-D Particle-Tracking Velocimetry systems (3-D PTV) have been developed and successfully applied during the last 15 years. For a review, see section 3.5. Previous work on 2-D and 3-D PTV algorithms carried out at the Interdisciplinary Center for Scientific Computing in Heidelberg is summarized in section 7.1.

1.2.3   Own contribution and objectives

This thesis presents the following new contributions:

- endoscopic 3-D PTV, i.e. the first implementation of highly resolved quantitative 3-D measurements of a dynamic process by endoscopic stereo imaging,

- enhanced image processing algorithms for 3-D PTV adapted to the particular needs of endoscopic imagery,

- an experimental setup enabling a synoptic view of turbulent open-channel flow above and pore flow within a rough and permeable wall (gravel layer),

- results of the first simultaneous 3-D flow measurements within and above a gravel layer.

With regard to the expected measurement results, the FERC proposal for the research project initiated by the BAW states the following primary goals, see also figure 1.1:

- vertical profiles of mean velocity and fluctuations in the boundary layer of the open-channel flow,

- vertical profiles of mean velocity and fluctuations within the filter layer,

- detection of potential grain movement at the interfaces water/gravel and gravel/sand,
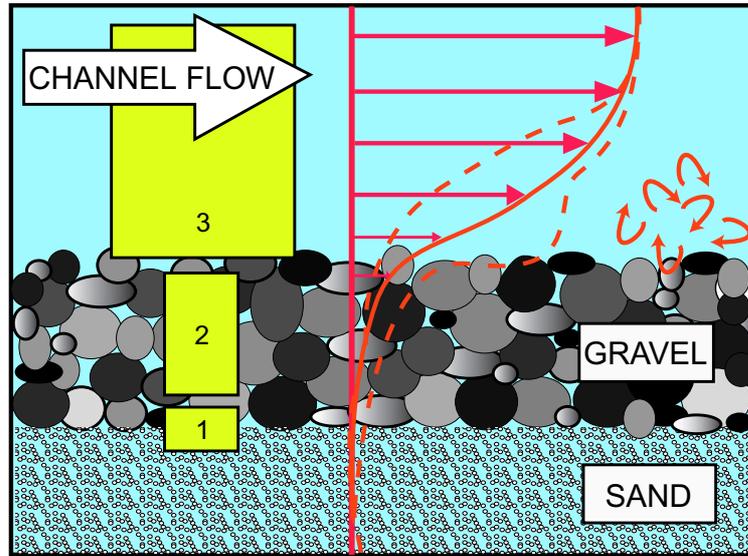
Figure 1.1: *Streamwise/wall-normal cross-section of the experimental observation area. The basic set-up under investigation is a horizontal gravel filter layer, a base layer of fine sand underneath it, and turbulent open-channel flow on top of the filter layer. The primary aim of the studies is to investigate the fluid velocity profiles through the filter layer as a result of the top flow characteristics (e.g. coherent turbulent eddy motions, indicated by the curved arrows). These profiles consist of average (continuous line) and fluctuations (dashed lines) as a function of the position in the filter. The experimental observation areas are indicated by the yellow boxes: (1) detection of sand motion, (2) endoscopic 3-D PTV, (3) 3-D PTV.*

- vertical profiles of pressure fluctuations,

- degree of fluctuations in the filter layer as a result of both uniform and non-uniform (extra turbulence due to waves or an obstacle) turbulent channel flow.

The major achievement of the thesis is the design of the endoscopic 3-D PTV system for pore flow measurements within the gravel layer (area 2 in figure 1.1). This includes both the experimental setup, the image processing algorithms, and a performance evaluation of the system. To achieve the goals stated by FERC, the experimental setup has been extended by another 3-D PTV system to observe the near-wall free surface flow (area 3 in figure 1.1) and additional endoscopic setups to observe the subsoil boundary layer and detect 2-D motion of sand grains (area 1 in figure 1.1). These additional setups are also presented in the thesis. The same 3-D PTV algorithm is used to process both the endoscopic images and the images of the free surface flow.

1.3   Thesis outline

The thesis is structured into three main parts: foundations (part I), digital image sequence analysis (part II), experiments and results (part III). Part I (chapters 2–3) is concerned with all the relevant background information on the subject. This includes both hydrodynamic resp. hydraulic engineering aspects and a review of quantitative flow visualization techniques. The development and extension of digital image processing techniques for endoscopic 3-D PTV is the topic of the second part (chapters 4–8). The algorithmic approaches

used in the different modules of the 3-D PTV are detailed and their performance with respect to accuracy, precision and efficiency is analyzed. Finally, the third part (chapters 9–10) covers the experimental setup for endoscopic 3-D PTV and its application within a series of investigations performed at the BAW Karlsruhe. The success of these experiments has been made possible by a careful re-design of the experimental setup used in previous investigations [Klar et al., 2002], together with the advances in image processing presented in part II. For the first time, results of 3-D flow measurements of the pore flow within a permeable wall, driven by turbulent open-channel flow on top, are presented.

In the following, the individual chapters of the thesis are outlined.

Part I: Foundations.

Chapter 2 reviews some basic theory of fluid flow, introducing the governing equations and some common notions that are used in the thesis. After the discussion of these general facts, the chapter specializes on some important types of flow, namely turbulent flow, open-channel flow, and flow in porous media. The chapter concludes by putting together all these components to end up with a description of the hydromechanic system that is studied in the thesis: turbulent open-channel flow above and pore flow within a rough and permeable wall.

Chapter 3 presents a survey of current state-of-the-art techniques for quantitative flow visualization, starting with some general considerations on motion estimation. First, 2-C/2-D methods like Particle Image Velocimetry (PIV) are explained. Extensions of these methods to measure three-component velocity vectors within a plane in space (3-C/2-D methods) or even within a volume in space (3-C/3-D methods) are briefly explained. A number of reasons are given, why PIV methods cannot be used for the measurement of flow within a gravel pore. We will see that another flow measurement technique, Particle-Tracking Velocimetry (PTV), is much better suited for this application. A specific implementation of 3-D PTV, adapted to the needs of endoscopic imaging, will be the tool that is used to obtain three-dimensional flow fields in part III of the thesis.

Part II: Digital image sequence analysis.

Part II of the thesis provides a detailed description of the algorithmic framework for three-dimensional Particle-Tracking Velocimetry that has been developed and implemented in this work. In the third part of the thesis, this framework is applied to compute three-dimensional flow trajectories, i.e. Lagrangian flow fields, from stereoscopic image sequences of tracer particles suspended in the flow. Towards this end, two prerequisites are necessary. First, in chapter 4 we take a look at the radiometric properties of the cameras that are used to acquire the image sequences. A method to correct for the strong fixed-pattern noise in the images is proposed. The efficiency of this method in reducing the noise level and thus improving subpixel accuracy is demonstrated.

The second prerequisite for the computation of 3-D trajectories is a geometric calibration of the stereo camera setup. It is the subject of chapter 5. In the proposed calibration module, the physical camera parameters are computed from images of feature points with known coordinates. The camera parameters define the relative orientation of the stereo setup and the perspective projections from 3-D space onto the image planes. Using this information, the 3-D coordinates of tracer particles are computed within a projective geometry framework. Under the difficult illumination conditions of endoscopic imaging, the probability for outliers

in the feature extraction is high. Thus, special attention is paid to achieve a robust calibration, i.e. one that computes the correct camera parameters also in the presence of data outliers.

Chapter 7 presents the 3-D PTV algorithm. The chapter starts with an overview of previous work, showing the capabilities as well as the limitations of former implementations. The following sections highlight the significant extensions that have been developed to make endoscopic particle-tracking feasible. The proposed algorithm consists of five modules, which are processed sequentially: image preprocessing, particle segmentation, particle tracking, stereoscopic correspondence analysis of 2-D trajectories and 3-D reconstruction. A separate section is devoted to each module. The first and second module are of particular importance: the particle segmentation and the tracking of the particles throughout an image sequence. The reason is that the probability for a unique 3-D reconstruction of a trajectory increases with the trajectory length. Thus, in the ideal case, the particles should be tracked over their complete length, i.e. starting when they enter and finishing when they leave the field of view. While previous implementations have been prone to segmentation failures, the new segmentation module achieves a much more robust segmentation by a combination of two steps. The first step is capable of segmenting even very faint particles, based on their feature 'motion'. Only a statistical significance level has to be specified by the user. A local segmentation threshold is computed, taking into account the results of the camera noise analysis in chapter 4. The second step is a separation of overlapping particle images by a watershed algorithm. The tracking module has also been completely redesigned. Within a Kalman filter framework, information from different low level motion estimators (temporal smoothing, optical flow and image cross-correlation) is combined to yield the optimal particle trajectory in a maximum likelihood sense.

The second part of the thesis is concluded with a performance evaluation of the 3-D PTV in chapter 8. The performance is tested both on synthetic and real image sequences. The analysis shows that in practical applications, typically 50% of the segmented particles can be reconstructed in 3-D with a relative velocity error below 10%.

Part III: Experiments and results.

The third part of the thesis is devoted to the experiments performed in a test flume at the BAW in Karlsruhe. In this flume, a sand layer is covered by a gravel layer. The flow in these sediment layers is driven by open-channel flow on top of the gravel layer. Chapter 9 presents the experimental setup. The setup enables the simultaneous acquisition of image sequences within the gravel layer and at its interfaces with the sand layer and the free surface flow. Several subsystems are used to achieve this, which are discussed in detail. New digital high-speed CMOS cameras are used for the image acquisition, which were among the first that have been commercially available. A major achievement was the realtime storage of the data streams from the cameras on RAID systems (requiring to write a sustained data rate of about $25MB/s$ to the disks). This is necessary to enable the acquisition of long image sequences for a statistical analysis of flow events.

Chapter 10 describes the application of the developed experimental setup in an extensive series of systematic measurements. These measurements have been carried out in cooperation with the BAW and the Institute for Hydromechanics of the University of Karlsruhe. For the first time, profiles of 3-D mean and fluctuating velocities within and above a permeable and rough wall, driven by turbulent open-channel flow, are presented.

The thesis concludes with a summary and an outlook on further research in chapter 11.

Part I

Foundations

Chapter 2

Hydrodynamic background

This chapter introduces the basic hydromechanic concepts and terminology that is relevant for this work. In section 2.1 we start with a description of the defining properties of the medium we are interested in, namely a **Newtonian fluid**. This definition also introduces the important notion of a **shear stress** (flow induced shear stresses are the driving forces for sediment motion). Section 2.2 continues with the basic equations, notions and parameters that describe different types of fluid flow. These equations may be formulated in different reference frames. In section 2.3 we show the two basic frameworks that are used to describe flowing fluids, the **Eulerian** and the **Lagrangian** flow field. While most of the common flow visualization techniques, in particular PIV (section 3.2), measure the Eulerian flow field, the 3-D PTV developed here yields the Lagrangian representation. Many geophysical and technical flows are turbulent. Thus, some important aspects of **fluid turbulence** are discussed in section 2.4. In the next two sections, we look at the two specific types of flow that are investigated in the experiments presented in this work, namely **open-channel flow** (section 2.5) and **flow in a porous medium** (section 2.6). Finally, section 2.7 combines the previous considerations and highlights some aspects of turbulent open-channel flow over a rough and permeable wall. Section 2.8 contains a compilation of some experimental approaches to quantitative flow measurements in porous media and over permeable walls. A summary of the most important facts of this chapter is given in section 2.9.

2.1 Newtonian fluids

We consider the flow of a Newtonian fluid, i.e. a fluid in which the viscous momentum flux (i.e. the **shear stress**) τ_{xy} across the xy-plane is proportional to the velocity gradient in the direction perpendicular to this plane. The constant of proportionality is the fluid *viscosity* μ. If the flow is in x-direction, the mathematical definition reads

$$\tau_{xy} = -\mu \frac{\partial v_x}{\partial y}. \tag{2.1}$$

Bird et al. [1960] give a useful interpretation of (2.1): the viscous momentum flux is in the direction of the negative velocity gradient. Thus, momentum tends to go in the direction of decreasing velocity. In analogy, the same behaviour is found e.g. in heat conduction, where heat flows from a hot region towards a colder one. In this case, the temperature gradient is responsible for the heat flux. Similarly, in a viscous fluid, a velocity gradient is the driving force for the transport of momentum.

2.2 Flow of fluids

2.2.1 Basic equations

The basic equations governing the motion of Newtonian fluids are the **Navier-Stokes equations**. They are derived from the balance of forces of a fluid element, using Newton's laws of motion, see e.g. [Tritton, 1977]. Together with the equation of continuity, these equations describe the conservation of mass and momentum of a moving fluid. If other processes are present, e.g. thermodynamic phenomena (heat conduction, convection, radiation, chemical reactions), the mathematical description is extended by a similar equation describing conservation of energy, see e.g. [Bird et al., 1960].

In the following, $\boldsymbol{x} = [x, y, z]^T$ is the position, t is the time, $\boldsymbol{v}(\boldsymbol{x}, t) = [v_x(\boldsymbol{x}, t), v_y(\boldsymbol{x}, t), v_z(\boldsymbol{x}, t)]^T$ is the flow velocity[1], $p(\boldsymbol{x}, t)$ the pressure, ρ the density, $\boldsymbol{g}$ the gravitational acceleration, μ the viscosity and $\nu = \mu/\rho$ the kinematic viscosity. We assume that the density and the temperature are constant throughout the flow. The Navier-Stokes equations and the continuity equation for incompressible, isothermal flow are

$$\frac{\partial \boldsymbol{v}}{\partial t} + (\boldsymbol{v} \cdot \nabla)\boldsymbol{v} = -\frac{1}{\rho}\nabla p + \nu \nabla^2 \boldsymbol{v} + \boldsymbol{g} + \frac{1}{\rho}\boldsymbol{F} \,, \tag{2.2}$$

$$\nabla \cdot \boldsymbol{v} = 0 \,. \tag{2.3}$$

Equation (2.2) describes the temporal change of the velocity field due to inertia, pressure, viscous, gravitational and external forces, with the corresponding terms read from left to right. $\boldsymbol{F}$ is a body force term (force per unit volume) containing external forces as e.g. a periodic forcing in a grid-stirred tank or in flows driven electromagnetically by periodic magnetic fields. Since in the type of flow studied in this work (open-channel flow, see section 2.5), gravity and pressure gradients are the only driving forces , we set $\boldsymbol{F} = 0$ in the following. Together with appropriate boundary conditions, equations (2.2) and (2.3) are a set of four coupled non-linear partial differential equations for the four unknown variables $\boldsymbol{v}(\boldsymbol{x}, t)$ and $p(\boldsymbol{x}, t)$.

2.2.2 Nondimensional parameters

Depending on the relative strength of inertial and viscous forces, the Navier-Stokes equations may be simplified, which results in two particular types of fluid flow, see section 2.2.3. A non-dimensional parameter that characterizes this proportion is the **Reynolds number** Re. If a typical velocity scale V_0 and a typical length scale L_0 are introduced for a given flow, the corresponding Reynolds number is

$$Re = \frac{V_0 L_0}{\nu} \,. \tag{2.4}$$

The Reynolds number is a measure of the relative strength of inertia and viscous forces. This can be seen from an estimation of the corresponding terms of (2.2) using the introduced scales: $\boldsymbol{v} \cdot \nabla \boldsymbol{v} \sim V_0^2/L_0$, $\nu \nabla^2 \boldsymbol{v} \sim \nu V_0/L_0{}^2$. The ratio of these two terms is the Reynolds number. Introducing non-dimensional quantities $\tilde{\mathbf{v}} = \boldsymbol{v}/V_0$, $\tilde{\mathbf{x}} = \boldsymbol{x}/L_0$, $\tilde{t} = tV_0/L_0$, and $\tilde{p} = p/(\rho V_0{}^2)$, the non-dimensional Navier-Stokes equations read

$$\frac{\partial \tilde{\mathbf{v}}}{\partial \tilde{t}} + (\tilde{\mathbf{v}} \cdot \tilde{\nabla})\tilde{\mathbf{v}} = -\tilde{\nabla}\tilde{p} + \frac{1}{Re}\tilde{\nabla}^2\tilde{\mathbf{v}} + \frac{1}{Fr}\boldsymbol{g}/g \,, \tag{2.5}$$

---

[1]Sometimes the equivalent notations $\boldsymbol{v} = [v_1, v_2, v_3]^T$ or $\boldsymbol{v} = [u, v, w]^T$ are more appropriate. We use these notations interchangeably in this chapter.

$$\tilde{\nabla} \cdot \tilde{\mathbf{v}} = 0 \,, \tag{2.6}$$

where

$$Fr = {V_0}^2/(L_0 g) \tag{2.7}$$

is the **Froude number** that quantifies the relative importance of inertia and gravity forces. The Froude number is sometimes also defined as $Fr_2 = \frac{V_0}{\sqrt{gh}}$, where h is the water depth. In this case, the Froude number is the ratio of the flow velocity to the speed of shallow water waves. The influence of gravity forces is decreasing with increasing Froude numbers. If the Froude number Fr_2 is smaller than one, the flow is called *subcritical*, otherwise it is called *supercritical*.

2.2.3 Basic types of flow

We briefly look at some important particular types of flow: **viscous flow**, **inviscid flow** and **potential flow**. For these flows, the Navier-Stokes equations may be simplified, based on the assumption of very small/large Reynolds numbers or zero vorticity. Obviously, such assumptions are not strictly valid for real flows, so the resulting ideal flows can only approximate real flows to some degree. Nevertheless, these concepts are also important in practical applications, since most real flows can be partitioned into areas of viscid and inviscid flow, separated by a so-called **boundary layer**. A similar separation is possible between rotational and irrotational flow areas. The concept of a boundary layer is discussed at the end of this section.

Viscous flow: $Re << 1$**.**

If the Reynolds number is small, viscous forces are dominating, resulting in a quite regular, steady flow (**laminar** flow). In this case, the inertia term in (2.2) may be dropped. The resulting equations are the (time-dependent) **Stokes equations** describing the slow flow of a viscous fluid ('creeping flow'):

$$\frac{\partial \boldsymbol{v}}{\partial t} = -\frac{1}{\rho}\nabla p + \nu\nabla^2\boldsymbol{v} + \mathbf{g}. \tag{2.8}$$

If in addition the flow is stationary, we have

$$0 = -\frac{1}{\rho}\nabla p + \nu\nabla^2\boldsymbol{v} + \mathbf{g}. \tag{2.9}$$

Equation (2.9) is also referred to as the Stokes equation. It is widely applied to describe slow flow through porous media [Sahimi, 1995].

Inviscid flow: $Re >> 1$**.**

In the limit of very large Reynolds number, the viscous term in (2.2) can be neglected (see (2.5)). The resulting flow is called **inviscid flow** or **Euler flow**, governed by the **Euler equation**:

$$\frac{\partial \boldsymbol{v}}{\partial t} + (\boldsymbol{v} \cdot \nabla)\boldsymbol{v} = -\frac{1}{\rho}\nabla p + \mathbf{g}. \tag{2.10}$$

Since viscosity is assumed to be zero, fluid undergoing Euler flow cannot sustain shear stress. Therefore, the pressure within it is isotropic everywhere. In steady flows, the relation between the pressure and the fluid velocity is given by **Bernoulli's theorem**, which states that

$$\frac{p}{\rho} + \frac{1}{2}v^2 = const. \tag{2.11}$$

along a line of flow. If in addition the flow is irrotational, the constant is the same for all lines of flow. Bernoulli's theorem is basically an expression of the conservation of energy in inviscid, incompressible, steady flows.

Potential flow, vorticity and Kelvin's theorem.

Potential flows are flows where the **vorticity**

$$\boldsymbol{\Omega} = \nabla \times \boldsymbol{v} \tag{2.12}$$

is zero everywhere throughout the flow. Since vorticity is associated with local rotation of the fluid, potential flows are also called *irrotational*. Because of their defining property, potential flow fields may be derived from a scalar **flow potential** ϕ by

$$\boldsymbol{v} = \pm \nabla \phi \tag{2.13}$$

If the flow is also incompressible ($\nabla \cdot \boldsymbol{v} = 0$), the potential satisfies the **Laplace equation**

$$\nabla^2 \phi = 0. \tag{2.14}$$

Scalar potentials that are solutions of the Laplace equation have been used for a long time to describe static magnetic fields in free space, hence many analytical and numerical methods are known to solve the Laplace equation subject to given boundary conditions. These methods can also be used to solve corresponding problems in fluid dynamics [Faber, 1995]. For example, Koenders et al. [2000] and Vollmer et al. [2002] use flow potentials to describe the flow within a gravel filter layer. Note that the theory of potential flow may be applied to both viscous and inviscid fluids. The only conditions that must be valid are zero vorticity $\nabla \times \boldsymbol{v} = 0$ and incompressibility $\nabla \cdot \boldsymbol{v} = 0$. In this case, the viscous term in the Navier-Stokes equation is also zero, since $\nabla \times (\nabla \times \boldsymbol{v}) = \nabla (\nabla \cdot \boldsymbol{v}) - \nabla^2 \boldsymbol{v}$.

If the vorticity is zero in a fluid of zero viscosity at one particular instant of time, then **Kelvin's circulation theorem** states that the fluid remains free of vorticity at all later times, both in steady and unsteady flows. For a proof of this theorem, which is valid for both incompressible and compressible flows, see e.g. [Faber, 1995]. The theorem is of rather theoretical importance, since there are many practical applications that violate it, for example the flow around an airfoil. The latter is steady in the beginning and becomes turbulent as the flow velocity is increased. The reason is that Kelvin's theorem does not take into account the effects that are caused by the boundaries of the fluid. In any practical application, fluids are bounded by solid walls or the flow is around some solid bodies. The flow within a permeable wall is a combination of both. The next subsection shows that in such **boundary layers** that develop along fluid-solid interfaces, viscosity may not be neglected and thus the assumption of potential flow breaks down.

Fluid-solid interfaces and boundary layers.

The so-called **no-slip boundary condition** states that where fluids meet solids, both must move at the same velocity, i.e. there can be no slip between them [Faber, 1995]. It has been verified in many experiments. In the case of flow along a solid, stationary boundary, as e.g. open-channel flow over a smooth wall, the no-slip boundary condition requires the flow velocity to drop to zero at the wall. Hence, there must be a near-wall region in the fluid where it is retarded and the flow velocity becomes small, so that viscous effects become important. This idea was first formulated by Prandtl [Schlichting and Gersten, 1997], who introduced the

notion of a **boundary layer**: in a potential flow over a solid surface, vorticity is created within a boundary layer next to the surface. In flows around solid obstacles, the boundary layer may separate from the obstacle (*flow separation*), and the vorticity generated in the boundary layer gets swept downstream in the *wake* created by the obstacle. Flow separation behind obstacles is one source of turbulent velocity and pressure fluctuations in an open-channel flow over a rough wall, where roughness elements (e.g. larger stones) create wakes downstream of them [Dittrich, 1997].

2.3 Description of fluid flow fields

The motion fields of flowing liquids may be described mathematically in two different ways. The first and most common way is to regard the physical variables describing the state of the liquid (velocity $\boldsymbol{v}$, pressure p, density ρ and temperature T) as functions of a 3-D coordinate vector $\boldsymbol{x}$ and time t. The resulting flow field is the **Eulerian flow field $\boldsymbol{v}_e$**:

$$\boldsymbol{v}_e = \boldsymbol{v}(\boldsymbol{x}, t). \tag{2.15}$$

The most common experimental approaches to flow measurement, Particle Image Velocimetry (see section 3.2) and Laser Doppler Anemometry (see e.g. [Lading et al., 1994]), provide velocity information within a fixed reference frame in space resp. at a fixed position in space and thus yield Eulerian flow velocities. In a Eulerian flow field, a **streamline** is defined as a line that is tangential to the flow vectors at a fixed time. Thus, if a streamline $\boldsymbol{x}(s, t = t_0)$ is parameterized by the parameter s at a fixed time $t = t_0$, it is defined by the equation

$$\frac{d\boldsymbol{x}}{ds} = \boldsymbol{v}_e(\boldsymbol{x}(s), t = t_0). \tag{2.16}$$

The second way to describe a flow field identifies a particular fluid element (a 'fluid particle') at position $\boldsymbol{x}_0 = \boldsymbol{x}(t = t_0)$ and time t_0 and follows this particle along its flow trajectory. In this case, flow velocities are given as time derivatives of the particle positions $\boldsymbol{x}_l = \boldsymbol{x}(\boldsymbol{x}_0, t - t_0)$, which corresponds to the **Lagrangian flow field $\boldsymbol{v}_l$**:

$$\boldsymbol{v}_l = \boldsymbol{v}(\boldsymbol{x}_0, t - t_0) = \frac{\partial}{\partial t} \boldsymbol{x}(\boldsymbol{x}_0, t - t_0). \tag{2.17}$$

A Lagrangian flow trajectory (a *pathline*) $\boldsymbol{x}(t)$ is a solution of the differential equation

$$\frac{d\boldsymbol{x}}{dt} = \boldsymbol{v}_e(\boldsymbol{x}(t), t). \tag{2.18}$$

Note the difference between streamlines (2.16) ('snapshot' of the flow field, t fixed) and pathlines (2.18) (time-dependent particle trajectories).

2.4 Turbulence

For higher Reynolds numbers, the smoothing influence of viscosity becomes weaker. At some critical Reynolds number depending on the flow under consideration, a transition to another flow regime takes place, which is unsteady and highly random in space and time: **turbulent** flow. Almost all geophysical flows as well as flows occuring in hydraulic engineering are turbulent. Turbulent flows have extremely complex spatio-temporal dynamics, with many motion patterns on a wide range of scales (vortices, 'eddies') super-imposed on each other and interacting. The reason for this rich structure is that at high Reynolds numbers
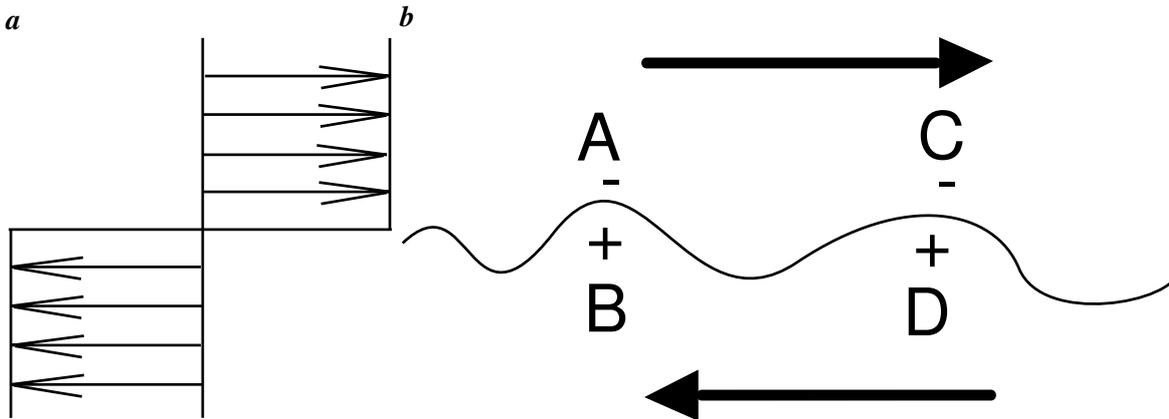
Figure 2.1: ***a*** *Idealized shear flow profile with discontinuity surface.* ***b*** *Disturbed discontinuity surface. +/- signs indicate high/low pressure.*

the nonlinear inertial terms dominate the Navier-Stokes equations (see (2.5)). The nonlinearity leads to different kinds of instabilities and a coupling of flow structures on different scales. Due to its complexity, fluid turbulence has been and still is the subject of extensive research [Hinze, 1975; Townsend, 1976; Frisch, 1995]. In this section, some important aspects of fluid turbulence are reviewed. Parts of the following sections are taken from [Klar, 2001].

2.4.1   Shear flow instability and transition to turbulence

In principle, for stationary boundary conditions, exact *steady* solutions of the Navier-Stokes equations should exist for arbitrary Reynolds numbers [Landau and Lifschitz, 1991]. But the formal mathematical existence is not sufficient for these solutions to appear in nature. A further necessary condition is that the solutions have to be stable against small perturbations, which are always present in real flows. If such perturbations grow in amplitude, the flow is unstable, and transition to another flow regime occurs, which may again be unstable. Hence, a series of transitions may take place, finally resulting in fully developed fluid turbulence, i.e. a turbulent flow field containing the full range of eddy scales with random behaviour in time and space. The different instability stages are connected to critical Reynolds numbers which are highly dependent on the details of the flow (initial and boundary conditions, geometry, surface roughness, level of preexisting disturbances in the flow). A well-known example of such a series of transitions is the 'Karman vortex street', which develops due to flow separation in the flow around an obstacle. With increasing flow velocity, several transitions occur: from laminar, stationary flow to quasi-periodic flow, intermittent flow with turbulent spots and finally fully turbulent flow [Merzkirch, 1987].

There is a number of hydrodynamical instability mechanisms that occur in different physical flow situations. Important in the present context is the instability of **shear flows**, the so-called **Kelvin-Helmholtz instability**. A shear flow is a flow in which the velocity varies principally in a direction at right angles to the flow direction, like e.g. open-channel flow. The Kelvin-Helmholtz instability may be explained using an idealized example of a flow with a finite discontinuity in the velocity profile [Tritton, 1977]. Such a flow is shown in figure 2.1a in the reference frame where the two velocities are equal and opposite.

We now suppose that a wavy disturbance is imposed on the discontinuity surface separating the opposed velocities, figure 2.1b. This means that the fluid on the convex sides (A,C) will move slightly faster than

on the concave sides (B,D). The indicated pressure changes follow from Bernoulli's equation (2.11), i.e. a pressure decrease where the fluid moves faster and vice versa. Thus, the disturbances will be amplified by the pressure forces. The flow is unstable.

A discontinuity surface is also referred to as a **vortex sheet** [Faber, 1995], because the vorticity $\mathbf{\Omega}$ is infinite on such surfaces. Of course, in reality, vortex sheets do not exist. Discontinuities are smoothed by viscosity, and vortex sheets become rather **vortex layers** or **shear layers**. Nevertheless, the above considerations deliver insight into the basic mechanisms. For example, boundary layers separating from smooth walls or from single grains of a rough gravel surface are examples of vortex layers.

A more precise mathematical treatment shows, that instability is coupled to the existence of an *inflectional point* in the velocity profile, see e.g. [Kundu, 1990]. Recent numerical research by Breugem [2004] suggests that turbulence near a (highly) permeable wall is dominated by vortical structures that originate possibly from a Kelvin-Helmholtz instability of the inflexional mean velocity profile.

2.4.2 Statistical approach

The irregularity and complexity of turbulent motion makes it impossible to describe turbulent flows in terms of their instantaneous and local flow and pressure fields resp. functions thereof. Both theoretically and experimentally, it is only practicable to consider mean values of functions of the instantaneous and local values of velocity and pressure. Hence, a **statistical approach** is necessary.

The statistical description of turbulence starts with the so-called **Reynolds decomposition** of the pressure and velocity fields into mean values and fluctuating parts:

$$p = \bar{p} + p', \; u = \bar{u} + u', \; v = \bar{v} + v', \; w = \bar{w} + w' \, , \tag{2.19}$$

where the bar indicates temporal average values and the primed quantities are the fluctuations. In this context, it is more convenient to use suffix notation for vectors, therefore the velocity is written as $v = [u, v, w]^T = [v_1, v_2, v_3]^T$ in the rest of this section.

In general, the mean values are ensemble averages, i.e. they are calculated by averaging over a large enough number of identical experiments. If the mean values are not time-dependent, i.e. for stationary statistical processes, the ensemble average may be replaced by a time average. Since in stationary open-channel flow ergodicity may also be assumed in the streamwise and spanwise direction, the averaging may also be carried out spatially over horizontal planes. This corresponds to the so-called **double-averaging** of the Navier-Stokes equations that is frequently applied in investigations of open-channel flow [Koll, 2002].

By substituting (2.19) into the continuity and Navier-Stokes equations, again carrying out the averaging process, and keeping in mind that, by definition, $\bar{u'} = \bar{v'} = \bar{w'} = \bar{p'} = 0$, one ends up with the **Reynolds equations** for a turbulent flow:

$$\bar{v}_j \frac{\partial \bar{v}_i}{\partial x_j} = -\frac{1}{\rho} \frac{\partial \bar{p}}{\partial x_i} + \nu \frac{\partial^2 \bar{v}_i}{\partial x_j^2} - \frac{\partial}{\partial x_j} (\overline{v_i' v_j'}) \, . \tag{2.20}$$

These equations differ from the Navier-Stokes equations (2.2) in the last term on the right handside, which is the gradient of the so-called **Reynolds stress tensor**

$$\tau_{ij} = -\rho (\overline{v_i' v_j'}) \, . \tag{2.21}$$

This tensor contains additional stresses exerted on the mean flow, and the gradient of this tensor represents additional accelerations of the fluid as compared to laminar flow. The origin of the Reynolds stresses is turbulence. Mathematically, the stresses are correlations of the fluctuating velocity components.

The diagonal elements of the Reynolds stress tensor are used to indicate the degree of turbulence of the flow. The square roots of these elements divided by the density, i.e. basically the standard deviations of the mean velocities, are called the **turbulence intensities** $v_{rms,i}$ ('root-mean-square' velocities).

The Reynolds stresses work against the mean velocity gradient and therefore transfer energy from the mean flow to the fluctuating field. They produce the turbulent kinetic energy maintaining the turbulent eddy motions. The eddies are steadily transferred to smaller and smaller scales (**eddy cascade**), and finally their energy is dissipated into heat by viscous forces, e.g. [Frisch, 1995].

2.5   Open-channel flow

In this section, some important features of open-channel flows are reviewed. More details can be found in [Nezu and Nakagawa, 1993].

2.5.1   Basic equations and parameters

Coordinate system.

The following coordinate system will be used. The origin in cross-stream direction is in the middle of the channel, the vertical zero-level is on the top of the gravel layer.

coordinate	velocity component	direction
x	u	streamwise, pointing in flow direction
y	v	vertical, pointing upwards
z	w	spanwise, pointing to the right

2-D open-channel flow.

Open-channel flow can be considered two-dimensional, i.e. $\frac{\partial}{\partial z} = 0$, if the ratio of channel width B to flow depth h exceeds a certain threshold α_c. Nezu and Nakagawa [1993] classify open-channel flows into two categories:

$$\textit{narrow open channels: } B/h \leq \alpha_c$$
$$\textit{wide open channels: } B/h > \alpha_c$$

with $\alpha_c \approx 5$ or slightly larger. In narrow open channels, side-wall and free-surface effects generate secondary currents, whereas in wide open channels, these effects can be neglected in most cases (at least in the central zone of the channel).

The following considerations are valid for a 2-D flow in a wide open channel. Further, $\bar{v} = \bar{w} = 0$ is assumed, and the mean velocity $\bar{u}$ is taken as a function of the vertical coordinate only: $\bar{u} = \bar{u}(y)$.

The logarithmic wall law.

The logarithmic wall law describes the vertical distribution of the mean velocity in the wall region ($y/h \leq 0.2$) of open-channel flow. It can be derived from the Reynolds equations using some further assumptions (a turbulence model), e.g. [Nezu and Nakagawa, 1993]. Landau and Lifschitz [1991] give a very concise derivation using simple physical arguments and dimensional analysis:
Open-channel flow is a quasi-steady turbulent shear flow, which means that

- the mean flow quantities are independent of time,

- the Reynolds number is high, thus viscosity may be neglected, except for a very thin layer near the wall (the **viscous sublayer**),

- there is a velocity gradient $d\bar{u}/dy$ in the vertical direction.

The velocity gradient causes a momentum flow through the fluid in the vertical direction. This momentum flow is equivalent to a friction force per unit area acting on the wall, the so-called **wall** or **bed shear stress** τ_0, see also section 2.1.

We are looking for a relation between the wall shear stress τ_0 and its cause, the velocity gradient. The only possible quantities that can enter this relation are the density ρ, the wall shear stress τ_0 and the vertical distance y (since viscosity is neglected). There is only one combination of these quantities with the dimension of a velocity gradient, namely $\sqrt{\tau_0/\rho}/y$. Therefore, a relation of the form

$$\frac{d\bar{u}}{dy} = \frac{1}{\kappa} \frac{\sqrt{\tau_0/\rho}}{y} \ , \tag{2.22}$$

must hold, where κ is a dimensionless constant, the so-called **von Karman constant**. The quantity $\sqrt{\tau_0/\rho}$ is defined as the **friction velocity** or **shear velocity** u_*:

$$u_* = \sqrt{\tau_0/\rho} \ . \tag{2.23}$$

The shear velocity is a measure for the typical velocity scales in the wall region of open-channel flow. The von Karman constant has to be determined experimentally and was found to be

$$\kappa = 0.4 \ . \tag{2.24}$$

Integration of (2.22) yields the logarithmic wall law

$$\bar{u} = \frac{1}{\kappa} u_* (ln(y) + c) \ , \tag{2.25}$$

where c is an integration constant. The determination of this integration constant requires some further physical reasoning. One of the prerequisites of the derivation so far was the disregard of viscosity. This is a valid assumption for the major area of the flow field, except for a thin layer near the wall, the so-called **viscous sublayer**. In this layer of thickness $\sim y_0$, viscosity plays a dominant role and cannot be neglected. The flow field in this layer is not dominated by Reynolds stresses but by viscous stresses. Thus, the logarithmic wall law is not valid for $y \leq y_0$. Instead, at $y \sim y_0$ it has to be matched to a different equation describing the velocity profile in the viscous sublayer.

The magnitude of y_0 may be determined in the following way. In the viscous sublayer, the characteristic length scale is $\sim y_0$, and the characteristic velocity scale is $\sim u_*$. Thus, the Reynolds number governing the flow at these scales is $Re_{sublayer} \sim y_0 u_*/\nu$. As explained in section 2.2.3, viscosity becomes important for Reynolds numbers of the order of unity and smaller. Therefore, from $Re_{sublayer} \sim 1$ one can determine the size of y_0:

$$y_0 \sim \frac{\nu}{u_*} \ . \tag{2.26}$$

y_0 is called the **inner length scale** or **viscous wall unit**, in contrast to the **outer length scale**, which is of the order of the flow depth.

The integration constant c has to be chosen in such a way that $\bar{u} \sim u_*$ for $y \sim y_0$, which yields $c = -ln(y_0)$. By introducing the dimensionless vertical coordinate (normalized by the inner length)

$$y^+ = \frac{y}{\nu/u_*} \ , \tag{2.27}$$

the logarithmic wall law may be re-written as

$$\frac{\bar{u}}{u_*} = \frac{1}{\kappa} ln(y^+) \ . \tag{2.28}$$

Experimental results have shown that the logarithmic wall law yields a more precise description of measurement data if empirical constants are added [Nezu and Nakagawa, 1993], e.g.

$$\frac{\bar{u}}{u_*} = \frac{1}{\kappa} ln(y^+) + 5.1 \ . \tag{2.29}$$

The vertical **shear stress profile** in turbulent open-channel flow can be derived from the Reynolds equations (2.20) and reads

$$\frac{\tau(y)}{\rho} = \frac{\tau_0}{\rho}\left(1 - \frac{y}{h}\right) = u_*^2\left(1 - \frac{y}{h}\right) = \nu\frac{d\bar{u}}{dy} - \overline{u'v'}. \tag{2.30}$$

The details of this derivation can be found e.g. in [Koll, 2002]. The first term on the right handside is the **viscous shear stress** τ_{vis},

$$\tau_{vis}/\rho = \nu\frac{d\bar{u}}{dy}, \tag{2.31}$$

compare section 2.1. The second term is the **turbulent shear stress** τ_{turb}, compare (2.21). The so-called 'closure problem' in turbulence modeling is to find a relation between the turbulent shear stresses and the mean flow velocities. A simple approach (the so-called *hypothesis of Boussinesq* [Schlichting and Gersten, 1997]) is to relate the turbulent shear stress to the gradient of the mean flow in the same way as in the viscous case (2.31):

$$\tau_{turb}/\rho = \nu_{turb}\frac{d\bar{u}}{dy}. \tag{2.32}$$

In (2.32), ν_{turb} is the so-called **eddy viscosity** or **Boussinesq viscosity**. A basic problem of turbulence modeling is now the modeling of ν_{turb} [Rodi, 1993], in particular within porous media, compare section 2.7.2 and [Shimizu et al., 1990; Hoffmans et al., 2000; Lesage et al., 2004].
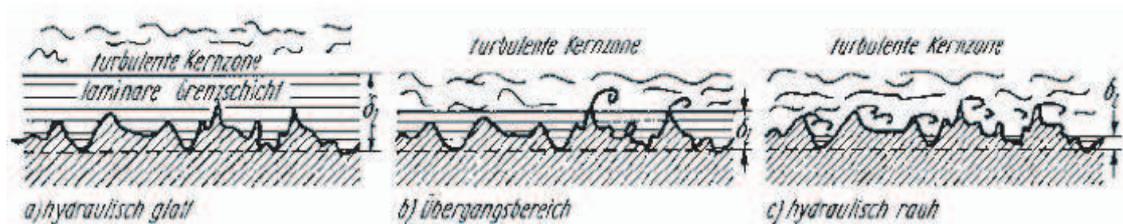
Figure 2.2: *Classification of flows by their hydraulic roughness. The images (from left to right) correspond to hydraulically smooth, transitional and fully rough beds. The viscous sublayer is indicated by straight lines. From [Bollrich, 1996].*

2.5.2 Influence of rough walls

All relations discussed so far have been derived for smooth open channels. However, most of them can also be used in a modified form for flows over rough surfaces. To establish the necessary modifications, one has to find a suitable parameter to represent the size of the roughness elements. Further, it is no longer clear where the zero-level of the vertical velocity profile ($y = 0$) is located. One has to apply some method of calculating a theoretical zero-level, which provides the best fit to experimental data. Most researchers find that the zero-level has to be shifted downwards from the top of the roughness elements by a value of $y' \approx (0.15 - 0.3)k$, where k is the height of the roughness elements [Nezu and Nakagawa, 1993; Dittrich, 1997].

Characterization of roughness.

Different parameters for roughness characterization can be found in the hydraulic engineering literature. The simplest one is the absolute height of the roughness elements k. The so-called **equivalent sand roughness** k_s is also frequently used. It was introduced in pipe flow experiments by Nikuradse [1933]. For rough beds consisting of one layer of closely-packed uniform sand grains, the equivalent sand roughness is equal to the absolute roughness height of the sand grains. For different roughness configurations, k_s has to be determined empirically or calculated theoretically, e.g. using the method described below. In contrast to the absolute roughness height k, which is a geometric quantity, the equivalent sand roughness reflects the *influence* of the roughness elements on the flow. This influence depends not only on the geometric roughness height, but also on the density, shape and surface properties of the roughness elements.

Flows over rough beds are further characterized by the so-called **roughness Reynolds number**, which is a measure of the **hydraulic roughness** of a flow over a rough surface. It is defined as

$$Re_* = \frac{ku_*}{\nu} = \frac{k}{y_0} \ , \tag{2.33}$$

which is the ratio of the absolute roughness height to the thickness of the viscous sublayer. Two important cases can be distinguished, see also figure 2.2:

1. Hydraulically smooth beds ($Re_* \leq 5$). In this case, the viscous sublayer is thick as compared to the roughness height. All roughness elements are within the viscous sublayer. Thus, roughness is not significant.

2. Completely rough beds ($Re_* \geq 70$). In this case, the viscous sublayer is very thin and roughness effects are dominant. This case is relevant for almost all natural rivers and waterways.

For completely rough beds, the logarithmic wall law (2.29) has to be modified in the following way. Since the viscous sublayer does not exist any more, the length scale y_0 cannot be used to normalize the wall distance. Instead of $y_0 = \nu/u_*$, for example $y_0 = \gamma k$ has to be introduced, since the length scale is now given by k (the viscous sublayer is replaced by the so-called **roughness sublayer**, [Dittrich, 1997]). Then the logarithmic wall law is

$$\frac{\bar{u}}{u_*} = \frac{1}{\kappa} ln(\frac{y}{k}) + B \; , \tag{2.34}$$

where the constant B (which is a function of the roughness Reynolds number) has to be determined from the best fit of experimental data.

In his experiments, Nikuradse [1933] found the following modified logarithmic wall law for completely rough beds ($Re_* \geq 70$):

$$\frac{\bar{u}}{u_*} = \frac{1}{\kappa} ln(\frac{y}{k_s}) + 8.5 \; . \tag{2.35}$$

Note that this version of the logarithmic wall law (i.e. $B = 8.5$) is only valid if the vertical distance is normalized by k_s, not by k, and in general $k \neq k_s$.

Determination of u_*, k_s and y'.

Since the shift of the zero-level y', the equivalent sand roughness k_s, and the friction velocity u_* are not known for a given flow over a natural rough surface, e.g. a gravel layer, these parameters have to be determined from the experimental velocity data and the geometric roughness height k. Dittrich [1997] gives the following method.

First, the zero-level of the logarithmic wall law is assumed to be at the top of the roughness elements. According to equation (2.34), a linear regression of the velocity $\bar{u}$ against the natural logarithm $ln((y + y')/k)$ is calculated. The best fit yields the value of y'. The friction velocity can then be calculated from the slope of the regression line. The constant B of (2.34) is determined from the y-axis intercept of the regression line and u_*. Subtracting (2.34) and (2.35) yields the relation

$$\frac{1}{\kappa} ln(\frac{k_s}{k}) = 8.5 - B \; . \tag{2.36}$$

Finally, the value of k_s/k resp. k_s is calculated from (2.36). Typical values for the ratio k_s/k are in the range of 0.5 to 4, depending on the roughness geometry (shape and density of the roughness elements) [Dittrich, 1997].

The details of the influence of wall roughness and permeability on the structure of near-wall turbulence and hence on the parameters of the logarithmic wall law are still subject of active research, see e.g. [Koll, 2002; Breugem, 2004].

2.5.3 Coherent structures

At the end of the 1960s, **coherent structures** like the **bursting phenomenon** have been discovered in turbulent boundary layers [Kline et al., 1967]. Unlike motions in homogeneous, isotropic turbulence, these
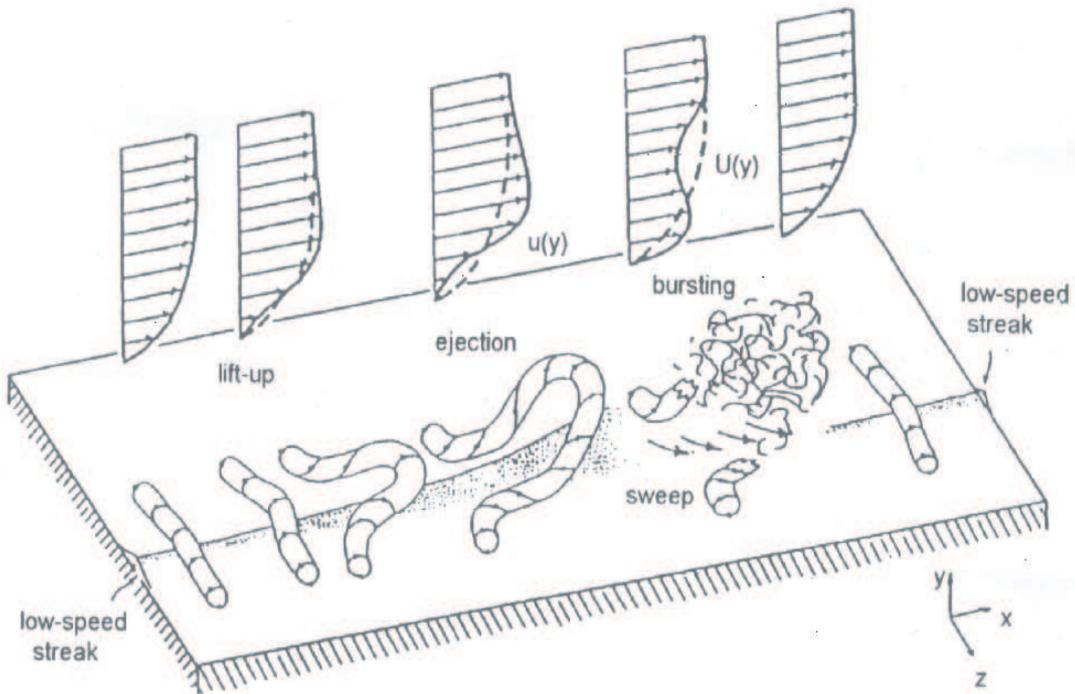
Figure 2.3: *Model of coherent burst event according to [Kline et al., 1967].*

vortex structures are not completely disorganized but appear quasi-periodically in space and time and exhibit a remarkable degree of coherence. It was found that such coherent motions dominate the production of turbulent kinetic energy in both smooth and rough [Grass, 1971] open-channel flows, i.e. the major part of the turbulent shear stress is accounted for by bursting motions. These motions are responsible for the transport of mass, momentum and heat across the boundary layer.

The presence of organized structures in the turbulent boundary layers near the walls of open-channel flows is confirmed by a large number of experimental and numerical studies carried out in the past three decades, e.g. [Kline et al., 1967; Corino and Brodkey, 1969; Grass, 1971; Grass et al., 1991; Robinson, 1991; Nezu and Nakagawa, 1993; Dittrich et al., 1996; Stoesser et al., 2003]. We briefly discuss two different conceptual models of burst motions. Many more have been proposed in the literature, see e.g. [Nezu and Nakagawa, 1993].

Burst events in boundary layers according to Kline.

Coherent structures in near-wall shear flow turbulence over a smooth surface have been reported first by Kline et al. [1967]. It was found that there exist surprisingly well-organized spatially and temporally dependent motions within the viscous sublayer ($y^+ <\simeq 5$). In particular, a quasi-periodic pattern of low-speed and high-speed streaks next to the wall has been found, i.e. the streamwise velocity changes quasi-periodically along a spanwise cross-section of the flow. The following series of events, which has later been summarized as a **burst event**, could be identified as the main source of turbulence production in open-channel flows:

- The low-speed streaks start to lift up randomly in space and time.
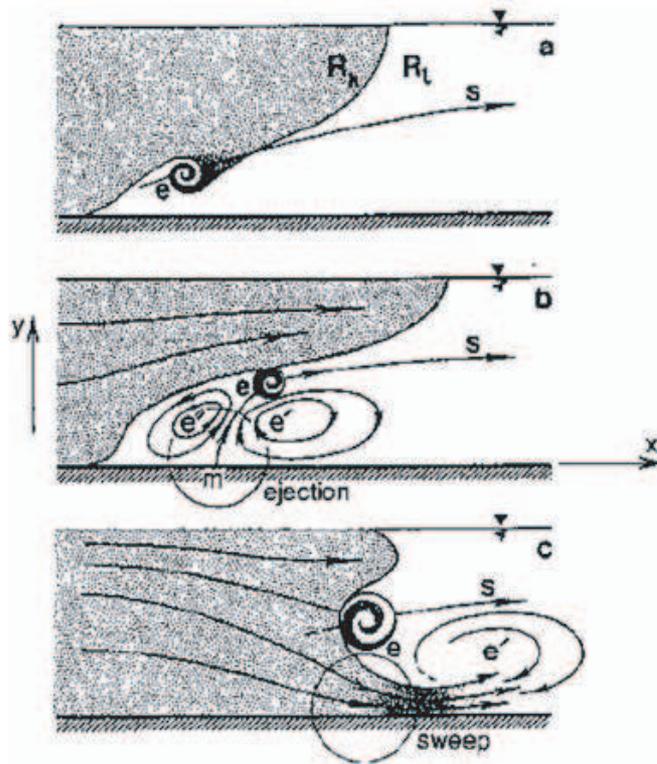
Figure 2.4: *Model of coherent burst event according to [Yalin, 1992].*

- At $y^+ \simeq 8 - 12$ they start to oscillate.

- At $10 < y^+ < 30$ the oscillating streaks suddenly break up due to local, short-duration dynamic instability of the instantaneous velocity profile near the wall.

The streak breakup results in a violent **ejection** of low-speed fluid from the viscous sublayer into the log-arithmic region. For the reason of continuity, the ejections are followed by an **inrush** of high-speed fluid from the logarithmic region into the viscous sublayer, which is also called a **sweep**. The series of ejection and sweep events is called a **burst event**. Figure 2.3 shows a model of such a bursting motion according to [Kline et al., 1967].

Burst events in rivers according to Yalin.

Yalin [1992] gives a somewhat different picture of a burst event. While the burst events as sketched by Kline et al. [1967] are rather microscopic (taking place in a layer up to about $y^+ = 100$, which corresponds to a physical distance to the wall of a few millimeters), Yalin [1992] describes a macroscopic effect and refers to it as 'burst', see figure 2.4:

- Small-scale eddies **e** are produced near the bed, caused by the large velocity gradients which are present in this area. The eddy **e** moves downstream along the path **s** and is growing in size (figure 2.4a).

- Because **e** is moving upwards and rotating at the same time, two further rotational movements **e'** and **e''** are generated. The eddies **e'** and **e''** in turn produce an ejection at the point **m** (figure 2.4b).

- Because **e** is further ascending, **e''** is swept away by the pressure of faster fluid elements moving from the central part of the flow field towards the bed. The eddy **e'** is finally neutralizing this 'sweep' motion by its counter-rotating movement (figure 2.4c).

This series of events may be termed 'coherent', because the origin of all successive processes of the series is the eddy **e**, and there is a definite phase relationship between the single processes. The eddy **e** grows further in size until it is reflected at the water surface. Then it breaks up into smaller and smaller eddies, the energy of which is finally turned into heat by viscous forces (eddy cascade).

Relation between small-scale bursts and large-scale vortical motions.

Different researchers conjecture, that macroscopic eddies similar to those described by Yalin [1992] are somehow linked to the small-scale bursts as described by Kline et al. [1967]. The details of these relations are still the subject of research. For example, there is a debate whether the large-scale vortices develop from the small-scale near-bed structures or whether the latter are triggered by the large-scale motions [Nezu and Nakagawa, 1993].

Adrian et al. [2000] study the structure of energy-containing turbulence in the outer region of a turbulent boundary layer by analysis of spatially high-resolved photographic PIV images obtained in wind tunnel experiments. They show that the turbulent boundary layer is densely populated by velocity fields associated with hairpin vortices (see figure 2.3). In the outer region of the flow, hairpin vortices occur in streamwise-aligned packets. The authors show how many features of turbulent boundary layers, e.g. bursts, can be explained in terms of hairpin vortices or coherent packets of hairpin vortices. Flow patterns that agree with the conceptual models of Adrian et al. [2000] have also been observed by Detert et al. [2005] in 1-D Acoustic Doppler Profiler Measurements from the flume experiments presented here.

Liu et al. [2001] carry out an analysis of 2-D PIV data of turbulent channel flow to investigate the scale and pattern of the eddies that contribute most to the total turbulent kinetic energy and the Reynolds shear stress. Their flow has a Reynolds number based on the channel half-height of ≈ 30000. They show that large-scale motions (with length scales of the order of the channel width) dominate turbulent transport in all parts of the channel except the buffer layer. These large eddies also contain a significant fraction (50% and more) of the total Reynolds shear stress. Flow patterns are extracted by proper orthogonal decomposition of the PIV vector fields. They are very similar to the well-known coherent structures in the near-wall boundary layer (hairpin or horseshoe vortices), which generate the typical sequence of sweeps (fourth-quadrant Q4 events) and ejections (second-quadrant Q2 events).

Shvidchenko and Pender [2001] study the macroturbulent structure of open-channel flow over gravel beds and find that the turbulent motion consists of large-scale eddies with a length about 4 to 5 times their height. The eddies originate in the outer flow. Similar observations have been made recently by Roy et al. [2004] in field studies in a natural gravel-bed river. They present a detailed investigation of the size, scale and dynamics of macroturbulent flow structures and conclude that the latter occupy the entire depth of the flow and scale with the flow depth. By flow visualization they study the dynamics and show how large-scale flow structures can generate intense near-bed turbulence.

2.5.4   Turbulence over rough beds

Grass [1971] and Grass et al. [1991] have found that the bursting motions as described in section 2.5.3 can also be found in open-channel turbulence over rough surfaces. The main difference is that there is no viscous sublayer in fully rough flows. Instead of the viscous sublayer, a retarded fluid layer forms due to penetration of the roughness elements into the turbulent flow field [Nezu and Nakagawa, 1993], the so-called **roughness sublayer** [Dittrich, 1997]. Turbulence in rough open channels is produced mainly in the roughness sublayer as a result of a complex interplay of two different phenomena:

- Coherent structures, i.e the bursting motions, are responsible for large fluctuations of velocity and pressure and therefore for turbulence production.

- Shear layers separating from single grains and penetrating into the flow field produce turbulent wakes downstream of the roughness elements. These wakes typically contain high-frequency velocity fluctuations due to eddy motions.

Further research has to be done in order to quantify the relative importance of these different turbulent structures in the context of river bed destabilization.

2.6   Flow in porous media

Flow through porous media is the subject of many textbooks, e.g. [Scheidegger, 1974; Bear, 1988; Dullien, 1992; Sahimi, 1995]. Most of these treatments are limited to steady viscous flow at small Reynolds numbers ($Re << 1$, Stokes flow). In this flow regime, it is possible to find analytical solutions to the Stokes equations (2.9), e.g. for boundary conditions given by a periodic array of cylinders [James and Davis, 2001]. Clearly, this situation is not applicable to most of the investigations in this thesis, where Reynolds numbers based on the grain diameter and the bed shear velocity usually are of the order of magnitude of 100 to 1000. However, in some of the experiments, Reynolds numbers based on the effective pore diameter (which is smaller than the grain diameter) and the volume-averaged velocity in the gravel layer (which is smaller than the bed shear velocity) are of the order of unity. In any case, the mathematical analysis of slow flow through porous media provides an instructive starting point for a theoretical investigation. A number of more or less empirical modifications of the theory for slow flow through porous media have been developed to extend the description towards higher Reynolds numbers, so-called **phenomenological flow models**. The Forchheimer equation described below is an important example in the context of flow through a gravel layer.

In this section, the defining parameters of a porous medium are introduced, followed by a review of the basic equations governing flow through porous media. We focus on unconsolidated porous media, i.e. random or regular packings of particles. A more detailed discussion of flow and transport in unconsolidated porous media is given e.g. in chapter 13 of [Sahimi, 1995].

2.6.1   Basic equations and parameters

In principle, flows in porous media are also governed by the incompressible Navier-Stokes equations (2.2) and (2.3). The basic difficulty in the analysis of pore flows (regardless if the approach is experimental, theoretical or numerical) stems from the **geometric complexity** of the flow domain. A porous medium

can be considered as a random system of separate pore volumes, which are more or less interconnected by smaller and narrower tube-like structures. An exact description of such a system, e.g. to prescribe exact boundary conditions in a numerical simulation, is not feasible. Thus, a first approach is to neglect the detailed microscopic structure by introducing **volume-averaged quantities**. The averaging is carried out over a spatial volume that is much larger than a single pore, but smaller than the overall domain where flow is described. In this way, macroscopic ('bulk') properties of the porous medium are derived. An important line of research on porous media is to find the correct relation between this large-scale macroscopic description and the microscopic structure on different scales [Cushman, 1990].

Geometry and length scales.

The geometry of a porous medium is related to the type of material and its arrangement in space. It is an important factor that influences the bulk properties and the flow through the medium. For example, fractured rock (a consolidated porous medium) has a geometry that is quite different from a loose packing of regular spheres (an unconsolidated porous medium), which in turn is different from a regular array of cubes. Regular arrays of uniform spheres or cubes are commonly used in numerical simulations, e.g. [Stoesser et al., 2004]. In theoretical studies, different arrangements of cylinders are often used as a model for a fibrous porous medium [James and Davis, 2001].

A gravel layer (or 'filter layer' in the context of hydraulic and geotechnical engineering) is a further example of an unconsolidated porous medium. Since size and shape of the gravels are not uniform, it is characterized by a **grain size distribution**. Usually, the mean grain diameter d_{mD} and different uniformity measures (e.g. the ratio d_{60}/d_{10}, where d_{60} is the mean diameter of the grains at 60% sieve fraction) are used to describe the material.

The definition of **appropriate length and velocity scales** in a porous medium is not trivial. To cite Dullien [1992], the notion of a mean pore diameter or grain diameter is 'an intuitive simplification of reality owing to the irregular variations of pore geometry'. Nevertheless, the **mean grain diameter** or **mean pore diameter** are frequently used to characterize a porous medium. Different formulas exist to relate the pore diameter to the grain diameter, see [Dullien, 1992]. In general, the pore diameter is smaller than the grain diameter, depending on the porosity. A third length scale describing a porous medium may be introduced as the square root of its permeability, see below. The latter is called the **effective pore diameter**.

Porosity.

The **porosity** ϕ of a porous medium is defined as the void fraction of the total volume,

$$\phi = \frac{V_{total} - V_{solid}}{V_{total}} = \frac{V_{void}}{V_{total}}, \tag{2.37}$$

whereas the **packing density** η is defined as the solid fraction of the total volume,

$$\eta = \frac{V_{total} - V_{void}}{V_{total}} = \frac{V_{solid}}{V_{total}}. \tag{2.38}$$

Thus, both quantities are related by $\phi + \eta = 1$. Typical values of porosity for natural materials cover a wide range, from ≈ 0.01 for solid granite to > 0.5 for clay and peat. The porosity of the gravel layer in the experiments described in chapter 10 is 0.38 [Daebel, 2001]. Hence, the solid fraction is 0.62, which is between the packing density of a hexagonal lattice of regular spheres ($\eta = 0.6046$) and the packing density of a random arrangement of regular spheres ($\eta = 0.64$ [Jaeger and Nagel, 1992]).

Characteristic velocities.

Akin to the definition of characteristic length scales, different characteristic velocities in a porous medium have to be distinguished.

The first is the so-called **superficial velocity** v_f, which is also called the **Darcy velocity** or the **filter velocity**. The superficial velocity is the volume-averaged velocity in a porous medium, where the averaging includes the solid phase (where the velocity is zero). The superficial velocity is equal to the flow rate per unit cross-sectional area.

The second characteristic velocity is the **intrinsic** velocity v_i, which is obtained as an average only over the fluid phase. Hence, superficial and intrinsic velocity are related by $v_f = \phi v_i$.

Hydraulic conductivity, permeability and Darcy's law.

Darcy's law relates the superficial velocity $\boldsymbol{v}_f$ to the gradient of the volume-averaged pressure:

$$\nabla \langle p \rangle = -\frac{\mu}{\kappa} \boldsymbol{v}_f. \tag{2.39}$$

Equation (2.39) has been derived empirically by Henry Darcy in 1855, but also follows rigorously from the volume-averaged Stokes equations (e.g. [Sahimi, 1995; Breugem, 2004]). Due to the averaging over the microscopic structure, it is a macroscopic equation, basically expressing the conservation of momentum. Similar transport equations are found in many other disciplines, e.g. Fourier's law in heat conduction or Fick's law in the theory of diffusion [Bird et al., 1960].

In (2.39), κ is the **permeability** of the porous medium, in units of m^2. The permeability is a measure of the ability of a material to transmit fluids through it. It is related to another characteristic quantity, the **hydraulic conductivity** k_f of a porous medium, by

$$k_f = \frac{\gamma \kappa}{\mu}, \tag{2.40}$$

where γ is the specific gravity of the liquid (in units of force per volume) and μ is the dynamic viscosity. The hydraulic conductivity is given in units of velocity (m/s). While the hydraulic conductivity takes into account the properties of the porous medium and the fluid that flows through it, the permeability is a property only of the solid phase. A natural length scale of the porous medium is given by the **effective pore diameter** $\sqrt{\kappa}$.

2.6.2   Transition to turbulence in porous media

The range of validity of Darcy's law is expressed in terms of the **pore Reynolds number** Re_p, defined by the effective pore diameter $\sqrt{\kappa}$ and the intrinsic velocity $v_i = v_f/\phi$:

$$Re_p = \frac{\sqrt{\kappa} v_f}{\phi \nu}. \tag{2.41}$$

Note that there are different possibilities to define such a Reynolds number based on the characteristic scales of the pore flow mentioned above. Many definitions of Reynolds numbers can be found in the literature [Wahyudi et al., 2002]. For example, some authors define the length scale by the effective pore diameter $\sqrt{\kappa}$, while others use a characteristic grain size. Both may differ by an order of magnitude.

Darcy's law is generally assumed to be valid for $Re_p < 10$ [Sahimi, 1995], in spite of the fact that it is derived from the Stokes equation (2.9), which assumes $Re_p << 1$. For $Re_p > 10$, a transition zone begins where inertial forces start to be important. The turbulent flow regime is reached at $Re_p \approx 100$. The strict validity of these numbers is questionable, since in general the transition to turbulence in a porous medium is much more gradual than e.g. in pipe flow [Dullien, 1992].

Phenomenological flow models have been developed to account for the deviations from Darcy's law in the transitional and turbulent flow regimes. One such modification is the **Forchheimer equation**, which simply adds a term quadratic in v_f to Darcy's law:

$$\nabla \langle p \rangle = -\frac{\mu}{\kappa} \mathbf{v_f} + a v_f \mathbf{v_f}, \tag{2.42}$$

where a is a constant that is typically determined from a fit to experimental data. Third or higher order terms may also be added.

2.7 Flow over a permeable wall

2.7.1 Laminar case

Another modification of Darcy's law is the **Brinkman equation**. It is used to account for large gradients in the superficial velocity, e.g. in the transitional flow field across the boundary of a porous medium, e.g. if the latter is in contact with an open-channel flow.

Due to the macroscopic approach, Darcy's law (2.39) is only sufficient for predicting bulk flow, i.e. the superficial velocity in the filter. However, detailed transport phenomena near the boundary of a permeable wall cannot be described by it. More sophisticated flow models are necessary for a precise description of this boundary layer. Towards this end, Brinkman [1947] proposed the following equation:

$$\nabla \langle p \rangle = \mu^* \nabla^2 \langle \mathbf{v} \rangle - \frac{\mu}{K} \langle \mathbf{v} \rangle, \tag{2.43}$$

where μ^* is an *effective viscosity* (or *renormalized viscosity*), which is larger than the actual fluid viscosity and depends on the porosity of the porous medium. The Brinkmann equation can be considered as a semi-empirical interpolation between the Stokes equation (2.9) and Darcy's law (2.39). It facilitates the matching of boundary conditions between the porous medium and the open-channel flow.

The boundary conditions at the interface are different for flows above permeable and impermeable walls. For impermeable walls, the usual 'no-slip' condition holds, which states that the velocity in the mean flow direction is zero at the wall. This is not the case at a permeable wall. At the interface with the open-channel flow, there is a so-called **slip-velocity**, which is larger than the superficial velocity. The first term on the right handside of (2.43) can be interpreted as the one which interpolates between the superficial velocity in the porous medium and the slip-velocity at the boundary. The corresponding solution of the Brinkmann equation is an **exponential decay** of the mean velocity with increasing depth in the porous wall.

2.7.2 Turbulent case

The boundary conditions at the interface between a porous medium and a turbulent flow are still the subject of research, e.g. [Jimenez et al., 2001; Hahn et al., 2002; Breugem, 2004]. Shimizu et al. [1990] present a
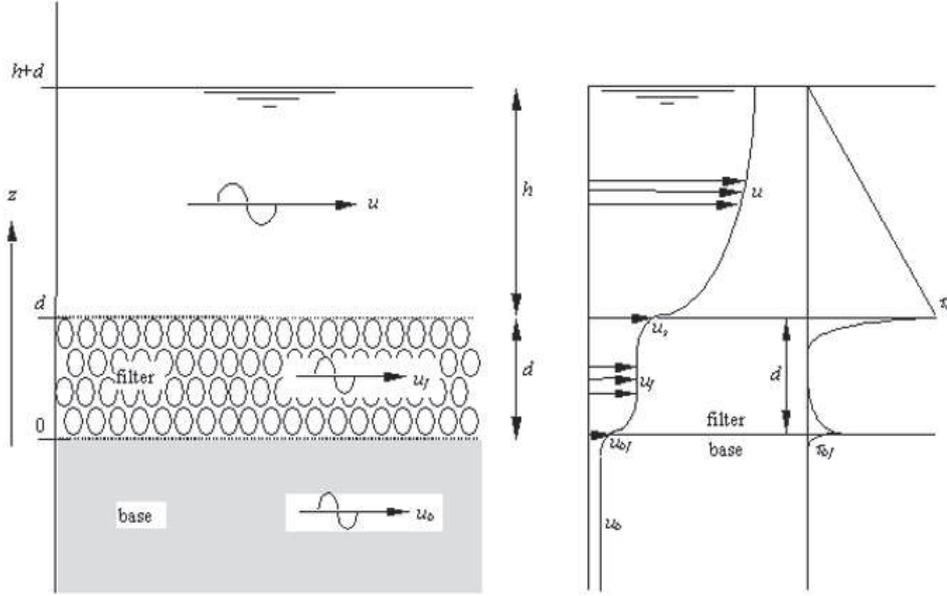
Figure 2.5: *Profiles of mean velocity through a filter layer according to [Hoffmans et al., 2000].*

one-dimensional macroscopic model of flow through a filter layer based on the Forchheimer equation (2.42) and a diffusive theory of fluctuations. The solution of their model equation is similar to that derived in the laminar case, namely an exponential decay of the mean velocity in the filter:

$$v(y) = v_f + v_s \exp(ay),\tag{2.44}$$

where $y < 0$, v is the volume-averaged velocity, v_f is the filter velocity and v_s is the slip-velocity. The parameter a in the exponential function is related to the characteristics of the porous medium and to the eddy viscosity within the filter. The model is verified by experimental measurements using a salt water tracer technique. The details of the phenomenological model by Shimizu et al. [1990], especially with regard to the formulation of the eddy viscosity within a filter layer, are still the subject of debate [Koenders et al., 2000; Hoffmans et al., 2000]. The eddy viscosity is an empirically introduced parameter that relates the turbulent shear stresses to the mean velocity gradient (according to the hypothesis of Boussinesq, compare section 2.5.1).

Hoffmans et al. [2000] carry out a very similar analysis and present some further considerations on the modelling of the eddy viscosity in a porous medium. They assume an eddy viscosity that is proportional to the filter velocity, while Shimizu et al. [1990] used a constant eddy viscosity. Profiles of mean velocity resulting from the analysis of Hoffmans et al. [2000] are shown in figure 2.5. To validate different models of the eddy viscosity in a filter layer, measurements of filter velocities in relation to loading parameters are necessary. Such measurements are obtained in this thesis.

Koenders et al. [2000] present an analytical analysis of turbulent flow through a filter layer. Their theory is two-dimensional and based on the Forchheimer and continuity equations. Their salient result is that *long* wavelength fluctuations in the pressure in the free surface flow are mostly responsible for variations in the superficial velocity near the *bottom* of the filter layer.

Breugem [2004] presents a detailed numerical analysis of laminar and turbulent flow over permeable walls. To classify peremeable walls, he introduces the **permeability Reynolds number**

$$Re_\kappa = \frac{\sqrt{\kappa}u_*}{\nu}, \tag{2.45}$$

where u_* is the shear velocity given by (2.23). The permeability Reynolds number can be interpreted as the ratio of the effective pore diameter to the typical length scale of near-wall turbulence. For $Re_\kappa << 1$, the wall is effectively impermeable. For highly permeable walls with $Re_\kappa >> 1$, turbulence penetrates the permeable wall. Vice versa, the permeable wall changes the structure of near-wall turbulence. Turbulent transport across the wall occurs. Breugem [2004] suggests that turbulence near a highly permeable wall is dominated by relatively large vortical structures, which originate possibly from a Kelvin-Helmholtz instability of the inflexional mean velocity profile. The structures are responsible for the exchange of mass and momentum between the top layer of the pore flow and the surface flow. For highly permeable walls, the parameters of the logarithmic wall law are influenced by permeability. In particular, the vertical coordinate offset is larger and the von Karman constant may be different from 0.4.

2.8 Flow measurements in/over porous media

This section reviews some results of experimental flow measurements in/over porous media. As described above, it is difficult to define flow rates and flow geometry in irregular porous media. Phenomenological flow models are based on volume averaging and describe uniform flow or uniform flow profiles of a macroscopically averaged velocity. Both the experimental studies reviewed in this section and the results of the measurements presented in section 10 show that this is a clear simplification. The flow in a porous medium below a turbulent channel flow is not uniform. Further, flow fields within local pores may significantly differ from the profile of the macroscopically averaged mean velocity.

Flow within porous media.

Yarlagadda and Yoganathan [1989] investigate steady flow inside a porous medium using refraction index matching and a 2-C Laser Doppler anemometer. They calculate the third velocity component by integrating the continuity equation. The measured flow is laminar and stable. Their general observations and conclusions are that the flow is sensitive to the local geometry. The flow direction is changing along the path of the flow. At all locations of the flow field, the flow is three-dimensional.

Saleh et al. [1992] describe 2-D velocity measurements using a PIV technique in a refractive index matched porous medium. The field of view is $\approx 5\ cm \times 5\ cm$, and the velocity range is $25\ \mu m/s$ up to $2.5\ mm/s$. Their spatial resolution is $0.5\ mm$. They are able to identify a number of features of the pore flow, such as separation and stagnation points, which can provide useful information about transport in porous materials. They also point out the three-dimensional character of the pore flow causing erroneous results of their 2-D PIV technique, see section 3.2.

Ogawa et al. [2001] investigate the local flow through a porous medium using nuclear magnetic resonance imaging. The porous medium consists of crushed glass and glass beads with a diameter between $2\ mm$ and $5\ mm$. The porosity is between 0.4 and 0.6, and mean velocities are in the range of $3\ mm/s$ to $13\ mm/s$. The field of view is $5\ cm \times 5\ cm$, the spatial resolution is $0.1\ mm \times 0.1\ mm \times 0.5\ mm$,

Table 2.1: *Important length and velocity scales at the interface of a porous medium and a free surface flow. The scales are used to define corresponding Reynolds numbers.*

region	length scale	velocity scale	Reynolds number
open-channel flow	h	U	$Re = \frac{Uh}{\nu}$
rough wall region	d_{mD}	u_*	roughness Reynolds number $Re_* = \frac{u_* d_{mD}}{\nu}$
permeable wall region	$\sqrt{\kappa}$	u_*	permeability Reynolds number $Re_\kappa = \frac{\sqrt{\kappa} u_*}{\nu}$
pore flow	d_p	u_i	pore Reynolds number $Re_p = \frac{u_i d_p}{\nu}$

and velocities are measured with an accuracy of about 20% of the mean velocity. We summarize their main findings. The flow is non-uniform, and the direction of flow depends strongly on the local geometry. Local flow velocities may be larger than 6 times the mean velocity. In places, reversed flow may be induced by the pore geometry. Flow patterns within single pores change in dependency of the mean flow velocity.

Flow over porous media.

Shimizu et al. [1990] study the flow in a porous filter that is driven by a turbulent channel flow on top by a salt water tracer technique. They show that an appreciable interaction between the pore flow and the surface flow takes place. This interaction results in the exchange of mass and momentum between the two flows. As a result, they confirm the finite slip-velocity at the boundary and show by mass dispersion tests that there is also transpiration at the boundary in the normal direction (injection and suction events). The filter velocity follows an exponential decay with increasing depth in the filter. There is a strong gradient in the filter velocity in the upper grain layers, and only in the deeper grain layers the velocity profile becomes uniform as given by Darcy's law.

Similar events have been observed by Klar [2001] and recently in the very detailed experimental studies of Vollmer [2005].

Lesage et al. [2004] present a new measurement setup for local flow measurements within a single pore. The system is based on an electrochemical technique and enables to study flow regimes within a pore by analysis of signal fluctuations. They identify three flow regimes and distinguish them by the pore Reynolds number. For $Re_p < 110$, the local pore flow is laminar and stable. For $Re_p > 280$, the local pore flow is 'turbulent-like' with a high fluctuation rate. Inbetween, there is a transient region showing a sharp increase of the fluctuation rate with increasing Re_p.

2.9   Summary

In this chapter, an overview of the current knowledge about turbulent flows over rough and permeable walls has been given, covering theoretical, numerical and experimental research. The basic equations and notions of turbulent open-channel flow and flow in a porous medium have been introduced. Clearly there is a mutual influence of both flows on each other in a configuration, where turbulent open-channel flow is in contact with a permeable and rough wall.

At the interface of a porous medium and a free surface flow, different important length and velocity scales may be identified, which are used to define different Reynolds numbers. These Reynolds numbers are summarized in table 2.1. They are used to characterize the influence of roughness and permeability of the wall on the turbulent near-wall flow field.

The near-wall flow field represents a complex interaction of many effects. Two important small-scale phenomena are turbulent burst motions (coherent structures) and fluctuations related to flow separation from roughness elements. Co-exisiting with these small-scale motions are the large-scale vortical motions in the outer flow. The relation between large-scale and small-scale structures is still the subject of research. However, it is clear that both play a dominant role in the interaction of the free surface flow with the pore flow and in the destabilization of gravel grains at a gravel surface [Hofland, 2004].

Turbulent flows over rough walls have been extensively studied experimentally, whereas much less data is available for permeable walls. Different phenomenological flow models have been reviewed, which describe flow profiles through a porous medium. Since these flow models also require turbulence modelling for closure of the equations, experimental data is necessary for verification, e.g. of the assumptions for the eddy viscosity in a porous medium.

Many researchers have shown that small-scale hydrodynamic processes in and near a porous wall in open-channel flow have a high spatial and temporal variability. This underlines the need for a 3-D measurement technique with a high temporal and spatial resolution. Such a technique is developed in this thesis.

Chapter 3

Quantitative flow visualization

This chapter presents an overview of current state-of-the-art techniques for quantitative flow visualization. The purpose of the chapter is to enable a comparison of the different approaches and to elaborate their strengths and limitations. The choice of a 3-D Particle-Tracking Velocimetry (3-D PTV) technique for the present application will be substantiated. A detailed description of the 3-D PTV algorithm is given in chapter 7.

The study of fluid flow is essentially the study of transport phenomena of mass, momentum and heat resp. energy. The task of flow visualization is to make these transport processes visible, which is achieved by introducing tracer substances into the flow, applying some kind of external illumination, and recording images or image sequences of the tracers following the flow. The basic principle of analyzing the light transmitted through or scattered by a fluid is illustrated in figure 3.1.

Flow visualization in a scientific sense can be traced back to the drawings of a turbulent jet made by Leonardo da Vinci in the Late Middle Ages. Since the end of the 19th century, many researchers have performed flow visualization experiments to study the dynamic behaviour of fluids, e.g. Reynolds, Prandtl and Mach, to name but a few. Many nice examples of flow visualization images can be found in [Van Dyke, 1982]. Merzkirch [1987] also gives many examples and discusses the basic physical principles of flow visualization methods. The handbook of Yang [1989] compiles all major techniques of qualitative flow visualization (e.g. smoke visualization and other tracer methods, shadowgraph and Schlieren methods, speckle photography) and demonstrates their applications in many fields of science and technology. This book was published in 1989, and the author states the beginning of 'a new era of *quantifying* the flow information', which has become possible due to advances in computer hardware. The increased computing power opens the way to a *completely automatic* processing of flow visualization images to extract *quantitative* information, i.e. the velocity field and other quantities derived from it[1]. Within the last 20 years, a large number of different methods for **automated quantitative flow visualization** have been developed, and there is an abundant literature on this topic (e.g. [Nieuwstadt, 1992; Lading et al., 1994; Raffel et al., 1998; Smits and Lim, 2000]). Some of these methods have reached the status of standard experimental tools which are commercially available and applied in many labs throughout the world. The most common examples are **Particle-Tracking Velocimetry** (PTV) and, above all, **Particle Image Velocimetry** (PIV).

---

[1]Besides the automatic image processing, the increased computer power has also pushed the feasibility of computational fluid dynamics (CFD). A large number of very illustrative animation sequences of different flows, which have been obtained by CFD, can be found in the 'Virtual Album of Fluid Motion': `http://www.featflow.de/album/index.html`.

incident light   fluid

transmitted light:
extinction measurements,
tomography,
shadowgraphs, schlieren method,
interferometry

scattered light:
Rayleigh scattering,
Raman scattering,
laser induced fluorescence,
LDA, PIV, PTV
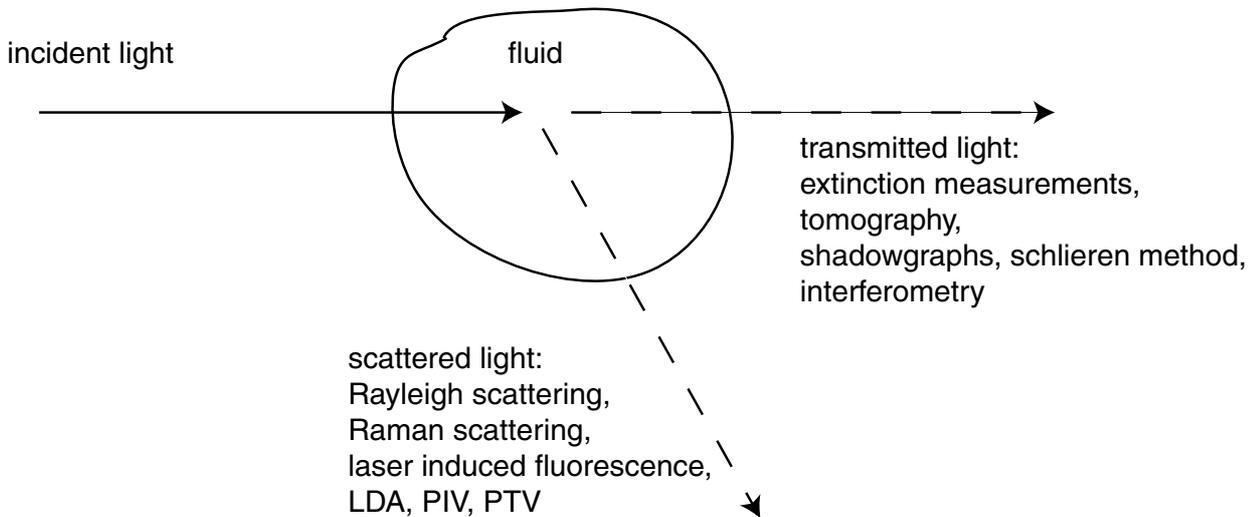
Figure 3.1: *Basic principle of flow visualization methods (after Merzkirch [1987]). A flow is visualized by tracer substances, illuminated by an incident light beam, light sheet or light cone (corresponding to a 1-D, 2-D or 3-D visualization). Information about the flow can be obtained either from the transmitted light or the scattered light. The methods described in this section are all based on light scattering.*

Different methods are distinguished by the number of components of the velocity vector they measure (1-C, 2-C or 3-C) and the dimensions of the spatial domain in which vectors are obtained (1-D, 2-D or 3-D, i.e. a point, plane or volume in space). Hence, a Laser Doppler Velocimeter observing the mean flow component at a single point in space is referred to as a 1-D/1-C method. In this chapter, the most frequently used field methods are reviewed, starting from simple 2-D/2-C approaches, and following the further development up to the most recent 3-D/3-C techniques. The focus of this chapter is on image evaluation techniques, i.e. techniques to obtain a flow *field* within a 2-D or 3-D domain (*multi-point* or *whole-field* techniques). Information on single-point techniques and other optical methods such as Laser Doppler Velocimetry or Interferometry can be found in e.g. [Merzkirch, 1987; Lading et al., 1994]. For information about the experimental setups of the different techniques, see the references mentioned above.

The chapter is outlined as follows. First, some general considerations on motion analysis are presented in section 3.1, including a classification of the different methods. The classification is based on the type of images that are acquired and on the type of evaluation algorithm that is used to process these images. The following sections explain the most important image-based velocimetry methods, namely PIV (section 3.2), least squares image matching (section 3.3), differential optical flow techniques (section 3.4), and PTV (section 3.5). In recent years, the most powerful methods have been obtained by combining the advantages of the different standard methods in so-called *hybrid* approaches (section 3.6). The algorithm proposed in this work (see chapter 7) also belongs to this class. Finally, section 3.7 summarizes the chapter and constitutes why 3-D PTV is the method of choice for the present application.

3.1   General considerations

Image-based whole-field velocimetry methods are used to measure the flow field of a fluid, based on the analysis of an image sequence visualizing the flow under consideration. Estimates of the flow velocity

are obtained by determining the displacements of some kind of image features in a number of successive frames (at least two). In a computer vision context, this displacement field in the image plane is called the **optical flow** $\boldsymbol{f}(\boldsymbol{x},t)$[2]. Computing optical flow resp. motion analysis in general is one of the major issues of computer vision (see e.g. [Murray and Buxton, 1990; Singh, 1991; Haußecker and Spies, 1999]), with applications not being restricted to fluid flow but including any kind of dynamic processes and scenes (e.g. [Frischholz, 1999]). Photogrammetrists also use image matching methods that are closely related to motion analysis, e.g. to establish correspondences between two stereo images in order to compute a disparity map or to locate target patterns within an image (e.g. [Baltsavias, 1991]). Accordingly, there is a huge amount of methods, algorithms and publications spread out through the computer vision, photogrammetry and fluid mechanics literature. The vast terminology for the different methods might confuse the unfamiliar reader, particularly because notions like 'optical flow', 'image matching', 'image correlation' and 'tracking' are not always used consistently by the different communities (or even within the same community).

In this section, a general overview on image-based velocimetry is given. The most important methods are introduced and classified according to the fundamental principles, assumptions and approximations they are based upon. In general, the task of all the methods is to compute the optical flow $\boldsymbol{f}(\boldsymbol{x},t)$, i.e. the apparent 2-D motion of image features in the image plane[3]. The methods differ in the following aspects:

A) What kind of image feature is used?

　　A1) single particles, i.e. discrete features

　　A2) particle patterns, i.e. patterns of discrete features

　　A3) continuous features

B) What kind of input data is used for the velocity estimation?

　　B1) spatial information, i.e. positions of features in the image plane

　　B2) temporal information, i.e. more than two frames are used

　　B3) intensity, i.e. gray values of features

　　B4) intensity gradients, i.e. local gray value differences

C) What is the computational approach to solve the motion correspondence problem ?

　　C1) cross-correlation

　　C2) least squares optimization (linear or nonlinear)

　　C3) discrete tracking techniques (Kalman filtering, combinatorial optimization)

Algorithms using almost any combination of image features, input data and computation method can be found in the literature. Some examples are compiled in table 3.1. A common classification is to divide the methods into two major groups: **region-based methods** and **feature-based methods**.

Region-based methods estimate the motion of gray value patterns within small image patches, so-called *interrogation areas* or *interrogation windows*. The most common region-based method used in fluid mechanic applications is **Particle Image Velocimetry** (PIV) (section 3.2), whereas in computer vision and photogrammetry, **differential optical flow** methods (section 3.4) and **least squares matching** (section 3.3) are frequently used. Since in all these methods the image is divided into a regularly spaced array of interrogation windows, the result is a displacement field on a regular grid. The interrogation windows are chosen

---

[2]Strictly speaking, the optical flow is the *continuous* apparent motion field in the image plane, and the displacement field is the sampled version of the latter that is obtained from the discrete pixel array of a camera sensor.

[3]Note that the term 'optical flow' is often used in a narrower sense, refering only to the methods described in Section 3.4

Table 3.1: *Examples of different approaches to image-based velocity analysis.*

method	reference	features	data	calculation
Standard PIV	Willert and Gharib [1991], Westerweel [1997]	A2	B1,B3	C1
correlation-based tracking, correlation imaging velocimetry	Fincham and Spedding [1997]	A2	B1,B3	C1
Image correlation velocimetry, Adaptive Least Squares Matching	Tokumaru and Dimotakis [1995], Gruen [1985]	A2,A3	B1,B3	C2
Multi-grid PIV with deformable windows	Scarano and Riethmuller [2000]	A2	B1,B3	C1
Hybrid PIV/PTV	Cowen and Monismith [1997], Bastiaans et al. [2002]	A1,A2	B1,B3	C1,C3
two-frame tracking	Baek and Lee [1996], Ohmi and Li [2000]	A1	B1, B3	C3
four-frame tracking	Hassan and Canaan [1991], Malik et al. [1993]	A1	B1,B2, B3	C3
Kalman Filtering	Takehara et al. [2000]	A1	B1,B2, B3	C3
optical flow techniques	Haußecker and Spies [1999]	A2,A3	B1,B2, B3,B4	C2

with a certain degree of overlap, typically 50% in PIV. In optical flow computations, a velocity vector is typically computed for every pixel in the image, i.e. the grid of velocity vectors is the pixel grid itself.

In contrast, feature-based methods (also known as *token-tracking* methods [Murray and Buxton, 1990; Faugeras, 1993]) try to identify single objects in the image, segment them from the background, and follow their motion throughout an image sequence. Thus, feature-based methods yield randomly spaced velocity information, depending on the distribution of moving objects in the image. The most important feature-based method for flow visualization is **Particle-Tracking Velocimetry** (PTV) (section 3.5), where individual tracer particle images are the objects to be tracked. In general, both classes of methods have advantages and disadvantages, which are outlined in the following sections. Hybrid methods (section 3.6) try to combine the advantages of region-based and feature-based approaches, and consequently have a better performance in many cases.

In the remainder of this section, some general aspects of motion analysis are outlined, which apply

equally to all the different approaches. A comparison of the methods is given in section 3.7.

3.1.1 Dynamic range, sampling theorem and subpixel accuracy

Dynamic range.

An important quantity characterizing the potential of a motion estimator is its **dynamic range** DR, which is defined as the ratio of the maximum to the minimum displacement that can be measured:

$$DR = \frac{\xi_{max}}{\xi_{min}}, \tag{3.1}$$

where ξ_{max} and ξ_{min} are the displacements measured in pixels/frame. Obviously, a dynamic range as high as possible is desirable, in particular with regard to the measurement of turbulent flows, which may contain strong velocity fluctuations.

The fundamental limits on the dynamic range of a digital imaging method are related to the discrete nature of the image data. The measurement of large displacements is limited by the **temporal sampling theorem** [Haußecker and Spies, 1999], while the measurement of small displacements is limited by the maximum **subpixel accuracy** that can be achieved. The latter in turn depends on the sampling and quantization of the image intensity [Jähne, 1999a]. Approaches to increase the dynamic range are outlined in section 3.1.2. If such an approach is used together with a subpixel-accurate determination of small displacements (see below), a dynamic range of 100-1000 can be achieved using standard equipment.

Sampling theorem and motion correspondence.

To estimate an object's velocity, given two successive image frames, the **motion correspondence problem** has to be solved, i.e. a unique correspondence between the two images of the same object in the two successive frames has to be established. This can only be achieved, if the conditions given by the temporal **sampling theorem** are valid. Put in simple words, the temporal sampling theorem (or *Nyquist criterion*) states, that an object's displacement between two images, i.e. the optical flow $\boldsymbol{f}$, should be less than half the smallest local spatial scale $\lambda_{g,min}$ of the image intensity $g(\boldsymbol{x}, t)$:

$$|\boldsymbol{f}|\Delta t = |\boldsymbol{f}| \cdot (1\ frame) < \frac{1}{2}\lambda_{g,min}, \tag{3.2}$$

where Δt is the time interval between two successive images, in units of frames (thus, $\Delta t = 1$), and $\boldsymbol{f}$ is the optical flow in units of pixels/frame. The sampling theorem imposes a fundamental limit on the relation between the size and intensity structure of an object and its motion, i.e. on the relation between the spatial and temporal intensity gradients. Given just two images and no further information, the motion of an object can only be recovered unambiguously if (3.2) is valid. In this case, the motion correspondence problem can be solved. Otherwise, temporal aliasing occurs [Jähne, 1999a], and the problem of low-level motion estimation becomes ill-posed, i.e. there is no unique solution.

As a consequence of the sampling theorem, there is a maximum allowed displacement that can be recovered by any region-based method. For example, consider a differential optical flow technique (section 3.4). In this case, there has to be a unique relation between the spatial and temporal gray value gradients. To estimate the motion of a single particle with a symmetrical Gaussian intensity distribution (as typical for

PIV and PTV particle images), the maximum allowed displacement corresponds roughly to the standard deviation of the Gaussian resp. radius of the particle. As a second example, for a quadratic PIV interrogation window of length L, the maximum displacement corresponds roughly to $L/4$, assuming that the intensity distribution resp. particle density within the window is homogeneous and sufficiently large. The latter result is known as the 'one-quarter-rule' in the PIV literature [Raffel et al., 1998]. Obviously, the smallest spatial scale within a PIV interrogation window depends on the particle distribution and density within that window. Generally, an optimal density of about 10 particles per interrogation window is recommended in the PIV literature (e.g. [Keane and Adrian, 1990]). Assuming a homogeneous distribution, the particles form a periodic intensity pattern of wavelength $\lambda = L/2$. Hence, the one-quarter-rule follows from the sampling theorem. Note that these limits are not strict but should be considered as more or less accurate estimates, since the actual spatial frequency content of the image depends on the particle distribution. The latter is a stochastic quantity, with varying values for different interrogation windows within one image.

The situation is a bit different for feature-based tracking methods. As a simple example, consider an image sequence with a single moving object. Its motion can be tracked with the only restriction that it stays within the field of view, since two successive images of the object can always be related to each other unambiguously. In this case, the wavelength of the spatial image structure corresponds to the size of the image (resp. twice this size). However, such a case is of limited practical importance in flow visualization, since the images contain a large number of particles. With increasing object resp. particle density, the spatial image scales become smaller, and the motion correspondence becomes more difficult, which again is a manifestation of the sampling theorem. Nevertheless, tracking algorithms are able to track motions violating the sampling theorem. But the latter is only possible, if further information is used (apart from two successive images). For example, a common assumption is that object trajectories are smooth, i.e. the direction and speed of an object does not change abruptly between two frames. In this case, it is possible to use information from previous frames in a motion model (section 3.1.3) and predict the position of the object in the next frame by extrapolation. If the model provides a good description of the actual motion, much larger displacements can be handled as compared to low-level approaches using only two frames.

Subpixel accuracy.

The subpixel accuracy of a velocity estimator determines the minimum displacement that can be measured. Since a digital image provides a sampled version of the original intensity distribution of the physical image, with gray values defined on an integer grid (pixel positions), the position of an object, e.g. a particle image or the correlation peak resulting from the cross-correlation of two PIV images (see section 3.2.1), can only be determined with an accuracy of ± 0.5 pixels. To achieve a higher accuracy, some kind of **subpixel interpolation** has to be carried out. One way to do this is to use a model of the intensity distribution of the object, and to determine the best fit of this model to the image data in a least squares sense. The most common model in PIV and PTV is a Gaussian distribution, since it provides a very good approximation to both the image of a single tracer particle and the displacement peak in a PIV correlation. The subpixel-accurate coordinates are introduced as parameters of the model and determined in a least squares minimization.

Another common approach to achieve subpixel accuracy in PIV and optical flow estimation is to warp the original images according to an estimated flow model. The warping is carried out iteratively, and a refined estimation of the velocity field is computed in each iteration. Since the warped image will generally be defined on non-integer pixel positions, warping requires a precise method to interpolate gray values, see

e.g. [Jähne, 1999c].

In applying subpixel interpolation, one should keep the following considerations in mind. To compute a subpixel-accurate position within an image, the information contained in the image intensity, i.e. in the gray values, is 'translated' to geometric information, i.e. position in the image. This 'translation' is based on certain assumptions concerning radiometric aspects of the imaging process. One such assumption is the Gaussian intensity distribution mentioned above. Further important assumptions, which are often taken for granted, are the linearity and homogeneity of the sensor, a pixel-synchronous transfer of the image data between sensor, frame grabber and memory, and a homogeneous illumination. If any of these assumptions is violated, subpixel accuracy will deteriorate or even become meaningless. Thus, it is very important to take into account the electronic and especially the **radiometric** properties of the cameras and illumination system, if very high accuracy is required. For example, if the cameras suffer from strong fixed pattern noise, a radiometric correction should be applied to the images. Such a correction is proposed in section 4.3. Even if the image data is 'perfect', and all the assumptions stated above are valid, the result of the subpixel interpolation may still be biased. For example, one source of bias in PIV evaluation is the so-called 'peak-locking' effect, see section 3.2.1. Peak-locking mainly occurs when the images of the tracer particles are too small, i.e. when the Gaussian intensity distribution is under-sampled.

As a general limit, for typical 8 bit images with 256 gray levels, one can expect a (theoretical) maximum subpixel accuracy of the order of magnitude of 0.01 pixels, given optimal image data, a good object model and an unbiased estimator [Wernet and Pline, 1993; Marxen et al., 2000]. Note that it may be very difficult to actually achieve such ideal circumstances in real PIV or PTV applications, where measurement errors are typically in the range of $0.05 - 0.2$ pixels [Westerweel, 2000].

3.1.2   Hierarchical multigrid approaches

As explained in section 3.1.1, the maximum displacement that can be determined by region-based approaches like PIV and differential optical flow methods is limited by the smallest spatial scales of the underlying image structure. However, images also contain information at larger scales than the neighbourhood size of the interrogation windows. The basic idea of iterative, hierarchical methods is to start the estimation of the optical flow at the largest image scales, which enable the determination of large displacements in a first iteration. This first estimation may be applied to warp the second image back along the estimated displacement field and refine the estimation at smaller spatial scales. An efficient implementation of such a coarse-to-fine strategy is a **Gaussian image pyramid** [Jähne, 1999b], which is basically a multigrid representation of an image at different spatial scales. The efficiency of Gaussian pyramids is due to the reduction of the linear image size by a factor of two at each level of the pyramid. This reduction makes the large-scale information in the image available to small filter masks. However, at the same time the image becomes more and more blurred. Hence, we have to take care in applying Gaussian pyramids to PIV images, since the small particle images may soon be completely smoothed out. Large-scale information can only be obtained if there is a certain fraction of larger particles in the image or the particle density varies locally. Hierarchical PIV approaches are often realized by starting with large interrogation windows and iteratively decreasing the size of the interrogation windows instead of decreasing the image size as in a Gaussian pyramid [Scarano and Riethmuller, 1999; Hart, 2000b].

3.1.3 Modeling of displacement fields

Given two successive images $g_0 = g(\boldsymbol{x}, t_0)$ and $g_1 = g(\boldsymbol{x}, t_1)$ of a flow field, the displacement field $\boldsymbol{\xi}(\boldsymbol{x}, t)$ can be thought of as the transformation, or mapping, of the spatial image intensity field from the first image to the second. The optical flow is the time derivative of this mapping: $\boldsymbol{f}(\boldsymbol{x}, t) = \partial_t \boldsymbol{\xi}(\boldsymbol{x}, t)$. Within a local neighbourhood N centered at $\boldsymbol{x}_0$, the displacement field may be approximated by a Taylor expansion:

$$\boldsymbol{\xi}(\boldsymbol{x}, t) = \boldsymbol{\xi}(\boldsymbol{x}_0, t) + (\boldsymbol{x} - \boldsymbol{x}_0)\nabla\boldsymbol{\xi}(\boldsymbol{x}_0, t) + \frac{1}{2!}\left[(\boldsymbol{x} - \boldsymbol{x}_0)\nabla\right]^2 \boldsymbol{\xi}(\boldsymbol{x}_0, t) + ... \qquad (3.3)$$

Taking into account only the first order terms, the equivalent formulation for the optical flow reads:

$$\boldsymbol{f}(\boldsymbol{x}, t) = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix} + \begin{bmatrix} a_5 \\ a_6 \end{bmatrix} = \boldsymbol{A}(\boldsymbol{x} - \boldsymbol{x}_0) + \boldsymbol{t} \qquad (3.4)$$

In this first order approximation, the displacement field consists of a constant shift $\boldsymbol{t}$ and a linear (affine) deformation of the local neighbourhood, described by the matrix $\boldsymbol{A}$. Note that in such a formulation, the spatial derivatives of the flow field are introduced as parameters:

$$a_1 = \frac{\partial f_x}{\partial x}, a_2 = \frac{\partial f_x}{\partial y}, a_3 = \frac{\partial f_y}{\partial x}, a_4 = \frac{\partial f_y}{\partial y}. \qquad (3.5)$$

This offers the possibility to estimate spatial velocity gradients without performing explicit differentiation of the velocity field. Thus, important hydromechanic quantities like the in-plane vorticity $\omega_z = \partial_x f_y - \partial_y f_x$ and the rate-of-strain tensor $\boldsymbol{S}$ can be directly estimated, since

$$\boldsymbol{A} = \begin{bmatrix} 0 & -\omega_z/2 \\ \omega_z/2 & 0 \end{bmatrix} + \boldsymbol{S}, \qquad (3.6)$$

with

$$\boldsymbol{S} = \begin{bmatrix} \partial_x f_x & \frac{1}{2}(\partial_y f_x + \partial_x f_y) \\ \frac{1}{2}(\partial_y f_x + \partial_x f_y) & \partial_y f_y \end{bmatrix}. \qquad (3.7)$$

Similar to this spatial modeling of displacement fields, the temporal evolution of the motion of a single particle along its Lagrangian trajectory (see section 2.3) around the point $\boldsymbol{x}_0$ may be approximated using a Taylor expansion in time:

$$\boldsymbol{\xi}(\boldsymbol{x}_0, t) = \boldsymbol{\xi}(\boldsymbol{x}_0, t_0) + \boldsymbol{v}(t - t_0) + \frac{1}{2}\boldsymbol{a}(t - t_0)^2 + ... \qquad (3.8)$$

This kind of modeling is frequently applied in particle-tracking algorithms, see section 3.5 and section 7.4.2.

A more detailed discussion of the modeling of flow fields is given by Haußecker and Spies [1999].

3.1.4 Confidence measures, validation and postprocessing

As any measurement technique, a velocity estimator should not only supply the velocity field, but also some kind of **confidence measure**. In order to enable a reliable interpretation of the velocity field, gross errors have to be detected and removed. The optical flow methods discussed in section 3.4 yield confidence measures as an integral part of the result. In PIV algorithms, typically the ratio of the tallest to the second tallest correlation peak is used to detect unreliable measurements [Keane and Adrian, 1990]. Based on such

confidence measures, questionable measurements are removed from the velocity field, which is typically done in a postprocessing step after the velocity field has been computed. However, in iterative methods, where the results strongly depend on the quality of the velocity estimates in previous iterations, the validation should be done after each iteration. Different validation methods are discussed by Westerweel [1994] and Hart [2000a].

After the erroneous vectors have been removed, the resulting gaps in the velocity field may be filled by applying an interpolation technique, e.g. *adaptive Gaussian windowing* (AGW) [Agui and Jimenez, 1987; Spedding and Rignot, 1993]. Such techniques can also be used to interpolate the randomly distributed velocity vectors resulting from a PTV technique to a regular grid. Basically, the interpolation corresponds to a convolution of the velocity field using a special convolution kernel, e.g. a Gaussian in the AGW. To account for the varying uncertainty of the computed velocity vectors, a *normalized convolution* may by computed, where pixels with suspicious information (as indicated by their confidence measure) get a low weighting factor in the convolution sum [Granlund and Knutsson, 1995]. For further information on interpolation and convolution techniques, refer to [Jähne, 1999c,d].

3.1.5   3-D motion estimation

Most of the methods discussed in this chapter refer to the case of 2-D motion estimation within a plane, namely the image plane. However, all of these methods can easily be extended to the case of 3-D motion estimation within a volume in space. From an algorithmic point of view, there is no principal difference between e.g. computing a cross-correlation in 2-D and in 3-D, see e.g. [Yamamoto et al., 1993; Pereirra et al., 2000; Schimpf et al., 2003] for examples of volumetric 3-D cross-correlation or [Deusch et al., 2000] for 3-D least squares matching of volumetric images obtained by Laser-Induced Fluorescence. Optical flow algorithms and tracking methods can also be applied to the 3-D case, simply by adding a further dimension. The challenge of 3-D motion estimation is rather a technological one: the acquistion of 3-D image data. Most approaches to 3-D velocity measurement are based on **stereoscopic** or **multi-view imaging** using two or more views of the same flow scene to recover the 3-D velocity field. The most prominent method applied to measure fluid flows is stereoscopic PIV [Prasad, 2000]. Some 3-D PTV approaches are discussed in section 3.5.3. The approach presented in this thesis also belongs to the stereoscopic techniques.

The basic new ingredient of 3-D methods as compared to 2-D methods is a **geometric camera calibration** (chapter 5). This calibration is necessary, because perspective effects have to be taken into account in the evaluation of stereo images. The task of the stereo evaluation is to establish stereoscopic correspondences between two different views of the same scene. Thus, in addition to the motion correspondence problem (temporal correspondence), the **stereo correspondence problem** (spatial correspondence) has to be solved (chapter 6): given two views of the same scene, e.g. a flow field visualized by tracer particles, a unique correspondence between the particle images in the two views has to be found. The camera calibration provides the geometric relationship between the two views, the so-called **epipolar geometry**. If this relationship is known, the stereo correspondence problem can be solved much easier and faster. Further, the calibration also provides the necessary geometric information to compute the 3-D position of an object by triangulation of two or more views. Stereo algorithms can be implemented very efficiently and transparently in terms of *projective geometry*. The (projective) geometry of multiple views and its implications for motion analysis are extensively discussed in the computer vision and photogrammetry literature, e.g. [Faugeras, 1993; Faugeras et al., 2001; Xu and Zhang, 1996; Hartley and Zisserman, 2000]. [Trucco
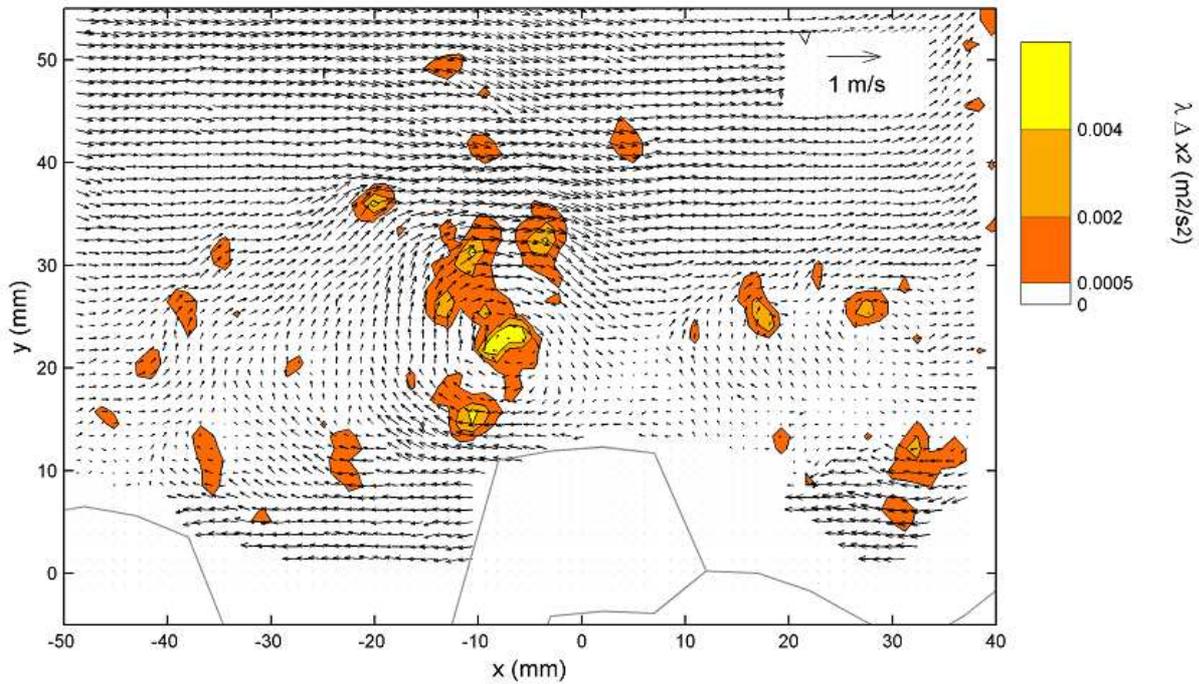
Figure 3.2: *Example of a 2-D PIV vector field, revealing the near-wall flow structures around a large stone embedded in a gravel layer. From [Hofland, 2004].*

and Verri, 1998] is an introductory text on 3-D computer vision, and Kanatani [1996] deals with geometric computations on a more advanced level. Camera calibration is a classic topic of photogrammetry [Slama, 1980; Atkinson, 1996; Gruen and Huang, 2001].

Another powerful but experimentally very elaborate 3-D approach is holographic imaging [Hinsch, 2002]. More information on this and other 3-D flow visualization methods can be found in [Dracos, 1996].

3.2   Particle Image Velocimetry

Since PIV is the most common flow visualization method, a large body of literature is available. Different PIV methods are reviewed by Adrian [1991]. Information on auto-correlation PIV including film-based acquisition and optical evaluation methods can be found in [Keane and Adrian, 1990, 1991]. The theory of cross-correlation PIV is developed in [Keane and Adrian, 1992]. The fundamentals of digital PIV are discussed by Willert and Gharib [1991]; Westerweel [1997] and in the books by Westerweel [1993] and Raffel et al. [1998]. The latter give a large number of references to further information on PIV.

3.2.1 Standard 2-D PIV

Basic principle.

Particle Image Velocimetry (PIV) is a technique to determine the 2-component displacement vectors of tracer particle patterns in a 2-D plane (light sheet) within a flow. The result is a 'snapshot' of the Eulerian flow field. An example is shown in figure 3.2. The displacements are found by partitioning two subsequent frames of a PIV sequence, $g_1 = g(\boldsymbol{x}, t_1)$ and $g_2 = g(\boldsymbol{x}, t_2)$ into interrogation windows, typically of a size of 16×16 or 32×32 pixels, and computing the cross-correlation coefficient $r(\boldsymbol{x}, \boldsymbol{s})$ of two corresponding windows:

$$r(\boldsymbol{x}, \boldsymbol{s}) = \frac{\int\limits_{-\infty}^{\infty} w(\boldsymbol{x} - \boldsymbol{x}')g_1(\boldsymbol{x}')g_2(\boldsymbol{x}' - \boldsymbol{s})d^2x'}{\left(\int\limits_{-\infty}^{\infty} w(\boldsymbol{x} - \boldsymbol{x}')g_1^2(\boldsymbol{x}')d^2x' \int\limits_{-\infty}^{\infty} w(\boldsymbol{x} - \boldsymbol{x}')g_2^2(\boldsymbol{x}' - \boldsymbol{s})d^2x' \right)^{1/2}}, \tag{3.9}$$

where the weight function $w(\boldsymbol{x} - \boldsymbol{x}')$ represents the size of the interrogation window, and it is assumed that the local mean gray values over the interrogation windows have been subtracted from g_1 and g_2. The correlation coefficient is computed for a given 2-D range of displacements $\boldsymbol{s}$ of the interrogation window, resulting in a so-called *correlation plane*.

Computation of velocity.

Because the direct evaluation of the cross-correlation coefficient (3.9) is computationally very expensive, it is usually computed using FFT-methods, because in Fourier space, the double summation is replaced by a simple pointwise multiplication. Once the correlation plane has been determined, the correct displacement is given by the maximum correlation peak. Thus, the optical flow is approximated as

$$\boldsymbol{f}(\boldsymbol{x}, t) \approx \frac{1}{\Delta t} \arg \max r(\boldsymbol{x}, \boldsymbol{s}), \tag{3.10}$$

where Δt is the time difference between the two successive images. Subpixel accuracy is achieved by computing the centroid ('center of gravity') or fitting a Gaussian to the correlation peak. Usually, in both methods only three neighbouring correlation values in each direction are used (three-point estimators). Depending on the image quality and the evaluation method, the accuracy of the displacement estimation is of the order of 0.01 to 0.1 pixels and the dynamic range of the method is of the order of 100 to 1000.

Velocity post-processing.

Since PIV is a statistical evaluation method, the obtained vector fields will contain a certain amount of spurious vectors ('outliers'), which result from interrogation windows containing an insufficient number of particle images or other shortcomings (see below). These outliers have to be removed prior to any further evaluation of the velocity field. Some methods for outlier removal and interpolation of the resulting gaps have been mentioned in section 3.1.4. After this post-processing, higher order quantities like vorticity, divergence or rate of strain may be computed.

Limitations.

The basic approach of PIV as discussed in this section suffers from a number of shortcomings, which limit the accuracy, dynamic range and spatial resolution of PIV. The main origin of these shortcomings is the fixed, finite size resp. shape of the interrogation window used in the correlation analysis, which effectively acts as a spatial low pass filter on the estimated velocity field. Another source of error stems from the spatial discretization of digital particle images. In detail, the following limitations exist:

- **In-plane loss-of-pairs:**

  Particles may enter or leave the finite interrogation window between subsequent frames, in particular those, that are faster than the mean velocity within the window. Thus, fast particles will not contribute to the correlation peak, since they do not have a matching partner within the interrogation window in the second frame. This results in a bias of the estimated velocity towards lower values.

- **Velocity gradients:**

  Spatial velocity gradients within the interrogation window also contribute to the in-plane loss-of-pairs and thereby reduce the signal-to-noise ratio in the correlation plane since not all particles within an interrogation window correlate equally well due to their nonuniform motion. As a rule of thumb, the degradation of the PIV result becomes significant if the displacement of tracer particles due to local flow gradients gets larger than the image diameter of the particles.

- **Out-of-plane loss-of-pairs:**

  Particles may enter or leave the light sheet along the optical axis during the time of two successive exposures. Such particles are only visible in one of the images and do not have a matching partner. Again, this results in a reduction of the signal-to-noise ratio. The out-of-plane loss-of-pairs is a principal physical limitation of PIV that can only be overcome by adjusting the experimental parameters, e.g. thickness of the light sheet or frame rate of the cameras.

- **Computational aspects:**

  To reduce the computational load, the correlation is often computed in the Fourier domain using FFT methods. However, the necessary assumption of the periodicity of the image data within the interrogation window introduces inaccuracies as compared to a direct spatial cross-correlation, which is principally more accurate [McKenna and McGillis, 2002].

- **Peak-locking resp. pixel-locking:**

  The discrete nature of the PIV images introduces a bias towards integer displacements in the subpixel evaluation of the displacement peak. Peak-locking is the result of a biased subpixel estimation, if the input data (i.e. correlation values) is distributed asymmetrically around the maximum peak. The degree of peak-locking depends on the size of the particle images. Peak-locking effects may dominate the velocity errors for very small (under-sampled) particle images.

Apart from peak-locking, all limitations are a consequence of the finite extent of the interrogation windows. The size of the window is given by a trade-off between dynamic range and accuracy on one hand and spatial resolution on the other hand. Large interrogation windows can resolve large motions and provide

good accuracy due to a high signal-to-noise ratio, given that the window contains only weak velocity gradients. Large windows are also more robust to outliers. On the other hand, smaller windows provide a better spatial resolution and are less affected by velocity gradients, e.g. shear flows or strong vortices. However, to enable a reliable evaluation of the cross-correlation, the windows must contain a sufficient number of particle images and thus must have a certain minimum size, which depends on the particle density.

The limitations imposed by fixed interrogation windows can also be explained by looking at the spatial Taylor expansion of the velocity field (3.4). The standard PIV approach can only compute a straight shift of the interrogation windows between two frames. The velocity field within the interrogation window is assumed to be constant. This corresponds to a zeroth order expansion of the velocity field. Linear effects like rotation, shear and dilation or higher order deformations are not accounted for. Due to the spatial averaging over the interrogation window, flow scales smaller than the window size cannot be recovered.

To summarize, accuracy, spatial resolution and dynamic range of the standard PIV method are coupled by the size of the interrogation window. The performance of PIV depends on three main factors: particle size and density, size of interrogation window and presence of velocity gradients. Particle size and density can be controlled during the setup of the experiment and are not discussed further. Recommendations for optimal settings are given in the PIV literature, e.g. [Raffel et al., 1998]. In the following sections, some advanced PIV methods are discussed. The goal of these methods is to overcome the limitations of the standard PIV approach to increase the accuracy, resolution and dynamic range. Towards this end, the latter three performance measures have to be decoupled. Most of the advanced methods rely on the following three major ideas: iterative methods instead of a single-pass evaluation to refine the solution, hierarchical multigrid approaches to resolve both large and small motions, and higher order models of the velocity field to account for velocity gradients and higher order deformations.

3.2.2 Advanced PIV image analysis

This section describes some advanced PIV image evaluation methods that aim to overcome some of the limitations mentioned in section 3.2.1.

Multiple pass interrogation with window shifting.

To reduce the in-plane loss-of-pairs, a discrete integer window offset determined in a first interrogation pass is introduced before doing a second interrogation using the shifted window [Westerweel et al., 1997; Cowen and Monismith, 1997]. The increased number of matched particle pairs results in an increased signal-to-noise ratio of the correlation peak. Iterations of the window shifting may be carried out until the displacement determined in the final iteration is below one pixel. Due to the discrete window shifting, the result still suffers from peak-locking, which can be reduced by applying a continuous window shifting, e.g. [Gui and Wereley, 2002].

Correlation-based tracking.

In the standard PIV approach, the interrogation windows in the first and second frame are of the same size and at the same location within the image. This is the major reason for the low velocity bias error due to the in-plane loss-of-pairs. A simple modification to eliminate this error is to use a larger interrogation window in

the second frame, centered around the smaller window in the first frame [Fincham and Spedding, 1997]. In this case, the correlation coefficient has to be computed directly in the spatial domain for all displacements of the small window within the large window. Fincham and Spedding [1997] have termed this approach 'correlation image velocimetry' to distinguish it from the standard 'correlation-based interrogation' using equally sized windows. In a number of more recent references, the approach using differently sized windows is referred to as 'correlation-based tracking', since the particle pattern defined by the small interrogation window is tracked within a search area defined by the large interrogation window.

Multiple pass with decreasing window size.

The optimal interrogation window size for PIV depends on the local flow conditions and the local seeding particle density, which means that it is rarely constant from one region of the flow to another. Thus, instead of using fixed window sizes, the size of the window should be dynamically adapted to the local flow conditions. A simple way to implement this idea is to refine the correlation interrogation in an iterative way by starting with large windows and decreasing the window size during the course of the iterations [Scarano and Riethmuller, 1999; Hart, 2000b]. In such a multigrid approach, the maximum in-plane displacement is decoupled from the interrogation window size, which increases the dynamic range without decreasing the spatial resolution. The displacements computed with larger interrogation windows can be used as predictions for further interrogations with smaller windows to shift the windows according to the prediction before the next interrogation is calculated. Thus, a high signal-to-noise ratio can be maintained also with small interrogation windows. The size of the interrogation windows may be decreased down to a correlation of single particles [Rehm and Clemens, 1999; Theunissen et al., 2004].

Since in such iterative methods, the quality of the final result depends on the results of previous iterations, in particular the first iteration, validation methods (section 3.1.4) should be applied after each iteration. Since the first iteration will generally be a standard PIV correlation and as such suffer from all the basic limitations mentioned in section 3.2.1, more sophisticated methods have been developed for the first iteration, e.g. [Lin and Perlin, 1998].

Deformable windows resp. higher order approximations of the displacement field.

In the standard PIV evaluation, interrogation windows with fixed size and shape are used, and the velocity field is assumed to be constant within the windows, which is a zeroth order approximation of the velocity field, see section 3.1.3. To account for spatial variations of the velocity within the windows due to velocity gradients and higher order effects, deformations of the interrogation windows resp. of the particle images have to be considered, corresponding to a higher order approximation of the velocity field. Towards this end, Huang et al. [1993a,b] introduced the 'Particle Image Distortion' technique: they use fixed interrogation windows, but apply an iterative deformation of the images to compensate for in-plane loss-of-pairs. In each iteration, the image area within the interrogation window is deformed ('warped') according to the displacement field calculated in the previous iteration. A similar approach is proposed by Nogueira et al. [1999]. The window deformation may also be combined with a multigrid approach [Scarano and Riethmuller, 2000]. To compute the deformed images, some kind of image interpolation has to be applied, e.g. bilinear interpolation or spline interpolation [Jähne, 1999c]. Care has to be taken in the interpolation step, in order not to spoil the accuracy gain due to the window deformation with an inaccurate interpolation scheme.

A further advantage of the window deformation using image interpolation is the possibility to introduce continuous window offsets, which reduces the peak-locking effect, see section 3.2.1.

Second-order correlation.

As an effective method to suppress false correlation peaks and amplify the correct one, Hart [2000a] introduced the 'second-order correlation', which is simply a multiplication of the correlation plane of an interrogation area by the correlation plane of one or more neighbouring interrogation areas (overlapping by e.g. 50%). Thus, it is a 'correlation of the correlation'. Since any peak that does not appear in both planes is eliminated, correlation anomalies are suppressed, resulting in more reliable and accurate velocity estimates. Unlike statistical PIV postprocessing methods to remove spurious vectors [Westerweel, 1994], which rely on the accuracy and similarity of neighbouring vectors, errors are directly eliminated in the correlation data. The second-order correlation may be applied together with any of the PIV methods discussed in this section to validate the results already during the computation step.

3.2.3 3-D/3-C PIV

Several methods have been proposed to extend the PIV technique towards measurements of full 3-component vectors resp. measurements within a three-dimensional volume in space. Stereoscopic PIV is the one that is applied most frequently.

Stereoscopic PIV.

Stereoscopic PIV enables the measurement of 3-C vectors within a plane in space. Hence, it is a 2-D/3-C method. For a review, see [Prasad, 2000]. The basic idea is to use two cameras observing the light sheet, and to compute the third velocity component (i.e. the out-of-plane motion) from the disparity map between the two particle images. Further, stereoscopic PIV also offers the possibility to eliminate perspective errors, which may contaminate the in-plane measurements if perspective effects are strong, i.e. when the lateral dimensions of the object plane are comparable to its depth. Since the latter situation is typical for endoscopic imaging, Reeves and Lawson [2004] suggested the use of a stereoscopic technique for quantitative endoscopic imaging.

Stereoscopic PIV systems can be arranged in two configurations. *Translational systems* have parallel optical axes, whereas in *rotational systems* the two optical axes are arranged enclosing a convergence angle α. Both arrangements have advantages and disadvantages, see [Prasad, 2000].

3-C vectors are obtained by mapping the displacements from each image plane to the object plane and combining them to obtain the third component. There are three different approaches [Prasad, 2000]:

1. **Geometric reconstruction**: A priori knowledge of the complete recording geometry is necessary. This information is used to perform an explicit ray tracing of the projection rays. This method is tedious and not very accurate, since the necessary geometric parameters (e.g. stereo baseline, depth of the measurement plane) often cannot be measured with a sufficient accuracy.

2. **2-D calibration**: A calibration is performed using one image of a calibration target, that has to coincide exactly with the plane of the light sheet during flow measurements. A general polynomial

transformation (typically up to second or third order to account for lens distortions) between the object plane and the image planes of the two cameras is estimated, based on the known correspondences between object and image points of the calibration target. Some authors estimate a plane homography between the object plane and the image plane, instead of a polynomial transformation (see also section 5.3.3). The final step of determining the 3-C velocity uses reconstruction equations that still require some knowledge of the geometry such as separation between the lenses, object distance or angular orientation of the cameras to the object plane.

3. **3-D calibration**: A full 3-D geometric camera calibration is performed, using several images of translated calibration planes. To compute 3-C vectors, explicit knowledge of the system geometry is not required. General higher-order polynomial transformations are also frequently applied in the 3-D calibration. Instabilities related to over-parameterization might be introduced [Hartley and Saxena, 1997] if the measurements are noisy, since typically ≈ 40 free parameters are calibrated for each camera. The application of photogrammetric pinhole camera models and self-calibration methods in stereoscopic PIV is a rather recent development [Wieneke, 2003, 2004].

Defocusing PIV.

[Willert and Gharib, 1992] introduced defocusing PIV as a method to obtain 3-D/3-C velocity fields. A volume illumination is applied, and the defocus principle is used to identify three-dimensional particle locations. Pereirra et al. [2000] use a similar technique to obtain full 3-D information. A volumetric cross-correlation is computed to estimate the velocity field.

Multiplane stereoscopic PIV.

The idea of multiplane stereoscopic PIV is to use several light sheets in different depths to obtain flow information from a number of different planes within a 3-D volume. The planes may be illuminated either simultaneously or sequentially. In the former case, several stereo camera setups are used to acquire the images. For details, see e.g. [Kähler and Kompenhans, 2000]. A recent variant of multiplane stereoscopic PIV is the 'XPIV' proposed by Liberzon et al. [2004]. It combines stereoscopic PIV, multi-sheet illumination and defocusing PIV. The latter is applied to separate the different depth planes which are all projected simultaneously into the same camera.

Photogrammetric PIV.

Schimpf et al. [2003] describe a 'photogrammetric PIV' system. The principle is similar to that of a 3-D Particle-Tracking Velocimetry setup (section 3.5.3). Three cameras are used to acquire images of the flow. The 3-D particle positions are reconstructed by triangulation, based on a geometric camera calibration that is performed prior to the flow measurements. The 3-D/3-C flow field is obtained by computing the volumetric cross-correlation of the particle positions in subsequent frames. The only difference between 3-D Particle-Tracking and the 'photogrammetric PIV' is that the former tracks single particles in 3-D, whereas the latter computes the cross-correlation of 3-D interrogation areas, i.e. volumetric, spatial particle patterns.

Holographic PIV.

In contrast to all the other methods discussed so far, holographic PIV [Hinsch, 2002] requires a volumetric illumination with *coherent* light to record holograms of the flow. The 3-D/3-C flow field is recovered by interrogating the holograms with coherent light beams. In principle, holographic PIV is superior to all the other methods, but the experimental setup and the data evaluation is very complex. For these reasons, holographic PIV does currently not provide the ability to collect large data bases for statistical analyses. Hence, the application of holographic PIV is limited to relatively simple flow configurations.

3.3   Least squares matching

3.3.1   Basic principle

Least squares matching is an alternative approach to maximizing the cross-correlation between two image patches to estimate the optical flow. Like correlation techniques, it also belongs to the region-based methods of motion estimation. Given two successive images $g_1 = g(\boldsymbol{x}, t_1)$ and $g_2 = g(\boldsymbol{x}, t_2)$, an interrogation window is selected in the first frame, and a larger search area centered around this interrogation window is selected in the second frame. The displacement of the interrogation window is calculated by *minimizing* a distance measure that quantifies the dissimilarity between the two image regions. This distance measure is given by the **sum-of-squared-differences** (SSD, 'least squares') of the gray values within the interrogation window between the first and second frame:

$$d(\boldsymbol{x}, \boldsymbol{s}) = \int\limits_{-\infty}^{\infty} w(\boldsymbol{x} - \boldsymbol{x}')[g_1(\boldsymbol{x}') - g_2(\boldsymbol{x}' - \boldsymbol{s})]^2 d^2 x', \qquad (3.11)$$

where the weight function $w(\boldsymbol{x} - \boldsymbol{x}')$ represents the size of the interrogation window. The optical flow is approximated as

$$\boldsymbol{f}(\boldsymbol{x}, t) \approx \frac{1}{\Delta t} \arg\min d(\boldsymbol{x}, \boldsymbol{s}), \qquad (3.12)$$

where Δt is the time difference between the two successive images. Subpixel precision may be achieved using the same methods as in correlation-based approaches, e.g. by fitting a Gaussian function to the (inverse) displacement peak.

Least squares matching is also referred to as 'image correlation velocimetry' [Tokumaru and Dimotakis, 1995], 'adaptive least squares correlation' [Gruen, 1985], 'MQD method' (minimum quadratic differences) [Gui and Merzkirch, 1996] and 'pattern tracking' [Maas et al., 1994; Deusch et al., 2000]. Least squares matching techniques have also been proposed in computer vision [Anandan, 1989; Singh, 1991].

3.3.2   Relation to other region-based approaches

Least squares matching is closely related to the other region-based approaches of motion estimation, namely optical flow techniques and correlation-based analysis.

The similarity between least squares matching and differential optical flow techniques (see section 3.4) is revealed by approximating $g_2(\boldsymbol{x}' - \boldsymbol{s})$ in (3.11) by a truncated Taylor expansion about $\boldsymbol{s} = 0$ and skipping all terms above first order. The resulting expression is the gradient-based formulation of the optical flow (see

(3.15)). For the case of subpixel motions, the equivalence of first order differential optical flow estimation and least squares matching using bilinear interpolation is shown by [Davis et al., 1995].

Gui and Merzkirch [1996] and Gui and Merzkirch [2000] discuss the relation between least squares matching and correlation techniques. Expanding the squared term in (3.11) shows that this expression contains the (negative) cross-correlation coefficient as used in a PIV evaluation. But in addition, there is a term accounting for non-uniformities in the particle image distribution and a non-uniform illumination. Gui and Merzkirch [2000] show that this term is responsible for the superiority of the least squares matching as compared to conventional correlation-based methods.

3.3.3   Advanced least squares matching

The real strength of least squares matching is revealed when it is combined with similar advanced evaluation methods as outlined in section 3.2.2. In particular, iterative approaches using a coarse-to-fine strategy (section 3.1.2) together with a higher order approximation of the displacement field (section 3.1.3) are widely used and show good performance [Gruen, 1985; Anandan, 1989; Tokumaru and Dimotakis, 1995]. Radiometric effects, i.e. intensity changes, may also be included in the model. The implementation of such methods within a general least squares parameter estimation framework provides the possibility to directly estimate higher order quantities like vorticity and rate of strain. Towards this end, these quantities are introduced as parameters to be estimated within the least squares optimization, see section 3.1.3. Thus, no explicit finite differences of the velocity field have to be calculated, which are very sensitive to numerical errors and noise. In addition, no explicit differentiation of the image data is required. In most cases, only a few parameters are extracted from the optimization (e.g. six parameters for a general 2-D affine transformation), but much more data points are used for the computation (e.g. 256 pixels in a 16×16 interrogation window). Due to this strong redundancy, the least squares matching is quite immune to image noise. Furthermore, the precision and the reliability of the estimated parameters can be easily assessed by a covariance analysis of the least squares result. Over-parameterization can be avoided by selecting a model of the displacement fields based on the significance of the computed parameters. Different models can be used for different interrogation windows, making the method adaptive to the local image structure [Gruen, 1985].

3.4   Optical flow methods in computer vision

The measurement of image velocity or optical flow (defined as the apparent motion of brightness patterns in an image sequence[4]) is also one of the fundamental problems of computer vision [Horn and Schunk, 1981]. Many different approaches to recover the optical flow exist. For reviews, see [Barron et al., 1994; Beauchemin and Barron, 1995]. In this section, we give a short description of a gradient-based technique to compute the optical flow, namely the **structure tensor technique** (section 3.4.1). We use this technique to compute the velocity of sand grains in the periscope images (section 9.2) and as a velocity estimator for the trajectory initialization in the particle-tracking algorithm (chapter 7).

The basic assumption of optical flow techniques is that the image brightness of an object point remains constant in a spatio-temporal neighbourhood:

$$g(\mathbf{x}(t), t) = const. \qquad (3.13)$$

---

[4]In the sense of this definition, the PIV methods of section 3.2 and the matching methods of section 3.3 are region-based methods to recover the optical flow, based on the computational approaches of cross-correlation resp. least squares minimization.

This means, that gray values changes in an image are only caused by the motion of objects. Mathematically, this assumption is expressed in the so-called **brightness change constraint equation** (BCCE), which is derived by a first-order Taylor expansion of (3.13):

$$\frac{dg}{dt} = g_x f_x + g_y f_y + g_t = 0, \tag{3.14}$$

where $\boldsymbol{f}(\boldsymbol{x}, t) = [f_x, f_y]^T = [dx/dt, dy/dt]^T$ is the optical flow, and g_x, g_y and g_t are the spatial and temporal derivatives of the image intensity. If we compute the optical flow based on (3.14), we assume that the apparent motion in the image corresponds to the physical motion of objects in the scene. This assumption may be violated, and we have to keep in mind the following aspects (for further details, see [Haußecker and Spies, 1999; Jähne, 2002]:

- **Aperture problem**: Equation (3.14) represents one equation in two unknowns. Hence, solving the BCCE is an ill-posed problem, and we cannot compute the full optical flow, but only the motion component in the direction of the local gray value gradient. Translations that are parallel to lines of constant gray values are not detectable.

- **Sampling theorem**: The solution for the optical flow from the BCCE is determined by relating spatial gray value gradients to temporal gray value gradients. This relation is only unique, if the temporal sampling theorem (3.2) is valid. Large-scale motions cannot be determined from small-scale gray value patterns.

- **Appropriate image texture**: The optical flow can only be computed in regions of the image that show sufficient spatial gray value gradients. In regions of constant gray value, we cannot determine the motion. On the other hand, gray value gradients caused by fixed pattern noise will bias the optical flow towards lower velocities, since image regions suffering from fixed pattern noise will be interpreted as areas of zero image motion. Hence, a correction of fixed pattern noise (section 4.3) improves optical flow estimations.

- **Illumination changes**: Variation in scene intensity that are not caused by motion violate the basic assumption of the BCCE. Intensity changes related to physical processes may be accounted for by an extended version of the BCCE, see e.g. [Haussecker and Fleet, 2001].

In general, the basic gradient-based optical flow techniques will yield good results for small motions in images with a sufficient texture and constant illumination. In applications with large motions and global illumination changes, correspondence-based methods like PIV (section 3.2) and gray value matching (section 3.3) will perform better.

3.4.1 Structure tensor method

The structure tensor method is one particular technique to compute the optical flow based on the BCCE (3.14). To make the problem well-posed, the motion patterns are assumed to be constant in a local neighbourhood. This assumption turns (3.14) into an over-determined system of linear equations, which can be solved by a Total Least Squares technique [van Huffel and Vandewalle, 1991]. The optical flow $\boldsymbol{f}(\boldsymbol{x}, t)$ is

estimated by minimizing an objective function pooling the BCCE constraints over a small neighbourhood, which is also extended into the time dimension to stabilize the solution [Jähne, 1997]:

$$\boldsymbol{f}(\boldsymbol{x}, t) = \arg\min \int_{-\infty}^{\infty} w(\boldsymbol{x} - \boldsymbol{x}', t - t')[f_x g_x(\boldsymbol{x}') + f_y g_y(\boldsymbol{x}') + g_t(\boldsymbol{x}')]^2 d^2 x' dt', \qquad (3.15)$$

Haußecker and Spies [1999] present a detailed discussion of the structure tensor technique, including its extended versions that take into account brightness variations and higher-order parametric flow models (see section 3.1.3). The extended structure tensor techniques have been successfully applied to estimate image velocities and physical parameters of dynamic processes in a number of scientific applications, e.g. [Haußecker et al., 1998; Jähne et al., 1998; Garbe, 2001; Spies, 2001; Dierig, 2002].

3.4.2   Confidence measures and multiple motion

The computation of the optical flow will fail if any of the underlying assumptions, e.g. sufficient variation in the gray values, constant motion within the local neighbourhood or constant illumination intensity, is violated. Hence, **confidence measures** are necessary that indicate whether the computed flow is reliable. The structure tensor technique yields such confidence measures as an integral part of the results, see [Haußecker and Spies, 1999].

In the context of particle-tracking images, the problem of **motion discontinuities** is of particular importance[5]. The latter may be caused by the volumetric illumination used in a 3-D PTV application, see section 3.5.3. The particles in an image sequence obtained under volume illumination represent more or less independently moving objects. For example, particles at large depths may move in opposite direction to particles that are closer to the lens. However, a local neighbourhood in the image may contain particles of both groups. Since multiple motions are present within such a neighbourhood, the assumption of constant flow is not valid, and the computation of the optical flow fails. Different methods have been proposed to overcome this problem. One approach is to use robust estimators to detect the dominant motion and reject the data stemming from other motions as outliers. By iterating this approach and subsequently removing the data that is in accordance with the actual estimation, multiple motions may be detected and separated [Bab-Hadiashar and Suter, 1998; Garbe, 2001]. Another approach relies on iterative coupling of optical flow estimation and object-based segmentation to extract the independently moving image regions [Memin and Perez, 1998].

The tracking algorithm presented in chapter 7 uses the basic version of the structure tensor technique to estimate the particle velocities. This estimation is only used to initialize the tracking of a particle at the beginning of its trajectory. To avoid motion discontinuities, the computation of the structure tensor is carried out on small neighbourhoods, corresponding roughly to the size of the particles. Obviously, discontinuities can also be present within small neighbourhoods. In this case, the confidence measure will be low and the estimation is rejected. The corresponding trajectory will be initialized in a later frame.

---

[5]This problem is not restricted to differential optical flow computation but is also encountered in other region-based methods like PIV and least squares matching.

3.5 Particle-Tracking Velocimetry

In this section, the feature-based approaches to flow visualization are discussed: the so-called particle-tracking or token-tracking techniques. Individual tracer particle images are the features that are tracked throughout an image sequence. The most important tracking method used to measure fluid flow is **Particle-Tracking Velocimetry** (PTV). Since we also develop a PTV technique in this work, the focus of this section is on PTV. A detailed description of the 3-D PTV algorithm is presented in chapter 7.

Besides PTV, some area-based methods are also referred to as 'tracking' techniques: 'correlation-based tracking', 'least squares tracking' or the 'KLT-tracker'. The latter methods consider the particle resp. gray value patterns within the interrogation windows as features that are tracked. The 'correlation-based tracking' (see [Fincham and Spedding, 1997] and section 3.2.2) is a special PIV evaluation mode, where the interrogation window in the second frame ('search window') is larger than that in the first frame, as opposed to the standard 'correlation-based interrogation' mode of PIV where both windows are of equal size. Thus, in this method, the correlation coefficient is used as a tracking criterion. 'Least squares tracking' is discussed in section 3.3. The 'KLT-tracker' [Shi and Tomasi, 1994] (Kanade-Lucas-tracker) is a differential optical flow method based on the early work of Lucas and Kanade [1981] (see section 3.4). This method performs a 'tracking' in the sense that individual image regions are automatically selected and tracked if the image structure within the regions is sufficient to compute the optical flow based on the BCCE (3.14).

Further information about tracking methods can be found in [Faugeras, 1993], [Murray and Buxton, 1990] and [Zhang and Faugeras, 1992] in the context of computer vision and in [Dracos, 1996] in the context of flow measurement. [Blackman and Popoli, 1999], [Stone et al., 1999] and [Brookner, 1998] are in-depth textbooks on Kalman filtering in the context of radar applications. Many ideas and techniques described there can also be successfully applied in visual tracking applications like PTV.

3.5.1 Standard 2-D PTV

Basic principle.

The basic idea of Particle-Tracking Velocimetry (PTV) is to identify single particle images within an image, segment them from the background, and track them along their trajectories throughout an image sequence. Thus, PTV belongs to the feature-based approaches to motion estimation. A PTV algorithm has to solve the following three tasks: **particle segmentation**, determination of **particle position** and **particle matching**, i.e. solving the motion correspondence problem. Since individual particle images are very similar to each other and cannot be distinguished reliably (e.g. by their shape or intensity), the latter task is the most difficult, due to ambiguities occuring especially for high particle densities. Thus, the particle density in PTV applications is generally lower than in PIV applications.

As a simple example, consider the following 'optimal' conditions for PTV: high image contrast (bright particles on a dark, noise-free background), low particle density, and small displacements. The latter two conditions imply that the mean distance between particles is much larger than their motion between two frames. In this case, a very simple PTV approach may be used: segment the particles by a global intensity threshold, determine their position by centroiding, and match each particle in the first frame to that particle in the second frame, which is closest to its position in the first frame ('nearest neighbour'). However, since the optimal conditions assumed above will rarely be given in real applications, more sophisticated algorithms are needed, in particular for the particle segmentation (see section 7.3) and the tracking (see section 3.5.2).

a   *b*
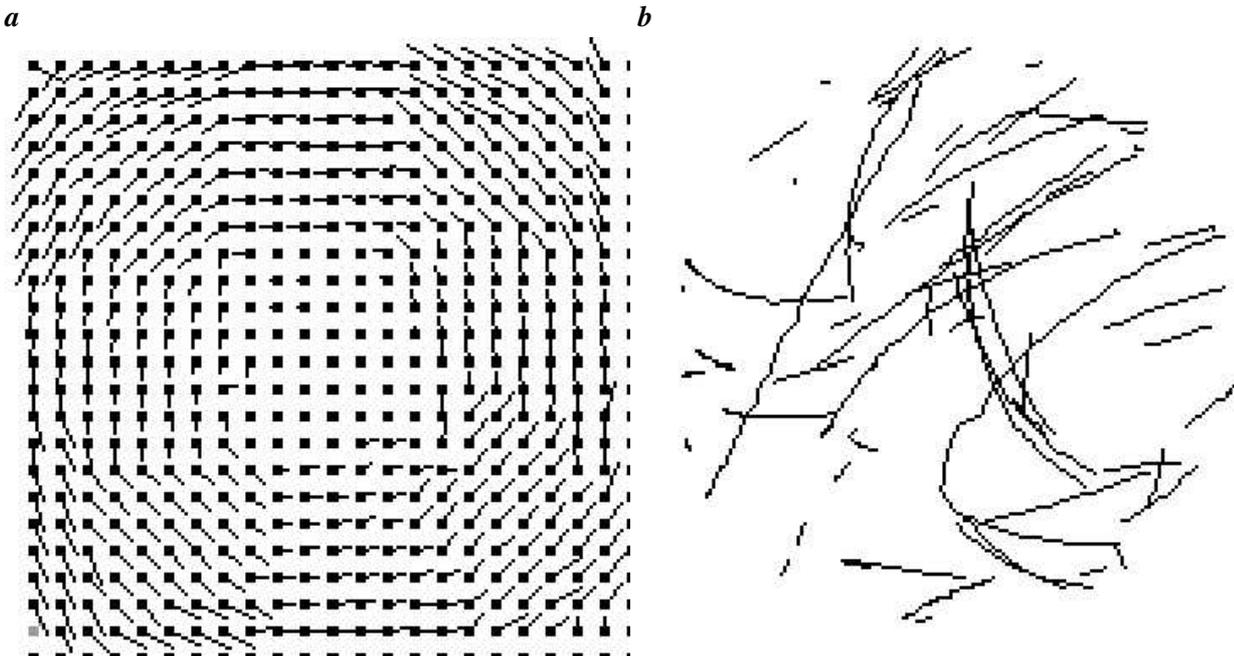
Figure 3.3: *Comparison of typical PIV and PTV vector fields.* ***a*** *PIV vector field of a simulated vortex. The vectors are distributed on a regular grid in space and correspond to a 'snapshot' of the flow field at a fixed point in time (Eulerian approach).* ***b*** *PTV vector field of the flow within a gravel pore. Flow vectors are distributed randomly in space. Note that the plot does not show streamlines (at a fixed point in time), but particle trajectories that correspond to an extended interval in time (Lagrangian approach).*

Differences to region-based approaches.

The major differences between tracking techniques and region-based motion estimation are the following:

- **Temporal scope:**

  To increase the reliability of the tracking, most PTV techniques use more than two successive frames to establish the temporal correspondences ('*multi-frame tracking*'): e.g. three-frame tracking [Papantoniou and Dracos, 1989], four-frame tracking [Hassan and Canaan, 1991], five-frame tracking [Wernet and Pline, 1993]. Note that there are also two-frame tracking techniques [Okamoto et al., 1995; Baek and Lee, 1996; Kim and Lee, 2002] as well as techniques trying to find the optimal set of trajectories by taking into account their complete (visible) length within a global optimization [Sethi and Jain, 1987; Salari and Sethi, 1990; Veenman et al., 2003]. Most correlation-based approaches and least squares techniques try to establish a matching between only two frames. Note that some optical flow techniques also operate on a temporal neighbourhood of more than two frames, e.g. for a more accurate computation of temporal gradients or to stabilize the results by temporal smoothing [Jähne, 1997].

- **Particle density and spatial resolution:**

  Since PTV aims at identifying individual particles and finding corresponding match partners in the next frame, the particle density is generally lower than in PIV, which results in a lower spatial resolution of the underlying flow field. On the other hand, the information given by the particle seeding is

a   *b*
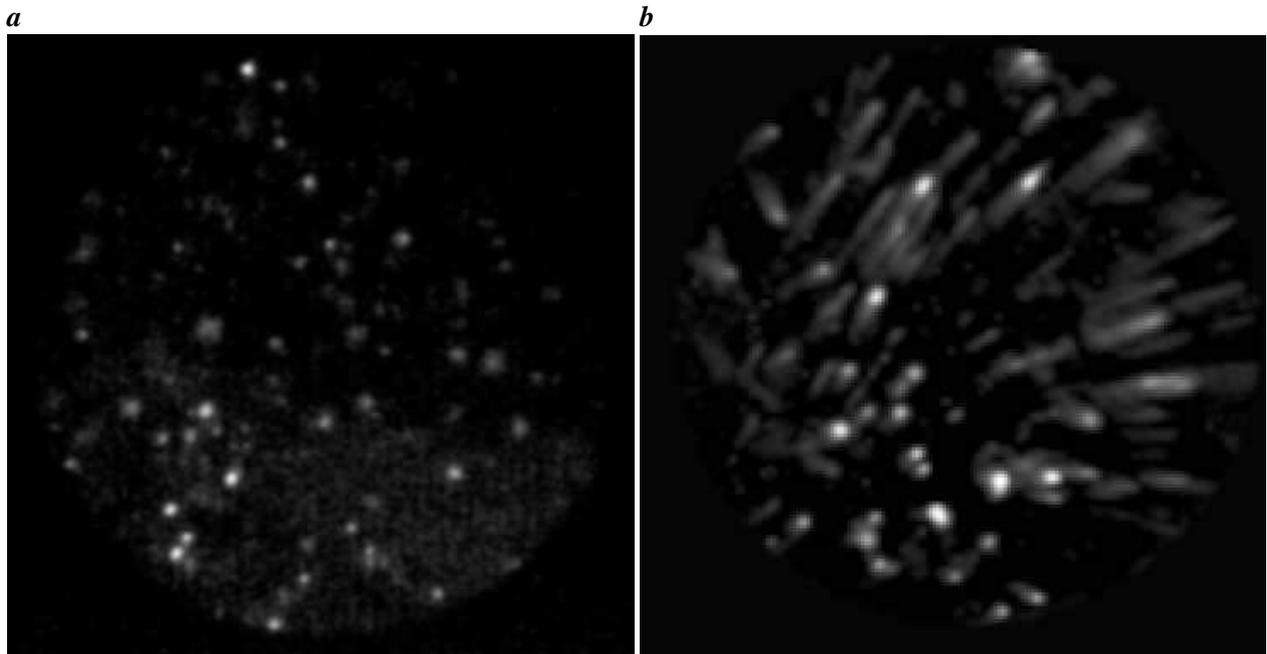
Figure 3.4: *Typical endoscopic tracer particles images of the flow within a gravel pore. These images show that the particle images do not have a homogeneous intensity but cover the complete gray value range. Further, the size of the particles varies considerably. Both effects are caused by the volumetric illumination that is necessary in a 3-D PTV technique.* ***a*** *Slow flow. The particles are rather symmetric.* ***b*** *Fast flow. Due to motion blur, the particles become larger, darker, and elongated along the direction of motion.*

fully exploited without any averaging effects as in PIV (see section 3.2.1), and therefore with a higher accuracy. For example, the motion of two particles located within one PIV interrogation window can be resolved individually. Thus, the local spatial resolution is higher than in PIV.

- **Spatio-temporal distribution of velocity vectors:**

  Feature-based approaches compute the **Lagrangian representation** of a flow field, i.e. the result of such algorithms is a set of trajectories of the tracked objects. Thus, the velocity information is given at random locations depending on the tracer distribution and density. In contrast to area-based methods, no dense, instantaneous flow field defined on a regular grid is computed (which is the **Eulerian representation** of a flow field), see figure 3.3. On the other hand, tracking methods allow to follow the motion of individual particles in time, enabling e.g. a Lagrangian study of diffusion, which is not possible using PIV results.

- **Large motions:**

  Since feature-based methods do not exploit the relation of spatial and temporal intensity gradients for velocity estimation, they are less sensitive to violations of the sampling theorem (see section 3.1.1). Thus, feature-based methods are more suitable for the handling of larger motions, given that the motion correspondence problem can be solved using one of the methods outlined in section 3.5.2.

Particle Segmentation.

In the particle segmentation step, a decision has to be made for each pixel of the image, whether it belongs to a particle or to the image background. The result of the segmentation is a binary image in which the particle images are marked with the value 1 and the background is marked with the value 0. Particle segmentation is rather difficult, because most particle images do not have a bimodal gray value histogram. Therefore, a simple segmentation by a global intensity threshold is not feasible. Particle images do not have a uniform (mean) gray value for the following reasons (see also figure 3.4):

- **Image noise:** Image noise introduces false positives and false negatives, especially in the segmentation of low-contrast particles.

- **Motion blur**: While small and slow particles are imaged as bright circular spots, the shape of the image of a faster particle is elongated in the direction of its motion due to the integration time of the camera. Faster particles cover a larger area in the image, i.e. the irradiance is distributed over a larger number of pixels. Thus, faster particles have a lower intensity and may appear as faint objects with gray values close to the image background.

- **Inhomogeneous illumination:** Several factors may contribute to an inhomogeneous illumination of the images, e.g. the general intensity distribution of the light source, glow of dirty water, or particles reflecting light to their neighbours. In particular, for 3-D applications using a volume illumination, inhomogeneities have to be taken into account.

- **Out-of-focus imaging:** In applications with volume illumination, particles may be out-of-focus, which reduces the image contrast.

The effects of high noise levels, inhomogeneous volume illumination and out-of-focus imaging are especially relevant for miniaturized applications like micro-PIV [Meinhart et al., 1999; Olsen and Adrian, 2000; Gui et al., 2002; Devasenathipathy et al., 2003] or the endoscopic imaging presented here. As a result of all these effects, the (mean) gray values of particle images vary locally (dependent on the depth of the particle) and may cover the complete range from the background noise level to the saturation of the sensor. Similarly, the size of the particle images is depth-dependent. This makes the choice of appropriate segmentation thresholds difficult. A good particle segmentation algorithm should be able to adapt to the local image contrast by using local and adaptive thresholds, e.g. [Kim and Lee, 2002]. Further, the additional feature of the particle *motion* can be exploited to improve the segmentation, see section 7.3.2. All kinds of thresholds should not be prescribed manually as constants determined empirically by the user. Instead, optimal thresholds should be computed automatically, taking into account the statistical properties of the image noise. Many authors have pointed out that the particle segmentation is one of the most crucial steps in a PTV algorithm, since it is the dominating factor that controls both the reliability and the accuracy of the tracking [Perkins and Hunt, 1989; Guezennec and Kiritsis, 1990]. Different methods for the segmentation of particle images are discussed in section 7.3.

Another difficulty is introduced by overlapping particles. Especially in 3-D applications using volume illumination, particle images may partly overlap, or particles may be completely occluded by others. Particle occlusion is a principal physical limitation of 3-D techniques using a volume illumination. A trade-off has to be made between the seeding density (i.e. spatial resolution of the flow field) and the depth of view. Different techniques to resolve overlapping particle images are discussed by [Perkins and Hunt, 1989; Maas

et al., 1993; Guezennec et al., 1994; Carosone et al., 1995]. In section 7.3.3, the method used in this work is described, which is a watershed algorithm from morphological image processing.

Particle position.

Once the particle images have been segmented from the background, different methods are available for the subpixel-accurate determination of the position of individual particles: centroiding [Guezennec and Kiritsis, 1990; Hassan and Canaan, 1991; Wernet and Pline, 1993], least squares fits, e.g. in the form of a Gaussian three-point estimator as applied in PIV [Marxen et al., 2000], fit of extended models taking into account the motion of the particles [Leue et al., 1996], or template matching by cross-correlation [Etoh et al., 1998].

The performance of the different methods depends on the properties of the images, e.g. particle size and density, particle contrast, and image noise. In general, errors in the determination of the particle positions are introduced by the discretization, noise in the signal (photon shot noise), noise related to the image acquisition and deviations from an assumed particle shape, e.g. a symmetric Gaussian. The latter is the most important factor. Jähne [1999a] shows that discretization errors for typical particle images are of the order of magnitude of $0.01 - 0.1$ pixels. Wernet and Pline [1993] show that the Cramer-Rao lower bound for the error in determining the position of a Gaussian-shaped particle is 0.015 pixels. In practical applications with additional noise sources related to the image acquisition and more or less asymmetric particle shapes, the mean centroid estimation error is generally larger ($0.1 - 0.2$ pixels). If highest accuracy is desired, the bias towards integer positions becomes important ('peak-locking', as discussed in section 3.2.1). In this case, iterative weighted least squares methods have to be applied to compute an unbiased estimate.

Computation of velocity.

If a unique particle match is found, a first order approximation of the particle's velocity may be computed by subtracting the positions of the particle in the two successive frames and dividing by the frame period of the camera. Assuming that the latter is given free of error, the absolute error in estimating the particle displacement is $\sqrt{2}$ times the particle position error. Therefore, errors of individual velocity vectors may be higher than in PIV. However, Wernet and Pline [1993] have shown that an accuracy superior to PIV can be achieved by averaging the PTV results over an area of the size of a typical PIV interrogation window. Higher order finite differences [Malik et al., 1993] or spline fits to particle trajectories [Willneff and Lüthi, 2003] may also be used to achieve a better quality of the velocity results. Further, the errors in the particle positions of two nearby particles in successive frames are typically highly correlated. The systematic error cancels out in the difference of the two positions. Hence, the error in the velocity estimation will be lower then the conservative estimate given above, see also section 8.4.

Velocity postprocessing.

Techniques to remove outliers in the velocity field and interpolate the resulting gaps as well as techniques to interpolate randomly spaced data to a regular grid have been mentioned in section 3.1.4. Similar techniqes may be applied to PTV data, where outliers have to be defined relative to a particle trajectory. Malik and Dracos [1995] present dedicated interpolation schemes for three-dimensional velocity fields from scattered data using Taylor expansions. Besides the interpolation to regular grids, the final PTV results (i.e. the flow trajectories) enable a Lagrangian analysis of the flow field, see e.g. [Virant and Dracos, 1997].

Limitations.

An important parameter describing the difficulty of particle tracking is the 'particle spacing displacement ratio' r_p [Malik et al., 1993]:

$$r_p = \frac{\Lambda_0}{\Lambda_t},\tag{3.16}$$

where Λ_0 is the average distance between the particles and Λ_t is the average particle displacement between two successive frames. Tracking is easy, if $r_p >> 1$. In this case, even simple nearest neighbour approaches like the one described in section 3.5.1 may yield good results. The tracking difficulty increases for $r_p \approx 1$, and tracking becomes virtually impossible for $r_p << 1$. In the latter case, the probability of ambiguities in the particle matching is very high. The motion correspondence problem cannot be solved if no additional information about the particles or their motion (e.g. size, shape, color, intensity, direction of motion) is available. The fundamental reason is the violation of the sampling theorem (section 3.1.1) in the case of high particle density and large motion.

Thus, for a given particle density, the tracking difficulty is related to the maximum possible displacement between two frames, which depends on the frame rate, image magnification and the flow field under investigation. For a reliable tracking, Λ_t should be small, i.e. the frame rate of the camera should be sufficiently high. On the other hand, a larger Λ_t yields a lower relative error of the velocity vector, since the absolute error in determining particle positions is independent of Λ_t. The basic idea to enable a reliable particle tracking for values of $r_p \approx 1$ and smaller is to take into account additional information about the flow and use this information to guide the particle matching. Towards this end, most PTV approaches introduce a motion model (section 3.1.3). Some advanced PTV techniques based on motion models are discussed in section 3.5.2.

Another difficulty in PTV stems from the fact, that the number of particles in two successive frames may not be equal. Even if the number of particles is equal, the assumption that the same set of physical particles is visible in both images is not valid. Instead, a PTV algorithm has to handle the following 'events':

- **Entry resp. trajectory initialization:** Particles may enter the field of view. These particles do not have a correspondence partner in the previous frame.

- **Exit resp. trajectory termination:** Particles may leave the field of view. These particles do not have a correspondence partner in the following frame.

- **False positives and false negatives:** Image noise may result in spurious particle images at locations where actually no particle is visible. Segmentation failures may result in a loss of particles, e.g. particles are not segmented due to their low intensity.

- **Overlapping particles and occlusion:** Especially for high particle densities, an overlap and occlusion handling has to be introduced.

Particle entry and exit may take place both in the lateral direction across the image border and in the depth direction along the optical axis. In the latter case, entry and exit occur gradually with an increasing resp. decreasing particle intensity. Since all these events and their consequences have to be considered for each particle in each image of a sequence, PTV algorithms tend to be quite complex. In the following sections, some advanced PTV techniques are discussed.

3.5.2 Advanced tracking techniques

As discussed in the previous section, the most difficult step of a PTV algorithm is to find the unique, correct solution to the **motion correspondence** problem: all particles in an image have to be associated to their correct matching partners in the next frame. This association is difficult for higher particle densities, since the probability of **ambiguities** becomes large, i.e. there will be several possible matching partners within a search region in the next frame. The basic idea of advanced PTV algorithms is to use additional a priori knowledge and assumptions about the flow field. Based on such information, a model of the flow can be introduced. Further, kinematic constraints can be used to reduce the number of corresponding particles and thus the number of ambiguities in the matching. The reduction of the number of ambiguities becomes possible for two reasons. First, using a model of the flow field enables the **prediction** of the future position of particles by extrapolation. Thus, the search region for the correspondence search may be located at this predicted position. Second, if the flow model is a good approximation of the actual flow field, there will be a high probability that the predicted position is already at the correct location of the match partner. Thus, the size of the search region may be reduced. If there is only one particle within the predicted search region, a unique match is established. If still several matching partners remain, one of them may be chosen as the correct match according to certain criteria, which are again based on additional information or assumptions on the flow.

The main assumption that is made in many tracking algorithms is the **smoothness of the flow field**. It is based on the physical principle of inertia. Due to inertia, the motion of an object will not change abruptly between two frames, given that the frame rate is sufficiently high. Inertia is related to the *temporal* smoothness of particle trajectories. In addition, *spatial* smoothness (or spatial *coherence*) of the velocity field may also be assumed. In particular, for incompressible, viscous flows, the velocity vectors within a spatial neighbourhood will vary smoothly. Thus, velocity vectors next to each other will be similar in speed and direction.

In addition to the basic smoothness assumptions, any a priori knowledge about the flow field may also be incorporated into the tracking. For example, if the flow is known to have a mean bulk velocity, this velocity can be used as an offset in the tracking algorithm. Search radii defining the area where matching particles are supposed to be found can be defined based on **hydromechanic knowledge**, e.g. maximum expected velocities or turbulence scales like the Kolmogorov scales or Taylor microscales [Malik et al., 1993; Virant, 1996]. Such considerations are particularly important for the initialization of trajectories, since for particles entering the field of view, no velocity vector is available that can be used to predict their next position.

Another approach to resolve ambiguities in the correspondence analysis is to take into account several possible matches and defer the decision of the correct match on later frames. Such approaches are realized using techniques of **statistical data association** or **combinatorial optimization**. These techniques can also deal with particle occlusion and thereby resolve crossing trajectories. This is possible, because a larger temporal scope is taken into account when solving the correspondence problem, e.g. a temporal neighbourhood of three previous and three future frames. The correspondence is solved by finding an optimal set of trajectories within this temporal neighbourhood, where optimality is expressed e.g. in terms of trajectory smoothness.

Finally, some remarks concerning the optimal choice of thresholds and other tracking parameters, e.g. size of search regions, shall be made. In many cases, the optimal parameters depend on the flow conditions, i.e. particle density and flow velocity. However, the latter are not constant throughout the whole image

resp. image sequence, but there may be significant variations of these quantities, even within a single image. Obviously, optimal performance of a tracking algorithm cannot be achieved using a fixed set of parameters. Instead, the parameters should adapt to the local flow conditions. For example, it does not make sense to use a search region based on a global maximum velocity constraint within a region of the image where the velocity is very small and the particle density is very high. In this case, a search region based on a maximum velocity constraint will be much too large. It may also be advantageous to adapt the shape of the search region to the flow conditions. For example, if there is a clear main flow direction, the search region should be elongated along this direction. Malik et al. [1993] state that the search radius is the most important and most sensitive parameter in a tracking scheme. Only an accurate choice of the search radius leads to efficient tracking. A PTV method using adaptive tracking parameters has been presented by Kim and Lee [2002].

In the following subsections, some implementations of the ideas developed in this section are reviewed.

Two-frame tracking.

The simplest two-frame tracking technique performs a nearest neighbour search based on a minimum-velocity constraint. Hering et al. [1997] use the spatial overlap of particle images to identify the nearest neighbour. The overlap is caused by overlapping integration times of the even and odd fields in an interlaced camera frame. If non-interlaced cameras are used, the overlap may be created artificially by a morphological dilation. The approach is only feasible for low particle densities.

The performance of two-frame tracking can be increased towards higher particle densities by including a spatial coherence constraint and requiring velocity vectors within a spatial neighbourhood to be similar. For example, Perkins and Hunt [1989] propose a nearest neighbour search for all particles within a certain neighbourhood around a 'main particle'. Corresponding particles are found by minimizing the sum of Euclidean nearest-neighbour distances of all particles within the neighbourhood. Other approaches based on spatial coherence are presented by Okamoto et al. [1995]; Baek and Lee [1996]; Ohmi and Li [2000].

Multi-frame tracking.

Multi-frame techniques are based on the assumption of trajectory smoothness. They use a model of the particle motion to predict the particle positions in the next frame. This model is given by the Taylor expansion of the particle trajectory (3.8). The techniques differ in the degree of approximation in the Taylor expansion. **Three-frame techniques** use the actual frame and one previous frame to compute a first order approximation of the particle velocity. The resulting velocity vector is used to predict the particle position in the next frame, assuming that the velocity stays constant. If several match candidates are found within a neighbourhood around the predicted particle position, the particle with the smallest distance to the predicted position is chosen. This choice corresponds to a *minimum acceleration constraint* on the particle motion. In a similar manner, higher order terms in the Taylor expansion may be taken into account to improve the accuracy of the predicted particle position, e.g. in **four-frame techniques** [Hassan and Canaan, 1991; Wernet and Pline, 1993; Malik et al., 1993].

From the point of view of implementation, imposing the temporal smoothness constraint is actually a matter of defining search areas around predicted particle positions. In most applications, circular areas are used. The radii of these areas are chosen according to the kinematic constraints imposed by the flow model or according to prior hydromechanic knowledge about the flow. For example, to initialize a trajectory,

the search radius in the second frame may be defined by the maximum expected velocity. The centers of the search areas in the third and fourth frame are predicted by extrapolating the model, while their radii may be chosen according to the expected fluctuations in the velocity. The latter can be estimated from the Kolmogorov scales of the flow and the imaging parameters [Malik et al., 1993].

Combinatorial optimization.

The motion correspondence problem between two sets of features in two successive frames may be formulated as a **combinatorial optimization problem**. Based on such a formulation, results and algorithms developed in graph theory and operations research may be applied in PTV.

Given are two sets of particle coordinates P_1 and P_2 (particle images in the first frame at $t = t_1$ and particle images in the second frame at $t = t_2 = t_1 + \Delta t$):

$$P_1 = \left\{ \boldsymbol{p}_{1,i}, i = 1..N_1 \right\}, \tag{3.17}$$

$$P_2 = \left\{ \boldsymbol{p}_{2,j}, j = 1..N_2 \right\}, \tag{3.18}$$

where $\boldsymbol{p}_{i,j} = (x_{i,j}, y_{i,j})$ are the coordinates of particle j in frame i. The particle matching between two successive frames may be described by an association matrix $\boldsymbol{\alpha} = (\alpha_{ij})$, with $\alpha_{ij} = 1$ if particle i in the first frame is matched with particle j in the second frame and $\alpha_{ij} = 0$ elsewhere. The task is to find an **optimal assignment** between the elements of the first and those of the second set.

Optimality is expressed in terms of an objective function d that is linear in the associations it includes between the two sets:

$$d = \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \alpha_{ij} c_{ij}, \tag{3.19}$$

where c_{ij} is the cost for the association α_{ij}. The costs c_{ij} are chosen according to the kinematic constraints discussed in the previous section, e.g. favouring smooth trajectories. The optimal assignment is determined by minimizing the objective function (3.19). This formulation of the two-frame tracking is equivalent to a so-called **bipartite graph matching** or **assignment problem**, which is a basic problem of combinatorial optimization occuring in many applications. Efficient algorithms to solve it have been developed, see e.g. [Cook et al., 1998; Nemhauser and Wolsey, 1999]. For an example of a PTV application based on bipartite graph matching, see [Dalziel, 1993]. Stellmacher and Obermayer [2000] and Krepki et al. [2000] present an interesting approach to simultaneously estimate particle correspondences and a local affine transformation by applying a combined discrete and continuous optimization method. This method has been originally proposed by Gold et al. [1998] for solving point matching problems in statistical pattern recognition.

If the temporal scope of the tracking is extended to three or more frames, the problem becomes a **multi-dimensional assignment problem**. Such problems are known to be NP-complete, i.e. there is no efficient algorithm to compute their solution [Nemhauser and Wolsey, 1999]. However, approximate solutions can be found using greedy search techniques and other heuristics: [Sethi and Jain, 1987; Salari and Sethi, 1990; Hwang, 1989; Chetverikov and Verestoy, 1999; Veenman et al., 2003]. The complexity of the correspondence analysis is significantly increased due to the extended temporal scope. Therefore, combinatorial techniques are computationally expensive. On the other hand, these techniques are able to resolve crossing trajectories and find an optimal set of trajectories also in the presence of particle occlusions and dropouts due to segmentation failures.

Statistical techniques.

Statistical approaches to solve the motion correspondence problem ('**statistical data association**') have been originally developed in the context of radar target tracking and surveillance, where particle tracking is referred to as 'multiple target tracking'. A large number of different approaches and algorithms has been published. For an introduction, see e.g [Brookner, 1998; Blackman and Popoli, 1999; Stone et al., 1999]. A review of statistical data association techniques in the context of computer vision is given by Cox [1993].

Statistical data association techniques are specifically developed to resolve ambiguities in motion correspondences, including events like track initiation and termination, particle occlusion and false positives resp. negatives. Like the other advanced techniques discussed so far, most of the statistical techniques are also based on the two paradigms of prediction of an optimal search region and postponing assignment decisions by examining subsequent frames. The difference of statistical approaches is that they model the motion of a particle over time as a **stochastic process**. The optimal particle state according to the measurements and the underlying model is then computed within the framework of Bayesian inference [Stone et al., 1999]. Within this framework, the **probability distribution of the particle state** is propagated over time and updated by the measurements in each frame. A (stochastic) state vector is introduced, which describes the actual state of a particle, e.g. its position, velocity and acceleration. This state vector follows a probability distribution, the so-called *prior distribution*. When measurements of the particle state become available (e.g. its position resulting from the particle segmentation), this measurement information is converted into a *likelihood function* defined on the particle state space. Likelihood functions are presumed to contain all the relevant information in the observed data. They provide a probabilistically correct method of combining all types of sensor information and incorporating it into a tracker's estimate of the particle state. Note that this may include the detailed physics of the camera response and its noise characteristics. Bayes rule is applied to combine all this information and compute the *posterior distribution* on the particle state by combining the prior distribution with the likelihood function. Finally, optimal particle associations are computed from these probability distributions. All statistical data association techniques may be formulated within this **Bayesian framework**.

The simplest statistical tracking technique is the **Kalman filter** [Welch and Bishop, 2001; Blackman and Popoli, 1999]. It is based on the assumptions of a linear model of particle motion (like (3.8)) and Gaussian distributions of particle state and measurement error. The Kalman filter computes the optimal position of a particle in the next frame by predicting its position according to a model and combining this prediction with a measurement. Note that the prediction also includes the precision of the particle location. Optimality is achieved by taking the errors of the prediction and the measurement into account. For example, the shape of the search region around the predicted position may be chosen according to the covariance of the predicted position, which typically results in elliptical search regions instead of circular ones. Examples for Kalman filters applied in PTV are [Yagoh et al., 1992; Ohba et al., 1992; Takehara et al., 2000]. In the algorithm developed here, we also apply a Kalman filter, which is described in section 7.4.2.

Simple Kalman filters have several drawbacks. They provide a statistically optimal estimate of particle position and thus can guide the tracking and implicitly reduce ambiguities. However, they cannot handle ambiguities explicitly. An extension of the Kalman filter with increased capabilities in the handling of ambiguities is the **Multiple Hypothesis Tracker** (MHT) [Cox and Hingorani, 1996; Blackman and Popoli, 1999]. In the case of ambiguities, the latter takes into account the k-best hypotheses (computed by a Kalman filter). Probabilities for all the hypotheses are computed, and the most probable hypothesis is chosen. Thus,

the MHT is a combined approach based on statistical reasoning and combinatorial optimization.

In complex tracking problems (like turbulent fluid motion), the assumptions of the Kalman filter, i.e. a linear model and Gaussian probability distributions, may be too simplistic. It is possible to relax these assumptions and work with general probability distributions of arbitrary shape as well as arbitrary nonlinear models. This general framework of **Bayesian Multiple Target Tracking** is discussed in the book by Stone et al. [1999].

3.5.3 3-D/3-C tracking

PTV can be extended to 3-D measurements using the same basic techniques as in 3-D PIV, namely stereo-scopic or multi-camera image acquisition. The stereoscopic 3-D PTV technique developed in this work is described in detail in chapter 7 (concerning the image processing) and chapter 9 (concerning the experimental setup).

In addition to solving the motion correspondence problem (temporal correspondence), a 3-D PTV algorithm has to solve the **stereo correspondence problem** (spatial correspondence), i.e. to find corresponding particles in two or more images taken from different viewpoints. If a unique correspondence is established, the 3-D position of the particle can be computed by triangulation of the projection rays. The first 3-D PTV systems have already been implemented between 1984 and 1989 [Chang et al., 1984, 1985a,b; Racca and Dewey, 1988; Adamczyk and Rimai, 1988]. The systems of Nishino et al. [1989] and Papantoniou and Dracos [1989] were the first that could track several hundred 3-D vectors per frame.

There are two different approaches to 3-D PTV. They differ in the order in which the spatial and temporal correspondences are solved. The first approach first performs a particle-tracking in the image planes, resulting in a set of 2-D particle trajectories. Thus, the motion correspondence is solved first. Then, the stereo correspondences of the particle trajectories are analyzed to find the 3-D coordinates [Guezennec et al., 1994; Engelmann et al., 1999]. This is also the approach of the algorithm presented in chapter 7. The other approach first solves the stereoscopic correspondences of single particles to compute their 3-D coordinates. Afterwards, the particle tracking is performed in 3-D space [Nishino et al., 1989; Maas et al., 1993; Malik et al., 1993]. For this approach, three or more views of the flow field are needed in order to resolve stereoscopic ambiguities in the determination of the particle positions. A further recent approach is to combine spatial and temporal information and solve the motion and stereo correspondences simultaneously ('spatio-temporal matching') [Willneff and Lüthi, 2003].

The goal of 3-D PTV is to measure the three components of the velocity vectors within a 3-D volume in space. Towards this end, a volume illumination has to be used instead of a light sheet that is typically used in 2-D applications. The volume illumination introduces a number of difficulties concerning the imaging of particles and the processing of such images:

- **Particle occlusion:** Particles located along the same optical ray occlude each other, which limits the particle density resp. the depth of view. A trade-off between these two parameters has to be chosen according to the goals of the measurement.

- **Projection of a 3-D scene:** A particle image obtained using volume illumination is the 2-D projection of a 3-D scene with a certain depth. Depending on the particle density, this will increase the probability of spatial overlap of the particle images resp. trajectories crossing each other in the image plane. In addition, the assumption of spatial coherence of the projection of the flow field is no longer

Table 3.2: *Comparison of different 3-D PTV implementations. $\sigma_X/\sigma_Y/\sigma_Z$ are the accuracies of the reconstructed 3-D coordinates in the corresponding directions. The tracking efficiency is the ratio of tracked particles to the total number of particles.*

	Tokyo	Zürich	Heidelberg	Eindhoven
cameras	3	$3-4$	2	3
frame rate $[Hz]$	30	$30, 60$	60	25
image resolution	512×512	640×480	640×480	512×512
$\sigma_X/\sigma_Y/\sigma_Z$ $[\mu m]$	$25/25/60$	$20-100$	$30/30/80$	$10/10/30$
vectors per frame	$400-800$	$500-1300$	$100-300$	500
tracking efficiency $[\%]$	$50-90$	$65-96$	$10-30$	60

valid, since particles that are actually far away from each other in the 3-D volume may be projected next to each other in the image. This may result in local motion discontinuities in the optical flow, see section 3.4.2.

- **Out-of-focus imaging:** If the illuminated volume is larger than the depth-of-field of the lens, some particles will be out of focus, resulting in a larger size and lower contrast of these particles. Such effects have to be considered in the particle segmentation, see also section 7.3.

Table 3.2 shows an overview of some 3-D PTV systems developed by different authors at the Universities of Tokyo [Nishino et al., 1989; Kasagi and Nishino, 1990], Zürich [Maas et al., 1993; Malik et al., 1993], Heidelberg [Engelmann et al., 1999] and Eindhoven [Kieft et al., 2002] during the past decade. The system of Engelmann et al. [1999] uses a previous version of the algorithm presented in chapter 7. The performance parameters compiled in the table give a general impression of the performance of typical 3-D PTV systems. Tracking efficiencies are typically in the range between 50% and 90%, and the relative spatial resolutions of the object volume are typically $1 : 4000$ in the lateral direction and $1 : 2000$ in the depth direction. Depending on the tracking approach, several hundred up to more than thousand 3-D vectors can be obtained per frame (using standard CCIR or RS170 cameras).

Further developments and applications of the system of Maas et al. [1993] have been carried out by Virant [1996], Stüer [1999], Lüthi [2002] and Willneff [2003]. The system of the University of Tokyo has been upgraded recently by three high-definition CCD cameras with a resolution of 1920×1024 pixels [Suzuki et al., 2000]. Other applications of 3-D PTV are presented by Hardalupas et al. [2000]; Ortiz-Villafuerte et al. [2000]; Doh et al. [2000]. La Porta et al. [2001] present a PTV system based on silicon strip detectors. This system achieves frame rates of up to $70000\,Hz$, enabling Lagrangian acceleration measurements in high Reynolds number turbulence.

3.6   Hybrid methods

Clearly, the maximum amount of information contained in a particle image is the motion of the individual particles. The number of particles within an image, i.e. the particle density, defines the maximum spatial resolution of the velocity field that can be achieved. The approach of Particle-Tracking Velocimetry (section 3.5), is to actually exploit the maximum resolution by identifying the individual particles and measuring

their motion. However, such an approach is not feasible in the evaluation of PIV images, since the particle density is much higher than in PTV images. A high particle density gives rise to ambiguities in the temporal correspondence analysis of the particle motion that can not be resolved without further information. The idea of super-resolution PIV resp. **Hybrid PIV/PTV** is to combine PIV and PTV [Guezennec and Kiritsis, 1990; Keane et al., 1995; Cowen and Monismith, 1997; Kim and Lee, 2002]. Stated more general, hybrid methods refine an initial region-based estimation with a subsequent feature-based estimation, with the goal of combining the advantages of both approaches to achieve high spatial resolution, accuracy and dynamic range at the same time. Typical problems of standard PIV like averaging and gradient-biasing effects are overcome by tracking the individual particles within the interrogation windows. The initial result of a coarse PIV interrogation is used in a predictor step to direct the particle matching algorithm in the right direction and thereby reduce the size of the search area. With a smaller search area located near the correct match partner, the probability of ambiguities is reduced. Bastiaans et al. [2002] present a hybrid method combining a PIV prediction step with a tracking step based on combinatorial optimization to resolve remaining ambiguities. Their method is able to track particles at r_p-values (see (3.16)) as low as 0.5.

The 2-D PTV algorithm presented in section 7.4 also belongs to the class of hybrid methods. It combines two region-based low-level motion estimators (PIV and optical flow) with a feature-based Kalman tracker.

3.7 Summary

In this chapter, a review of state-of-the-art methods to compute the apparent motion in an image sequence, i.e. the **optical flow**, has been presented, focusing on quantitative flow visualization applications. In a strict sense, the term optical flow refers to the continuous velocity field of the apparent motion in a scene that is projected onto an image plane. The sampled and discrete version of the optical flow is the **displacement vector field**. We have discussed different approaches to compute displacement vector fields. They are basically categorized into region-based and feature-based methods. The common assumption of **region-based** methods is the local spatial coherence of the apparent flow field. The three main computational approaches have been discussed: image cross-correlation (PIV), minimization of gray value differences (least squares matching) and gradient-based optical flow techniques from computer vision. Optical flow techniques are generally a good choice to estimate small motions, whereas correlation and matching techniques perform better in the case of large motions and small image structures. In contrast to region-based methods, the **feature-based** techniques do not necessarily assume a coherent flow field, but estimate the motion of independently moving objects. A number of different approaches to track particles throughout an image sequence have been reviewed.

Neither region-based nor feature-based approaches can be considered as perfect choices that can solve any motion estimation problem. They should rather be considered as complementary to each other. This is the basic idea of **hybrid methods**, which try to combine the advantages of region-based and feature-based techniques. Currently, hybrid methods achieve the highest spatial resolution, accuracy and dynamic range.

In their basic 2-D implementations, all the methods discussed in this chapter acquire information from a planar slice of the flow field, that is illuminated by a light sheet. Strictly speaking, if we want to infer quantitative information about a physical flow field from a series of 2-D images, we have to assume that the flow is two-dimensional and that the images are created by orthographic projection, so that there are no perspective distortions. Only in this case will the optical flow represent a scaled version of the original flow. If there is significant out-of-plane motion in the flow or the images show strong perspective effects or
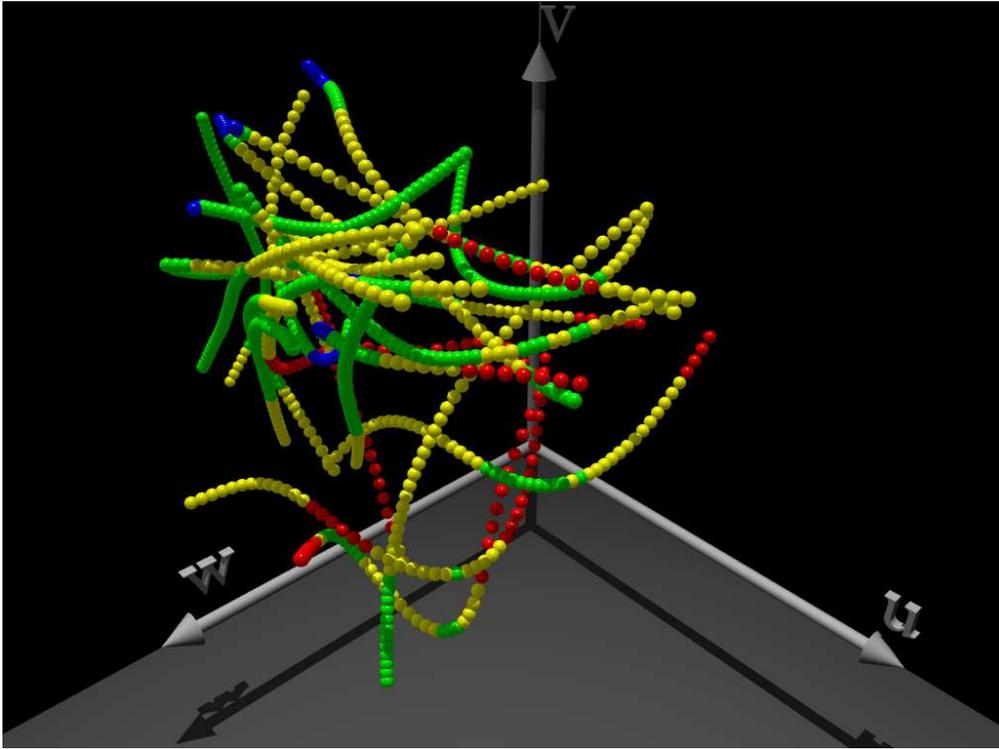
Figure 3.5: *Lagrangian flow trajectories of the pore flow within a gravel pore, obtained from the measurements performed in this thesis using the algorithm of chapter 7. The different colors represent the speed of the particles: from blue (slow) via green and yellow to red (fast). The trajectories show that the flow field is three-dimensional. Hence, a 3-D/3-C technique is necessary to perform quantitative measurements.*

other distortions, large errors will be introduced if we estimate the real flow velocities just by scaling the image velocities. The endoscopic imaging of the confined flow in a gravel pore presented in this thesis is an example, where both out-of-plane motion and perspective effects are significant. Because of the interaction of the erratic pore geometry and the intrinsic (turbulent or intermittent) fluctuations of the flow, the flow field in the pore is expected to be strongly three-dimensional. This has been confirmed by the measurement results presented in section 10.3, see also figure 3.5. Perspective effects will be strong because of the short focal lengths of the endoscopes. Both effects underline the need for a 3-D/3-C technique to perform accurate quantitative flow measurements.

From an algorithmic point of view, all the methods discussed in this chapter can easily be extended to compute three-dimensional velocity vectors within a volume in space. The challenge of 3-D motion estimation is rather a technological one: the acquistion of 3-D image data. The state-of-the-art technique used for 3-D flow visualization is stereoscopic resp. multi-view imaging. Both 3-D PIV and 3-D PTV approaches have been discussed. While most 3-D PIV methods are still restricted to measurements within a plane (or multiple planes if a scanning light sheet is used), 3-D PTV is the only method that can be realized at relatively moderate expense in a miniaturized implementation and is capable of performing quantitative volumetric measurements. Hence, it is the method of choice for the present application.

Part II

Digital image sequence analysis

Chapter 4

Radiometric image analysis

Digital images can be considered as discrete intensity values (gray values) defined on a rectangular spatial grid. Thus, they provide a combination of **radiometric** and **geometric** measurement data. The aim of image processing is to extract the information of interest, e.g. the 3-D coordinates of tracer particles, from these measurements. In doing so, one has to keep in mind that digital images are **sampled** and **noisy** versions of the underlying continuous intensity distribution of the physical image. In other words, as any measurement data, image data are subject to statistical and systematic errors. In order to ensure an optimal performance of any image processing algorithm in terms of precision and accuracy, it is necessary to know the **statistical** properties of the measurement errors as well as the physical origin of **systematic errors**. Given this knowledge, it is possible to take into account the error statistics in the evaluation (e.g. in a weighted least squares fit) resp. to include terms correcting for systematic errors in the mathematical model.

In this chapter, we analyze the radiometric properties of the cameras, focusing on image noise, i.e. any deviations of the gray values from their 'true' values. In section 4.1, the noise characteristics of the CMOS cameras are analyzed, showing that there is a significant fixed pattern noise in the images. A comparison with CCD cameras is presented in section 4.2. Section 4.3 proposes a method to correct for fixed pattern noise. In section 4.4, we discuss a method to reduce the inhomogeneity of the illumination, both in calibration and particle images. The chapter is summarized in section 4.5.

4.1   Analysis of CMOS camera noise

The goal of this section is to quantify the various noise sources that corrupt the digital gray values in the images, so that

1. systematic (spatial) noise components can be removed by a radiometric correction,

2. the remaining (temporal) noise can be quantified accurately and accounted for in the image processing algorithms.

We have to distinguish between two different types of noise. The first is **fixed pattern noise (FPN)**, which is the spatial inhomogeneity of the gray values in an image acquired under homogeneous illumination. Spatial inhomogeneities of the sensor are caused by technological variations at the fabrication step of the sensor, resulting in slightly different properties (gain, size resp. capacity) of the individual sensor pixels. Since FPN
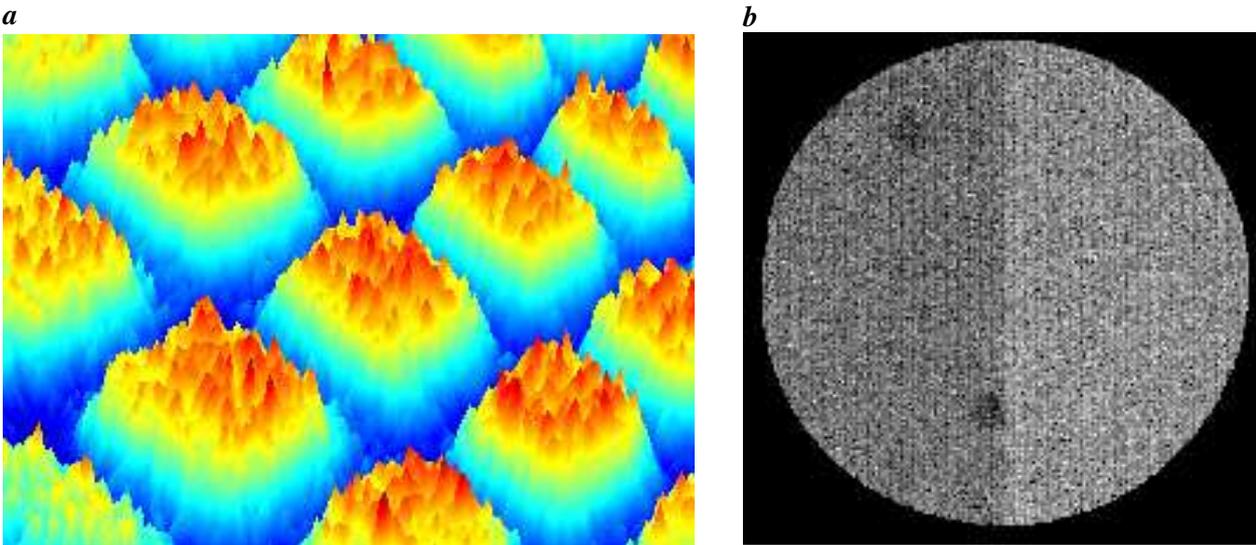
a

b

Figure 4.1: *Typical noise phenomena of CMOS cameras:* **a** *An image of a calibration grid showing strong photoresponse-nonuniformity (PRNU). The gray values are shown as a colored 3-D surface plot. Note that this is a temporally averaged image, thus the PRNU is the only noise source.* **b** *Temporal noise variance in an image taken at homogeneous illumination, coded in gray values. The variance is larger in the right part of the image, probably due to different electronic amplifier or readout circuits. Vertical stripes can be seen, which are introduced by the column amplifiers of the camera. The two darker spots may be caused by dust particles on the sensor.*

is a systematic effect producing a 'fixed pattern' that does not change with time, it can be removed (or at least reduced) by a radiometric correction. We distinguish between fixed pattern noise in dark images taken without any illumination, which is also called *dark-response-nonuniformity* (DRNU), and fixed pattern noise in images taken at non-zero illumination, which is also called *photo-response-nonuniformity* (PRNU). An example of PRNU is shown in figure 4.1a.

The second type of noise is the **temporal gray value variance** of a single pixel. This temporal noise in turn has two different origins, namely the **photon shot noise** and different types of **electronic noise** of the cameras (readout noise, amplifier noise, thermal noise, dark current and quantization noise) [Janesick, 2001; Kamberova, 1996; Healey and Kondepudy, 1994]. An example of electronic noise of a CMOS camera is shown in figure 4.1b. The electronic design of the different amplifiers and readout circuits of a CMOS camera (e.g. column amplifiers) typically causes correlations between pixels.

As we can see from the last example, camera noise strongly depends on the camera type. We have to distinguish between analog and digital cameras, and between CCD and CMOS cameras. For example, one prominent noise source of analog cameras is the so-called *line-jitter*, which is caused by inaccurate synchronization of the sampling frequencies of the camera and the frame grabber (sensor clock resp. pixel clock) [Beyer, 1992]. Line-jitter results in random horizontal shifts of the individual lines of an image, with displacements of up to about 0.5 pixels. Thus, line-jitter seriously jeopardizes subpixel-accuracy if there is no possibility to reduce it by temporal averaging of several frames. Fortunately, there is no line-jitter in digital cameras. However, CMOS cameras typically show a larger PRNU than CCD cameras, because each individual pixel has its own electronic gain circuit, independent of the other pixels. In some CMOS cameras, this pixel mismatch in gain (and also in offset) is corrected already in hardware, e.g. [Loose et al., 1998].
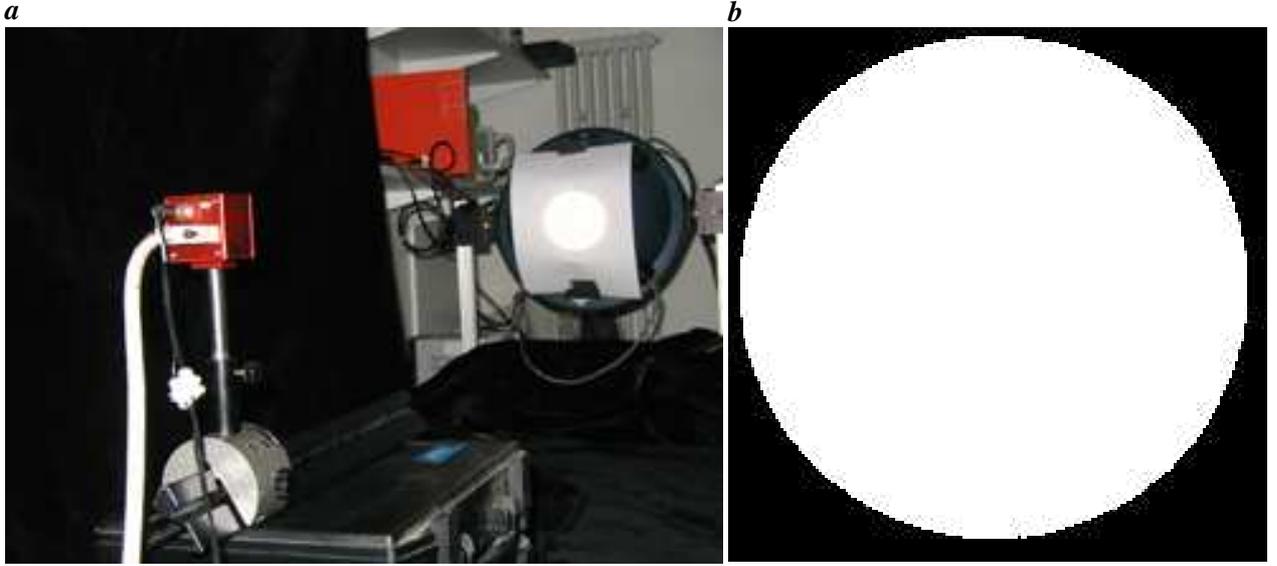
***Figure 4.2: a** Ulbricht sphere used to obtain a homogeneous illumination of the CMOS sensor. **b** Circular field of view of an endoscope.*

In this work, three types of cameras are used: digital CMOS cameras, analog interline transfer CCD cameras and analog progressive scan CCD cameras. In this chapter, the focus is on the CMOS cameras, which are the most important part of the experimental setup. A short comparison with the CCD cameras is carried out in section 4.2. Further information can be found in [Klar, 2004a].

4.1.1 Radiometric sensor model

The noise analysis is based on the following linear sensor model [Jähne, 2002; Jähne and Herrmann, 2003; Gröning, 2003]. As already mentioned, there are two principally different sources of temporal noise: photon shot noise, which depends on the gray value, and electronic noise, which is independent of the gray value. We assume that all the electronic noise sources are independent of each other and normally distributed. They are summed up in a single (temporal) variance $\sigma_0^2(i,j)$.

The gray value dependent noise component has its origin in the Poisson-distributed photon flux. Except for very small fluxes, when only single photons reach the sensor, the Poisson distribution is well-approximated by a normal distribution $N(N_e, \sqrt{N_e})$ (i.e. the variance is equal to the mean value), where N_e is the number of photoelectrons (photon-generated electrons) collected at a single pixel. Thus, the measured gray value $g(i,j)$ and its variance $\sigma_g^2(i,j)$ can be written in the following way:

$$g(i,j) = g_0(i,j) + \alpha(i,j)N_e(i,j), \tag{4.1}$$

$$\sigma_g^2(i,j) = \sigma_0^2(i,j) + \alpha^2(i,j)N_e(i,j) = \sigma_0^2(i,j) + \alpha(i,j)[g(i,j) - g_0(i,j)]. \tag{4.2}$$

The constant $\alpha(i,j)$ is the gain factor of the pixel, in units of gray values per electron. As can be seen from (4.2), the noise variance increases linearly with the digital gray value.

There are basically three different methods to analyze image noise statistically [Beyer, 1992]:

- **Single Pixel Method**: The statistics of single pixels (temporal mean and variance) are collected using a large number of frames.

Table 4.1: *Electronic noise $\langle\sigma_0(i,j)\rangle$ and total noise $\langle\sigma_1(i,j)\rangle$ at $\langle g_1(i,j)\rangle \approx 150$. of the CMOS cameras. To identify the different cameras, we use the name of the corresponding workstation that was used to acquire the data of that camera.*

	Höllbach, left	Höllbach, right	Kocher, left	Kocher, right
$\langle\sigma_0(i,j)\rangle_{min}$	0.0	0.0	0.0	0.0
$\langle\sigma_0(i,j)\rangle$	0.003	0.38	0.08	0.002
$\langle\sigma_0(i,j)\rangle_{max}$	1.21	0.97	1.71	1.23
$\sigma_0(i,j) = 0$	99.4%	22.6%	81.2%	99.5%
$\langle\sigma_1(i,j)\rangle_{min}$	0.3	1.4	0.7	1.4
$\langle\sigma_1(i,j)\rangle$	1.8	1.7	1.3	1.6
$\langle\sigma_1(i,j)\rangle_{max}$	3.6	3.2	5.0	3.3

- **Patch Method**: Subareas of about 20 by 20 pixels are used to analyze spatial and/or temporal statistics.

- **Frame Averaging Method**: This method uses a large number of frames with uniform illumination and computes the statistics from all pixels over the frames.

The basic mathematical operation in all methods is temporal and/or spatial averaging. However, the results of each method clearly depend on the kind of data that is available (images under uniform illumination, static scenes, sequences or single images) and on the assumptions that are made. For example, if only one image with uniform illumination is available and all pixels share the same properties, the patch method may be used to estimate the image noise by spatial averaging. If a sequence of a static scene is available, the single pixel method and temporal averaging may be applied. However, if there are strong gray value gradients in the scene, and the sequence has been acquired with an analog camera, line-jitter may introduce additional large temporal noise components at the positions of large gray value gradients. One has to keep such effects in mind when carrying out a noise analysis.

In the following, we use the single pixel method to obtain statistics for individual pixels by pixelwise temporal averaging of image sequences acquired under constant, uniform illumination, so-called **flatfields**. Such sequences have been obtained by removing the lens from the camera to rule out any vignetting and radiometric falloff introduced by the lens and using an Ulbricht sphere as light source, see figure 4.2a. Since the endoscopes' field of view is circular (see figure 4.2b) and does not completely cover the quadratic image area of the sensor, all black pixels are masked out and only the pixels within the circular endoscope image are used in the image processing.

4.1.2 Temporal noise

Electronic noise $\sigma_0^2(i,j)$.

To measure the electronic noise, sequences of dark images with closed lens cap are acquired. From these sequences, the pixelwise mean and variance is calculated. The variance is equal to $\sigma_0^2(i,j)$ in (4.2). The spatial average of the gray value variances is used as a measure of the electronic noise. The results are
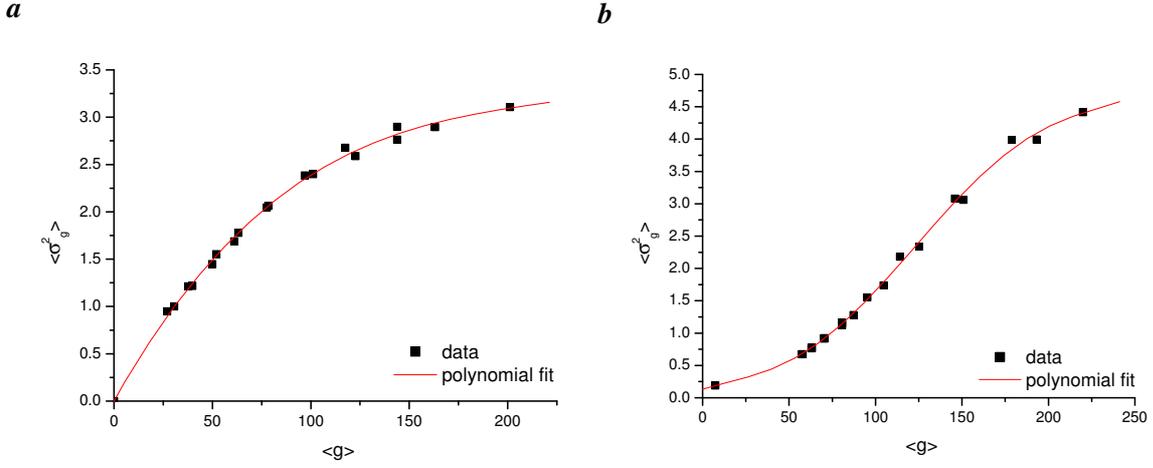
a   **b**

Figure 4.3: *Dependence of the total noise variance on the mean gray value. **a** Höllbach, left. **b** Höllbach, right.*

shown in table 4.1. The electronic noise is quite low, with average values of below 0.5 gray values, which is the manufacturer's specification [Photonfocus, 2003]. Further, we see that there are pronounced differences between the cameras, although they are of the same type and the same camera configuration has been used for all cameras. While the electronic noise is nearly zero for the cameras (Höllbach, left) and (Kocher, right), it is much larger for the cameras (Höllbach, right) and (Kocher, left).

Total noise variance $\sigma_g^2(i,j)$.

The total noise variance is analyzed using sequences of flatfields. Again, we use the spatial average of the variance as a global noise measure of the sensor. Figure 4.3 shows the dependence of the total noise variance on the mean gray value for the Höllbach-cameras. The dependence is not linear as (4.2) suggests. Thus, the gain factor $\alpha(i,j)$ is not constant, but depends on the gray value. This effect is introduced by the *skimming* mode of the cameras, which has been used to amplify the gray values on-chip before readout and other electronic noise adds to the output, see section 9.4.3 and [Photonfocus, 2003].

Note, that again there is a different behaviour of the two cameras, as can be seen from the different polynomial fits to the data in figure 4.3.

4.1.3 Dark-response nonuniformity (DRNU)

Gray value offset g_0.

DRNU is the inhomogeneity of the gray value offsets in a dark image. DRNU is caused by processing errors during sensor fabrication, which introduce small variations in quantum efficiency and charge collection volume [Janesick, 2001]. The DRNU is characterized by a spatial standard deviation $\sigma_{spatial}(g_0)$ of the gray values in a dark image. In figure 4.4, the DRNU is shown, coded in three intensities: black for all pixels with deviations below $\sigma_{spatial}$, gray for all pixels with deviations between $\sigma_{spatial}$ and $2\sigma_{spatial}$, and white for all pixels with deviations larger than $2\sigma_{spatial}$. Both the inhomogeneity and the absolute values of the offset (see table 4.2) are larger for the cameras (Höllbach, right) and (Kocher,left). The offset can be neglected for the other two cameras, since the average offset is well below one gray value. However, there
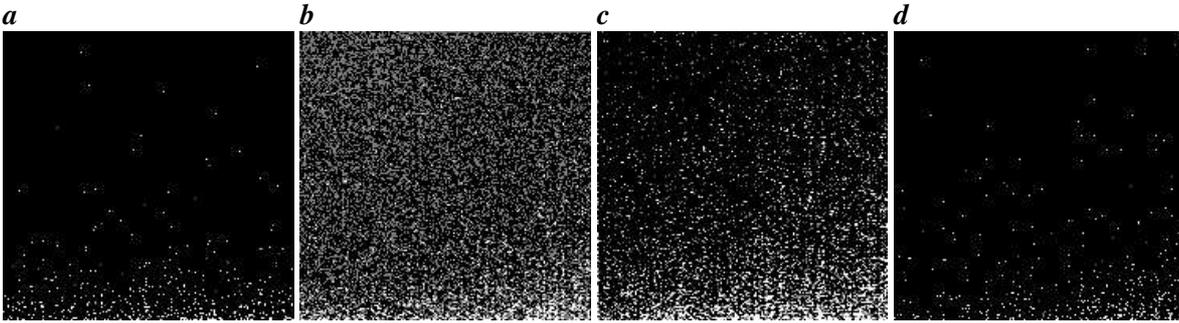
Figure 4.4: DRNU in the dark images: **a** Höllbach, left. **b** Höllbach, right. **c** Kocher, left. **d** Kocher, right.

Table 4.2: Dark-response nonuniformity (DRNU) of the CMOS-cameras: gray value offset $\langle g_0 \rangle$ and its spatial inhomogeneity $\sigma_{spatial}(g_0)$. In the first three lines, both the values obtained right after switching on the cameras and after a period of 48 hours of operation are given. The latter are significantly higher. The last three lines show the percentage of pixels within the intervals $g_0 \pm n\sigma_{spatial}$ with $n = 1, 2$ and outside these intervals.

	Höllbach, left	Höllbach, right	Kocher, left	Kocher, right
$\langle g_0 \rangle$	0.01-0.1	5-12	1-3	0.01-0.1
$\langle g_0 \rangle_{max}$	19-45	49-68	100-158	69-126
$\sigma_{spatial}(g_0)$	0.2-1.0 (0.4%)	6.4-9.0 (3.5%)	2.3-4.8 (1.9%)	0.4-1.0 (0.4%)
$\pm\sigma_{spatial}(g_0)$	98.6%	64.0%	85.3%	98.8%
$\pm2\sigma_{spatial}(g_0)$	0.2%	32.4%	8.7%	0.2%
$> 2\sigma_{spatial}(g_0)$	1.2%	3.6%	6.0%	1.0%

are some pixels with significantly larger offsets, with maximum values of over 100 gray values. Obviously, these are defect pixels or so-called *hotpixels*. They are excluded from all further calculations and their gray values are interpolated from the neighbouring pixels.

Warm-up effects and temporal drift.

In an analysis of camera noise it is important to take warm-up effects into account [Beyer, 1992] in order to ensure that the noise characteristics are constant throughout different measurements, e.g. carried out at different days. The DRNU analysis has been carried out several times during a period of 48 hours of continuous operation of the cameras. Figure 4.5 shows the results of the Höllbach-cameras. Even after two days, there is still a small drift in the mean gray value offset in the dark images. As mentioned above, the offset is very small in the left camera, so we may also neglect the small drift. However, for the right camera, the gray value offset is quite large already after switching on the camera (about 5 gray values). During 48 hours, it increases to about 12 gray values, which is a significant effect, e.g. when subtracting a dark image. Thus, care has to be taken that all cameras are warmed up sufficiently before the acquisition of image sequences.
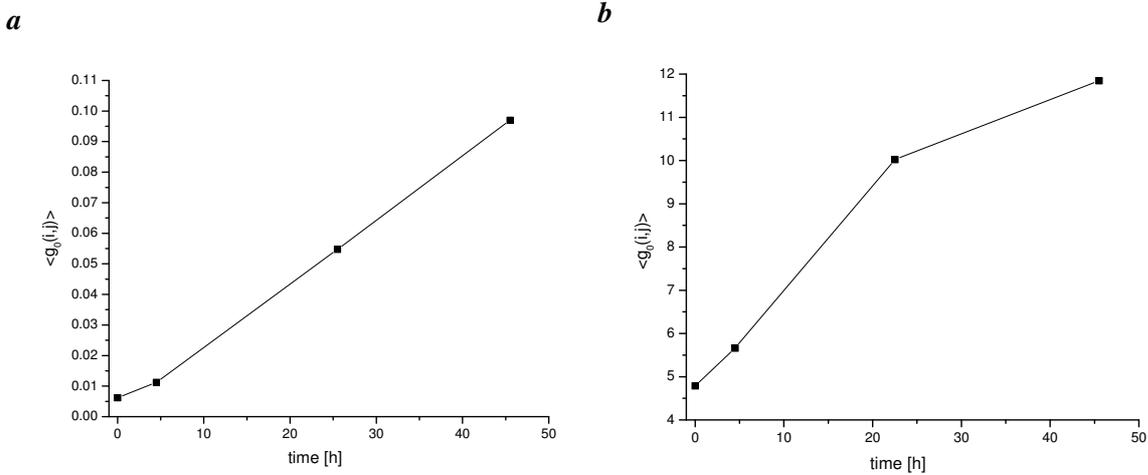
a **b**

Figure 4.5: *Temporal drift in the mean gray value offset (dark image).* **a** *Höllbach, left.* **b** *Höllbach, right.*

a **b**
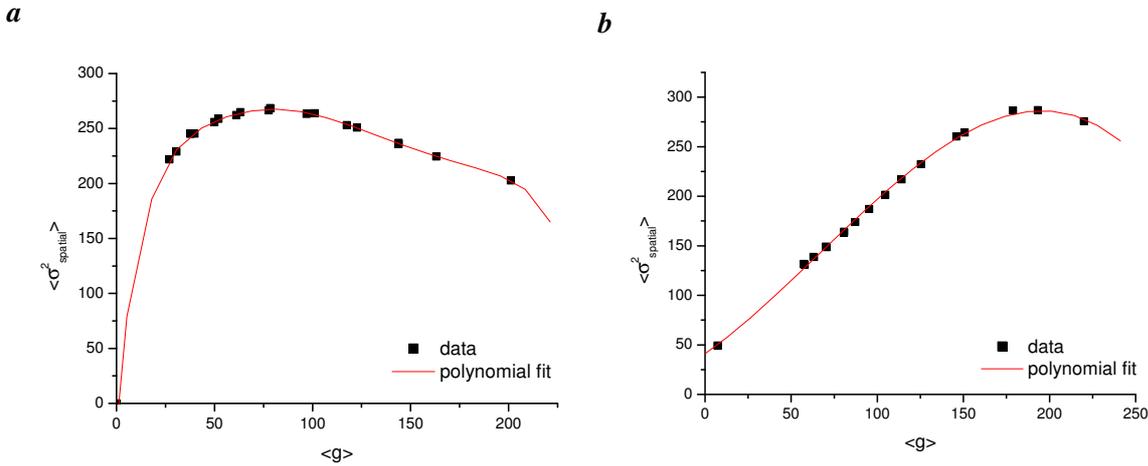
Figure 4.6: *Dependence of the PRNU on the mean gray value.* **a** *Höllbach, left.* **b** *Höllbach, right.*

4.1.4 Photo-response nonuniformity (PRNU)

PRNU is the inhomogeneity of the sensitivity of individual pixels, caused by variations in the gain factor. Even with a perfectly homogeneous illumination, the gray values in an image will not be constant but show a certain variation, which can be characterized by a spatial variance $\sigma^2_{spatial}$. Figure 4.6 shows the dependence of $\sigma^2_{spatial}$ on the mean gray value for the Höllbach-cameras. These data have been obtained from flatfield sequences acquired using the Ulbricht sphere, see figure 4.2a. There are two important conclusions which can be drawn from these plots. First, the PRNU also reflects the difference in the cameras (compare the different polynomial fits). Second, and more important, in a gray value range of about $100 - 150$, which is typical for endoscope calibration images, the spatial inhomogeneity $\sigma_{spatial}$ is about 15 gray values or about 6% of the maximum gray value, which is much larger than the temporal noise. Thus, the main noise contribution in the images is the PRNU. Since this is a systematic effect, we can apply a radiometric correction method to reduce the inhomogeneity. The benefit of such a correction for the performance of a stereo matching algorithm is shown by Kamberova and Bajcsy [1998], where a radiometric correction

a

b
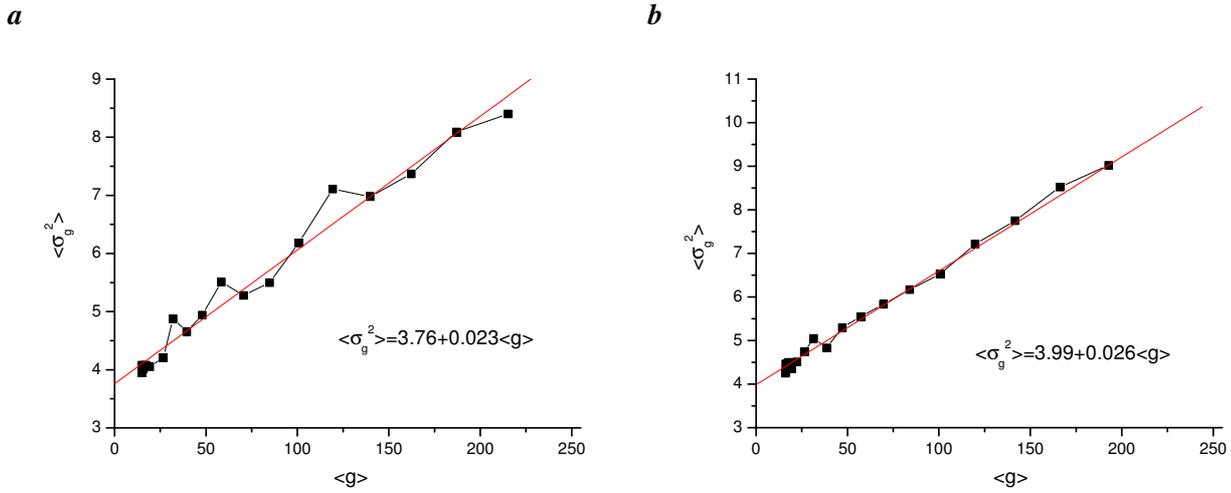
Figure 4.7: *Temporal noise of the Pulnix cameras at a frame rate of* $130\ Hz$. ***a*** *Left camera.* ***b*** *Right camera.*

a

b

c

d
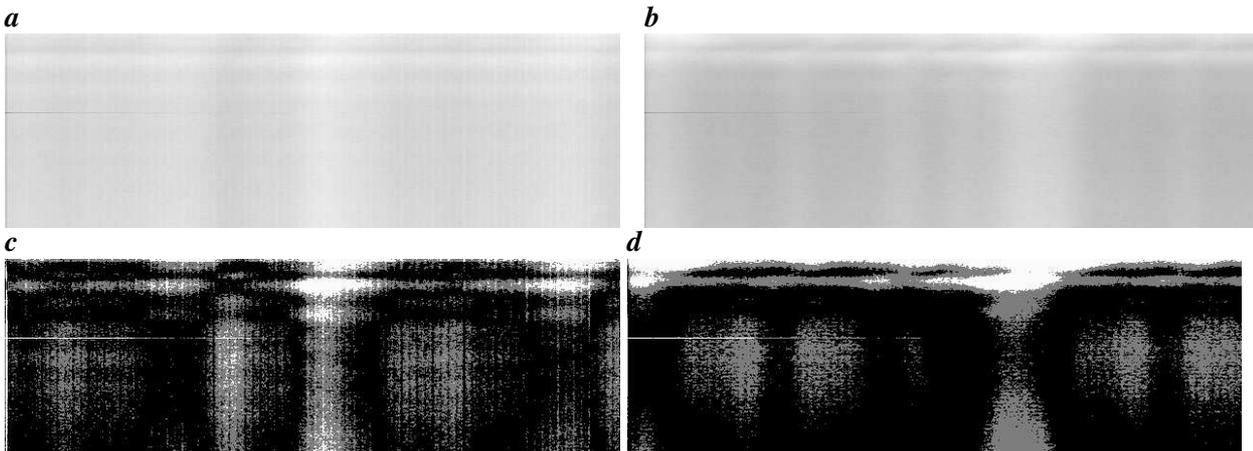
Figure 4.8: *Dark images (**a** left, **b** right, contrast enhanced) and DRNU (**c** left, **d** right) of the Pulnix cameras.*

reduces the relative error in the stereo reconstruction of a 3-D plane by 59%. The correction method applied in this work is described in section 4.3.

4.2   Comparison with CCD cameras

Two different types of CCD cameras are used: Pulnix TM6701AN progressive scan cameras in the Neckar-setup (free surface flow) and Sony XC73-CE interline transfer cameras in the Jagst- and Elbe-setups. All are analog cameras, which are synchronized with the frame grabbers by PLL [Beyer, 1992]. Hence, line-jitter has to be expected as additional noise component. In this section, we take a look at the Pulnix cameras to provide a comparison with the effects of the CMOS cameras.

Figure 4.7 shows the temporal noise of the Pulnix cameras. The variance of the electronic noise is about 4 (gray values)^2, which is significantly higher than that of the CMOS cameras. The noise variance increases linearly with the mean gray value, as expected from (4.2). At large gray values, a maximum noise variance

Table 4.3: *Electronic noise of the CCD cameras: gray value offset g_0, temporal noise σ_0 and DRNU $\sigma_{spatial}$. The last three lines show the percentage of pixels within the intervals $g_0 \pm n\sigma_{spatial}$ with $n = 1, 2$ and outside these intervals.*

	Pulnix, left	Pulnix, right	Sony, left	Sony, right
$\langle g_0 \rangle$	14.9	16.1	16.3	16.5
$\langle g_0 \rangle_{max}$	27	30	22	22
$\langle \sigma_0 \rangle$	2.0	2.1	0.9	0.8
$\langle \sigma_0 \rangle_{max}$	6.1	6.5	1.0	0.8
$\sigma_{spatial}$	0.5 (0.2%)	1.0 (0.4%)	0.4 (0.2%)	0.5 (0.2%)
$\pm\sigma_{spatial}$	70.8%	73.8%	84.4%	84.2%
$\pm 2\sigma_{spatial}$	24.8%	21.0%	8.3%	8.5%
$> 2\sigma_{spatial}$	4.4%	5.2%	7.3%	7.3%

of about 11 (gray values)^2 is reached. This value is roughly a factor 2 to 3 larger than the maximum noise variance of the CMOS cameras.

Figure 4.8 shows the DRNU of the Pulnix cameras. These images show typical electronic interference effects of analog CCD cameras. The systematic patterns can be considered as an electronic 'fingerprint' of the camera/frame grabber configuration [Beyer, 1992]. The mean gray value offset is about 15 gray values, with a spatial standard deviation of 0.5 to 1 gray values. Hence, the offset is much larger than that of the CMOS cameras. Table 4.3 summarizes the results concerning the DRNU of both the Pulnix and the Sony cameras.

In contrast to the CMOS cameras, the Pulnix cameras show a very good homogeneity of the pixel gains. The spatial standard deviation in flatfield images (not shown here) is always below one gray value throughout the whole dynamic range. Hence, PRNU can be neglected.

To summarize, the dominant noise source of the CMOS cameras is fixed pattern noise, whereas CCD cameras suffer from much larger electronic noise. Fixed pattern noise can be neglected for CCD cameras.

4.3   Correction of fixed pattern noise

The basic physical quantity which determines a pixel's gray value is the number of photoelectrons $N_e(i, j)$ that have been generated at that pixel. However, the measurement data obtained from a camera are the gray values $g(i, j)$ of the pixels. From (4.1) we have

$$N_e(i, j) = \frac{g(i, j) - g_0(i, j)}{\alpha(i, j)}.$$ (4.3)

Thus, if we assume $g_0(i, j) = 0$ and $\alpha(i, j) = \alpha_0 = const.$, the gray values and the number of photoelectrons are proportional to each other, with the same proportionality factor at every pixel. In real images, these assumptions are not true, which is the reason for fixed pattern noise. In particular,

- $g_0(i, j) \neq 0$, which is the reason for DRNU or inhomogeneous offsets, and

- $\alpha(i, j) \neq const.$, which is the main reason for PRNU or inhomogeneous gains.

Further, there may also be some deviation from linearity, which means that the gain factors $\alpha(i,j)$ are not independent of the gray value. For example, most CMOS cameras have a logarithmic response. The skimming mode of the Photonfocus cameras is also non-linear, since it amplifies low gray values stronger than high gray values.

If we assume that the camera response is linear, combining (4.1) and (4.2) yields a simple method to determine the gain factors:

$$\alpha(i,j) = \frac{\sigma_g^2(i,j) - \sigma_0^2(i,j)}{\widehat{g}(i,j) - \widehat{g_0}(i,j)}, \tag{4.4}$$

where $\widehat{g}(i,j)$ denotes the temporal average of a flatfield sequence. Thus, taking sequences of identical flatfield exposures at various illumination levels and plotting the variance against the mean, the gain factor for every pixel can be determined as the slope of the resulting curve. This method is called 'photon-transfer technique' and is frequently applied [Mortara and Fowler, 1981; Janesick et al., 1987; Healey and Kondepudy, 1994; Kamberova, 1997]. If only one global gain factor for the whole sensor is to be determined, the spatial variance of the difference of two flatfields can be plotted against the sum of the spatial means of the two flatfields. The differencing removes pixel to pixel sensitivity variations, and the variance of the difference is twice the variance of the shot noise of a single frame. For a linear sensor, the point at which the curve deviates from a straight line is the full well capacity (saturation level) of the sensor. The y-intercept is the electronic noise variance. However, electronic noise is more accurately determined from dark images.

To perform a radiometric correction, the mathematical function describing a sensor's response to an incident illumination has to be known. Basically, this function is given by $\alpha(i,j)$, described by a number of parameters, which can be estimated for each pixel separately, e.g. using the mean and variance data as in the photon-transfer technique. Once the parameters are estimated, we can correct a given image using Equation (4.3). In general, this approach may be difficult, because the sensor response may be a complicated function which is not known a priori. A common assumption for CCD cameras is $\alpha(i,j) = const.$, i.e. the sensor response is linear. Many CMOS cameras have a logarithmic response. To correct the PRNU of such cameras, the photon-transfer technique has to be adapted to the logarithmic response [Tian and Gamal, 2001; Joseph and Collins, 2002]. Gamal et al. [1998] also take into account correlations between noise terms, which are typical for CMOS cameras.

For the CMOS cameras used in this work, attempts to reduce PRNU by modeling the sensor response were not successful, mainly due to the lack of a good model describing the behaviour of all pixels equally well. Thus, a different approach has been taken. In this approach, the model parameters are not recovered explicitly, but the images are preprocessed so that the offset is zero and the photoresponse is uniform. Unlike the photon-transfer technique, this **flatfielding** or **shading correction** does not make explicit use of the Poisson statistics of the photon shot noise. It is explained in the remainder of this section. First, we describe the simple (linear) flatfielding (or two-point calibration), then we extend it to non-linear sensors by using a piecewise linear approximation of the sensor response. A similar approach is taken by Fowler et al. [1998].

PRNU is caused by the spatial variation of the gain factors $\alpha(i,j)$, which we write as

$$\alpha(i,j) = \alpha_0(1 + K(i,j)), \tag{4.5}$$

with $\langle K(i,j) \rangle = 0$. Plugging (4.5) into (4.1) we get

$$1 + K(i,j) = \frac{\widehat{g_1}(i,j) - \widehat{g_0}(i,j)}{\alpha_0 \widehat{N_e}(i,j)} = \frac{\widehat{g_1}(i,j) - \widehat{g_0}(i,j)}{\widehat{g_{ideal}}(i,j)}, \tag{4.6}$$

where $\widehat{g_{ideal}}(i,j)$ is the ideal gray value of the pixel at (i,j) as it would be without any inhomogeneity ($K(i,j) = 0$). To estimate the PRNU-factor $K(i,j)$, we need an estimation of $\widehat{g_{ideal}}(i,j)$. Such an estimate may be calculated by spatial averaging of $\widehat{g_1}(i,j) - \widehat{g_0}(i,j)$ using a binomial filter. The filter mask has to be large enough, so that the mean of $K(i,j)$ over the filter mask is approximately zero. On the other hand, the mask has to be small enough to avoid any influence of an inhomogeneous illumination. Thus, we get an approximation of $K(i,j)$ by subtracting a dark image from a flatfield, smoothing the difference image with a binomial filter of appropriate size and dividing the difference image by the smoothed version of it.

In summary, we have the following linear radiometric correction method:

Given an image $g(i,j)$, a dark image $\widehat{g_0}(i,j)$ and a flatfield $\widehat{g_1}(i,j)$, preferably at maximum illumination, but without reaching saturation:

1. Subtract the dark image from the image:

$$g'(i,j) = g(i,j) - \widehat{g_0}(i,j). \tag{4.7}$$

2. Compute the PRNU correction factor:

$$c(i,j) = \frac{\mathcal{B}(\widehat{g_1} - \widehat{g_0})(i,j)}{(\widehat{g_1} - \widehat{g_0})(i,j)} \tag{4.8}$$

3. Compute the corrected image:

$$g_{corr}(i,j) = g'(i,j)c(i,j) \tag{4.9}$$

In the plots of the noise variance against the mean gray value shown in figure 4.3, the slope of the curves is not constant, i.e. there are significant deviations from linearity in the sensor response. Furthermore, the two curves for the two cameras are qualitatively different, although the cameras are of the same type. Thus, a simple flatfielding correction will not yield good results in general. The correction will only be valid for gray values around the mean gray value of the flatfield used to compute the correction factor. To overcome this problem, we approximate the sensor response resp. the noise variance function by a piecewise linear function. Then, the linear flatfielding can be applied separately for each line segment, using several corresponding flatfields acquired at different illumination intensities to compute the correction factors. A similar method to parameterize the noise variance function has been proposed by Förstner [1998]. The procedure is summarized as follows:

Given an image $g(i,j)$, a dark image $\widehat{g_0}(i,j)$ and several flatfields $\widehat{g_k}(i,j), k = 1..n$, acquired at increasing illumination intensities:

1. Set $lower = 0$. Compute $g'(i,j)$ as in (4.7). Set $g_{corr}(i,j) = 0$.

2. For $k = 1..n$ do:

    (a) Compute the mean gray value $\langle f_k \rangle$ of the flatfield. Set $upper = \langle f_k \rangle$. Compute the correction factor $c_k(i,j)$ as in (4.8).

    (b) Segment all pixels with gray values in the interval $[lower, upper]$ from $g'(i,j)$.

    (c) Correct the segmented pixels by multiplying with $c_k(i,j)$. Add the corrected pixels to $g_{corr}(i,j)$.

    (d) Set $lower = upper$.

a   *b*
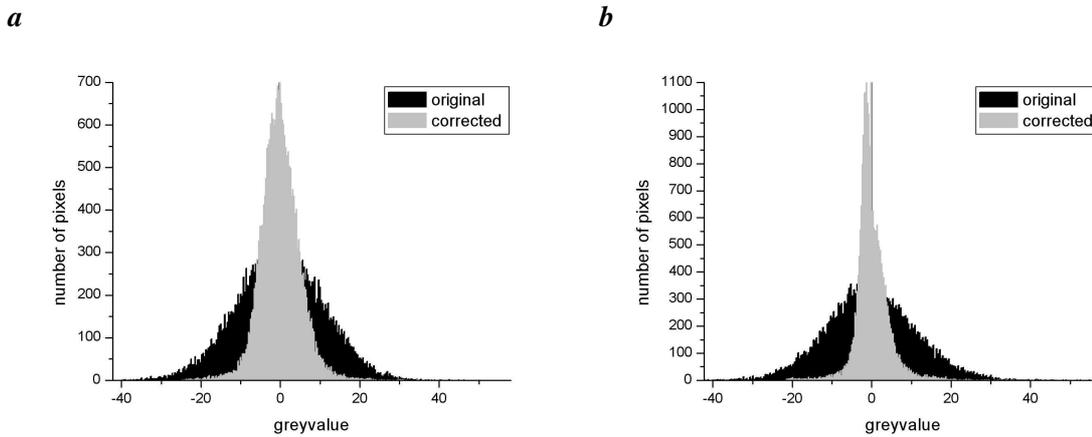

Figure 4.9: *Histograms of the high-frequency gray value noise in a typical endoscope calibration image. The histogram of the original image is shown in black, the histogram of the corrected image is shown in gray. The reduction of the PRNU, i.e. the FWHM of the histograms, can be clearly observed.* **a** *Höllbach, left. Reduction of $\sigma_{spatial}$ from 10.8 to 4.3.* **b** *Höllbach, right. Reduction of $\sigma_{spatial}$ from 11.0 to 3.5.*

a   *b*   *c*
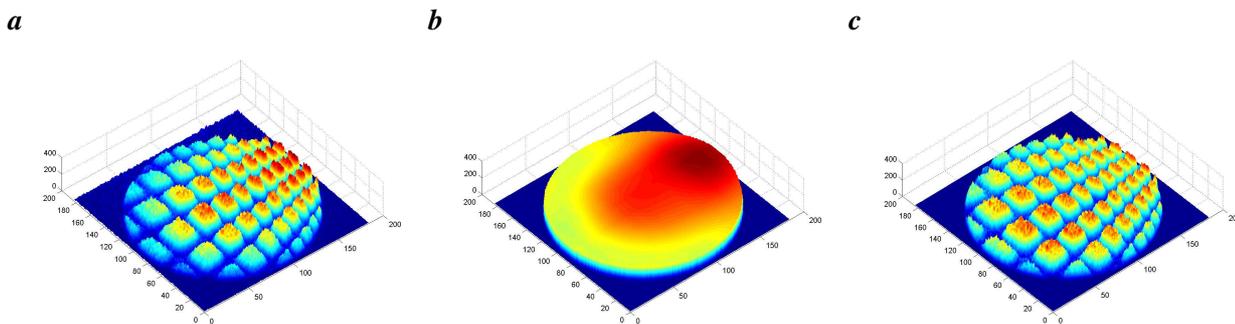
Figure 4.10: *Correction of inhomogeneous background, shown as colored surface plots.* **a** *Original image with inhomogeneous illumination. The average gray value increases towards the right image border.* **b** *Correction (background) image obtained by repeated smoothing and subsampling using an image pyramid.* **c** *Corrected image obtained by division of the original image by the normalized correction image. The inhomogeneity is significantly reduced.*

The results of this flatfielding correction are shown in figure 4.9. The input image is an image of a calibration grid. To extract the high-frequency noise components (corresponding to the PRNU), the images have been smoothed using a 3-tap binomial mask and the smoothed images have been subtracted from the original images. Figure 4.9 shows the gray value histograms of the resulting images. The correction reduces the spatial standard deviation of the noise to less than half of the initial values. The remaining noise amounts to about one to two percent of the dynamic range. Thus, a significant improvement can be achieved if a radiometric correction is applied. The application of the correction to particle images is shown in section 7.3.1.

4.4   Correction of inhomogeneous illumination

In practical applications, it is often quite difficult to obtain a homogeneous illumination of the images. The (average) gray values are higher in some areas of the image than in others. In the case of the endoscope

Figure 4.11: *Result of the radiometric flatfielding correction. **a** Original image showing strong fixed pattern noise. **b** Corrected image. The noise is significantly reduced.*

calibration images, the viewing direction is not perpendicular to the calibration grid, but tilted at an angle of about $45°$, whi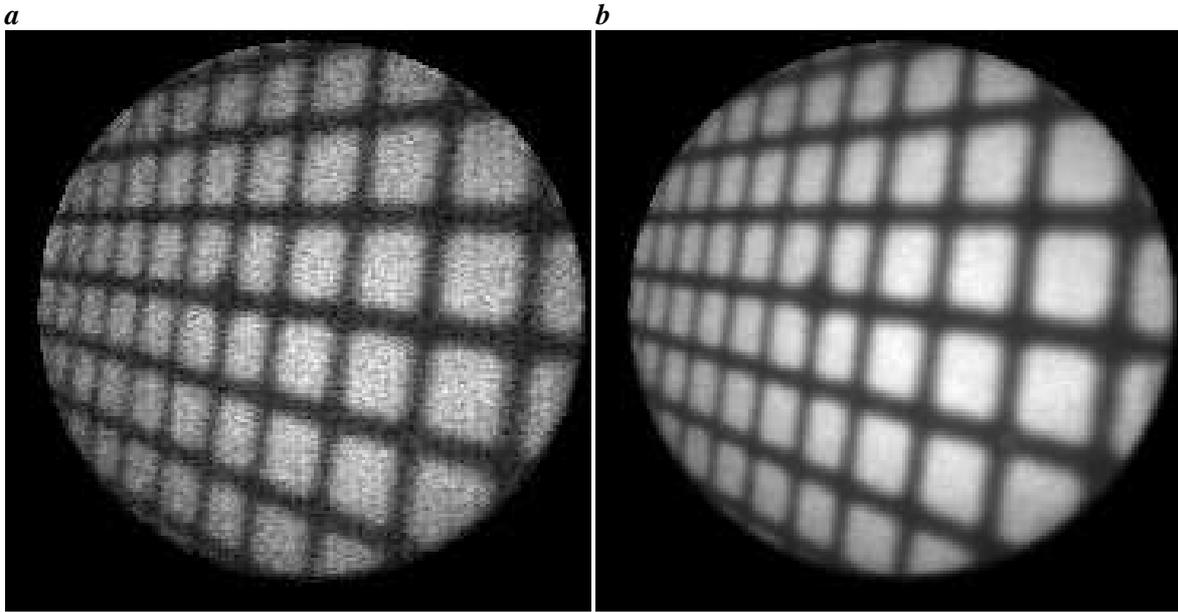ch is the reason for the inhomogeneous gray values, see figure 4.10a. One can compensate for the inhomogeneity to some degree by applying an illumination correction. A low pass filtered image (figure 4.10b) is used as a correction image. It only contains the global gray value structures (i.e. the inhomogeneous background illumination), the fine structures of the original image (grid lines) have been smoothed out. The low pass filter can be realized very efficiently using the highest level of a Gaussian pyramid [Jähne, 1997]. A division of the original image by the correction image (normalized to a mean gray value of one) reduces the illumination inhomogeneity:

$$g_{corr}(i,j) = \frac{1}{c}\frac{g(i,j)}{\mathcal{B}^6 g(i,j)},\qquad(4.10)$$

where c is the normalization constant and $\mathcal{B}^6 g(i,j)$ is the 6th level of the Gaussian pyramid (interpolated to the original size). The result is a more homogeneous gray value distribution in the image, see figure 4.10c. This correction is important, since the model of the grid crossings used to extract their positions by a gray value fit assumes a homogeneous background, see section 5.1. Any gray value gradients violate this assumption and may yield biases in the resulting grid line positions [Steger, 1998].

4.5 Summary

In this chapter, the noise characteristics of the CMOS cameras used in this work have been analyzed, including both temporal noise (photon shot noise and electronic noise) and spatial inhomogeneities (fixed pattern noise). The most important results are summarized as follows:

- Although the four CMOS cameras are of the same type and running under the same configuration, they have significantly different noise characteristics. Thus, each camera's special characteristics have

to be taken into account when evaluating images acquired by that camera.

- After switching on the cameras, the noise characteristics change with time due to warm-up effects, which e.g. cause a drift in the mean gray value offset by up to seven gray values. Before calibrating and performing measurements, the cameras have to be warmed up for a sufficiently long time (at least 24 hours).

- The temporal noise is dominated by the photon shot noise, the electronic noise can be neglected. The gray value dependence of the photon shot noise has been approximated by a polynomial. Thus, we can estimate the noise variance of a pixel, given its gray value, to assess its uncertainty. The explicit estimation of each pixel's expected noise variance will improve thresholding operations, e.g. in the particle segmentation (section 7.3). However, the estimation can only provide a lower bound on the noise, since during the measurements there are additional effects contributing to the temporal gray value variance, e.g. flickering in the illumination, reflections and glow from dirty water or dirt particles. The temporal noise estimated using the Ulbricht sphere is quite low, with a maximum standard deviation of about $\sqrt{5}$ gray values. An alternative to taking into account the explicit gray value dependence of the noise variance is to perform a gray value transformation of the image to make the noise variance equal for all pixels. Such a *variance equalization* has been proposed e.g. by Förstner [1998].

- The most important noise component of the CMOS cameras is the fixed pattern noise. Without any correction, the spatial standard deviation of the gray values in an image acquired at homogeneous illumination can reach values of up to seven percent of the dynamic range. Using the flatfielding correction described in this chapter, the standard deviation is reduced to about one to two percent. A visual impression of this correction is shown in figure 4.11. Since any subpixel-accurate determination of object positions in an image is based on some form of gray value interpolation, systematic errors in the gray values will be translated to geometric errors. Hence, to obtain the highest possible subpixel-accuracy, a radiometric correction is necessary.

Chapter 5

Geometric camera calibration

In order to extract metric 3-D information from image sequences, the relation between image coordinates (i.e. pixel positions in an image) and world coordinates (i.e. positions in 3-D space) has to be known. This relation is given by a model function $\mathbf{f}$ that maps 3-D world points $\mathbf{P}_w$ (in metric units) to their corresponding 2-D image points $\mathbf{p}_p$ (in pixel units):

$$\mathbf{P}_w \mapsto \mathbf{p}_p = \mathbf{f}(\mathbf{P}_w; \theta). \tag{5.1}$$

The camera model is defined by set of **camera parameters** θ, which are a priori unknown. Geometric camera calibration is essentially the task of estimating these parameters.

All calibration methods require the image coordinates $\mathbf{p}_{p,i}$ of feature points in a set of calibration images to be precisely known. Section 5.1 discusses the **feature extraction**, i.e. the precise determination of the position of grid points in the image of a calibration grid. The most common **camera model** is a perspective projection by a **pinhole camera**. This model is introduced within a projective geometry framework in section 5.2. Different **estimation methods** are reviewed in section 5.3, focusing on the methods applied in this work. Calibration results of the different experimental setups are presented in section 5.4. Section 5.5 summarizes this chapter.

5.1 Feature extraction

In any camera calibration method, the image coordinates of a sufficient number of feature points have to be measured in the images. Circular targets, the line crossings of rectangular grids or the corners of checkerboard patterns are used as feature points. To obtain the highest possible quality of the camera parameters, the feature points have to be measured with subpixel accuracy [Lavest et al., 1998]. The most common techniques used for subpixel-accurate feature detection are [West and Clarke, 1990; Luhmann, 2000]:

- **Centroiding**:

  The weighted mean of the pixel positions of an isolated object is calculated, where the weights are the gray values of the individual pixels ('center of gravity'). Best results are obtained for symmetric blob-like objects.

- **Shape-based methods**:

  These methods include line, edge or corner detection as well as extraction of more complicated contours, if the shape of the objects in the image is known, e.g. ellipse fitting.

- **Correlation template matching**:

  The maximum of the cross-correlation coefficient between an image area containing the feature and a given template of the feature is determined. This method is only suitable for the determination of a known target in an image, since good results are only achieved if the shape of the template provides a good match of the shape of the image feature.

- **(Adaptive) least squares template matching**:

  The gray value residual of the image feature and a given template is minimized iteratively. During the iterations, both geometric and radiometric transformations of the template are performed to ensure the best possible match [Gruen, 1985].

- **Model-based least squares matching**:

  The position of the feature is determined by fitting a 2-D model function to the image data. The model describes the spatial gray value distribution of the feature. It is defined in terms of a parameter vector, which includes two parameters describing the exact position of the feature relative to a given, local coordinate origin. A model-based method is used in this work.

The performance of the different methods strongly depends on the application situation (e.g. image quality, size, shape and contrast of the features or properties of the hardware like synchronization of camera and frame grabber). Typical feature extraction accuracies reported in the literature are in the range of 0.01 to 0.1 pixels [West and Clarke, 1990; Beyer, 1992; Gruen, 1997; Steger, 1998; Luhmann, 2000]. Under optimal conditions (homogeneous illumination, symmetric features of high contrast and a diameter of at least 6 pixels, pixel-synchronous sampling, warmed-up cameras), an extremely high precision of 0.004 pixels has been attained by Beyer [1992]. This precision reaches the theoretical precision limit of feature extraction by centroiding in 8-bit digital images [Wernet and Pline, 1993; Schaum, 1993]. It is interesting to note that even on discrete binary images, geometric information like position and length of objects can be determined with statistical errors of only about 0.01 pixels [Jähne, 1999a].

5.1.1   Model-based cross detection

In this work, a planar grid target is used for calibration, together with a model-based feature extraction algorithm. To obtain subpixel-accurate positions of the grid crossings, a model of the image intensity distribution $F(x, y)$ resulting from the perspective projection of two crossing grid lines is needed. Such a model has been proposed by Peuchot [1993]. We use a slightly modified version, which reads

$$f_{cross}(x, y; \theta_{\mathbf{cross}}) = a - d \cdot (1 - e^{-w_1((y-l_2)cos\theta_1 + (x-l_1)sin\theta_1)^2}) \cdot (1 - e^{-w_2((y-l_2)cos\theta_2 + (x-l_1)sin\theta_2)^2}). \quad (5.2)$$

The parameter vector $\theta_{\mathbf{cross}} = [a, d, w_1, w_2, l_1, l_2, \theta_1, \theta_2]^T$ contains eight parameters which are described in table 5.1. An example plot of this model function is shown in figure 5.1a.

Due to the convergent stereo setup, the calibration planes are not parallel to the image plane but are tilted at an angle of about $45°$. Because of the strong perspective distortion, the size, shape and orientation of the

parameter	meaning
a	gray value minimum at the center of the crossing
d	gray value difference between grid lines and background
w_1, w_2	width of the grid lines
l_1, l_2	subpixel displacement relative to local origin
θ_1, θ_2	angles between the grid lines and the sensor lines

a
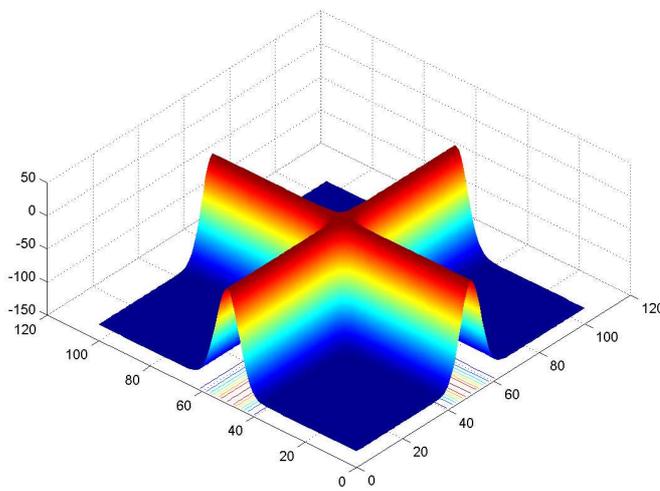
b
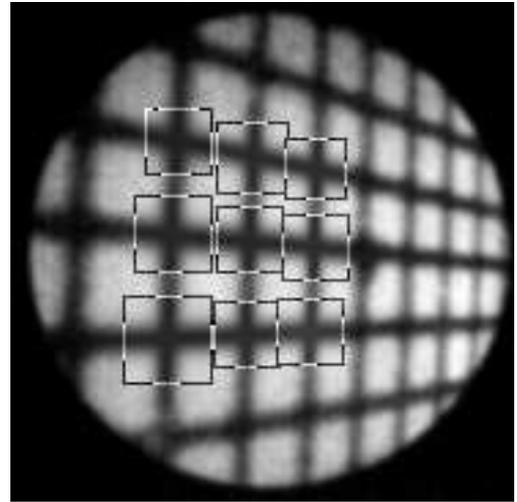
Figure 5.1: *Subpixel-accurate detection of grid line crossings.* **a** *Surface plot of the model function.* **b** *Endoscope calibration image. The result of the subpixel-fit of nine crossings is shown inside the indicated boxes.*

grid line crossings changes significantly even within one calibration image, see figure 5.1b. This makes e.g. a simple template matching infeasible. The model-based approach is more flexible, because the model function adapts to the different size and orientation of the crossings via the parameters w_1, w_2 and θ_1, θ_2. The subpixel-accurate position of the grid crossings relative to a local origin (which also defines the origin of the fitted image area) is given by the parameters l_1, l_2.

The first step of the algorithm is to find the positions of the grid crossings to an accuracy of about one pixel, as well as approximate values for the other fit parameters in (5.2). The user has to specify the origin and the basis vectors of the calibration grid interactively. Given this information, all grid crossings visible in the image are detected automatically by an interest operator similar to the ring-operator used in photogrammetry [Luhmann, 2000]. Due to the quite inferior quality of the endoscope images, the automatic method sometimes detects false crossings. To remove such outliers, we perform a **robust estimation of a homography** (a homogeneous, non-singular 3×3 matrix) between the calibration grid and its image, see section 5.3.3. Robustness is achieved by the Least Median Of Squares estimator [Zhang, 1995], see section 5.3.2. Applications have shown that this method (combining an interest operator with a robust homography fit) is a very stable and reliable method to compute the initial positions of the grid crossings.

a

| σ=5.0 | σ=10.0 | σ=30.0 | σ=50.0 | σ=80.0 | σ=100.0 |

b   *c*
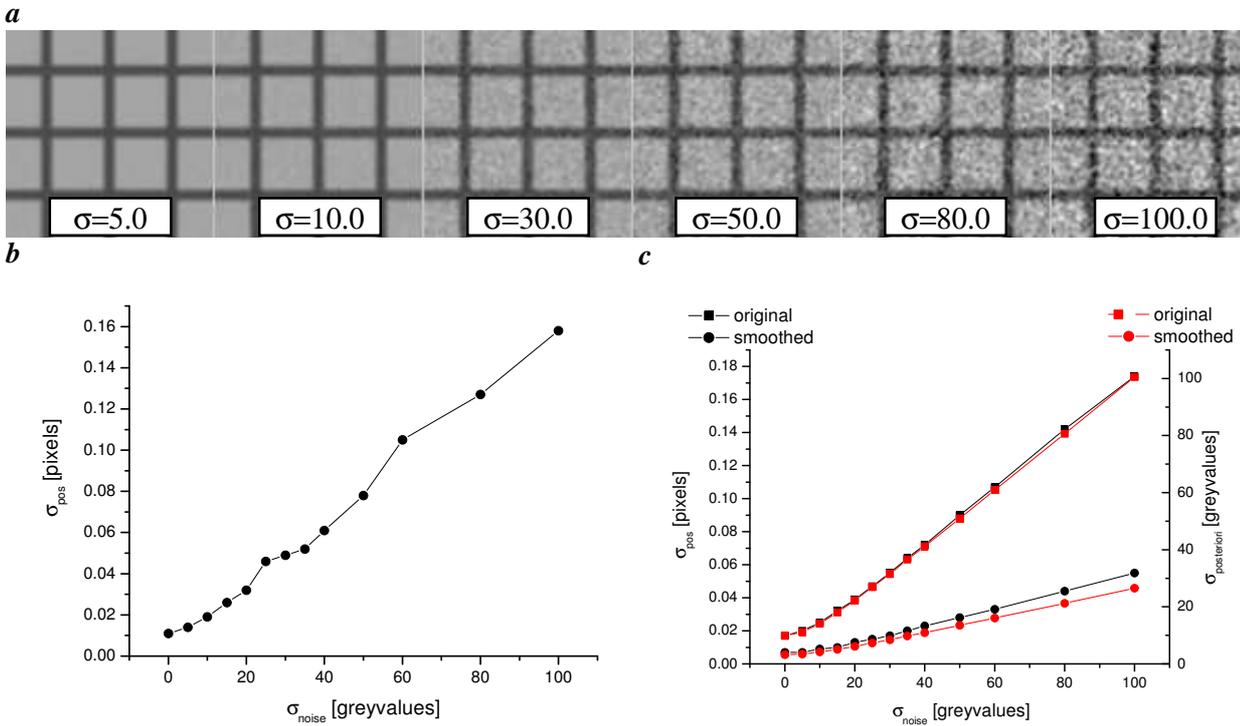
Figure 5.2: *Accuracy and precision of the feature detection for synthetic images with different noise levels. **a** Smoothed example images. **b** Standard deviation of the image residuals in x-direction, computed using ground truth. **c** Left ordinate: Standard deviation of the statistical position errors, estimated by the covariance matrix. Right ordinate: a posteriori standard deviation of the fit result (in gray values).*

A further refinement is carried out by the subsequent nonlinear fit of (5.2). A Levenberg-Marquardt algorithm based on the implementations of Schultz [1997] and Garbe [1998] is used. The result of the algorithm is a list of the image coordinates of the grid crossings, their precision as estimated by the standard error and the covariance matrix of the fit parameters, and their corresponding world coordinates in the target plane. A visual impression of the performance of the method is given in the endoscope calibration image figure 5.1b, where the gray values given by the estimated model are plotted inside the indicated boxes.

5.1.2   Accuracy assessment

To characterize the performance of any computer vision algorithm, the following terms should be distinguished:

- **Precision**:

  The precision is given by the standard deviation of the *statistical* measurement errors. It can be estimated by the covariance matrix of the fit parameters. It is only a measure of the internal errors of the method, which cannot be used to describe the absolute accuracy of the method (e.g. accuracy of the coordinates of a calculated world point in 3-D space). If there are systematic errors in the algorithm, the results have a low accuracy, but still might have a high precision.

- **Accuracy**: Accuracy is a measure of *systematic* errors. It can only be estimated by a comparison of

the results of an algorithm to known ground truth.

- **Reliability**: The reliability of a method describes its robustness against gross errors in the measurements.

Synthetic images.

In the following, the accuracy, precision and reliability of the model-based cross-detection is analyzed. Subpixel accuracy can only be achieved if the image of two crossing grid lines is accurately modeled by (5.2). Otherwise, systematic errors will result in a low accuracy. A number of previous applications has shown that (5.2) is well-suited to model grid targets [Li and Lavest, 1996; Schultz, 1997; Garbe, 1998; Stöhr, 1998; Engelmann, 2000; Klar, 2001]. In [Schultz, 1997], a very high accuracy of 0.002 pixels is reported for synthetic images without noise. In [Garbe, 1998], the accuracy has been tested both with synthetic and real images. For synthetic images, an accuracy between 0.01 pixels and 0.05 pixels is found for noise levels (standard deviation of gray values) between 2 and 10 gray values. The accuracy for real images has been investigated using a precise translation stage. The reported accuracy is 0.03 pixels.

We investigate the accuracy for synthetic images, but applying higher noise levels than in [Garbe, 1998]. Synthetic grid images are generated and Gaussian noise with standard deviations of up to 100 gray values (corresponding to a lowest signal-to-noise ratio of 3.5 dB) is added. As a simple modeling of the imaging process, the grid images are slightly blurred using a 5×5 binomial filter before adding the noise. The grid crossings are extracted from these images using the method described in section 5.1.1. The resulting image coordinates are compared to the known ground truth by calculating the mean differences of the extracted coordinates and their true values. The standard deviations of these differences are shown in figure 5.2. This plot confirms the previous results of Garbe [1998]. Further, the residuals stay below 0.1 pixels for noise levels of up to about 60 gray values. The maximum residuum for a noise level of 100 gray values is 0.16 pixels. For real images of the Photonfocus CMOS cameras (see section 9.3.4), a spatial noise level of up to ≈ 20 gray values has to be expected (resulting from the PRNU, see Section 4.1.4). Thus, the maximum achievable accuracy for the given cameras can be estimated to about 0.03 pixels, which again confirms the results of Garbe [1998] (however, different cameras have been used there).

The accuracy analysis has shown that there are no major systematic errors in the model function. To investigate the statistical errors, we use the covariance matrix of the estimated parameters as a measure of the internal precision of the fit. Figure 5.2c shows the estimated standard deviation of the x-coordinate on the left ordinate, together with the a posteriori standard deviation of the fit result on the right ordinate. To reduce the influence of noise, all calculations have also been done after smoothing the simulated images with a 5-tap binomial filter. The plot shows that the precision is very similar to the accuracy shown in figure 5.2b. In the noisy images, the a posteriori standard deviation of the gray values reflects the applied noise level. The simple smoothing results in a significant reduction of the gray value residuals and hence a higher precision of the estimated positions.

In summary, the analysis of synthetic images shows that both subpixel accuracy and precision are in the range of 0.01 to 0.05 pixels for typical noise levels of real applications.
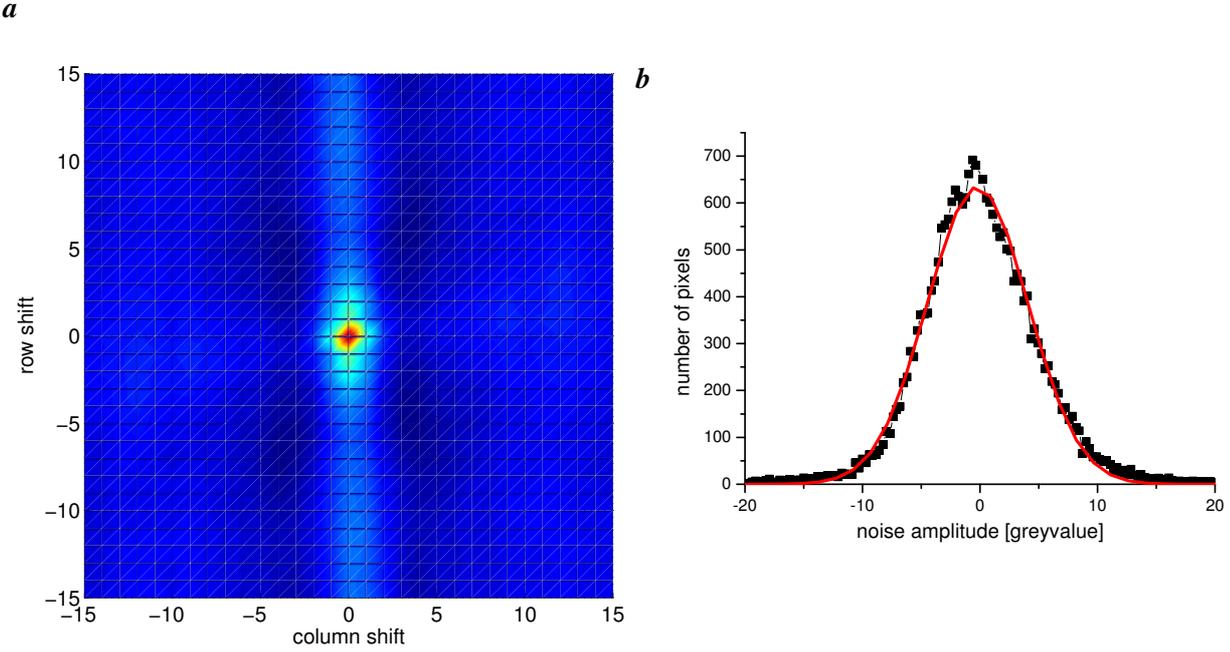
a

*Figure 5.3: **a** Autocovariance function of the remaining image noise after radiometric correction, assuming that the noise in the corrected images can be modeled as a homogenous random field. In this case, the 2-D autocovariance function only depends on distance. **b** Histogram of the image noise after correction. The noise distribution is Gaussian.*

Real images.

Compared to synthetic images with Gaussian noise, there are some additional error sources for real images, including geometric and radiometric errors. Geometric errors may result from electronic effects (e.g. line-jitter for analog cameras) or optical effects of the lens system (e.g. lens distortion). Radiometric errors (i.e. noise in the pixel intensities resp. gray values) arise from inhomogeneous illumination or PRNU of the sensor (see section 4.1.4 and section 4.4). It is important to take the radiometric errors into account [Beyer, 1992], since in any subpixel-accurate feature extraction (which is always based on interpolation of image intensities resp. gray values), radiometric irregularities directly translate into geometric imprecision of the feature positions. As an example, Steger [1998] points out the influence of gray value gradients on the subpixel-accurate extraction of lines and edges. Gradients resulting from an inhomogeneous background introduce a bias into the extracted positions. In [Steger, 1998], a method for unbiased line detection is developed. A subpixel accuracy of about 0.033 pixels is reported for typical industrial inspection tasks.

In this section, we analyze the precision of the feature extraction on real images. To estimate the precision, we use the covariance matrix of the estimated shift parameters l_1 and l_2 of (5.2). To ensure that this is a valid measure, we first take a closer look at the parameter estimation and the statistical properties of the image noise. A Levenberg-Marquardt algorithm is used to estimate the parameters [Press et al., 1992]. This method minimizes the following sum of squared residuals:

$$\chi^2 = \sum_{i,j \in \mathcal{N}} \left(\frac{g(i,j) - f_{cross}(i,j,\theta_{\mathbf{cross}})}{\sigma(i,j)} \right)^2, \tag{5.3}$$

where $x = j\Delta x$, $y = i\Delta y$ are the pixel positions, which are assumed to be free of error. $\mathcal{N}$ is the fit area (neighbourhood of a grid crossing). Each term in the sum is weighted by the pixel's gray value

Table 5.2: *Precision of the feature extraction using different preprocessing methods. For a discussion, see the text.*

method	number of pts.	σ_0 [gray values]	σ_x [pix]	σ_y [pix]	σ_{xy} [pix^2]
raw	32	13.54	0.10	0.10	0.0015
smo	36	7.05	0.07	0.07	0.0010
illu	37	14.08	0.11	0.11	0.0013
illu+smo	38	7.30	0.07	0.07	0.0016
fpn	38	6.29	0.04	0.04	0.0005
fpn+illu	41	6.23	0.04	0.04	0.0006

standard deviation $\sigma(i,j)$. If we assume that the errors of single pixels are independent of each other and follow a Gaussian distribution, the result of minimzing (5.3) will be optimal in the sense of maximum-likelihood estimation, i.e. the probability of the data given the estimated parameters is maximized. A further simplifying assumption that is frequently used is to use only one constant standard deviation for all pixels: $\sigma(i,j) = \sigma_g = const.$. Thus, we have to check the validity of these assumptions.

The noise analysis in section 4.1 has shown that the noise variance is not the same for all pixels. First, the temporal noise depends on the gray value. However, this is not a problem, since temporally averaged images are used in the geometric calibration. Second, and more important, the dominant noise source is the fixed pattern noise. FPN introduces errors, which are also gray value dependent, see figure 4.6. However, this is a systematic effect which can be corrected using the method described in section 4.3. This correction method takes the gray value dependence into account. Since no further information on the remaining noise in the corrected images is available, the assumption of a constant standard deviation is justified. To check the independence of the pixels, the autocovariance function [Jähne, 1997] of a corrected calibration image has been calculated and is shown in figure 5.3a. This figure shows, that the pixel noise is not independent but correlated along columns. This is a typical effect of CMOS cameras, probably introduced by columnwise electronic amplification. However, since the correlation decreases quite quickly, we may neglect it. Attempts to take into account the correlations by minimizing the Mahalanobis distance instead of (5.3) did not improve the results. Finally, assuming independent identically distributed noise, a single pixel's pdf may be estimated by using all the pixels in an image as an ensemble and determining the ensemble's pdf. Figure 5.3b shows that a Gaussian distribution is a good approximation.

To summarize, we compute the maximum-likelihood estimate, assuming independent and identically distributed Gaussian image noise (setting $\sigma(i,j) = 1$). The **standard error** σ_0 of the estimator (i.e. the a posteriori standard deviation of the fit) is used as a measure of the goodness-of-fit:

$$\sigma_0^2 = \frac{1}{N-M} \sum_{i,j \in \mathcal{N}} [g(i,j) - f_{cross}(i,j,\tilde{\theta}_{\mathbf{cross}})]^2, \qquad (5.4)$$

where N is the number of data points (pixels in the fit area), M is the number of fit parameters (here: $M = 8$), and $\tilde{\theta}_{\mathbf{cross}}$ is the estimated parameter vector. Further, we estimate the precision of the estimated cross locations by the corresponding elements of the covariance matrix, i.e. the standard error ellipses, see appendix A.

The results of the precision analysis are shown in table 5.2. In all cases shown in the table, the calibration image is the same. However, different preprocessing steps have been applied, corresponding to the lines of the table: no preprocessing ('raw'), smoothing using a 5-tap binomial filter ('smo'), correction of inhomogeneous illumination as outlined in section 4.4 ('illu') and radiometric correction as explained in section 4.3 ('fpn'). The columns of the table show the number of extracted grid crossings, the standard error σ_0, the mean errors in the grid position σ_x and σ_y and the mean covariance of the x- and y-components. As expected, best results are achieved by applying the radiometric correction followed by the illumination correction. Using this approach, the errors are minimum while the number of extracted crossings is maximum. The position errors are slightly larger than those estimated from the synthetic images, but still quite close to them.

To summarize, this section has shown that the proposed feature extraction method is able to determine the positions of grid line crossings in calibration images with a subpixel accuracy of about $1/25$ pixel. To achieve this precision, it is necessary to examine carefully the noise statistics in the images. Correction procedures have to be applied in order to make the assumptions of the applied estimator valid. Such procedures have been proposed in chapter 4.

The feature extraction algorithm yields a list of the extracted grid positions $\boldsymbol{x}_i$, as well as their covariance matrices $\boldsymbol{\Sigma}_i$, which are needed in the further processing to assess the precision of the grid points.

5.2 Camera model

5.2.1 Linear pinhole camera model

In the pinhole camera model, the image $\mathbf{p}_p$ of a world point $\mathbf{P}_w$ in the scene is determined by the intersection of the image plane and the line connecting $\mathbf{P}_w$ with the **center of projection** or **camera center** $\mathbf{C}$. In the following, we use homogeneous coordinates $\mathbf{P}_w = [X_w, Y_w, Z_w, 1]^T$ and $\mathbf{p}_p = [x_p, y_p, w_p]^T$ (see [Hartley and Zisserman, 2000] and [Faugeras et al., 2001] for an in-depth introduction to projective geometry in the context of computer vision). The Euclidean coordinates of the corresponding 3-D world point are $\boldsymbol{P}_w = [X_w, Y_w, Z_w]^T$, and the Euclidean pixel coordinates of the 2-D image point are $\boldsymbol{p}_p = [x_p/w_p, y_p/w_p]^T$. In homogeneous coordinates, the perspective projection $\mathbf{f}$ of a pinhole camera is given by a linear mapping $\mathbf{f} : \mathbb{R}^4 \mapsto \mathbb{R}^3$ and can be written as

$$\mathbf{p}_p = \boldsymbol{K}\left[\boldsymbol{R}|\boldsymbol{T}\right]\mathbf{P}_w \tag{5.5}$$

with

$$\boldsymbol{K} = \begin{pmatrix} \alpha & s & c_x \\ 0 & \beta & c_y \\ 0 & 0 & 1 \end{pmatrix}, \tag{5.6}$$

$$\boldsymbol{T} = [T_x, T_y, T_z]^T, \tag{5.7}$$

$$
\begin{aligned}
\boldsymbol{R} &= \begin{pmatrix} \cos\kappa & -\sin\kappa & 0 \\ \sin\kappa & \cos\kappa & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\phi & 0 & \sin\phi \\ 0 & 1 & 0 \\ -\sin\phi & 0 & \cos\phi \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\omega & -\sin\omega \\ 0 & \sin\omega & \cos\omega \end{pmatrix} \\[2mm]
&= \begin{pmatrix} \cos\kappa\cos\phi & \cos\kappa\sin\phi\sin\omega - \sin\kappa\cos\omega & \sin\kappa\sin\omega + \cos\kappa\sin\phi\cos\omega \\ \sin\kappa\cos\phi & \cos\kappa\cos\omega + \sin\kappa\sin\phi\sin\omega & \sin\kappa\sin\phi\cos\omega - \cos\kappa\sin\omega \\ -\sin\phi & \cos\phi\sin\omega & \cos\phi\cos\omega \end{pmatrix} \\[2mm]
&= \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix}.
\end{aligned}
\tag{5.8}
$$

The mapping is decomposed into a 3-D rotation $\boldsymbol{R}$ (parameters are the rotation angles κ, ϕ, ω), a 3-D translation $\boldsymbol{T}$ (parameters T_x, T_y, T_z) and a **camera calibration matrix $\boldsymbol{K}$** (parameters $\alpha, \beta, s, c_x, c_y$). Hence, the parameter vector to be estimated is

$$
\theta_{\mathbf{cam}} = [\alpha, \beta, s, c_x, c_y, \kappa, \phi, \omega, T_x, T_y, T_z].
\tag{5.9}
$$

The parameters $(T_x, T_y, T_z, \kappa, \phi, \omega)$ are called the *external camera parameters* or *exterior orientation*, the parameters $(\alpha, \beta, s, c_x, c_y)$ are called *internal camera parameters* or *interior orientation.*

The **external parameters** describe the change of coordinates between the world coordinate system and the camera coordinate system by a rotation and a translation. They contain the information about the relative position and orientation of the camera with respect to a given world coordinate system. In (5.8), the rotation is parameterized by three Euler angles. First, a rotation around the $X-$axis by an angle ω is carried out. Second, a rotation around the (already once rotated) $Y-$axis by an angle ϕ follows. Finally, a rotation around the (twice rotated) $Z-$axis by an angle ω is applied. From (5.8), we have

$$
\tan\omega = \frac{r_{32}}{r_{33}}, \qquad \sin\phi = -r_{31}, \qquad \tan\kappa = \frac{r_{21}}{r_{11}}.
\tag{5.10}
$$

The latter equations show that this parameterization might introduce singularities if $r_{11} \approx 0$ or $r_{33} \approx 0$, which corresponds to the angles $\omega, \phi, \kappa \approx (\pm\pi/2, \pm3\pi/2, ...)$. However, in all experimental setups used in this work, the angles will be approximately $\omega \approx 0°, \phi \approx \pm45°, \kappa \approx 0°$. These angles are far away from the singularities in parameter space, so the chosen parameterization is numerically well-conditioned.

The **internal parameters** describe the change of coordinates between the camera coordinate system and the image coodinate system, i.e. the imaging by a perspective projection. The internal parameters of a basic pinhole camera consist of the coordinates $\boldsymbol{c} = [c_x, c_y]^T$ of the **principal point** and the **focal length** f. The principal point is the point of intersection of the optical axis and the focal plane. Equation (5.6) corresponds to a slight modification of the basic pinhole model. This modification is frequently used in CCD or CMOS camera applications, where the image is produced by a discrete pixel array. In this case, the parameters α and β give the focal length in units of pixels. To account for **different scale factors** in the x- and y-direction, i.e. possible deviations from a quadratic pixel shape, two focal lengths are introduced. If f is the focal length of the lens in units of mm and p_x and p_y are the corresponding pixel sizes, we have $\alpha = f/p_x$ and $\beta = f/p_y$. The parameter s is called **skew** and accounts for a possible non-orthogonality of the axes in the image plane. This might be the case if the rows and columns of pixels on the sensor are not perpendicular to each other. However, for most CCD and CMOS cameras one can assume that $s = 0$.
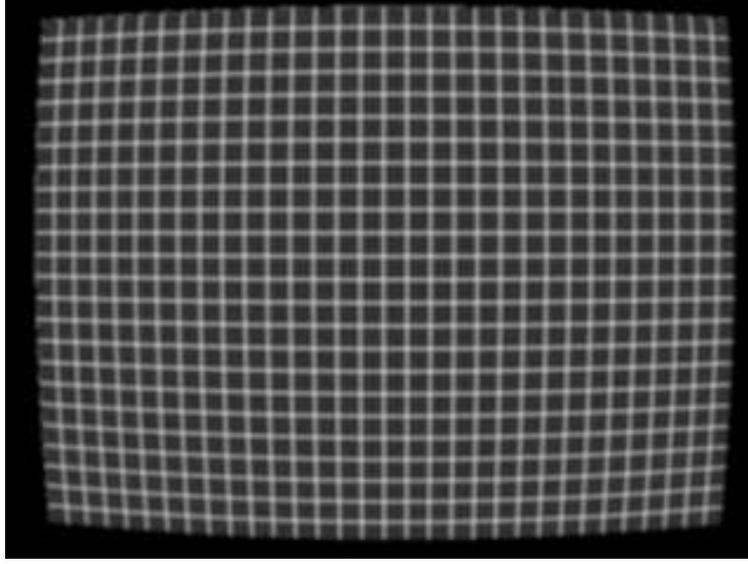
Figure 5.4: *Typical pincushion distortion introduced by the wide-angle lenses of rigid endoscopes (see Section 9.2.1).*

If $s = 0$ and $\alpha = \beta = f_{eff}$ is the effective focal length, (5.5) is equivalent to the well-known **collinearity equations** of photogrammetry [Slama, 1980; Luhmann, 2000]:

$$
x - c_x = f_{eff} \frac{r_{11}X + r_{12}Y + r_{13}Z + T_x}{r_{31}X + r_{32}Y + r_{33}Z + T_z}, \tag{5.11}
$$
$$
y - c_y = f_{eff} \frac{r_{21}X + r_{22}Y + r_{23}Z + T_y}{r_{31}X + r_{32}Y + r_{33}Z + T_z}.
$$

These equations may be derived from the perspective geometry of the imaging process. They express that an object point, its image point and the camera center are collinear, i.e. they lie on a straight line in space (the **projection ray**). In the parameterization (5.5), the coordinates of the **camera center** in the world coordinate frame are given by

$$
\boldsymbol{C} = [C_x, C_y, C_z]^T = -\boldsymbol{R}^{-1}\boldsymbol{T}. \tag{5.12}
$$

Introducing the homogeneous (3×4)-projection matrix

$$
\mathsf{P} = \boldsymbol{K}\left[\boldsymbol{R}|\boldsymbol{T}\right], \tag{5.13}
$$

the mapping from world to image coordinates can be written as a simple linear mapping (a projective linear transformation):

$$
\mathbf{p}_p = \mathsf{P}\mathbf{P}_w. \tag{5.14}
$$

Like any homogeneous quantity, the (3×4)-matrix P is only defined up to an overall scale factor, so there are only 11 independent parameters in P, corresponding to the 6 parameters of the exterior orientation and the 5 parameters of the interior orientation.

5.2.2   Nonlinear lens distortion

Due to its linearity, the pinhole camera model introduced in Section 5.2.1 provides a very simple and practical formulation of the imaging process. However, it is only an approximation of real cameras, since it

does not account for **lens distortion** effects exhibited by real lenses. Lens distortion changes the position of image points from their 'true' location as given by a linear pinhole camera model (5.14). Obviously, it also affects distances between image points and leads to a variation of scale of an image as a function of position in the image plane, since scale at any image point is related to the ratio of the rate of change of position in the image plane to change of position in the object plane. Two types of lens distortion are commonly distinguished [Slama, 1980]: (symmetric) *radial distortion* and (asymmetric) *decentering distortion*, where symmetry means point symmetry with respect to the *center of radial distortion*, which is often assumed to be the principal point.

Radial distortion is not the result of any imperfection in lens manufacturing, but a fundamental property of any lens with a finite aperture ring. The latter results in variations in angular magnification with angle of incidence, see e.g. [Thormählen et al., 2003]. If image points are displaced radially farther from the distortion center, image scale is increased in the outer portion of the image plane (*pincushion distortion*), see Figure 5.4. If image points are displaced radially closer to the distortion center, image scale is decreased (*barrel distortion*). Radial distortion may be as large as 300 μm (corresponding to $\approx$ 30 pixels) at the edges of the image and is usually an order of magnitude larger than decentering distortion [Atkinson, 1996; Gruen, 1997]. Hence, many camera calibration methods applied in computer vision only use radial distortion parameters, e.g. [Tsai, 1987; Lenz, 1987; Lenz and Tsai, 1989; Fryer et al., 1994].

Decentering distortion has a tangential (normal to the radial direction) and an asymmetric radial component. Both result from the decentration of the lens and other optical elements. Any vertical displacement or rotation of a lens element from a perfect alignment collinear to the optical axis will cause a displacement of image points. Hence, decentering distortion might be of importance for compound systems consisting of several lenses and other optical elements, like the endoscopes used in this work. Since decentering distortion may be created by placing an appropriately oriented thin prism in front of a perfectly centered lens, it is sometimes also called *thin-prism distortion*.

To achieve highest accuracy in image measurement, the systematic errors introduced by lens distortion have to be corrected for in the mathematical model. Modeling of lens distortion has been introduced into analytical photogrammetry by Brown [1966, 1971]. The set of camera parameters (5.9) is extended by an additional set of distortion parameters $\theta_\mathbf{d}$, which describe the distorted image locations in terms of a polynomial series.

In the calibration of CCD cameras, where possibly $\alpha \neq \beta$ and $s \neq 0$, the terms modeling lens distortion have to be put in the right place in the projection process. To clarify this, we again decompose the projection matrix P, following Hartley and Zisserman [2000]. The starting point of the projection is the homogeneous vector of world coordinates of an object point $\mathbf{P}_w = [X_w, Y_w, Z_w, 1]^T$, where $[X_w, Y_w, Z_w]^T$ are the Euclidean coordinates of the object point in a given Euclidean world coordinate frame. The end point of the projection is the homogeneous vector of pixel coordinates $\mathbf{p}_p = [x_p, y_p, 1]^T$, with the Euclidean coordinates of the image point (in pixels) given by $[x_p, y_p]^T$. The projection process may be considered as a series of successive transformations of the world coordinates:

$$\mathbf{P}_w \longrightarrow \mathbf{p}_n \longrightarrow \mathbf{p}_u \longrightarrow \mathbf{p}_d \longrightarrow \mathbf{p}_p. \tag{5.15}$$

In (5.15), the four transformations are indicated by arrows, corresponding to the following steps:

1. a rigid body transformation from world coordinates to the so-called **normalized image coordinates**

$$\mathbf{p}_n = [x_n, y_n, w_n]^T = [\boldsymbol{R}|\boldsymbol{T}] \, \mathbf{P}_w, \tag{5.16}$$

2. the projection, resulting in the ideal, undistorted image coordinates (in units of focal lengths), as computed by a linear pinhole model

$$\mathbf{p}_u = \frac{1}{w_n}\mathbf{p}_n = [x_n/w_n, y_n/w_n, 1]^T = [x_u, y_u, 1]^T, \qquad (5.17)$$

3. the computation of the real, distorted image coordinates, as observed in an image of a real camera, using a **distortion model** $\mathbf{d}_+(\mathbf{p}_u, \theta_\mathbf{d})$[1]

$$\mathbf{p}_d = [x_d, y_d, 1]^T = \mathbf{d}_+(\mathbf{p}_u, \theta_\mathbf{d}), \qquad (5.18)$$

4. the transformation of distorted image coordinates to pixel coordinates, modeling the effects of the discrete pixel sensor

$$\mathbf{p}_p = [x_p, y_p, 1]^T = \boldsymbol{K}\mathbf{p}_d. \qquad (5.19)$$

Lens distortion takes place during the second transformation, i.e. the projection onto the image plane. Subsequently, the camera calibration matrix (5.6) describes a choice of new affine coordinates, translating physical locations in the image plane to pixel coordinates. The distorted image coordinates are computed as a function of the normalized image coordinates in the form (5.17) (i.e. with the third component normalized to unity). The advantage of this method is the resulting independence of the distortion parameters $\theta_\mathbf{d}$ of the focal length.

The distortion model $\mathbf{d}_+$ in (5.18) computes distorted coordinates from undistorted coordinates. However, Brown [1971] originally introduced the following **undistortion model** $\mathbf{p}_u = \mathbf{d}_-(\mathbf{p}_d, \theta_\mathbf{u})$, which enables the computation of ideal, undistorted coordinates given the measured, distorted ones:

$$
\begin{aligned}
x_u &= x_d + \Delta x_d, \\
y_u &= y_d + \Delta y_d,
\end{aligned}
\qquad (5.20)
$$

with

$$
\begin{aligned}
\Delta x_d &= x_d(k_1 r_d^2 + k_2 r_d^4 + k_3 r_d^6) \\
&\quad + [t_1(r_d^2 + 2x_d^2) + 2t_2 x_d y_d][1 + t_3 r_d^2], \\
\Delta y_d &= y_d(k_1 r_d^2 + k_2 r_d^4 + k_3 r_d^6) \\
&\quad + [t_2(r_d^2 + 2y_d^2 + 2t_1 x_d y_d][1 + t_3 r_d^2],
\end{aligned}
\qquad (5.21)
$$

and

$$r_d = \sqrt{x_d^2 + y_d^2}. \qquad (5.22)$$

The radial distortion parameters are (k_1, k_2, k_3), the decentering distortion parameters are (t_1, t_2, t_3). We call the set of parameters $\boldsymbol{\theta}_u = [k_1, k_2, k_3, t_1, t_2, t_3]^T$ *undistortion parameters* to indicate that in general they are different from the parameters $\boldsymbol{\theta}_d$ in (5.18).

This undistortion model is widely applied in photogrammetric applications [Faig, 1975; Slama, 1980; Luhmann, 2000]. However, there seems to be some confusion about distortion and undistortion models in the computer vision literature. For example, Zhang [2000] uses the same functions as in (5.21) as a distortion

---

[1] the subscript '+' indicates that distortion is added to an ideal, undistorted point

model, i.e. to compute distorted from undistorted coordinates. The same approach is taken in many other publications, e.g. [Heikkilä and Silven, 1996, 1997; Altunbasak et al., 2003]. Tamaki et al. [2002] have recently clarified this, showing that in almost all practical applications the same model performs equally well both to distort and undistort image points, given that the parameters have been estimated correspondingly (see Section 5.3.6).

Whether lens distortion should be considered and how many distortion parameters should be used in the camera calibration depends on three aspects: the accuracy of the 3-D world coordinates of the calibration points, the image noise in the feature detection of these points and the goodness of the distortion model. Shih et al. [1995] derive error envelopes, comparing the errors introduced by lens distortion to the image noise. Florou and Mohr [1996] use statistical tests to assess the significance of the distortion parameters. If lens distortions are small, the image noise has to be smaller to enable a reliable estimation of the distortion parameters. If errors introduced by lens distortion are comparable in size to errors due to noise in the feature detection or inaccuracy of the calibration marks, the calibration results might be better without any distortion modeling, since the nonlinear parameter estimaton is very sensitive to noise, and instabilities might be introduced if higher order distortion models are used [Tsai, 1987; Florou and Mohr, 1996; Shih et al., 1995].

Finally, we have to take into account the goodness of the distortion model. In photogrammetric applications, typically high-quality equipment is used under highly controlled conditions. In this case, the higher-order distortion models provide a good fit to the behaviour of real lenses. On the other hand, for example, it is known that radial distortion varies both with focus setting (magnification, focal length) and with depth in the scene [Brown, 1971; Atkinson, 1996; Godding, 1998], the latter being especially important for camera to object distances of under 30 focal lengths and considerable variation of depth in the scene. This is the typical situation for endoscopic imaging. Modeling changes in lens distortion with object distance may be significant to achieve highest subpixel accuracy (e.g. Robson et al. [1993] found differences of up to ≈ 0.3 pixels in images of a Pulnix camera equipped with a standard lens). Hence, since the imaging behaviour of endoscopes is complex and difficult to be predicted, it might not be possible to achieve highest accuracy with the traditional models of the image-formation process. The same argument holds for many lenses used in computer vision applications which do not conform with the high standards of photogrammetry. Hence, in the computer vision literature, many authors propose to use only the first radial distortion parameter, e.g. [Tsai, 1987; Fryer et al., 1994]. In this case, there is the additional advantage of an analytical inverse of the distortion model [Lenz, 1987; Lenz and Tsai, 1989]. Note that also Brown [1971] originally found that only k_1 contributed significantly to the distortion correction.

5.2.3 Effects of multimedia geometry

3-D Particle-Tracking Velocimetry is a typical application where the projection ray from an object (a particle) to the sensor passes several optical media with different refractive indices. In such a **multimedia environment**, additional deviations from the standard pinhole camera model are introduced by the **refraction** of the projection rays at the interfaces of the different media. PTV corresponds to the standard case, where the object is situated in a liquid, the sensor is positioned in air, and a plane glass window separates these two media. Maas [1995] presents an algorithm for the strict geometric modeling of the twice broken projection rays. This ray-tracing module based on Snell's law can be integrated into any standard calibration procedure. Maas [1995] shows that using this module it is even possible to estimate the refractive indices

Table 5.3: *Comparison of the internal camera parameters for the stereo rig 'Höllbach' (left camera), resulting from calibration in air and calibration in water. $\langle \Delta x \rangle$ resp. $\langle \Delta y \rangle$ are the mean image residuals (mean reprojection errors) in the x- resp. y-direction (mean $\pm$ standard deviation), $\langle \Delta r \rangle$ is the mean distance between the measured and projected calibration points in the image plane.*

	water, no dist.	air, no dist.	water, k_1	air, k_1
f_x $[cm]$	2.4193	1.8068	2.4902	1.8864
f_y $[cm]$	2.4589	1.8023	2.5052	1.8755
s $[-]$	-0.0154	-0.0113	-0.0051	-0.0034
c_x $[pix]$	83.5436	83.2335	93.8934	95.8277
c_y $[pix]$	93.4073	89.1152	90.3013	94.4576
k_1 $[-]$	0	0	0.0816	0.1712
$\langle \Delta x \rangle$ $[pix]$	-0.02 ± 0.56	-0.01 ± 0.60	0.00 ± 0.48	0.00 ± 0.36
$\langle \Delta y \rangle$ $[pix]$	0.00 ± 0.54	0.00 ± 0.93	0.00 ± 0.48	0.00 ± 0.36
$\langle \Delta r \rangle$ $[pix]$	0.68 ± 0.36	0.97 ± 0.52	0.62 ± 0.27	0.45 ± 0.24

of the liquid and the glass by introducing them as unknown parameters in the estimation. Since the refraction correction depends on the depth of the object point, it has to be computed iteratively. To speed up the computations, the refraction correction for a given experimental configuration may be computed once and stored in a look-up table.

In the present application, we do not apply a multimedia modeling for the following reasons. Two multimedia environments have to be distinguished. The first is that of the 'Neckar'-setup used to acquire the free surface flow. Instead of modeling the multimedia effects, we aim for a physical reduction of the refraction at the interfaces using liquid prisms, see section 9.4.1. Liquid prisms are also frequently used in stereoscopic PIV applications and have been shown to be extremely efficient at reducing radial distortions arising from a water-air interface [Prasad, 2000; van Doorne et al., 2003]. Maas [1995] mentions a number of further effects of the multimedia environment causing a degradation of image quality and hence larger errors in the coordinate determination of underwater objects. He concludes that image residuals of camera models including a multimedia correction still are typically a factor two to four larger than those of applications in air. This is also the order of magnitude of the residuals obtained using liquid prisms and a standard pinhole model, see section 5.4.3.

The second case is endoscopic underwater imaging, where the multimedia environment is slightly different. Since the endoscopes are completely submerged in the water, the projection rays pass through the media water - glass (lens) - air instead of air - glass - air for lenses used in air. Lavest et al. [2000] show that immersing the camera in water simply results in an increase of the effective focal length: $f_{water} \approx n_{water} f_{air}$, where $n_{water} \approx 1.33$ is the refraction index of water. The increase of the focal length yields a decrease of the field of view in water, so that the image in water is magnified by a factor of ≈ 1.33. Further, the radial distortion in water is smaller because $1.33(x + \Delta x)_{water} \approx (x + \Delta x)_{air}$. Lavest et al. [2000] conclude that also for underwater applications the camera may be calibrated in air. Afterwards, the underwater camera parameters can be computed simply by the above multiplications. Using this method together with a bundle adjustment on real underwater images, they are able to achieve image residuals as low as 0.04

pixels, which is comparable to the highest accuracies reported in the photogrammetric and computer vision literature [Lavest et al., 1998; Heikkilä, 2000; Clarke and Wang, 2000; Hastedt et al., 2002].

Table 5.3 shows the results of the calibration of a flexible endoscope, both in air and in water (see section 5.4.2 for details of the calibration procedure). Two calibrations have been carried out: the first without any distortion parameters and the second including the first radial distortion parameter k_1 (the other distortion parameters have also been included in further tests, but were found to have no significant effect). The results confirm the findings of Lavest et al. [2000], in particular:

- The ratio of the focal lengths f_{water}/f_{air} is almost exactly equal to 1.33.

- If no radial distortion parameter is included in the model, the residuals in air are larger than those in water. Hence, the radial distortion in water is smaller. The radial distortion parameter k_1 estimated in air is about twice as large as that estimated in water.

- For calibration in air, the inclusion of k_1 in the camera model reduces the image residuals roughly by a factor of two. For calibration in water, the reduction of the residuals is much smaller. Hence, the residuals in water are not caused by radial lens distortion.

5.3 Estimation of camera parameters

The task of the geometric camera calibration is to estimate the camera parameters θ and possibly the additional parameters of lens distortion $\theta_{\mathbf{d}}$ resp. θ_u. A large variety of estimation methods is available, see [Clarke and Fryer, 1998] and [Armangue et al., 2002] for reviews or [Atkinson, 1996] or [Hartley and Zisserman, 2000] for in-depth introductions.

The basic principle of all methods is to derive constraints on the camera parameters using a sufficient number of known correspondences between features in the 3-D scene and features in the image. These constraints represent an over-determined system of equations that is solved by a least squares method. Towards this end, a **cost function** is introduced, which is a measure of the discrepancy between the camera model and the observations. Depending on the parameterization of the problem, the cost function may be expressed as an **algebraic**, **geometric** or **statistical error**, with different error measures possibly resulting in different parameter estimates. Additional terms may be added to the cost function to regularize the results or to constrain the parameters, e.g. to a certain range of values.

The minimum of an algebraic error can be found using *linear least squares methods*, while geometric or statistical errors have to be minimized by *nonlinear iterative methods*. An example for minimizing algebraic error is given by the *Direct Linear Transformation* (DLT) calibration method explained in section 5.3.1. The calibration methods discussed in section 5.3.4 and section 5.3.5 are nonlinear approaches based on minimizing geometric errors. For more information on parameter estimation techniques in general, see [Press et al., 1992; Gill et al., 1981; Nocedal and Wright, 1999]. Zhang [1995] gives an overview of parameter estimation for computer vision. Textbooks on parameter estimation in the context of computer vision are e.g. [Weng et al., 1993b; Kanatani, 1996; Hartley and Zisserman, 2000]. See especially [Hartley and Zisserman, 2000] for a detailed discussion of algebraic, geometric and statistical errors and their relations.

The most common way of calibrating cameras in laboratory experiments is to acquire images of a dedicated calibration target, which is put in the observation area of the experiment. The target is either a

three-dimensional target, e.g. a cube, or a planar target. The calibration points are usually the edges of checkerboard patterns, the crossings of grid lines or the centers of circular markers. In this work, a **planar grid target** is applied. For a reliable calibration of the complete camera model (5.5), the calibration points have to be distributed in 3-D throughout the whole observation volume. Hence, if a planar target is used, there are two possibilities to achieve this:

1. **'Simulated 3-D target'**: A 3-D target is created artificially by fixing the planar target on a linear positioner, moving it in precisely known steps along an axis perpendicular to the calibration plane and acquiring several images successively. This method is applied to calibrate the endoscope stereo rigs, see section 5.4.2. The positions of the calibration points are assumed to be precisely known (both in the images and the world) and well distributed throughout the whole observation volume. If lens distortion is neglected, the corresponding estimation method (DLT, see section 5.3.1) is linear and known as **camera resectioning** in photogrammetry (the related problem of estimating the 3-D structure of a scene given the camera parameters and several views is called **intersection**).

2. **'Multi-plane calibration**: The second possibility is to acquire several images of the calibration plane in different positions and orientations. To create the different views, the plane is freely moved in space, without the 3-D coordinates of the calibration points being known. Only the 2-D relative position of the points within the plane is known. Hence, in addition to the camera parameters, the positions and orientations of the calibration planes also have to be estimated in the calibration, which calls for a nonlinear estimation method (Section 5.3.4). The multi-plane method is used to calibrate the stereo rig 'Neckar' (free surface flow), see Section 5.4.3.

If at least two different views of a scene are available (as e.g. in a stereo camera setup), the most general method to calibrate a camera is to estimate both the camera parameters and the 3-D coordinates of the calibration points within a simultaneous full-scale nonlinear optimization. This method is called a **bundle adjustment** or **joint bundle adjustment with self-calibration** in photogrammetry, where it has been successfully applied for a long time [Brown, 1966; Faig, 1975; Bopp and Krauss, 1978; Slama, 1980; Luhmann, 2000]. In computer vision, the application of bundle adjustments to compute structure and motion parameters has been proposed much later, e.g. [Lavest et al., 1998; Triggs et al., 2000]. Under the assumption of independent Gaussian image noise, bundle adjustment corresponds to the *maximum-likelihood estimation* (MLE) of the unknown parameters (camera calibration and 3-D structure). In this case, it is the method obtaining the highest possible precision. Due to the well-defined statistical basis, reliable RMS errors of all estimated quantities can be obtained, given enough redundancy in the measurements. Bundle adjustment is a very general and flexible method, since any geometric constraints or a priori knowledge can be incorporated in the estimation easily (e.g. the constraint that certain world points lie on a common plane or line in space). If no metric information is available, a bundle adjustment can also be computed using only implicit geometric information, i.e. corresponding image points of a static scene in at least three views. In the computer vision literature, this approach is known as **self-calibration** [Faugeras et al., 1992]. In this case, the 3-D structure of the scene can only be estimated up to a projective transformation or a similarity tranformation, depending on the number of cameras and views [Hartley and Zisserman, 2000].

Since any type of bundle adjustment requires a nonlinear, iterative estimation method, it has to be applied with some precautions [Luhmann, 2000]:

- Bundle adjustment requires a 'strong network geometry', i.e. a careful arrangement of the camera views: Several views (at least two) of the same 3-D points are required to get a high redundancy.

Table 5.4: *Overview of the different methods that are applied to calibrate the experimental subsystems (DLT: Direct Linear Transformation, LM: nonlinear minimization by a Levenberg-Marquardt algorithm).*

subsystem	Jagst	Elbe-u	Neckar
	Kocher	Elbe-m	
	Höllbach	Elbe-d	
optics	flexible endoscopes	rigid endoscopes ('periscopes')	C-mount lenses
location	pore flow	sand layer	free surface flow
observation area	3-D volume	2-D cross section	3-D volume
measurement	3-C vectors	2-C vectors	3-C vectors
setup	stereoscopic	monoscopic	stereoscopic
camera model	linear pinhole	linear pinhole + k_1	linear pinhole
approach	3-D target	single-plane	multi-plane
estimation	projection	homography	projection
	(3D-2D)	(2D-2D)	(3D-2D)
	robust DLT	robust DLT	robust DLT + LM

The points have to be distributed throughout the whole observation volume to stabilize the estimation. To avoid correlations in the camera parameters, different views with the camera rotated around the optical axis should be available.

- To start the iterative estimation, initial guesses of all parameters are required. The latter should be close to the true values to avoid getting stuck in local minima during the minimization of the cost function.

- Degenerate geometric configurations (e.g. all calibration points lie within the same plane in space) have to be avoided, since they result in singularities in the estimation.

- Since all unknowns are estimated simultaneously, bundle adjustment is computationally very expensive for large numbers of views resp. 3-D points. Efficient computational methods like factorization [Sturm and Triggs, 1996] or sparse Levenberg-Marquardt optimizers [Hartley and Zisserman, 2000] are necessary.

A very efficient variant of bundle adjustment is *interleaved bundle adjustment*. Instead of the simultaneous full-scale nonlinear estimation of all parameters, groups of parameters (typically camera parameters and 3-D coordinates) are estimated separately within an iterated resection-intersection loop [Weng et al., 1993a,b; Wang and Clarke, 1996; Bartoli, 2002]. The advantages are faster computation and reduced correlations between parameters [Clarke et al., 1998].

The multi-plane approach to camera calibration mentioned above lies between the methods of resectioning and bundle adjustment, since it uses 2-D metric information (2-D relative position of the points within the calibration planes) rather than full 3-D information (like resectioning using a 3-D target) or only implicit geometric information (like self-calibration).

In this work, three calibration methods are applied, see table 5.4. The reasons for using three different methods are related to the particular conditions of the experimental setups, see section 10.2.1. The first is resectioning by a robust DLT, outlined in sections 5.3.1–5.3.2. The second method is a 2-D calibration using a single calibration plane, described in section 5.3.3. It is also a DLT approach. The differences to the first method are that a plane homography is estimated, i.e. a 3×3-matrix instead of a 3×4-projection matrix. Further, a radial distortion parameter is included in the estimation. The third method is a multi-plane approach proposed by Zhang [2000], based on an implementation by Heikkilä [2000] (sections 5.3.4–5.3.5). Results of the three methods including their precision are shown in section 5.4.

5.3.1 DLT by mixed OLS-TLS estimation

A very simple approach to camera calibration is the *Direct Linear Transformation* (DLT) [Abdel-Aziz and Karara, 1971; Luhmann, 2000]. Since it is a linear method, it cannot account for lens distortion. The DLT computes the parameters of the projection matrix P in (5.14) in the following way. A list of $n \geq 6$ corresponding world points $\mathbf{P}_{w,i}$ and image points $\mathbf{p}_{p,i}$ has to be given. We re-write (5.14) as

$$\lambda \begin{bmatrix} x_p \\ y_p \\ 1 \end{bmatrix} = \begin{bmatrix} H_{11} & H_{12} & H_{13} & H_{14} \\ H_{21} & H_{22} & H_{23} & H_{24} \\ H_{31} & H_{32} & H_{33} & H_{34} \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}. \tag{5.23}$$

Eliminating λ from (5.23) yields

$$\begin{aligned} x_p &= \frac{H_{11}X_w + H_{12}Y_w + H_{13}Z_w + H_{14}}{H_{31}X_w + H_{32}Y_w + H_{33}Z_w + H_{34}}, \\ y_p &= \frac{H_{21}X_w + H_{22}Y_w + H_{23}Z_w + H_{24}}{H_{31}X_w + H_{32}Y_w + H_{33}Z_w + H_{34}}. \end{aligned} \tag{5.24}$$

Re-arranging (5.24) shows, that each pair of corresponding world and image points gives two constraints on the twelve elements of the projection matrix P:

$$\begin{aligned} X_w H_{11} + Y_w H_{12} + Z_w H_{13} + H_{14} - x_p X_w H_{31} - x_p Y_w H_{32} - x_p Z_w H_{33} - x_p H_{34} &= 0, \\ X_w H_{21} + Y_w H_{22} + Z_w H_{23} + H_{24} - y_p X_w H_{31} - y_p Y_w H_{32} - y_p Z_w H_{33} - y_p H_{34} &= 0. \end{aligned} \tag{5.25}$$

To reduce the degree of freedom in P to eleven, we set $H_{34} = 1$. This might introduce singularities if the actual value of H_{34} is close to zero, but it has been checked that this is not the case in the geometries of the stereo rigs used in this work. We obtain the following equations:

$$\begin{aligned} X_w H_{11} + Y_w H_{12} + Z_w H_{13} + H_{14} - x_p X_w H_{31} - x_p Y_w H_{32} - x_p Z_w H_{33} &= x_p, \\ X_w H_{21} + Y_w H_{22} + Z_w H_{23} + H_{24} - y_p X_w H_{31} - y_p Y_w H_{32} - y_p Z_w H_{33} &= y_p. \end{aligned} \tag{5.26}$$

Stacking up the equations for a set of n point correspondences, we obtain an over-determined system of the form

$$\boldsymbol{A}\boldsymbol{h}' = \boldsymbol{b}, \tag{5.27}$$

where $\boldsymbol{A}$ is the $(2n \times 11)$-matrix

$$\boldsymbol{A} = \begin{bmatrix} X_{w,1} & Y_{w,1} & Z_{w,1} & 1 & 0 & 0 & 0 & 0 & -x_{p,1}X_{w,1} & -x_{p,1}Y_{w,1} & -x_{p,1}Z_{w,1} \\ 0 & 0 & 0 & 0 & X_{w,1} & Y_{w,1} & Z_{w,1} & 1 & -y_{p,1}X_{w,1} & -y_{p,1}Y_{w,1} & -y_{p,1}Z_{w,1} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ X_{w,n} & Y_{w,n} & Z_{w,n} & 1 & 0 & 0 & 0 & 0 & -x_{p,n}X_{w,n} & -x_{p,n}Y_{w,n} & -x_{p,n}Z_{w,n} \\ 0 & 0 & 0 & 0 & X_{w,n} & Y_{w,n} & Z_{w,n} & 1 & -y_{p,n}X_{w,n} & -y_{p,n}Y_{w,n} & -y_{p,n}Z_{w,n} \end{bmatrix}, \tag{5.28}$$

and $\boldsymbol{h}'$ is the vector of the remaining eleven parameters of P. Hence, we have to solve a linear least-squares problem, where errors occur both in the data matrix $\boldsymbol{A}$ and in the right-hand-side vector $\boldsymbol{b}$. Problems of this kind are known as **Total Least Squares**[2] (TLS) problems [van Huffel and Vandewalle, 1991]. In a TLS estimation, the problem (5.27) is re-formulated as

$$[\boldsymbol{A}|\boldsymbol{b}] \begin{bmatrix} \boldsymbol{h}' \\ -1 \end{bmatrix} = \boldsymbol{0}. \tag{5.29}$$

The basic idea of TLS is to modify the noisy data matrix $\boldsymbol{D} = [\boldsymbol{A}|\boldsymbol{b}]$ into a matrix $\tilde{\boldsymbol{D}} = \left[\tilde{\boldsymbol{A}}|\tilde{\boldsymbol{b}}\right]$ that is close to the original matrix in the Frobenius norm, subject to the constraint that $\tilde{\boldsymbol{b}}$ is in the range of $\tilde{\boldsymbol{A}}$. In this approach, all the data are modified in contrast to an ordinary least squares (OLS) approach, where only one column of $\boldsymbol{D}$ is modified. Since all the columns of $\boldsymbol{D}$ contain noisy data, OLS estimation generally yields biased results, whereas TLS does not.

A closer look at (5.28) reveals, that not all columns in the matrix $\boldsymbol{A}$ are subject to noise. The fourth and the eighth column contain exactly known constants. In this case, the TLS estimator produces suboptimal results. This problem can be overcome by partitioning the matrix into an OLS and a TLS subproblem in a **mixed OLS-TLS** approach. Towards this end, a QR-decomposition of $\boldsymbol{A}$ is calculated. For the details of this approach, see [van Huffel and Vandewalle, 1991]. The superior performance of mixed OLS-TLS estimation in the case of exactly known columns in the data matrix has been shown by Garbe [2001] and Mühlich and Mester [1998]; Mühlich et al. [1999]. The latter apply mixed OLS-TLS (among other applications) for the computation of a DLT and the estimation of the fundamental matrix (see section 7.5), where a similarly structured data matrix occurs. Garbe [2001] presents a detailed study of mixed OLS-TLS estimation in the context of optical flow computations. We use the mixed OLS-TLS estimator as described in [Garbe, 2001] to compute the DLT.

An important topic in parameter estimation is the **scaling** or **normalization** of the input data prior to estimation. Scaling is necessary for two reasons:

1. Both OLS and TLS only produce optimal results if the noise in the data matrix is independent, of zero mean and identically distributed (i.i.d. noise). In OLS estimation, the distribution additionally has to be Gaussian, whereas in TLS the distribution is arbitrary. In this case, OLS and TLS are maximum-likelihood estimators. In general, the noise in the data matrix will not be i.i.d. However, this noise structure can be obtained by applying an appropriate scaling of the data matrix (also known as *equilibration* [Mühlich and Mester, 2001].

---

[2]also known as orthogonal least squares regression or errors-in-variables regression
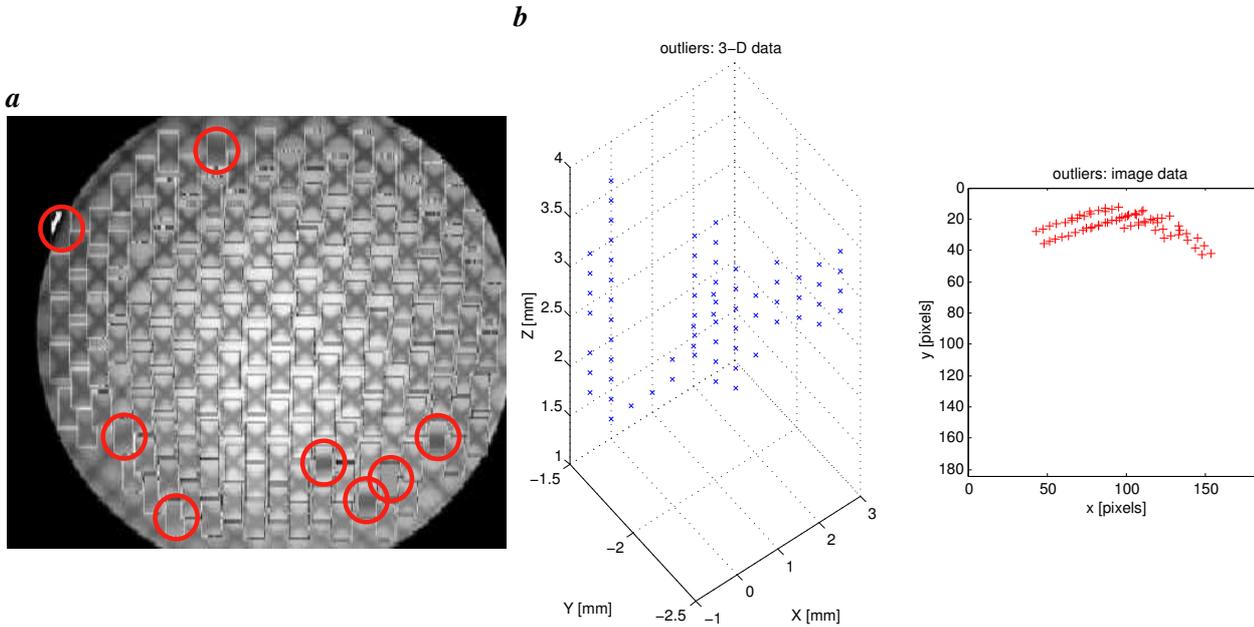
b

Figure 5.5: *a Outliers in the feature extraction of the calibration points (red circles). The positions of the grid crossings are determined by a nonlinear gray value fit (section 5.1), which may sometimes fail due to the limited quality of the endoscopic images. Further, the fit does not work at the border of the circular field of view (see* **b** *), where the grid crossings are only partially visible. In both cases, the fit gets stuck in a local minimum of the cost function and yields completely wrong parameters.* **b** *All outliers of an endoscope calibration shown in one plot (left: 3-D view of the corresponding world points, right: image coordinates of the outliers).*

2. If the numerical range of the input data covers several orders of magnitude, unscaled data may introduce numerical instabilities in the TLS algorithms, where typically matrix factorizations like QR-decomposition, Cholesky decomposition or singular value decomposition are computed [Golub and van Loan, 1996].

Hartley [1997a] proposes the following normalization scheme: a translation of the image coordinates to their center of gravity, followed by a scaling to make the mean square of the coordinates equal to unity. The scaled data have zero mean and unit variance. Hartley shows the efficiency of this scaling by a number of examples and gives some motivation for it, but does not analyze the statistical input-output relation in terms of bias, variance and consistency. The latter is carried out by Mühlich and Mester [1998] and Mühlich et al. [1999]. They show that the scaling prescriptions proposed by Hartley [1997a] can be derived by a statistical analysis of the error structure in the TLS estimation and that it leads to unbiased TLS estimates. Hence, we apply this scaling in the DLT computations in this work.

After the computation of the projection matrix P by the mixed OLS-TLS algorithm, a QR-decomposition of P is calculated to recover the physical (internal and external) camera parameters [Hartley and Zisserman, 2000]. An equivalent decomposition may be derived by purely geometric considerations [Strat, 1984].

5.3.2   Robust DLT

In practical imaging applications, input data is often corrupted by outliers, i.e. gross errors resp. 'wrong' observations, which do not conform to the estimated model. An example is shown in figure 5.5, where

Table 5.5: *Comparison of calibration results using robust and non-robust estimation (calibration data hmb01r, right camera of Höllbach-setup). The first column shows the correct results. The results of the second and third columns have been obtained from a dataset contaminated with outliers. While the non-robust approach yields completely wrong results, LMedS is able to remove the outliers and recover the correct parameter set.*

	correct	non-robust	LMedS
f_x $[mm]$	2.5490	1.2828	2.4869
f_y $[mm]$	2.5234	1.2985	2.4499
s $[-]$	0.0751	0.1389	0.0751
c_x $[pix]$	96.1666	96.1600	90.3455
c_y $[pix]$	86.3097	93.4385	88.4236
T_x $[mm]$	0.0212	-0.3596	0.1505
T_y $[mm]$	-0.0113	0.0857	-0.0536
T_z $[mm]$	5.2200	1.6930	5.0222
ω $[°]$	-0.6911	13.4740	0.3596
ϕ $[°]$	47.2855	54.9056	48.5539
κ $[°]$	4.5645	18.3554	5.3670
$\langle \Delta x \rangle$ $[pix]$	-0.00 ± 0.35	-8.33 ± 116.7	0.00 ± 0.28
$\langle \Delta y \rangle$ $[pix]$	-0.00 ± 0.35	9.58 ± 164.8	0.00 ± 0.39
$\langle \Delta r \rangle$ $[pix]$	0.42 ± 0.26	25.9 ± 200.7	0.40 ± 0.26

outliers occur in the extraction of the grid points in an endoscopic calibration image. Estimation methods like the mixed OLS-TLS estimator of section 5.3.1 assume that all data can be described by one model, i.e. by one set of parameters to be estimated. All least squares estimators based on the L_2-norm are very vulnerable to the violation of this assumption [Zhang, 1995]. In some cases, even one single outlier may completely perturb the results. Hence, estimation techniques for computer vision should be robust to outliers, i.e. they should be able to identify bad data points automatically and remove them from the computations, so that the resulting estimates are unaffected by outliers. Many robust estimators have been proposed in the context of computer vision, for reviews see [Meer et al., 1991] and [Stewart, 1999]. An introduction to robust statistics is given by Huber [1981].

The series of measurements described in chapter 10 has been carried out during a time period of about six months. All experimental setups had to be re-calibrated before each experiment to account for possible changes in the geometry of the stereo rigs which may evolve gradually over time. As a result, more than 2500 calibration images had to be processed. Hence, manual inspection to remove outliers was not feasible, making a robust technique absolutely necessary. A robust DLT estimation has been implemented, based on a random sampling technique known as **Least Median of Squares**[3] (LMedS). The implementation is based on the description given by Zhang [1995].

In the LMedS approach, a number of randomly selected subsamples is drawn from the data. The size of the subsamples is defined by the minimal solution of the problem. Since there are eleven parameters

---

[3]also known as Least Median of Squares of Orthogonal Distances (LMSOD) if the estimator is based on TLS

in the projection matrix of the DLT and each point correspondence yields two constraints, each subsample consists of six correspondences. For every subsample, the parameter vector is estimated, and the residuum is computed using the complete data set. The parameter vector corresponding to the minimum median of the residuals is chosen as the correct one. Outliers are detected using a statistically robust threshold and removed. Finally, the model is estimated again (by mixed OLS-TLS) using only the inliers. The number of subsamples one has to draw to find the correct solution can also be estimated by statistical considerations, assuming a certain outlier probability. For further details, see [Zhang, 1995].

We demonstrate the effectiveness of the LMedS approach by the following experiment. In a set of 18 endoscope calibration images, we exchange the X- and Y-components of the world coordinates in one image, whereby a number of gross outliers are introduced. We estimate the camera parameters both using a non-robust DLT and using LMedS. The results are shown in table 5.5. LMedS is able to recover the correct camera model also in the presence of outliers. Figure 5.5b shows a further example of outlier detection in endoscopic calibration images using the LMedS approach. In this case, the outliers are 'real' outliers occuring due to a failure of the feature extraction (see section 5.1) at the border of the endoscopic field of view.

5.3.3   Estimation of a planar homography with lens distortion

In the context of projective geometry, a nonsingular linear transformation in a projective space is called a **homography**. An example is the projective transformation between points $\mathbf{p}_w = [X_p, Y_p, 1]^T$ *on a plane in 3-D space* (where X_p and Y_p are Euclidean 2-D coordinates in a frame of reference defined within this plane) and their (ideal, undistorted) image points $\mathbf{p}_u = [x_u, y_u, 1]^T$ (which lie also within a plane, namely the image plane). In homogeneous coordinates, this homography is described by a (3×3)-matrix H:

$$\mathbf{p}' = \mathsf{H}\mathbf{p}. \tag{5.30}$$

In this thesis, homographies within $\mathbb{R}^3$ are important in two respects:

1. They define the perspective projection for the periscopes (section 9.2), where observations are restricted to a plane in space. In this case, the geometric camera calibration is reduced to the estimation of a (3×3)-homography.

2. Homographies between planes are computed in the initialization phase of multi-plane calibration algorithms like the one outlined in section 5.3.4.

The estimation of a (3×3)-homography is carried out in complete analogy to the estimation of the projection matrix P, as described in section 5.3.1, including the robustification by LMedS shown in section 5.3.2. The only difference is that we estimate a homogeneous (3×3)-matrix with eight degrees of freedom instead of a (3×4)-matrix with eleven degrees of freedom.

Singularities and degenerate configurations.

For the geometric calibration of the periscopes, only one image of a calibration grid is necessary, since the observations are restricted to a plane. Hence, we do not have to compute the full projection matrix P, but only a (3×3)-homography H. In such a case, we have to take special care of singularities and degenerate configurations. Both are the result of an insufficient[4] geometrical distribution of calibration points.

---

[4]with respect to the chosen camera model

Singularities are cases where the estimation completely fails because certain parameters are not estimable. Degenerate geometries do not determine a *unique* solution, because there are correlations between certain parameters. Note that the definitions of both singularities and degeneracies involve both the geometric configuration and the type of camera model to be estimated.

Sturm and Maybank [1999] discuss minimal cases in plane-based calibration algorithms. They show that if only one plane is available, the following two cases have to be distinguished:

- If the calibration plane is not parallel to the image plane, the focal lengths f_x and f_y can be calibrated, provided that the principal point is given. Hence, this case corresponds to a degenerate configuration where the principal point is correlated with exterior orientation parameters, see section 5.3.6 and [Clarke et al., 1998]. Hartley and Kaucic [2002] show that the principal point is also correlated with the focal length.

- If the calibration plane is parallel to the image plane, a singularity is introduced. The only parameter that can be estimated is the aspect ratio f_x/f_y. The focal lengths cannot be estimated, since they are correlated with T_z (the Z-component of the translation vector).

Strictly speaking, the periscope calibration corresponds to the singular case, where the calibration plane is parallel to the image plane. However, tests have shown that the deviations from parallelism introduced by the experimental imprecision together with the strong radial distortion of the wide-angle lenses enable a determination of both the focal length and the principal point. The latter is estimated as the center of the radial distortion. For results, see section 5.4.1.

Including lens distortion.

The wide-angle lenses of the rigid endoscopes introduce a significant amount of radial distortion, see the examples of calibration images in figure 5.6. The straight lines of the calibration grid appear more or less curved in the images. This curvature cannot be described by a homography and hence will reduce the accuracy of the camera model. To account for the curvature, we introduce the model (5.20) as a distortion model, i.e. with (x_u, y_u) and (x_d, y_d) interchanged. This approach is frequently used in computer vision, see section 5.2.2 and [Tamaki et al., 2002]. Only the parameter k_1 is used, the other distortion parameters are set to zero.

To avoid a nonlinear estimation method, we divide the model parameters into distortion (k_1) and non-distortion (H) parameters and estimate each set separately while keeping the other set fixed, using only linear total least squares methods. The estimations are iterated until convergence. A similar multi-step iterative estimation procedure has been proposed recently by Cornelis et al. [2002].

5.3.4 Zhang's method for multi-plane calibration

Zhang [2000] proposed the following multi-plane calibration method (other plane-based calibration methods are [Tsai, 1987; Lenz and Tsai, 1988; Wei and Ma, 1994; Sturm and Maybank, 1999]). Images of a planar calibration grid are acquired from several different orientations. Either the camera or the planar pattern can be freely moved between the acquisitions, the motion does not have to be known.

The parameter estimation is carried out in two steps. First, the homography between each calibration plane and the image plane is computed. Towards this end, a linear method similar to that in section 5.3.1 is

used, followed by a nonlinear optimization of the parameters for each homography. As shown in [Zhang, 2000], each homography gives two constraints on the intrinsic parameters. Hence, the five parameters of the calibration matrix $\boldsymbol{K}$ can be computed given at least three homographies, i.e. three views of the calibration plane in different orientations. As a result of the first step, initial values for the intrinsic (without distortion) and all extrinsic parameters (orientations of all planes relative to the first plane) are available. The second step of the algorithm is a refinement of the initial parameters obtained in the first step, using a nonlinear optimization (Levenberg-Marquardt algorithm). Further, a lens distortion model is included in the second step. For details, see [Zhang, 2000]. A Matlab implementation of this method due to J.-Y. Bouguet is available at `http://www.vision.caltech.edu/~bouguetj/calib_doc/index.html`.

Zhang [2000] also carries out a performance analysis of his method. He shows that errors in the estimated parameters decrease with increasing number of planes. At least three views should be used. The average error decreases until a number of about ten views is reached. Typical RMS image residuals of 0.3 to 0.4 pixels are reported. Zhang's method assumes that the exact 2-D positions of the calibration points within the planes are known. The sensitivity of the method with respect to violations of this assumption is checked, both for random noise in the calibration points and systematic deviations from planarity. The method is relatively robust to random noise. Only the radial distortion parameters become less useful for higher noise levels. This behaviour is expected, see the discussion in section 5.3.6. Systematic non-planarity of the calibration planes has more effect on the calibration precision than random noise. However, the results are still reasonable for systematic non-planarity less than 3% (measured as the ratio of the maximum displacement perpendicular to the plane to the size of the calibration pattern).

5.3.5 Heikkilä's method

Heikkilä [2000] presents a nonlinear calibration method that can handle both multi-plane calibration and calibration using a 3-D target. The multi-plane version is similar to Zhang's method. However, instead of computing the initial values for the interior orientation based on the constraints given by the homographies, Heikkilä [2000] simply uses the nominal values as initial guess (focal length given by the manufacturer, unity aspect ratio, zero skew, center of the image as the principal point). We also use the nominal values, since it was found that this approach is more stable and reliable than estimating the initial values.

Further advantages of Heikkilä's method are the correction of biases introduced by circular calibration points and a method to reverse the distortion model and obtain an accurate mapping in both directions (distorted-undistorted and vice versa). However, these features were not used (because the calibration points are crosses and distortion is neglected). Using his full camera model including the bias correction, Heikkilä [2000] achieves rms image residuals as low as ≈ 0.05 pixels on real images obtained with standard equipment. A remarkable result reported by Heikkilä is that errors of up to 0.5 pixels may be introduced if different light sources are used for calibration and measurement (as a result of chromatic aberrations). This underlines the importance of the radiometric information in the estimation of precise geometric positions.

At `http://www.ee.oulu.fi/~jth/calibr/`, a Matlab implementation of Heikkilä's method is available. The nonlinear parameter optimization of all setups in this work is based on this implementation because of its flexibility to handle both multi-plane calibration and 3-D targets.

5.3.6 Estimating lens distortion

In this section, some further important aspects related to the estimation of lens distortion parameters are mentioned. The discussion will substantiate the decision not to use distortion models in this work, except for the wide-angle lenses of the periscopes.

Metric and non-metric approaches.

Methods to estimate lens distortion can be classified into two major groups: *metric* and *non-metric* approaches. In the metric approaches, the distortion parameters are calibrated along with the standard external and internal parameters of a pinhole camera, e.g. [Weng et al., 1992; Zhang, 2000; Heikkilä, 2000]. Usually, a dedicated calibration target with precisely known 3-D points is used. Due to the nonlinearity of the resulting camera model, an iterative optimization is necessary. The simultaneous estimation of all parameters may introduce correlations between distortion and non-distortion parameters in case of degenerate configurations or weak geometry of the calibration points, see below.

Akin to self-calibration methods, the non-metric approaches to distortion estimation only use implicit geometric information. They are based on the fact that straight lines in the scene should also be straight in the image of an ideal pinhole camera. In images of real lenses, straight lines in the world appear as more or less curved lines in the images. Hence, a distortion model may be estimated by requiring these curved image lines to become straight if transformed with the model. This approach was originally introduced into analytical photogrammetry as the 'plumbline-method' by Brown [1971]. The plumbline-method estimates a set of undistortion parameters independent of the other camera parameters. Hence, correlations are avoided. The method has been re-discovered recently in the computer vision community [Devernay and Faugeras, 2001]. Since the probability of data outliers is high in typical computer vision applications, it is important to use robust estimation methods, e.g. Least Median of Squares [El-Melegy and Farag, 2003] or RANSAC [Thormählen et al., 2003]. A further approach similar to the plumbline-method has been proposed for stereo applications by Zhang [1996] and Stein [1997]. They estimate radial distortion from point correspondences in two views by minimizing the distance of points to their corresponding epipolar lines.

Correlations between parameters.

In the case of (nearly) degenerate configurations or weak network geometry of the calibration points resp. camera views, correlations between the following sets of camera parameters may show up in a simultaneous least squares estimation:

- principal point c_x, c_y and decentering distortion parameters t_1, t_2, t_3, coupled via the external parameters ω and ϕ,

- radial distortion parameters k_1, k_2, k_3,

- radial distortion parameters k_1, k_2, k_3 and focal length f.

Clarke et al. [1998] give a detailed discussion and explain the origin of these correlations. The position of the principal point is especially sensitive. It is coupled to both the focal length and the exterior orientation [Hartley and Kaucic, 2002; Ruiz et al., 2002]. If there is a discrepancy in the modeling between the exterior

orientation and the principal point, it can be reduced by t_1 and t_2 to a significant extent. For example, if the principal point is not estimated but set to a fixed (but possibly wrong) value, the estimated exterior orientation can compensate for image residuals created by the wrong principal point. The remaining residuals can be compensated by t_1 and t_2 [Clarke et al., 1998]. El-Melegy and Farag [2003] also estimate distortion parameters using the plumbline-method while keeping the principal point fixed at the image center. They show that deviations of the principal point from its true location under both radial and decentering distortion are equivalent to adding two decentering distortion terms. We confirm this in the calibration of the periscopes in section 5.4.1. Due to the high correlations, Fryer et al. [1994] recommend not to estimate both decentering distortion parameters and the principal point simultaneously in a convergent bundle adjustment. They show that over-parameterization can lead to poor conditioning in the numerical optimization, with adverse effects on the finally computed 3-D coordinates of object points.

Different strategies may be applied to avoid or reduce correlations:

- Correlations between the focal length and the radial distortion parameters are reduced in the so-called *balanced* distortion model [Atkinson, 1996]. In this model, a constant term $\propto k_0$ is added to the radial distortion function, which is equivalent to specifying a second zero-crossing of the distortion function. The resulting change is compensated by a shift in the focal length.

- Correlations between the principal point and the decentering distortion parameters in a bundle adjustment are reduced by including additional calibration images with the camera rotated around the optical axis [Luhmann, 2000].

Clarke et al. [1998] conclude that the full parameter set including distortion can only be reasonably estimated provided that

1. the functional model represents a good approximation of the physical effects,

2. the image coordinates are subject to only random errors,

3. a reasonably large number of well distributed 3-D targets is available and

4. several views in a highly convergent network are used.

If there are deficiencies in any of these requirements, poor parameter estimation may introduce errors in 3-D reconstructions based on the estimated model, and alternative strategies might be more successful.

Definition of the cost function.

As explained in Section 5.2.2, there are two approaches to model lens distortion, resulting in two different ways to compute image residuals. The difference in the approaches is that either a *distortion* or an *undistortion* model is used. The distortion model $\mathbf{d}_+$ adds distortion to ideal undistorted points, whereas the undistortion model $\mathbf{d}_-$ computes ideal points from distorted ones[5]. The undistortion model corresponds to the usual approach initially proposed in the photogrammetric literature [Brown, 1971]. Because analytical inversion is not possible, the parameter vectors in the two models are not the same and have to be estimated separately by the following procedures (we denote the ideal points as computed by a pinhole model (5.14) with $\boldsymbol{p}_{i,i}$):

---

[5]Note that the distortion model $\mathbf{d}_+$ corresponds to the 'reverse' model of [Heikkilä, 2000].

- **Distortion model**: Take the points of an ideal pinhole model $\boldsymbol{p}_{i,i}$ and use a distortion model to calculate the distorted points $\boldsymbol{p}_{d,i} = \mathbf{d}_+(\boldsymbol{p}_{i,i}; \boldsymbol{\theta}_d)$. The cost function C is given by the sum of squares of the differences between the distorted points $\boldsymbol{p}_{d,i}$ and the measured points $\boldsymbol{p}_{m,i}$:

$$C = \sum_i \left\| \boldsymbol{p}_{d,i} - \boldsymbol{p}_{m,i} \right\|^2. \tag{5.31}$$

- **Undistortion model**: Take the measured (distorted) points $\boldsymbol{p}_{m,i}$ and use an undistortion model to calculate the undistorted points $\boldsymbol{p}_{u,i} = \mathbf{d}_-(\boldsymbol{p}_{m,i}; \boldsymbol{\theta}_u)$. The cost function C' is given by the sum of squares of the differences between the undistorted points $\boldsymbol{p}_{u,i}$ and the ideal points given by the pinhole model $\boldsymbol{p}_{i,i}$:

$$C' = \sum_i \left\| \boldsymbol{p}_{u,i} - \boldsymbol{p}_{i,i} \right\|^2. \tag{5.32}$$

Hence, the cost function may be defined either in the distorted or in the undistorted image plane. Examples for both approaches are found in the computer vision literature. Zhang [2000] uses the first approach and minimizes an objective function defined in terms of distorted image points. On the other hand, Willson [1994] uses the second approach and minimizes the sum of squared errors in the undistorted image plane. Swaminathan and Nayar [2000] also use an undistortion model in a plumbline calibration. However, they point out that image data is always noisy, resulting in the possibility of nonlinear biases inherent to errors in the estimated undistortion model (amplification of image noise due to the higher-order terms in the undistortion model). Hence, they prefer to define the objective function in terms of the measured distorted points, which corresponds to a maximum-likelihood estimation of the undistortion parameters together with the 'true' distorted image points. For details, see [Swaminathan and Nayar, 2000].

Conclusion.

As the discussions in this section show, estimation of lens distortion is a tricky business. The reason is that distortion is modeled by higher order polynomials, which are very sensitive to noise. The parameter estimation may become unstable already at moderate noise levels, where 'noise' means both random noise and systematic deviations from the applied distortion model. Tests using both endoscopic and ordinary calibration images have shown that instabilities indeed occur quite often. On the other hand, if the iterations converged, the estimated distortion parameters did not significantly reduce the image residuals (see e.g. table 5.7), except in the case of the wide-angle lenses of the periscopes (see section 9.2 and section 5.4.1). Several reasons might be responsible for the bad performance of the distortion modeling (see also the conclusions of Clarke et al. [1998] mentioned above):

- A multi-plane method is applied to calibrate the Neckar-setup (free surface flow). However, rotations of the cameras or the calibration plane around the optical axis are not possible due to limitations of the experimental setup. Hence, correlations resp. over-parameterization may occur.

- In the endoscopic pore flow setups, radial distortion is reduced because the endoscopes are submerged in water (section 5.2.3). The 3-D accuracy of the calibration points might be too small to enable an estimation of the remaining small distortions.

- The 'noise' in the calibration points is not random. Figure 5.8 shows that there are systematic patterns in the image residuals, which are not described by radial or tangential distortion parameters. Since
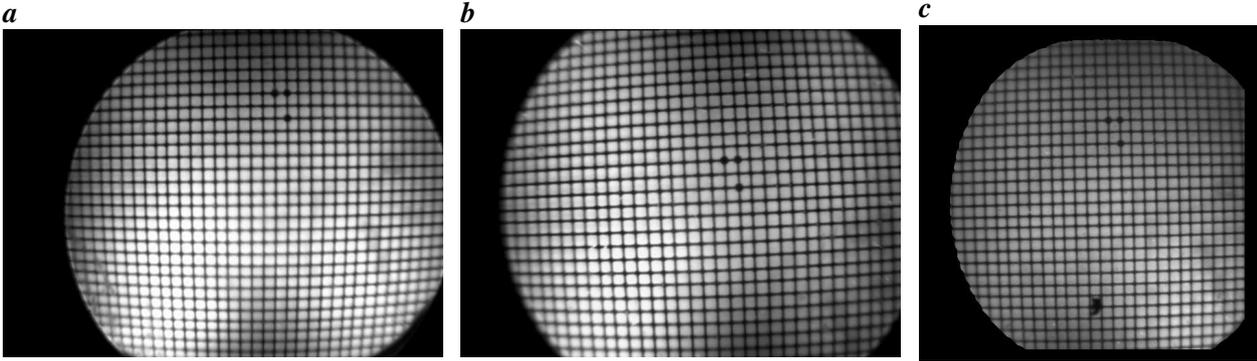
Figure 5.6: *Calibration images of the periscopes.* ***a*** *Elbe-d.* ***b*** *Elbe-m.* ***c*** *Elbe-u.*

these patterns move with the translation of the calibration grid, they are most probably related to systematic inaccuracy of the calibration grid, see section 5.4.2.

- For the flexible endoscopes, inaccuracies comparable to those introduced by lens distortion might also be caused by small systematic deviations in the arrangement of the hexagonal fiber array between the entry and exit side. In this case, the image projected onto the entry side of the fiber bundle is systematically distorted at the exit side. A global modeling of this distortion might be possible if the deviations are caused e.g. by a shear or twist of the whole bundle. However, deviations may also be individual for every fiber. In this case, a global physical modeling is not possible.

For these reasons, lens distortion parameters have only been estimated for the periscopes. For the flexible endoscopes and the free surface flow setup, standard pinhole camera models without distortion parameters are used.

5.4 Calibration results

In this section, representative calibration results of the three different experimental setups (rigid endoscopes, flexible endoscopes and free surface flow) are discussed, including an analysis of the calibration accuracy.

5.4.1 Subsystem 'Elbe' (rigid endoscopes)

Setup.

A planar grid target is used for calibration. The grid has been printed on a transparency, using a laser printer with a resolution of $1200\ dpi$. The distance of the grid lines is $0.5\ mm$, and the width of the grid lines is $0.09\ mm$. We estimate the accuracy of the printer by the variance of its 'quantization' noise, which is $(1\ inch/1200)^2/12 = (21.2\ \mu m)^2/12$, hence the standard deviation is $6.1\ \mu m$, which is assumed to be the accuracy of the printed grid[6]. The diameter of the field of view is $\approx 15\ mm$, hence the relative accuracy of the given world points is approximately $6\mu m/15\ mm \approx 1 : 2500$.

---

[6]This assumption is possibly too optimistic, since there might be larger errors introduced by non-planarity and distortions of the transparency during the printing process.

Table 5.6: *Results of the distortion estimation and mean image residuals ϵ_{rms} for the periscopes. The principal point has been fixed to $(c_x, c_y) = (369.0, 250.0)$ (center of the circular field of view) in the first four lines. In the fifth line, the principal point has been estimated together with the first radial distortion parameter $((c_x, c_y) = (332.65, 194.89))$.*

k_1	k_2	t_1	t_2	$\epsilon_{rms}\ [pix]$	
-	-	-	-	1.365	no dist.
-0.020301	-	-	-	0.412	k_1
-0.020315	0.000006	-	-	0.412	k_1, k_2
-0.020251	0.000003	-1.159842	-0.714596	0.193	k_1, k_2, t_1, t_2
-0.019664	-	-	-	0.194	k_1, c_x, c_y

To realize a planar surface and a high contrast of the grid lines, the transparency is attached to a white glass plate. The glass plate is fixed in the observation plane in front of the periscopes, and calibration images are acquired by temporally averaging 100 frames. Example images are shown in figure 5.6. The strong radial distortion of the wide-angle periscope lenses is evident from these images. Prior to the extraction of the grid crossings, the illumination inhomogeneities are reduced using the method outlined in section 4.4.

Method.

To compute the camera parameters, we estimate a planar homography including a radial distortion parameter as described in section 5.3.3. The homography and the distortion model are estimated in separate linear optimizations, which are iterated until convergence.

Results.

The monoscopic image sequences obtained from the periscopes are processed by an optical flow algorithm to compute 2-D displacement vector fields. In this case, the relevant information from the geometric calibration is only the distortion model, since it can be used to account for the scale variations of the displacements introduced by lens distortion. Table 5.6 shows the results of the distortion estimation, together with the image residuals Δr. The image residuals are given as the rms errors of the Euclidean distance between the extracted feature points in the calibration images and the projected points using the estimated homography and distortion model. We obtain the following results:

- Without distortion model, the rms error is more than one pixel, which is very large. The simple homography cannot account for the curvature introduced by lens distortion.

- Using only k_1, the rms error is significantly reduced to ≈ 0.4 pixels. However, the next order distortion term (k_2) does not yield a smaller residuum.

- Including tangential distortion terms, the rms error is further reduced to ≈ 0.2 pixels. The same rms error can also be obtained by using only k_1, but allowing also the principal point to vary. This confirms the correlation between the principal point and the tangential distortion parameters, as mentioned in section 5.3.6.
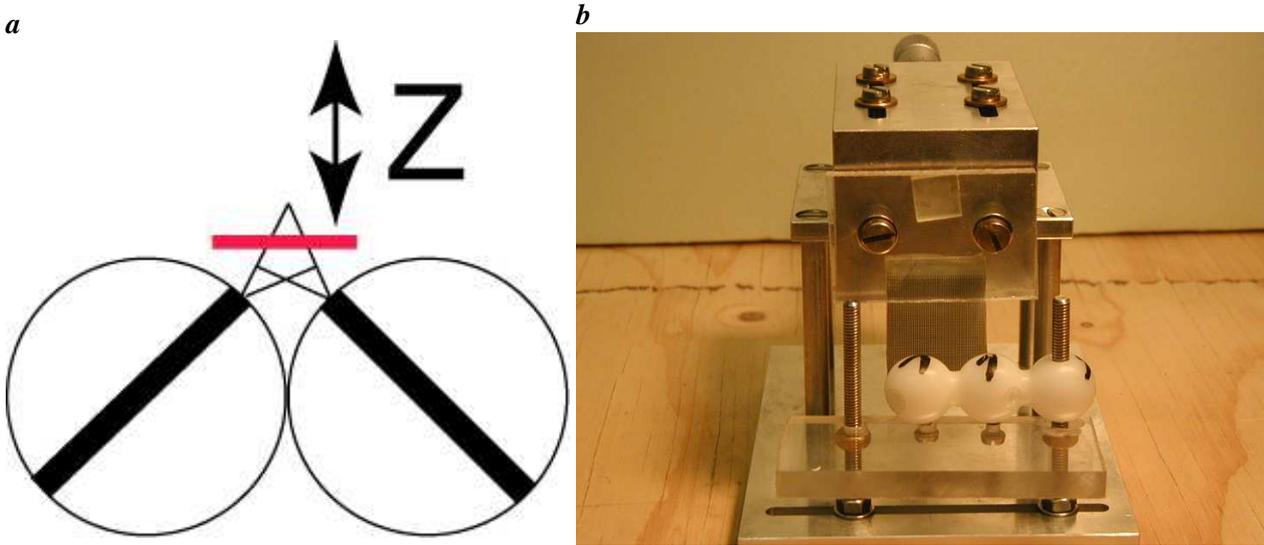
a

b

Figure 5.7: *Calibration by 'simulating' a 3-D object using a linear translation stage.* ***a*** *Sketch of the stereo rig (top view). The calibration plane is indicated by the red line.* ***b*** *Front view of the translation stage. Both the stereo rig and the translation stage are fixed to a ground plate made of steel.*

- The image diameter of the circular field of view is ≈ 520 pixels. Hence, the rms error of 0.2 pixels yields a relative accuracy of $\approx 1 : 2600$, which is in good agreement with the relative accuracy of the calibration grid.

5.4.2   Subsystems 'Jagst', 'Kocher', 'Höllbach' (flexible endoscopes)

Setup.

In his preliminary studies, Janßen [2000] mentions the infeasibility of available calibration methods for endoscopic setups, which is mainly caused by the small dimensions. He proposes to use a plane-based calibration method, but reports difficulties with the fixation of the calibration planes in arbitrary orientations. With the present setup, the problems become even more serious, because the convergence angle of the stereo rig has been increased (see section 9.3.2), and the observation volume is located very close to the endoscope lenses. Hence, a calibration grid has to be placed very close to the stereo rig, which limits rotations of the plane to very small angles. Larger rotation angles are necessary to obtain a stable multi-plane calibration. For these reasons, we do not apply a multi-plane method, but use a (simulated) 3-D target instead.

A 3-D target is created by fixing a planar calibration grid to a small translation stage, shifting the target in precisely known small steps along the Z-axis and thereby acquiring several images sequentially, see figure 5.7 and section 10.2.1. The Z-axis is perpendicular to the calibration plane. Both the stereo rig and the translation stage are fixed to a common ground plate made of steel, because the relative position of the stereo rig and the translation stage has to be absolutely stable during the acquisition of the images. Images of the same calibration points are acquired simultaneously by both endoscopes, and the origin of the common metric coordinate system for the later 3-D reconstruction is defined in the first calibration plane. The translation stage enables to shift the plane in minimum increments of $10~\mu m$, with a sensitivity of $\pm 1\mu m$. We use Z-increments of $0.2~mm$, which typically requires 15-20 calibration images to cover the stereo volume.
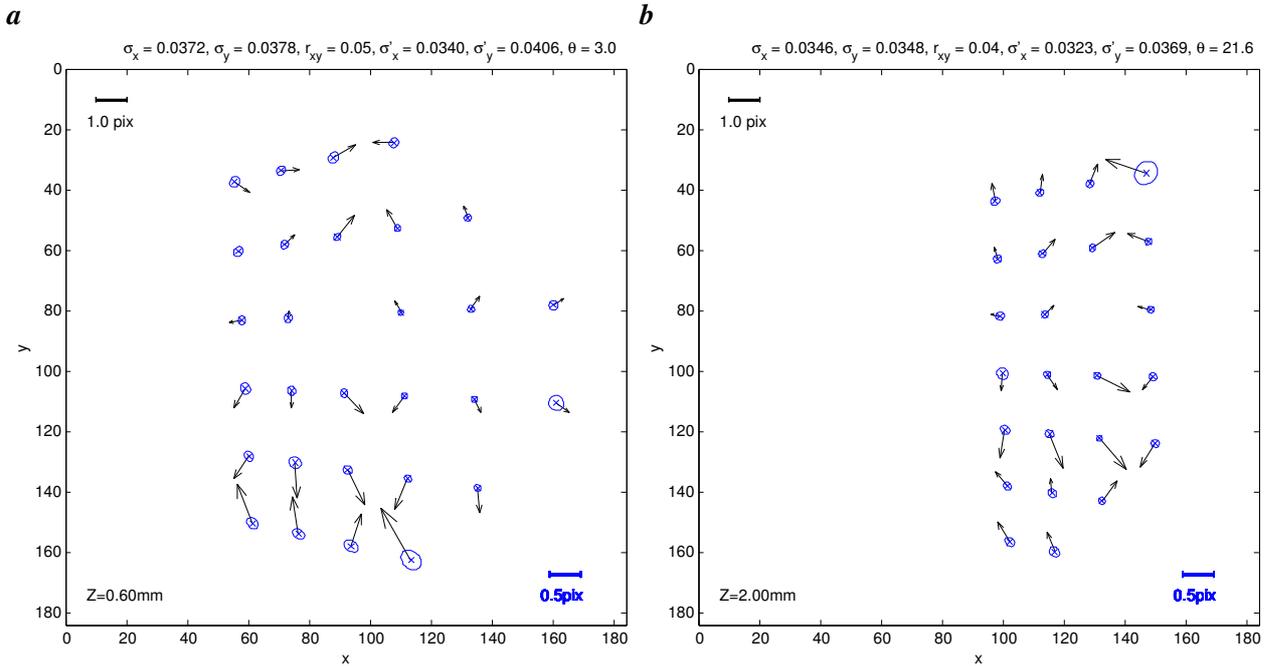
a   *b*

Figure 5.8: *Image residuals of the calibration of the Kocher-setup (calibration kmf01, left camera). Black arrows: reprojection errors of the estimated camera model. Blue ellipses: standard error ellipses of the feature extraction algorithm, see section 5.1 and figure A.1 in appendix A for explanations. Note the different scales. The reprojection errors are much larger than the feature extraction errors.* **a** *Calibration plane at $Z_w = 0.6$ mm.* **b** *Calibration plane at $Z_w = 2.0$ mm.*

The planar grid target is produced in the same way as that used for the periscopes, see section 5.4.1. Hence, we assume a world point accuracy of $\approx 6\ \mu m$ in the X- and Y-direction and of $\approx 1\ \mu m$ in the Z-direction. These values have to be considered as optimistic lower bounds, since additional errors are introduced by non-planarity and non-orthogonality of the grid and deviations from straigthness and flatness of the translation stage.

Method.

We assume that the world coordinates of the calibration points can be considered free of error and apply the resection method (robust DLT) outlined in sections 5.3.1–5.3.2 to compute the camera parameters. Since the introduction of lens distortion parameters (see table 5.7) did not reduce the image residuals, we do not use any distortion modeling.

Results.

Figure 5.8 shows the image residuals for two different calibration planes at $Z_w = 0.6$ mm and $Z_w = 2.0$ mm. The errors seem to be correlated along columns and lines of the calibration grid. There are systematic patterns in the error vectors, which do not correspond to radial or tangential distortion patterns. Since these patterns move over the image plane with the translation of the calibration plane along the Z-axis, they are most probably related to inaccuracies in the world coordinates of the calibration points (which have

Table 5.7: *Calibration results for the stereo rig 'Höllbach' (calibration hm08, left camera), using different estimation methods and camera models. DLT = Direct Linear Transformation. NL = Nonlinear optimization. The table is partitioned into four parts: internal camera parameters, external camera parameters and camera center (C_x, C_y, C_z), error measures in image space and error measures in 3-D space. For a discussion, see text.*

	DLT no dist.	NL no dist.	NL k_1	NL k_1, k_2	NL k_1, k_2, t_1, t_2
$f_x \, [mm]$	2.4729	2.4711	2.4894	2.4706	2.4731
$f_y \, [mm]$	2.4660	2.4623	2.4850	2.4653	2.4684
$s \, [-]$	0.0133	0.0155	0.0132	0.0134	0.0139
$c_x \, [pix]$	87.9705	93.6124	90.8360	91.5201	93.6917
$c_y \, [pix]$	86.1348	86.7716	86.5935	86.4175	90.5759
$k_1 \, [-]$	0	0	0.0234	-0.0222	-0.0219
$k_2 \, [-]$	0	0	0	0.0524	0.0555
$t_1 \, [-]$	0	0	0	0	-0.0024
$t_2 \, [-]$	0	0	0	0	-0.0017
$T_x \, [mm]$	0.4625	0.3655	0.4144	0.4030	0.3633
$T_y \, [mm]$	0.1748	0.1628	0.1660	0.1697	0.0941
$T_z \, [mm]$	4.1352	4.1406	4.1310	4.1323	4.1406
$\omega \, [°]$	-2.7977	-2.7841	-2.7730	-2.8354	-1.4901
$\phi \, [°]$	-41.568	-43.009	-42.321	-42.5053	-43.0977
$\kappa \, [°]$	4.5107	4.5788	4.5202	4.5688	3.6617
$C_x \, [mm]$	-3.0990	-3.1003	-3.0965	-3.0981	-3.0943
$C_y \, [mm]$	-0.0021	0.0016	0.0012	-0.0000	0.0023
$C_z \, [mm]$	-2.7822	-2.7736	-2.7706	-2.7692	-2.7702
$\langle \epsilon_x \rangle \, [pix]$	-0.00 ± 0.45	0.00 ± 0.46	0.00 ± 0.44	0.00 ± 0.44	-0.00 ± 0.44
$\langle \epsilon_y \rangle \, [pix]$	-0.00 ± 0.40	0.00 ± 0.47	0.00 ± 0.44	-0.00 ± 0.44	0.00 ± 0.44
$\langle \epsilon \rangle \, [pix]$	0.53 ± 0.27	0.57 ± 0.32	0.54 ± 0.31	0.54 ± 0.31	0.54 ± 0.31
$(\Delta X)_{rms} \, [\mu m]$	15.5	13.5	12.7	12.8	instab.
$(\Delta Y)_{rms} \, [\mu m]$	10.4	10.0	11.0	10.8	instab.
$(\Delta Z)_{rms} \, [\mu m]$	10.2	8.3	5.2	6.0	instab.
$\langle D_\perp \rangle \, [\mu m]$	5.1 ± 4.2	5.1 ± 4.1	2.7 ± 2.5	3.2 ± 2.9	instab.
NSCE [-]	1.95	1.72	1.77	1.75	119.78

been assumed to be free of error). Note that the largest errors are aligned along particular columns and rows of the calibration grid.

To check the performance of the calibration including distortion modeling, we compare results obtained with different numbers of distortion parameters in table 5.7. This table shows the obtained camera parame-

ters together with error measures both in image space and 3-D space for linear and nonlinear estimation. We compare the following approaches: robust DLT in the first column, and different versions of the nonlinear optimization (NL) (section 5.3.5) in columns two to five.

The following **error measures** are used. In *image space*, $\langle \epsilon_x \rangle$ and $\langle \epsilon_x \rangle$ are the mean differences (x- resp. y-component) between the measured feature points and the projected points using the estimated camera model, given as mean value and standard deviation over all calibration points. Further, we use the mean Euclidean distance between measured and projected points $\langle \epsilon \rangle$. In *3-D space*, we compute three different error measures. First, $(\Delta X)_{rms}$, $(\Delta Y)_{rms}$, $(\Delta Z)_{rms}$ are the rms component errors between the given 3-D coordinates of the calibration target (assumed free of error) and the reconstructed 3-D coordinates using the calibrated stereo rig (see section 6.2.2 for the reconstruction method). A second 3-D error measure is computed by fitting planes to the reconstructed 3-D points within one calibration plane. The fit planes are used as reference planes, and the mean distance $\langle D_\perp \rangle$ of the calibration points to their reference plane is computed. Finally, we compute the so-called *normalized stereo camera error* (NSCE) introduced by Weng et al. [1992]:

$$ NSCE = \frac{1}{n} \sum_{i=1}^{n} \left[\frac{(\hat{X}_i - X_i)^2 + (\hat{Y}_i - Y_i)^2}{\hat{Z}_i^2 (\alpha^{-2} + \beta^{-2})/12} \right]^{1/2}, \tag{5.33} $$

where $(\hat{X}_i, \hat{Y}_i, \hat{Z}_i)$ are the reconstructed 3-D points, (X_i, Y_i, Z_i) is the known 3-D ground truth, and α, β are the effective focal lengths in x- and y-direction in units of pixels. Looking at (5.33), we see that the NSCE is the mean of the ratio of the lateral triangulation error to the lateral standard deviation of the pixel digitization noise, back-projected to the depth of the corresponding 3-D point. The underlying idea of the NSCE is that the basic factor limiting 3-D accuracy is the pixel resolution of the digital images. The NSCE compares the obtained 3-D accuracy to the potential provided by the given image resolution. An NSCE value of unity indicates a good calibration. Values larger than one indicate that for some reasons the highest possible accuracy has not been achieved. The NSCE is an error measure that is independent of the stereo rig geometry (field of view, depth range, convergence angle, length of baseline).

To compute the 3-D error measures, the reconstructed world points have been computed using the method described in section 6.2.2. Since the image points in the left and the right camera belong to the same calibration target with known world coordinates, the stereoscopic correspondences that are necessary to compute 3-D points are trivially known.

We can draw the following conclusions from table 5.7:

- The image residuals (rms errors) are $\approx 0.4 - 0.5$ pixels. While being only half as large as those obtained by Janßen [2000], they are still quite large as compared to typical calibrations of C-mount lenses in air (factor 5-10 larger). Lens distortion parameters do not reduce the residuals. Hence, the limiting factors might be the accuracy of the 3-D world points (which are assumed free of error in the calibration) and the limited geometric precision and image resolution of the fiber bundles in the flexible endoscopes.

- The computed camera center is very stable throughout all methods. However, we can see slightly different values of the principal point and the angle ϕ, which reveals the correlations between these parameters.

- The estimated 3-D accuracy is $\approx 10 \ \mu m$. It is larger for the X-component ($\approx 15 \ \mu m$). The reason is that the X- and Z-component of a 3-D point are particularly affected by the perspective projection,

since the optical axis is aligned with the diagonal between the X- and Z-axis of the world coordinate frame. The smaller error in the Z-component is probably due to the higher Z-accuracy of the target points (which is $1~\mu m$, the sensitivity of the translation stage).

- The mean distance between the reconstructed 3-D points and their estimated reference planes is $\approx 5~\mu m$.

- Although the lens distortion parameters do no reduce the residuals in image space, they do reduce the 3-D errors (at least in the Z-component). This means that even very small changes in the image coordinates may yield better 3-D estimates. However, distortion parameters also introduce instabilities in the 3-D reconstruction, since the latter is carried out numerically in the present implementation (see section 6.2.2). For example, if the full distortion model was used, the 3-D reconstruction did not work at all because of singularities in the estimation. Hence, we still do not use distortion parameters.

- The NSCE is always larger than one, indicating that the maximum 3-D accuracy (limited by the image resolution) is not achieved. The reason may again be the inaccuracy introduced by endoscopic imaging. Note also that in the computation of the NSCE we assume the given 3-D coordinates as error-free ground truth, which is not true.

- The image diameter of the circular field of view of the endoscopes is 170 pixels, hence the relative calibration accuracy in image space is $\approx 1 : 400$. Note that the relative accuracy based on the feature extraction accuracy (0.05 to 0.1 pixels) is much larger. The relative accuracy in object space varies with depth, see also figure 6.4, and is approximately $1 : 200$ to $1 : 400$. It is lower than in image space, which is also expressed by the NSCE values larger than one. The lower accuracy in object space indicates shortcomings in the performance of the camera model, which are supposed to be caused by inaccuracies of the fiber bundle.

5.4.3   Subsystem 'Neckar' (free surface flow)

Setup.

For the geometric camera calibration of the Neckar-setup, a planar calibration target of side length $\approx 15~cm$ is necessary. The target is made of a sheet of opalescent glass, on which a regular quadratic grid of grid spacing $1~cm$ has been imprinted (accuracy $\approx 10~\mu m$). To calibrate the cameras, this grid is put inside the stereo volume and viewed by both cameras. The calibration is carried out with the target submerged in water. Multiple images (≈ 15) of this grid in different orientations and positions are acquired. In the first view, the grid position and orientation relative to the channel walls and the gravel bed has to be measured, because the first calibration plane defines the coordinate system of the measurement results. In particular, the experimental zero-level of the logarithmic velocity profile is determined by the first calibration plane. For an exact determination of the zero-level, the calibration grid is mounted on a plane steel plate of about $25 \times 50~cm^2$ size. The gravel surface in the measurement area is leveled using this steel plate and a spirit-level. The vertical position (i.e. the zero-level) is measured by dropping a perpendicular on the plate.
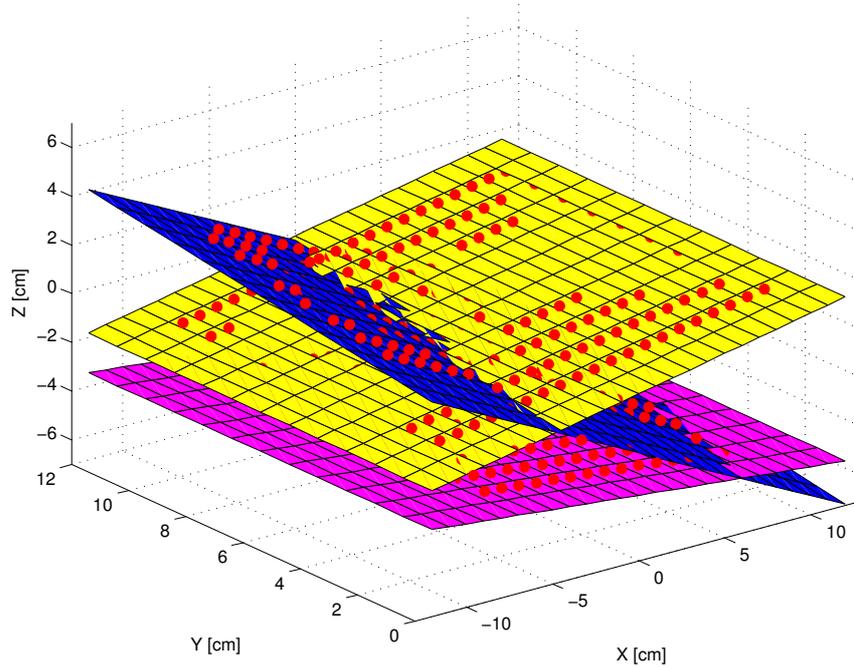
Figure 5.9: *Three references planes obtained by fitting a plane to the estimated 3-D points of the planar calibration target in three different orientations. The red markers show the reconstructed 3-D points, computed by the method explained in section 6.2.2.*

Method.

Camera parameters are computed by the multi-plane method explained in section 5.3.4, using the implementation of Heikkilä [2000] (section 5.3.5). As explained in section 9.4.3, image sequences of the free surface flow are acquired in the partial scan mode, reading out only the lower 200 lines of the sensor. The reason is the increased frame rate of $130 \, Hz$. However, the calibration images are acquired in full scan mode because the pinhole camera model describes the imaging process onto the complete sensor. Hence, a good calibration can only be achieved by taking into account the whole sensor and using calibration points distributed throughout the whole image plane. To use the full-frame calibration also for partial scan images, we have to check that there is a unique and stable correspondence between the image lines of images captured in full resp. partial scan mode. This has been confirmed by acquiring images of the same calibration grid both in full scan and partial scan mode without changing the relative position of the camera and the grid. A constant offset of 269 image lines has been found between both scan modes. Hence, row numbers in full scan mode are equal to row numbers in partial scan mode plus the constant offset.

Results.

We use the same error measures as in section 5.4.2 to assess the quality of the calibration results. However, since a multi-plane method is used here, no ground truth is available, since the orientations of the planes are also estimated. The 3-D coordinates are known only for the calibration points in the first plane, since this plane defines the origin of the world coordinate system, hence $Z_w = 0$ for all points within this plane. We assume the world coordinates of this plane to be free of error and use only the points of the first plane to

Table 5.8: *Calibration results for the stereo rig 'Neckar', using different camera models and the estimation method of Section 5.3.5 (right camera of calibration nm08).*

	NL	NL (k_1)	NL (k_1, k_2)	NL (k_1, k_2, t_1, t_2)
f_x [mm]	31.0442	31.0069	30.9745	31.1771
f_y [mm]	31.1231	31.1122	31.0405	31.3388
s [−]	−0.0643	0.3543	0.3213	−0.1587
c_x [pix]	303.6818	320.0515	329.6208	415.6912
c_y [pix]	275.8413	309.7256	314.2721	48.6522
k_1 [−]	0	−0.00022627	−0.00019847	−0.00019188
k_2 [−]	0	0	0.00000002	0.00000351
t_1 [−]	0	0	0	0.00112635
t_2 [−]	0	0	0	−0.000441
T_x [cm]	−0.1316	−0.5547	−0.7974	−2.9602
T_y [cm]	1.0133	0.1516	0.0362	6.7907
T_z [cm]	88.0622	88.0690	87.8702	87.7557
ω [°]	−3.0376	−3.3517	−3.1543	−8.0946
ϕ [°]	43.9025	43.6553	43.5783	41.87377
κ [°]	−2.9052	−2.3320	−2.2534	−6.4123
C_x [cm]	61.1970	61.2011	61.1512	61.3313
C_y [cm]	2.3517	3.5739	3.4676	2.4996
C_z [cm]	−63.2882	−63.2307	−63.0104	−63.1520
$\langle \epsilon_x \rangle$ [pix]	0.00 ± 0.16	0.00 ± 0.16	0.00 ± 0.16	−0.00 ± 0.16
$\langle \epsilon_y \rangle$ [pix]	0.00 ± 0.16	0.00 ± 0.16	0.00 ± 0.16	0.00 ± 0.16
$\langle \epsilon \rangle$ [pix]	0.19 ± 0.12	0.19 ± 0.12	0.18 ± 0.12	0.19 ± 0.11
$(\Delta X)_{rms}$ [μm]	90.9	75.0	77.5	70.0
$(\Delta Y)_{rms}$ [μm]	47.1	48.8	49.5	25.4
$(\Delta Z)_{rms}$ [μm]	31.0	48.6	53.5	40.5
$\langle D_\perp \rangle$ [μm]	26.8 ± 23.2	79.8 ± 63.8	37.4 ± 29.2	128.7 ± 828.3
NSCE [-]	0.84	0.68	0.69	0.61

compute the 3-D error measures. Further, we also use the method of fitting reference planes and computing the mean distance of calibration points to their reference planes. An example is shown in figure 5.9.

Table 5.8 shows the calibration results, in particular:

- The image residuals (rms errors) are $\approx 0.1 - 0.2$ pixels. They are comparable to typical residuals that are obtained using similar equipment, e.g. [Clarke et al., 1993; Wei and Ma, 1994; Willson and Shafer, 1994; Wunsch et al., 1996; Devernay and Faugeras, 2001]. Lens distortion parameters do not reduce the residuals. Hence, no distortion modeling is used.

- The 3-D accuracy estimated from the plane fits is $\approx 30\ \mu m$. Lens distortion parameters do no significantly improve the 3-D accuracy, but again introduce instabilities, especially if the full distortion model is used.

- The NSCE is always lower than one, indicating that the maximum 3-D accuracy (limited by the image resolution) is achieved.

- The relative calibration accuracy in image space is $\approx 1 : 4000$ (based on the image diagonal), which is comparable to the relative accuracy in object space and the relative accuracy of the feature extraction. Hence, the estimated camera model provides an accurate description of the imaging process.

5.5   Summary

In this chapter, a detailed description of the camera calibration algorithms used in this work has been given. All approaches are based on images of a planar calibration grid. The grid line crossings serve as calibration markers. A method has been described to determine the positions of the crossings with a high subpixel accuracy of below 0.1 pixels. The well-known linear pinhole camera is used to model both the endoscopic imaging and the C-mount lenses. The camera model is defined in terms of a set of camera paramters, which are estimated by least squares algorithms. The over-determined equation system that has to be solved is given by a number of corresponding world and image points. The world coordinates are those of the calibration points on the target, which are assumed to be free of error. The corresponding image coordinates are determined by the feature extraction algorithm.

Different methods to estimate camera parameters have been discussed. The simplest approaches estimate a perspective projection matrix by linear least squares methods. The criticism of these methods is that they do not achieve the best possible accuracy because they minimize an algebraic error, which has no direct physical meaning. However, it has been shown in the literature [Hartley, 1997b; Liu and Männer, 2003] and confirmed here that in absence of degenerate configurations and using an appropriate data normalization, linear methods can obtain results close to optimal. This has been shown by comparing results obtained by linear estimation to those obtained by a nonlinear Levenberg-Marquardt optimizer. The latter minimizes a physically-based geometric error, namely the sum of squared differences between the points measured in the images and the corresponding projected points using the estimated camera model. The geometric error is also only an approximation[7] to the cost function that should actually be minimized to obtain the best possible results. This cost function is given by the statistical error, which takes into account the statistical error structure of the given data. Minimizing the statistical error corresponds to a maximum-likelihood estimation of the camera parameters *and* the object points (which are assumed known and free of error in the other methods), known as bundle adjustment in photogrammetry.

The problems related to estimating lens distortion parameters have been discussed in some detail. If lens distortion is small and there are other systematic and/or random noise sources that are responsible for deviations comparable in magnitude to those of lens distortion, distortion models should not be used, because of the danger of modeling the noise instead of the lens distortion. The latter may result in a worse performance of the 3-D coordinate estimation. It has been shown that only the wide-angle lenses of the rigid

---

[7]The approximation is the assumption that the noise is Gaussian on each image coordinate with zero mean and uniform standard deviation.

endoscopes show a relatively large radial distortion that can be reliably modeled. In the other setups, lens distortions are very small. In addition, there are other systematic error sources, namely small deviations in the arrangement of the fiber bundle between entry and exit side in the case of the flexible endoscopes and effects related to the multimedia geometry in the case of the free surface flow setup. For these reasons, lens distortion is not modeled.

The accuracy of the calibration has been estimated using different error measures. With the endoscopic stereo setups, the rms error in image space is ≈ 0.4 pixels, and the rms error in object space is $\approx 5 - 10 \ \mu m$. The best possible accuracy (limited by the image resolution) is not achieved due to the effects of the fiber bundle mentioned above. In the free surface flow setup, the image residuals are ≈ 0.2 pixels, and the 3-D error is $\approx 30 \ \mu m$. Here, a better accuracy as that limited by the pixel discretization has been achieved.

Chapter 6

Stereo vision

In this chapter, the geometric concepts related to a stereo camera setup are discussed. The task of a stereo system is to reconstruct the 3-D coordinates of object points in a scene, given two different views of the scene (in the following called the *left image* and the *right image*). Towards this end, the first problem to be solved is to establish **stereoscopic correspondences**: for each image point in the left image, the corresponding image point in the right image has to be found. Many applications aim for a reconstruction of 3-D surfaces. They rely on a **region-based matching** of a stereoscopic image pair. To obtain dense disparity maps, corresponding image points are identified by the similarity of the gray value patterns in the left and right image. These approaches are very similar to the region-based velocity estimators described in chapters 3.2–3.4. The only difference is that a disparity map between two stereo images is estimated instead of a displacement field between two subsequent images of a monoscopic sequence. For recent examples, see e.g. [Mühlmann, 2002] (reconstruction of static surfaces, including an endoscopic application) or [Hilsenstein, 2004; Fuß, 2004] (reconstruction of dynamic water surfaces).

In the present application, we use a **feature-based matching**, since we want to estimate single 3-D object points, namely the positions of the tracer particles. The geometric constraint that makes the correspondence search feasible is given by the **epipolar geometry** of a stereo rig, which is discussed in section 6.1. Once the correspondence between two points in both images has been found, the 3-D position of the corresponding object point can be computed by **triangulation** of the projection rays, which is the subject of section 6.2. For in-depth discussions of stereo vision, see e.g. [Xu and Zhang, 1996; Trucco and Verri, 1998; Hartley and Zisserman, 2000]. A recent review of stereo methods is given by Brown et al. [2003].

6.1 Stereo correspondence and epipolar geometry

Given two views of a set of scene points and no further information, the search for corresponding points has to be carried out throughout the whole image plane. However, if the epipolar geometry of the stereo setup is known, the **epipolar constraint** can be exploited to reduce the search space to a line in the second image, the so-called **epipolar line** (section 6.1.1). Hence, the search space is reduced from 2-D to 1-D. The algebraic expression of the epipolar constraint is formulated with the **fundamental matrix**, which is a (3×3)-matrix of rank 2. It is possible to estimate the fundamental matrix only from a number of corresponding image points (section 6.1.2). A metric calibration of the stereo rig is not necessary. However, since in general the correspondences are a priori unknown, estimation of the epipolar geometry (i.e. the fundamental matrix)
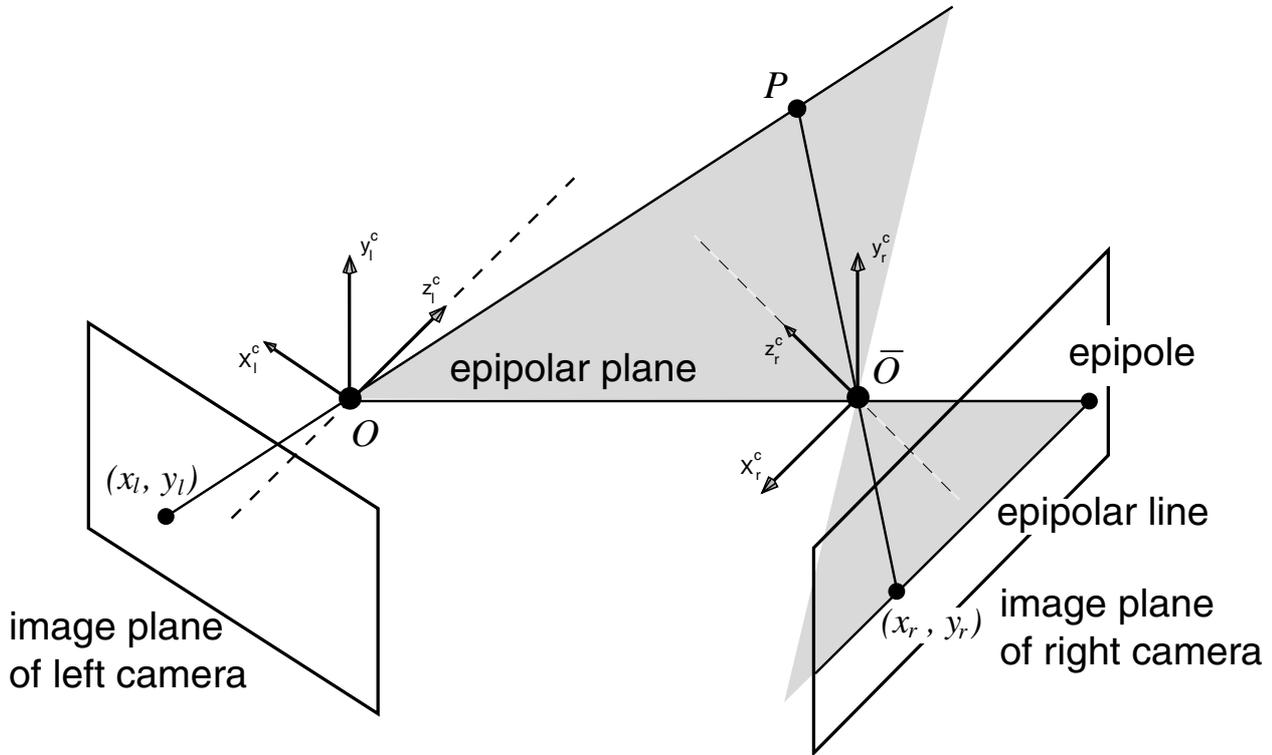
Figure 6.1: *Epipolar geometry of a stereo camera setup. A world point* **P** *is projected into both cameras. Corresponding image points are constrained to lie on the so-called epipolar lines, which are the projections of the line of sights connecting the image point, the optical center and the world point.*

is a kind of chicken-egg problem. Usually, robust methods have to be applied, which iteratively estimate correspondences and the epipolar geometry, e.g. within a Least Median of Squares framework [Zhang et al., 1995].

6.1.1   Epipolar lines and the fundamental matrix

Epipolar geometry.

Given a stereo pair of cameras and a 3-D point $\mathbf{P}_w$[1], this point and the two projection centers of the cameras define a 3-D plane, the so-called **epipolar plane**, see figure 6.1. The image lines, where the epipolar plane intersects the image planes are called **epipolar lines**. The epipolar lines in the right camera are the images of the 3-D projection rays of the left camera and vice versa. The image in one camera of the projection center of the other camera is called **epipole**. All epipolar lines of one camera intersect in the camera's epipole. Corresponding image points in the left and the right camera are constrained to lie on conjugated epipolar lines. This reduces the stereo correspondence problem to a 1-D search.

---
[1]Akin to chapter 5, we apply a projective framework and use homogeneous coordinates.

Essential matrix.

Consider a 3-D point $\mathbf{P}_w$ and its two image points in the left and the right camera, $\mathbf{p}_{n,l}$ and $\mathbf{p}_{n,r}$, given in normalized coordinates (see section 5.2.2). Then, the two projection centers, the 3-D point and the two image points all lie within the epipolar plane. Mathematically, this coplanarity is expressed as

$$\mathbf{p}_{n,r}^T \boldsymbol{E} \mathbf{p}_{n,l} = 0, \tag{6.1}$$

with

$$\boldsymbol{E} = \boldsymbol{RS} = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix} \begin{pmatrix} 0 & -T_z & T_y \\ T_z & 0 & -T_x \\ -T_y & T_x & 0 \end{pmatrix}, \tag{6.2}$$

where the matrices $\boldsymbol{R}$ (rotation) and $\boldsymbol{S}$ (translation) define the 3-D rigid transformation between the left and right camera coordinate system.

The matrix $\boldsymbol{E}$ is called the **essential matrix**. Note that computations using the essential matrix are carried out in normalized coordinates. Since the coordinates actually measured from images are image coordinates (or pixel coordinates), it is necessary to know the transformation from pixel coordinates to normalized coordinates. This transformation is given by the inverse of the calibration matrix $\boldsymbol{K}$ (intrinsic camera parameters). Thus, calculations involving the essential matrix can only be done for calibrated cameras. Estimation of the essential matrix is closely related to the structure-from-motion problem in computer vision, e.g. [Weng et al., 1993b]. In the case of uncalibrated cameras, the essential matrix is replaced by the **fundamental matrix**, see below.

In summary, the essential matrix has the following properties:

- Since the matrix $\boldsymbol{S}$ is singular, the essential matrix is also singular.

- It is a homogeneous matrix, hence only defined up to scale.

- It encodes information on the six extrinsic parameters only.

- Its two non-zero singular values are equal.

- Thus, it only has five degrees of freedom (three parameters for rotation and two parameters defining the direction of translation between the two projective centers).

- It is defined in terms of normalized coordinates, restricting its application to the case of calibrated cameras.

Fundamental matrix.

Let $\boldsymbol{K}_l$ and $\boldsymbol{K}_r$ be the calibration matrices of the internal camera parameters of the two cameras, according to (5.6). The pixel coordinates of the two image points of $\mathbf{P}_w$ are given by $\mathbf{p}_{p,l}$ and $\mathbf{p}_{p,r}$. Then, the transformation from pixel coordinates to normalized coordinates is given by

$$\mathbf{p}_{n,l} = \boldsymbol{K}_l^{-1} \mathbf{p}_{p,l} \tag{6.3}$$

and

$$\mathbf{p}_{n,r} = \boldsymbol{K}_r^{-1} \mathbf{p}_{p,r}. \tag{6.4}$$

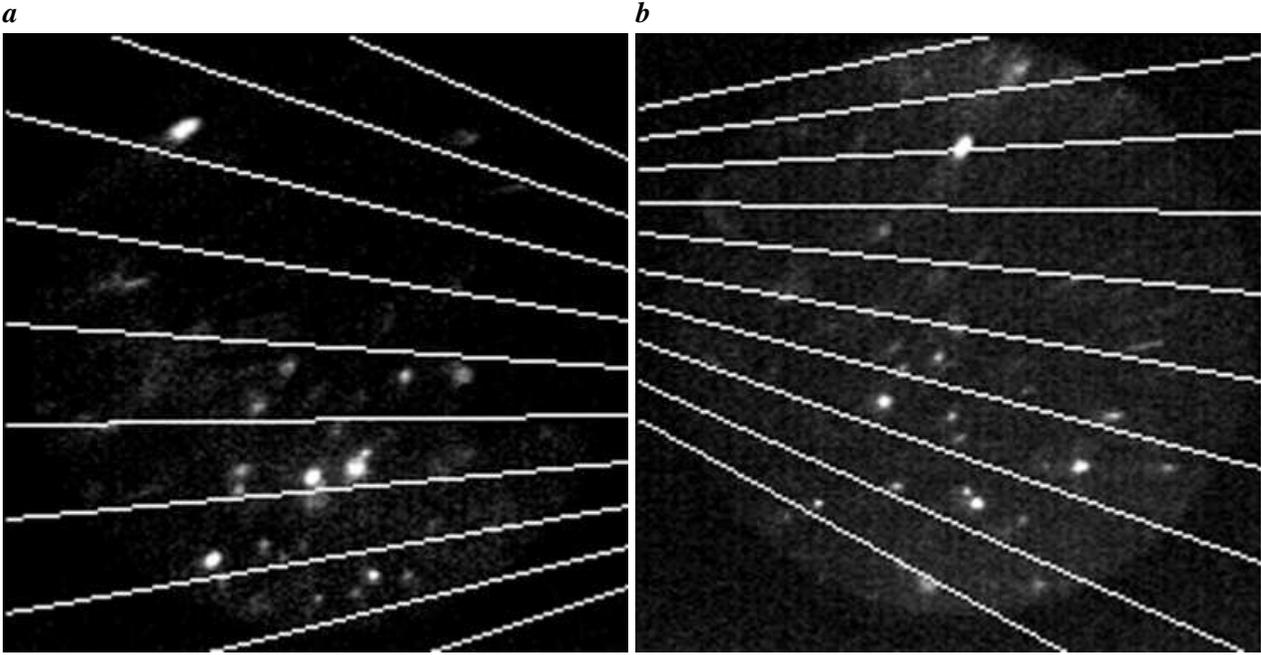a   b

Figure 6.2: *Epipolar geometry of the Höllbach setup. A number of arbitrarily selected epipolar lines is drawn in two corresponding particle images. The lines in the left and right image do not correspond to each other. **a** Left image. **b** Right image. The epipoles are located outside the image borders.*

Substituting (6.3) and (6.4) into (6.1), we have

$$\mathbf{p}_{p,l}\boldsymbol{F}\mathbf{p}_{p,r} = 0, \tag{6.5}$$

with

$$\boldsymbol{F} = \boldsymbol{K}_r^{-T}\boldsymbol{E}\boldsymbol{K}_l^{-1}. \tag{6.6}$$

The matrix $\boldsymbol{F}$ is called the **fundamental matrix**. It has the following properties:

- Since the matrix $\boldsymbol{E}$ is singular, the fundamental matrix is also singular.

- It is only defined up to scale.

- It encodes information on the extrinsic parameters *and* on the intrinsic parameters.

- It has seven degrees of freedom.

- It is defined in terms of pixel coordinates, thus its application is not restricted to the case of calibrated cameras. In fact, by estimating the fundamental matrix one can reconstruct the epipolar geometry of two cameras without any knowledge of the extrinsic and intrinsic parameters of the cameras.

Note that the fundamental matrix provides an easy way of calculating the epipolar line $\mathbf{l}_r$ corresponding to a given image point $\mathbf{p}_{p,l}$:

$$\mathbf{l}_r = \boldsymbol{F}\mathbf{p}_{p,l}, \tag{6.7}$$

and similarly

$$\mathbf{l}_l = \boldsymbol{F}^T\mathbf{p}_{p,r}, \tag{6.8}$$

Table 6.1: *Comparison of two different methods (metric and non-metric) to compute the fundamental matrix. The accuracy of both methods is estimated by computing the rms distance of points to their corresponding epipolar lines. The rms distance is computed symmetrically, i.e. both in the left and in the right image plane.*

setup	Höllbach	Jagst	Neckar
d_{rms} [pix] (metric)	0.25	0.26	0.25
d_{rms} [pix] (non-metric)	0.21	0.23	0.19

where $\mathbf{l}_r$ and $\mathbf{l}_l$ are the homogeneous representations of the epipolar lines. Given a line $\mathbf{l} = (l_1, l_2, l_3)^T$ and a point $\mathbf{p} = (x, y, 1)^T$, the distance of the point to the line is

$$d(\mathbf{p}, \mathbf{l}) = \frac{l_1 x + l_2 y + l_3}{\sqrt{l_1^2 + l_2^2}}. \tag{6.9}$$

The epipolar constraint can be implemented by computing the epipolar line for a given image point in the left image (using the fundamental matrix) and selecting a correspondence candidate in the right image by imposing a threshold on (6.9). The size of the threshold has to be chosen according to the accuracy of the feature extraction and the estimated epipolar lines. An example of the epipolar geometry of an endoscopic stereo rig is shown by a number of epipolar lines in figure 6.2.

6.1.2 Estimation of the fundamental matrix

The computation of the epipolar geometry resp. the fundamental matrix is an important step in many structure and motion algorithms. Hence, a lot of research on this topic has been carried out in the computer vision community. Zhang [1998] and Torr and Murray [1997] give recent reviews, where the latter focuses on robust estimation methods. The books by Zhang and Faugeras [1992] and Weng et al. [1993b] give detailed analyses of recovering the epipolar geometry in dynamic scenes, i.e. from image sequences.

A very simple method to estimate the fundamental matrix has been put forward by Hartley [1997a]. While most methods require large-scale nonlinear optimization and a careful parameterization of the fundamental matrix (to account for its singularity), the method of [Hartley, 1997a] only uses linear least squares. Hence, it enjoys some popularity due to its ease of implementation. If the data is properly scaled, very good results that are close to the optimum values of the nonlinear algorithms can be obtained (see also [Hartley, 1997b, 1998] and the discussion of linear camera calibration in section 5.3.1). Recently, another (iterative) linear method has been proposed by Liu and Männer [2003] and shown to perform equally well as nonlinear Levenberg-Marquardt optimizers. However, several aspects have to be kept in mind, both concerning the numerical stability of the estimation and the ill-posedness of the problem in certain geometric configurations. Given that the configuration allows for a unique solution, usually linear algorithms are sufficient (e.g. that of [Hartley, 1997a]). To achieve the highest possible accuracy, a linear estimation may be followed by iterative nonlinear refinement. The latter might be necessary because the linear algorithms do not take into account the constraints of the fundamental matrix, which is a singular, homogeneous (3×3)-matrix. Hence, it has only seven degrees of freedom. To account for these constraints, special parameterizations of the fundamental matrix may be introduced, which turn the estimation into a nonlinear least squares problem.

In this work, we use two different methods to compute the fundamental matrix. The first is the direct computation of $\boldsymbol{F}$, using the known camera parameters obtained from the geometric camera calibration

(see section 5.4). Since both cameras are calibrated within the same metric world coordinate system, the fundamental matrix can be computed analytically from the two projection matrices. For details of the computation, see [Xu and Zhang, 1996]. We refer to this method as the **metric approach**. In the 3-D PTV implementation, we do not compute the fundamental matrix, but instead perform an explicit determination of the 3-D projection rays and hence the epipolar lines, see section 6.2.2. The results are equivalent.

The second method to compute the fundamental matrix provides a further check of the geometric camera calibration, since the epipolar geometry is estimated independently of the calibration results. The fundamental matrix is computed using only implicit geometric information, i.e. a sufficient number of image point correspondences without any further information on the 3-D geometry of a calibration object or properties of the cameras (e.g. an underlying pinhole model). In the present applications, image point correspondences are trivially given, since images of the same calibration target with known world-coordinates are available for both cameras. We use these correspondences to estimate the fundamental matrix by the method proposed by Boufama and Mohr [1998]. The details of the method are given there. We refer to this method as the **non-metric approach**.

Table 6.1 shows that both approaches achieve the same accuracy. The rms distance of points to their corresponding epipolar lines is ≈ 0.25 pixels. This distance is quite close to the rms image residuals of the camera calibration. In the non-metric approach, no world coordinates are used in the computation, hence errors in world coordinates do not influence the results of the non-metric approach. Since both metric and non-metric approach achieve the same accuracy, the limiting factor in the metric approach is not the accuracy of the given world points, but deviations of the imaging process from an ideal pinhole camera. The latter are introduced by the fiber bundles in case of endoscopic imaging and by multimedia effects in case of the C-mount lenses observing the free surface flow.

6.1.3   Calibration of a stereo rig

The most important parameters of a stereo camera system are those describing the *relative orientation* of the cameras with respect to each other. The relative orientation is required for the calculation of 3-D world coordinates by triangulation (section 6.2) and allows for a simple determination of the fundamental matrix (see the 'metric approach' in section 6.1.2).

There are two options to compute the relative orientation of a stereo rig. If the two cameras are calibrated with respect to the same metric world coordinate system, the relative orientation is easily computed from the external camera parameters, see Xu and Zhang [1996]. The second option is to perform an initial calibration of each camera independent of the other as a first step. In a second step, the camera parameters of both cameras are refined simultaneously in a further nonlinear optimization, taking into account the additional constraint of the fixed relative orientation between the two cameras. Both approaches have been tested in this work. Since the second option did not improve the final accuracy, we only use the first option. In all stereo calibrations, the same calibration planes are projected simultaneously into both cameras. The global world coordinate system used for both cameras is defined by the first calibration plane (with the Z-axis perpendicular to the plane).

6.2 3-D reconstruction

In order to reconstruct a 3-D point, we have to solve the back-projection problem: given an image point (possibly affected by lens distortion), the projection ray corresponding to this point has to be found. If two projection rays of the two different views are available, the corresponding 3-D point is given by the intersection of the two rays (**triangulation**). We discuss our triangulation approach in section 6.2.1. In section 6.2.2, the explicit computation of the projection rays is outlined, including a method to compute curved epipolar lines in the presence of lens distortion. An accuracy assessment is given in section 6.2.3.

6.2.1   Triangulation method

The two estimated projection rays will only meet exactly in 3-D space in the noise-free case of exactly known camera parameters and image points. Obviously, both the camera parameters and the image points will be noisy in practical applications. Hence, one has to find an optimum solution under the given noise level. One common method is to choose the midpoint of the perpendicular distance between the two projection rays. Hartley and Sturm [1997] discuss several triangulation methods. They show that the midpoint method and linear triangulation methods do not perform well in cases where the projection matrices are only known up to a projective ambiguity (i.e. for *projective reconstructions*). The reason is that geometric relations like distance and orthogonality have no meaning in a projective reference frame. However, since we work in a calibrated metric reference frame, this is not a critical issue here. Further, Hartley and Sturm [1997] show that there is virtually no difference in the obtained 3-D accuracy between the midpoint method and other more sophisticated triangulation methods for image noise levels of up to four pixels (rms). Only for higher noise levels, the other methods perform better. Hence, we use the midpoint method due to its ease of implementation.

6.2.2   Explicit determination of projection rays and epipolar lines

Linear case.

The first step of the triangulation is to compute the projection rays. Towards this end, we can use the pseudo-inverse of the projection matrix P, which maps image points to their optical rays, see [Hartley and Zisserman, 2000]. However, we choose a different approach. In a 3-D PTV application, we can choose two planes in the world coordinate system, $Z = Z_{min} = const.$ and $Z = Z_{max} = const.$, corresponding to the minimum and maximum Z-coordinate of the stereo volume, see figure 6.3. Given an image point $\mathbf{p} = [x_p, y_p, 1]^T$ in the left camera, we can compute the two world points $\mathbf{P}_0 = [X_0, Y_0, Z_0 = Z_{min}, 1]^T$ and $\mathbf{P}_1 = [X_1, Y_1, Z_1 = Z_{max}, 1]^T$ by solving two systems of two linear equations in the two unknowns X_0, Y_0 resp. X_1, Y_1. The equations are given by inserting $\mathbf{P}_0$ resp. $\mathbf{P}_1$ and $\mathbf{p}$ into (5.14). The result is an algebraic representation of the projection ray, based on the two points $\mathbf{P}_0$ and $\mathbf{P}_1$. A similar representation of the corresponding epipolar line segment in the right camera can simply be computed by projecting $\mathbf{P}_0$ and $\mathbf{P}_1$ into the right camera.

Including lens distortion.

The backprojection problem is more difficult if a distortion model is used. The reason is that it is not possible to compute an analytical inverse of the distortion model, which is a multi-variable polynomial
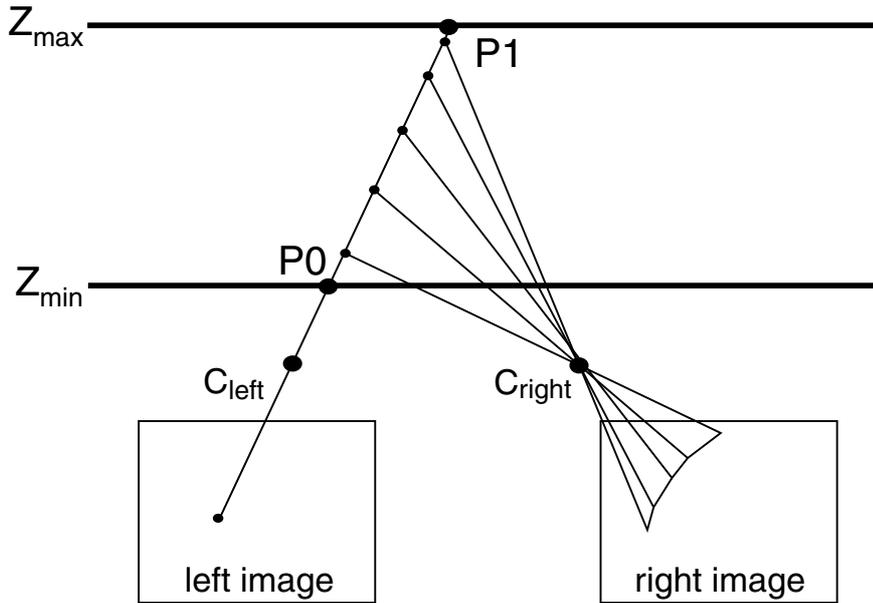
Figure 6.3: *In a PTV application, the observation volume is known to lie between two planes $Z = Z_{min} = const.$ and $Z = Z_{max} = const.$. Given an image point in the left camera, we can compute the two world points $\mathbf{P}_0$ and $\mathbf{P}_1$, which yield a direct geometric representation of the projection ray. If lens distortion has to be considered, we can obtain a piecewise linear parameterization of the corresponding curved epipolar line by sampling the projection ray.*

function. However, to compute the projection rays linearly as outlined above, we first have to compute the ideal undistorted image points, given the measured distorted ones. Two solutions are possible:

1. **Application of an undistortion model**: In addition to the distortion model that is estimated along with the other camera parameters, an undistortion model is estimated subsequently. Towards this end, a regular array of image points is distorted using the distortion model obtained in the camera calibration. These point sets are then used to estimate the parameters of the undistortion model, which maps distorted image points to undistorted ones. This approach is used e.g. by Heikkilä [2000]. Note that it is possible to use the same function (with different numerical values of the parameters) both as distortion and undistortion model (see the discussion in section 5.2.2).

2. **Numerical inversion of the camera model**: The second solution does not require an additional undistortion model. An initial (distortion-free) solution for the projection ray is computed using the linear method explained above, yielding the points $\mathbf{P}_0$ and $\mathbf{P}_1$. Subsequently, we refine these points by iterative minimization of the following cost function c:

$$c = \left(\mathbf{p} - \mathbf{d}_+ \left(\mathsf{P}\mathbf{P}_0 |_{Z_0 = Z_{min}} \right) \right)^2 , \tag{6.10}$$

where $\mathbf{p}$ is the (distorted) image points, $\mathbf{d}_+$ is the distortion model, P is the projection matrix and $\mathbf{P}_0$ is the sought world point (with Z-coordinate fixed to $Z_0 = Z_{min}$). Minimization of c yields refined coordinates X_0' and Y_0' of the world point $\mathbf{P}_0' = [X_0', Y_0', Z_{min}, 1]^T$ which is projected to $\mathbf{p}$ by P followed by the lens distortion model. A similar minimzation is carried out to find the refined point $\mathbf{P}_1'$.

We use the second method, since it also provides a simple way of computing curved epipolar lines. The minimization is carried out using Powell's method [Press et al., 1992].

Epipolar lines under lens distortion.

If lens distortion cannot be neglected, it also affects the epipolar geometry, since the epipolar lines are no longer straight but become curved [Zhang, 1996]. Since we compute the projection rays explicitly as shown above, we can easily take lens distortion into account in the following way. A 3-D projection ray is sampled by a sufficient number of 3-D points along the ray. The number of points is chosen according to the desired accuracy. A piecewise linear approximation of the curved epipolar line is obtained by projecting the sampling points into the images, see figure 6.3. Hence, we can compute the distance of points to curved epipolar lines by computing the distance of a point to the closest line segment. This method is also used by Maas [1992] and Virant [1996] to account for influences of the multimedia environment on the epipolar geometry.

6.2.3   3-D accuracy

In this section, we assess the accuracy of the 3-D reconstruction. Towards this end, we compute component-wise rms errors of reconstructed calibration points, using their known 3-D coordinates as ground truth. Further, we use the rms distance of points to their estimated reference planes as a 3-D error measure. The detailed analysis of these error measures has already been presented in sections 5.4.2–5.4.3. Here, we carry out an additional comparison of the results with theoretical error bounds derived in the literature. First, we give a short description of these error bounds.

Tsai [1987] proposed a theoretical **upper bound** for the triangulation error (for one component of a 3-D point) of a stereo camera setup. It is given by

$$\epsilon_{total} = \epsilon_{calib} + \epsilon_{noncalib}, \tag{6.11}$$

with

$$\epsilon_{calib} = \left[\frac{1}{\sqrt{6N_0 N_f}} \frac{Z}{f} + \frac{|T_s|}{L\sqrt{6N_0}} \left(1 + \frac{1}{N_f} \right) + \frac{1}{2\sqrt{6N_0}} + \frac{1}{2\sqrt{6N_0}} \frac{Z}{f} \right] \frac{Z}{|T_s|} \cdot \delta, \tag{6.12}$$

and

$$\epsilon_{noncalib} = \frac{Z}{f} \frac{Z}{|T_s|} \delta + \Delta q, \tag{6.13}$$

where δ is the precision of the feature extraction in mm, N_0 is the total number of points used for calibration, N_f is the number of views used for calibration, $|T_s|$ is the baseline length of the stereo rig in mm, L is the size of the active area in the image plane in mm, Δq is the precision of the known 3-D ground truth in mm, Z is the depth of the measured 3-D points in mm and f is the focal length of the lens in mm. Typical values for the stereo rigs used in this work are shown in table 6.2. As (6.11) shows, ϵ_{total} is the sum of two terms, where the first can be made arbitrarily small by choosing a large number of calibration points resp. views. In our calibrations, the ratio $\epsilon_{calib}/\epsilon_{noncalib}$ is always below 3%, indicating that the number of points is large enough.

As a **lower bound** on the 3-D error, we use the denominator of the NSCE, that has been introduced in section 5.4.2. This expression gives the lateral standard deviation of the pixel digitization noise, back-projected to the depth of the corresponding 3-D point. Hence, we assume that the best-possible 3-D accuracy is limited by the resolution of the images.

The comparison of the estimated 3-D rms errors with the theoretical error bounds is shown in figure 6.4 (**a** : Höllbach-setup, **b** : Jagst-setup). For the Neckar-setup, a comparison is not possible, since no ground

Table 6.2: *Comparison of geometry and 3-D accuracy of three different stereo rigs. For discussion, see text.*

	Neckar	Jagst	Höllbach
f_x $[mm]$	31.0442	4.5584	2.4729
f_y $[mm]$	31.1231	4.7560	2.4660
T_x $[cm]$	-0.1316	0.05744	0.04625
T_y $[cm]$	1.0133	0.07036	0.01748
T_z $[cm]$	88.0622	0.6581	0.41352
ω $[°]$	-3.0376	0.1732	-2.7977
ϕ $[°]$	43.9025	33.1282	41.568
κ $[°]$	-2.9052	1.2982	4.5107
C_x $[cm]$	61.1970	-0.4091	-0.3099
C_y $[cm]$	2.3517	-0.0706	-0.0002
C_z $[cm]$	-63.2882	-0.5187	-0.2782
$\langle \epsilon_x \rangle$ $[pix]$	0.00 ± 0.16	0.00 ± 0.46	0.00 ± 0.45
$\langle \epsilon_y \rangle$ $[pix]$	0.00 ± 0.16	0.00 ± 0.53	0.00 ± 0.40
$\langle \epsilon \rangle$ $[pix]$	0.19 ± 0.12	0.63 ± 0.32	0.53 ± 0.27
$(\Delta X)_{rms}$ $[\mu m]$	90.9	15.8	15.5
$(\Delta Y)_{rms}$ $[\mu m]$	47.1	14.1	10.4
$(\Delta Z)_{rms}$ $[\mu m]$	31.0	11.7	10.2
$\langle D_\perp \rangle$ $[\mu m]$	26.8 ± 23.2	6.7 ± 8.1	5.1 ± 4.2
NSCE [-]	0.84	1.94	1.95
δ $[mm]$	0.005	0.01	0.01
N_0	1500	2500	700
N_f	13	1	1
$\|T_s\|$ $[cm]$	100	0.5	0.5
L $[mm]$	5.96	3.25	1.9
Δq $[mm]$	0.01	0.006	0.006

truth is available. The plots show that all estimated errors are approximately within the theoretical error bounds[2]. We can draw the following conclusions from the plots (see also table 6.2):

- In both endoscopic setups, the error in the Z-component is lower than in the other components. Theoretically, the errors should be equal in a setup with $90°$ convergence angle (like the Höllbach-setup). The higher Z-accuracy is probably related to the better precision of the calibration points in the Z-direction, i.e. it reflects imprecisions in the calibration points perpendicular to the translation axis of

---

[2]In the discussion of these results, we have to keep in mind that both the error bounds and the experimental rms errors are only approximate estimations since there might be systematic errors in the ground truth due to effects of non-planarity and non-orthogonality of the calibration target.
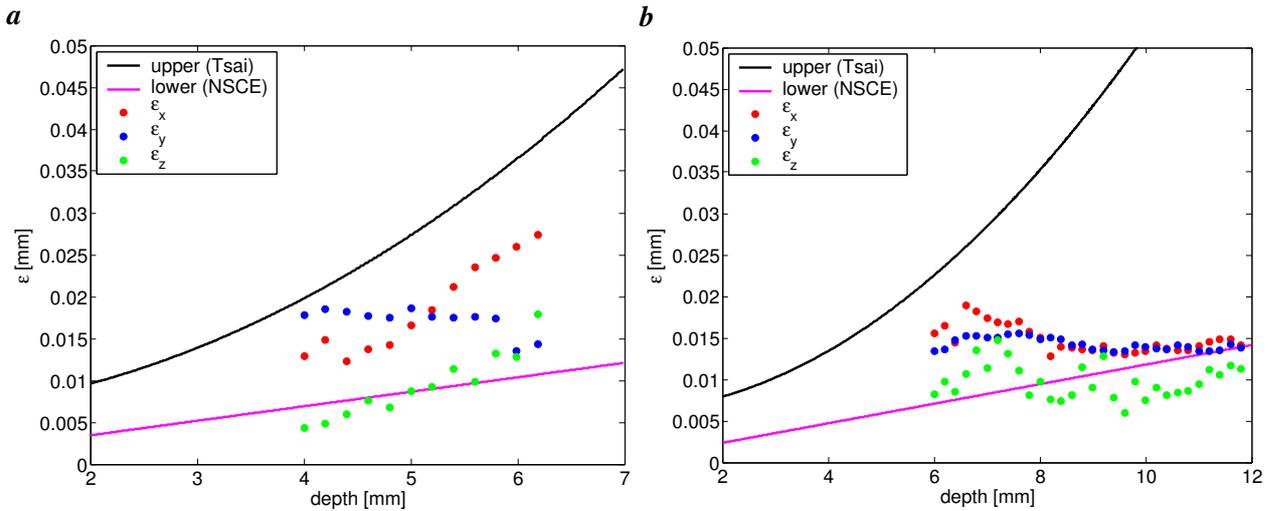
Figure 6.4: *Comparison of rms 3-D triangulation errors with theoretical error bounds. The curves show the upper and lower error bounds. The dots are the component-wise rms errors, computed using the triangulation results and the known ground truth. The errors are shown in dependency of the scene depth Z. **a** Höllbach-setup. **b** Jagst-setup.*

the linear stage.

- In the Höllbach-setup, both the Z-error and the X-error increase with depth. This increase is caused by the strongly convergent stereo setup. The optical axes (along which the depth is defined) enclose an angle of $45°$ with the world coordinate plane $Z_w = 0$. Hence, depth-related errors are spread into the X- and Z-components of the 3-D points.

- In the Jagst-setup, the errors are approximately constant throughout the whole depth, because the convergence angle is smaller $(30°)$.

6.3   Summary

In this chapter, the geometric concepts of a stereo camera setup have been explained. To compute 3-D co-ordinates from stereoscopic image pairs, two tasks have to be solved: the matching of corresponding image points in the left and right image (stereo correspondence analysis) and the triangulation of the corresponding projection rays.

The stereo correspondence analysis is based on the epipolar geometry, which constrains corresponding image points to lie on conjugated epipolar lines. Mathematically, the epipolar geometry is parameterized in terms of the fundamental matrix. If the cameras are calibrated, their relative orientation is known, and the fundamental matrix can be computed from the camera parameters. However, metric information is generally not necessary to estimate the fundamental matrix. An estimation that is based only on image point correspondences is also possible. Such a non-metric estimation has been carried out and compared to the results obtained from the calibrated camera parameters. This provides an additional, independent check of the camera calibration, since no world coordinates are used in the estimation of the fundamental matrix. The same accuracy has been obtained with both methods. This indicates that the main accuracy limitation is not given by inaccurate world coordinates of calibration points, but rather by deviations from an ideal pinhole model. The latter result from imprecisions due to endoscopic imaging and multimedia effects.

A simple linear triangulation method is used to compute 3-D coordinates. As already shown in section 5.4.3, the Neckar-setup achieves a 3-D accuracy that is better than the theoretical lower bound given by the pixel discretization (corresponding to a subpixel accuracy better than the discretization error of $\sqrt{1/12} \approx 0.29$ pixels). A comparison with theoretical error measures has shown that the 3-D errors in the endoscopic setups are larger. This is expected because of degradations related to the transmission of the images through a fiber bundle. Nevertheless, the 3-D errors are well below the theoretical upper bound.

Chapter 7

3-D Particle-Tracking Velocimetry

This chapter presents a detailed description of the 3-D PTV algorithm, taking a close look at all of its components. We assume that the information from the radiometric camera analysis (chapter 4) as well as the geometric camera calibration (chapter 5) is available and can be used as input to the 3-D PTV algorithm. In section 7.1, we start with a review of previous work on PIV and PTV carried out in the image processing research group at the IWR. Parts of this work provide the basis for the present implementation. Section 7.2 shows an overview of this implementation. The algorithm consists of three major parts, which are discussed in the following sections in the order of their computation: particle segmentation (section 7.3), 2-D particle tracking (section 7.4) and stereoscopic 3-D reconstruction (section 7.5). A summary of the algorithm is given in section 7.6. The performance of the algorithm is analyzed in chapter 8.

7.1 Previous work at the IWR

During the last 15 years, research work on PIV and PTV has been carried out in the image processing research group at the Interdisciplinary Center for Scientific Computing (IWR) in Heidelberg. Both experimental setups and image processing algorithms have been developed and successfully applied in different scientific applications. Some components of the current implementation are based on this previous work, which is reviewed in this section.

Wierzimok [1990] presents an early application of 2-D particle-tracking to study transport of mass and momentum beneath a wind-stressed wavy water surface. He uses the PTV results to compute Lagrangian drift velocities, shear velocities at the water surface, depth-dependent velocity profiles, and turbulent kinetic energy. Using a realtime-segmentation by a global threshold, his system is able to record a maximum of 32 binary images, resulting in a sequence length of $1.28 \ s$. Velocities are measured with an accuracy of $\pm 3 \ mm/s$ and resolving fluctuation frequencies of up to $12.5 \ Hz$. To establish the temporal correspondences in the tracking, the disparity between the two fields of the interlaced images is used as a velocity estimator to define a search window in the next frame.

Hering [1996] develops the next generation of the 2-D PTV, which is also applied to study the flow field beneath a wind-driven water surface in experimental facilities in Delft, Scripps and Heidelberg [Hering et al., 1995a,b; Hering, 1996; Hering et al., 1997, 1998]. A laser light sheet of 1 to $5 \ cm$ thickness is used for illumination, the size of the observation area is $14 \ cm \times 10 \ cm$. Image sequences of 100 to 500 frames, with a resolution of 512×480 pixels and a frame rate of $30 \ Hz$ can be recorded. The velocity accuracy

is $\pm 0.8 \ mm/s$. The system is able to track up to 1000 particles per frame (corresponding to a particle density of ≈ 0.0041 particles per pixel), with displacements of up to 12 pixels per frame. The particle segmentation is done by a region growing algorithm, and the tracking algorithm again uses the field overlap of interlaced images to establish correspondences. For non-interlaced cameras, the overlap is generated artificially by a morphological dilation of the particle images. Since unique particle overlaps are only found for small particle densities (less than 100 particles in an image of size 256×256, additional features are needed to track particles at higher densities. Hering et al. [1997] combines different features (normalized sum of gray values, area, velocity, distance of candidates) in a fuzzy-logic approach. Further, a constant acceleration model is used to predict the particle position in the next frame. In addition to the Lagrangian flow trajectories, the Eulerian velocity field is also estimated by Adaptive Gaussian Windowing [Agui and Jimenez, 1987] (which is basically a weighted averaging resp. a normalized convolution [Granlund and Knutsson, 1995; Jähne, 2002]). The relative accuracy of this interpolation is tested by simulations and found to be below 10% for sufficiently high particle densities (at least 300 vectors in an image of size 512×480).

To perform spatially high resolved measurements within the viscous sublayer at the water surface, Dieter [1994] develops a 2-D PIV system. The size of the observation area in these experiments is only $4 \ mm \times 4 \ mm$. Such an investigation of small scales at large velocities requires a high frame rate and an algorithm that can handle very large displacements (up to about 50 pixels). Thus, a camera operating at $200 \ Hz$ is used, and the sequences are evaluated by a multigrid PIV algorithm (see section 3.2.2), employing the first two levels of an image pyramid [Jähne, 2002]. The image correlation is computed in Fourier space using an FFT algorithm.

Further work has been done in the field of particle segmentation, since this is an important step in a PTV system. The region growing algorithm of Hering [1996] tends to separate particle images into several smaller ones if the particle velocities are large and the images become elongated (and hence larger and darker) due to motion blur. One way to achieve a reliable segmentation of such particle streaks is to take into account prior knowledge of the structure of the particle images. This is the approach of Leue et al. [1996], who implement a model-based segmentation. In the first step, streaks are detected based on the analysis of the local orientation [Jähne, 2002] in the image, resulting in a list of areas of interest. In the second step, a model function describing the gray value distribution of a particle streak is fitted to these areas of interest. The shape of this function (a modified Gaussian) takes into account that the underlying physical process is the imaging of a moving particle. The model function is defined by six parameters: the sum of the gray values, position of the particle center, streak length and width, and streak orientation. A Levenberg-Marquardt method is used to solve the nonlinear optimization problem. Initial values for the fit parameters are given by the first and second order gray value moments of the particle images. A further benefit of this method is that the information of the streak length and orientation can be used to predict the streak position in the next frame, allowing to track large displacements of up to 50 pixels. However, due to the nonlinear, iterative parameter estimation (which has to be carried out for each particle), the method is computationally intensive.

Borchers [1997] presents another approach to segmentation and shape analysis of elongated particle streaks. He uses a local gray value threshold to segment the particles. The thresholding is followed by morphological operations to fill gaps in the segmented image and to separate overlapping streaks. The latter becomes especially important for high particle densities, which cause a high probability of overlapping particle images. The streak separation is optimized for streaks with a pronounced elongation along their
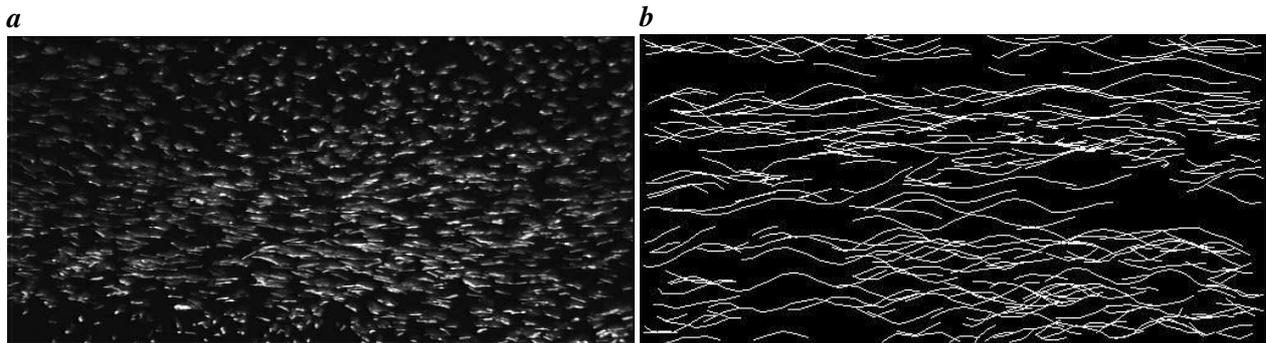
Figure 7.1: *Application of 2-D PTV to track rising bubbles in a liquid column.* **a** *Streak images of the rising bubbles. Due to the integration time of the camera, the streaks are elongated along their direction of motion. The bubbles move from right to left.* **b** *Corresponding trajectories of the bubbles, showing an oscillatory motion perpendicular to the rising velocity. From [Borchers, 1997].*

direction of motion, see figure 7.1a. The method is applied to perform 2-D tracking of rising bubbles in a liquid column and to analyze the flow field around the bubbles, see figure 7.1b. von Busse [1997] also applies the method to track rising bubbles. In his application, bubble mediated gas transfer at the air-water interface is studied.

Netsch [1995] introduces the first basic algorithm for 3-D PTV based on stereoscopic correspondences of 2-D trajectories. Two cameras are used to record stereoscopic image sequences, and the 2-D methods mentioned above are used to track the particles in each sequence. Two extensions are developed that are necessary to reconstruct the 3-D trajectories: a module for geometric camera calibration and a module to establish stereoscopic correspondences of trajectories and perform a triangulation in 3-D space. The efficiency of this method is shown using synthetic test images. Garbe [1998] carries out precise evaluations of all components of the 3-D PTV and presents a first application on real image sequences from a wind-wave-flume. Engelmann [2000] continues these analyses of flow fields induced by wind-driven water waves. Stöhr [1998] applies the 3-D PTV to study rising bubbles in a liquid column, and Klar [2001] uses it to investigate open-channel flow above a gravel layer.

A first prototype of a miniaturized PTV setup to study the pore flow within a gravel layer is developed by Spies [1998]. He shows the general feasibility of using fiberoptic endoscopes to record particle image sequences from the flow field within an artificial gravel pore, and to apply a tracking algorithm to estimate the particle velocites. Ehrbächer [1999] carries out a study of LED illuminations to create a miniaturized light sheet. However, the light intensity of the available LEDs is found to be too low for an endoscopic application. Janßen [2000] returns to a conventional cold light illumination using a fiber bundle and achieves a further miniaturization. He extends the setup by a second endoscope to enable stereoscopic measurements. However, the signal level in the images still is very low, and a significant amount of preprocessing has to be carried out to denoise the images and compensate artefacts introduced by the fiber structure of the endoscopes. To reduce the noise level, a computationally very intensive anisotropic diffusion algorithm [Scharr, 2000] is applied. Still, the image quality is too poor, and while Janßen [2000] shows some 2-D tracking results, the available 3-D PTV algorithm is not able to compute 3-D velocities.

Stybalkowski [2001] and Klar [2001] carry on with the work on endoscopic 3-D PTV and present the first application of this method to perform systematic measurements in an open-channel flow. To obtain a higher signal-to-noise ratio in the endoscopic image sequences, they use a second illumination fiber, see
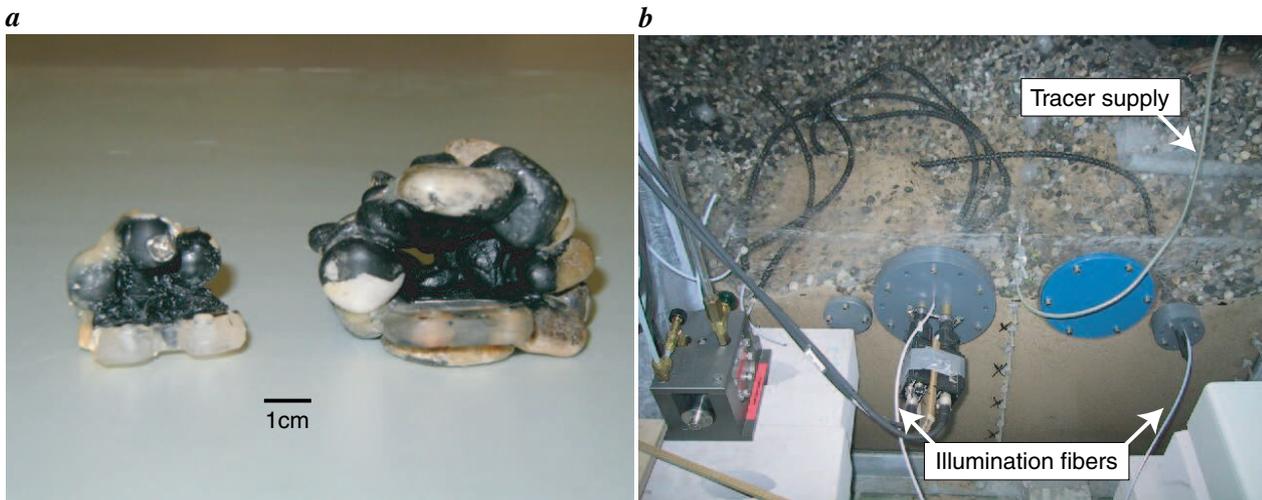
a   *b*

Figure 7.2: *3-D Particle-Tracking within a gravel layer using artificial gravel pores.* **a** *Comparison of the artificial pore used by Stybalkowski [2001] (right) and one of the pores used in this work (left). By re-arranging the stereoscopic geometry, it was possible to reduce the pore volume significantly, making the artifical pore more representative of the natural pore spaces within the gravel layer.* **b** *View of the setup of Stybalkowski [2001] and Klar et al. [2002] during installation in the test flume at BAW Karlsruhe. To enhance the image quality and avoid the computationally very intensive denoising applied by Janßen [2000], an additional illumination fiber is used.*

figure 7.2. With this setup, a denoising of the image sequences is not necessary any more. The first 3-D velocity measurements of the flow within the artificial pore are presented.

7.2   Overview of the 3-D PTV

In the previous section, the major milestones in the past development of the 3-D PTV have been reviewed. The current version that is realized in this work could benefit from these developments in many respects. The general organization of the algorithm has not been changed. A flow diagram of the algorithm is shown in figure 7.3. A preliminary step before the acquisition of flow sequences is the geometric camera calibration (chapter 5) that is necessary to determine the geometric structure of the stereo setup (position and orientation of the cameras relative to each other resp. relative to a given world coordinate frame). After the calibration, the setup is installed in the gravel layer, and flow measurements are carried out (chapter 10).

The first step of the 3-D PTV is the particle segmentation and 2-D tracking throughout the image sequences of both cameras. The results are two lists of 2-D trajectories, one for each camera. The next module performs the stereoscopic correspondence analysis of the 2-D trajectories. Finally, for all trajectories with unique correspondence partners the 3-D coordinates are reconstructed.

In summary, the 3-D PTV algorithm has a modular structure, comprising the following six modules:

1. geometric camera calibration,

2. image preprocessing,

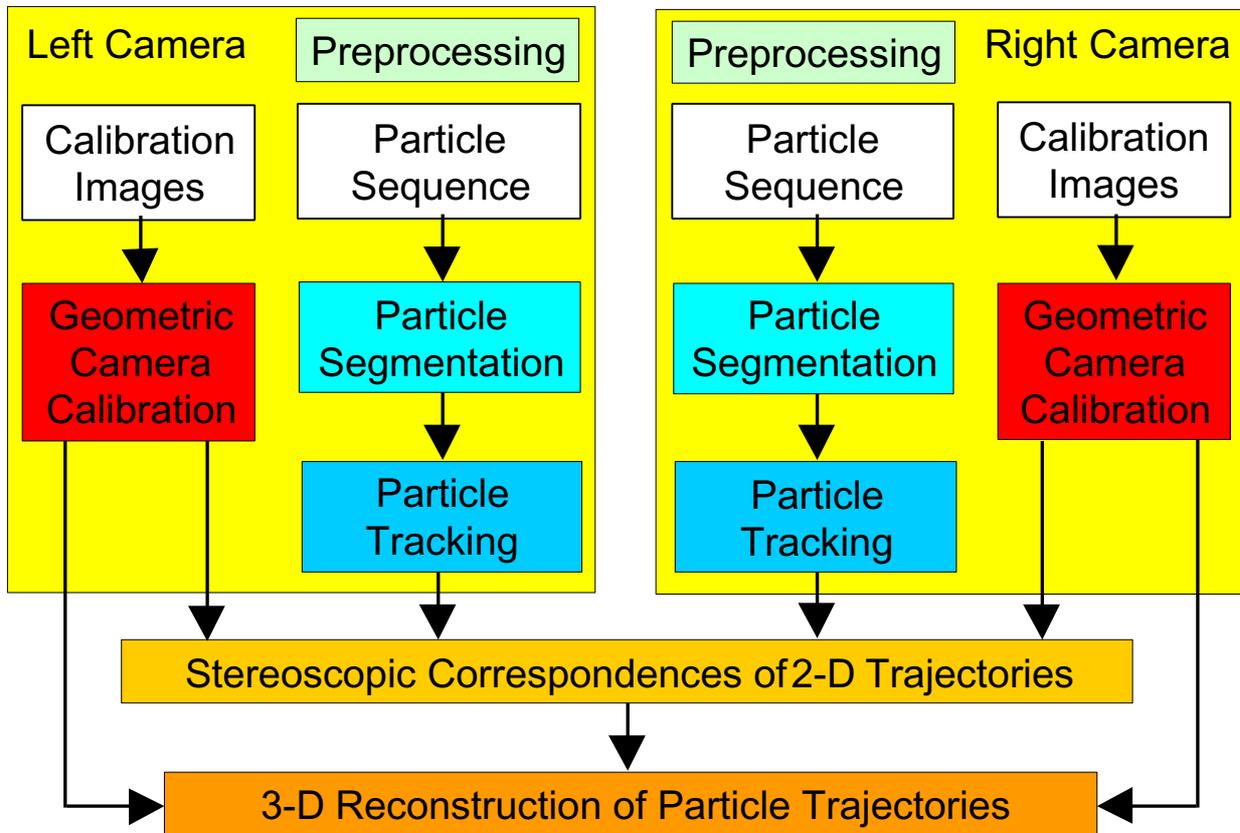3. particle segmentation,

4. 2-D particle tracking,

Figure 7.3: *Flowchart of the 3-D PTV algorithm. The algorithm consists of six major modules, which are processed sequentially: image preprocessing, geometric camera calibration, particle segmentation, 2-D particle tracking, stereoscopic correspondence analysis and 3-D reconstruction.*

5. analysis of stereoscopic correspondences,

6. 3-D reconstruction.

A description of the geometric camera calibration has been given in chapter 5. The image preprocessing consists of a radiometric correction of the gray values that suffer from fixed pattern noise of the CMOS sensors. It has been discussed in chapter 4. For the remaining modules, some parts of previous implementations mentioned in section 7.1 could be used without relevant changes, in particular the data structures of Hering [1996] for the efficient storage management of particle image sequences and trajectories. However, it was also necessary to modify or extend other components significantly to adapt them to the special conditions of endoscopic imaging. In particular, the following improvements were necessary:

- **Reliable and fast particle segmentation:** The segmentation algorithm should be able to segment all kinds of particle images (including small and dark ones) reliably. Previous implementations are sensitive to noise, since they perform a segmentation of single frames, based on some form of intensity information. A new segmentation module has been implemented that takes into account a temporal neighbourhood within the image sequence and uses the feature 'motion' as a segmentation criterion.

- **Separation of overlapping particle images:** Previous algorithms to separate particle images assumed that the images are elongated along their direction of motion. However, also slow particles

may overlap, especially in setups with volumetric illumination like the present one. Here, particle overlaps are created by the projection of particles in different depths, but with a lateral position close to a common projection ray. A watershed algorithm has been implemented that is able to separate both slow and fast particles.

- **Reliable tracking of long trajectories:** The probability of unique trajectory correspondences between the left and the right camera increases with the trajectory length. Hence, in the ideal case, the particles should be tracked throughout the whole field of view. Previous implementations tend to interrupt trajectories in case of segmentation failures or crossing trajectories. Further, they use the image field overlap (resp. overlap created by morphological dilations) as the primary feature for solving the temporal correspondence problem. This approach is only feasible for low particle densities and small displacements. A Kalman tracker has been implemented that achieves a more reliable tracking by predicting a reasonable search region in the next frame. To initialize the tracking for new particles entering the field of view, different low-level motion estimators are integrated into the Kalman tracker.

The following sections present a detailed description of the 2-D particle-tracking algorithm implemented in this work.

7.3   Particle segmentation

In the particle segmentation step, the individual particle images are extracted. Thus, for every pixel in the image, a decision has to be made whether it belongs to a particle or to the background. For an image $g(i, j)$, the segmented image $g_s(i, j)$ is given by the following operation:

$$g_s(i, j) = \begin{cases} 1 & : \quad g(i, j) \in \text{object (particle image)} \\ 0 & : \quad g(i, j) \in \text{image background .} \end{cases} \tag{7.1}$$

Thus, the result of the segmentation is a binary image in which the particles are marked with the value one and the background is marked with the value zero. Using these particle masks, the second step is to compute the position (pixel coordinates) and additional features (e.g. area, mean or maximum intensity, shape parameters) for each particle.

Segmentation approaches can generally be classified into **pixel-based**, **region-oriented**, **edge-oriented** and **model-based** methods [Jähne, 1997]. The particular problems related to the segmentation of tracer particle images have already been discussed in section 3.5.1. As a result of the latter, the methods to segment particle images are mostly region-oriented or model-based, e.g. region-growing methods [Maas, 1992; Hering et al., 1997], template matching [Etoh et al., 1998] or Gaussian gray value fits [Perkins and Hunt, 1989; Marxen et al., 2000]. The fitting areas are selected by searching local gray value maxima using a global or local threshold [Borchers, 1997], or by analyzing the local orientation [Jähne, 1997] of the images [Leue et al., 1996]. For a comparison of different particle segmentation approaches, see also [Klar, 2001].

Two basic features can be exploited to segment particle images. The first is the **intensity** (i.e. the gray value difference between particle and background), since simply speaking the particles appear as bright spots on a dark background. Most segmentation algorithms use only this feature. The second feature that may also be applied is the **motion** of the particles. In other words, we may not only use the spatial information in a single frame, but also the temporal information in a short sequence of frames to segment particles.

In this thesis, a region-oriented segmentation has been implemented. In contrast to most other segmentations, we do not only consider spatial regions, but also temporal regions: in a first step, moving objects within a sequence of images are identified. In a second step, the moving objects are separated from the background and from each other by a watershed algorithm, which operates on the gray values of a single frame.

The single steps of the segmentation algorithm are outlined as follows:

1. **Image preprocessing**: correction of fixed pattern noise, spatial and/or temporal smoothing, illumination correction (section 7.3.1)

2. **Detection of moving objects**: temporal high pass filter (section 7.3.2)

3. **Separation of particle images**: watershed algorithm (section 7.3.3)

4. **Determination of particle position and shape**: gray value moments (section 7.3.4)

7.3.1   Preprocessing

Clearly, low-level image preprocessing can only destroy or rearrange information, but cannot add new information to the images. Hence, the (ideal) basic goal of image preprocessing is a reduction of the noise level to yield an increased signal-to-noise ratio (with unchanged signal levels). Towards this end, isotropic smoothing filters are the simplest approach. However, they also affect the signal, since they do not only blur noise but also small-scale image structures and edges. Hence, only very small filter kernels should be applied to flow visualization images, since larger kernels may also destroy particle images. The problems of isotropic smoothing can be overcome by nonlinear diffusion filters [Weickert, 1998; Scharr, 2000]. However, the latter are computationally very intensive and have not been applied here.

We use the following image preprocessing steps:

- If necessary, a correction of inhomogeneous illumination is computed, as explained in section 4.4.

- In the images of the CMOS cameras, we perform a correction of the fixed pattern noise as outlined in section 4.3. An example is shown in figure 7.4. Note that this correction is a point-operation, i.e. no smoothing over extended neighbourhoods is involved.

- For further noise reduction, we smooth the images using a (3×3)-binomial filter. This filter mask reduces single-pixel noise, while being small enough to keep small particle images unaffected. Marxen et al. [2000] reported better results for particle images smoothed by the same filter. If the particles do not move by more than their diameter, temporal smoothing may also be applied to enhance the signal-to-noise ratio. An example is shown in figure 7.5.

- In the Jagst-setup, the images are over-sampled, since the endoscopes contain only 10000 fibers. The relatively thick cladding of the fibers is visible in the images and introduces a hexagonal lattice structure, see figure 7.6. Hence, we do not work with the original images, but compute the first level of a Gaussian image pyramid [Jähne, 1997]. The smoothing involved in the pyramid computation reduces the effects of the lattice structure and adapts the effective image resolution to the number of fibers. A further preprocessing step for images of the Jagst-setup (working with interlaced analog cameras) is the separation of the images into even and odd fields to increase the temporal resolution from $25\,Hz$ to $50\,Hz$. The missing lines are linearly interpolated.
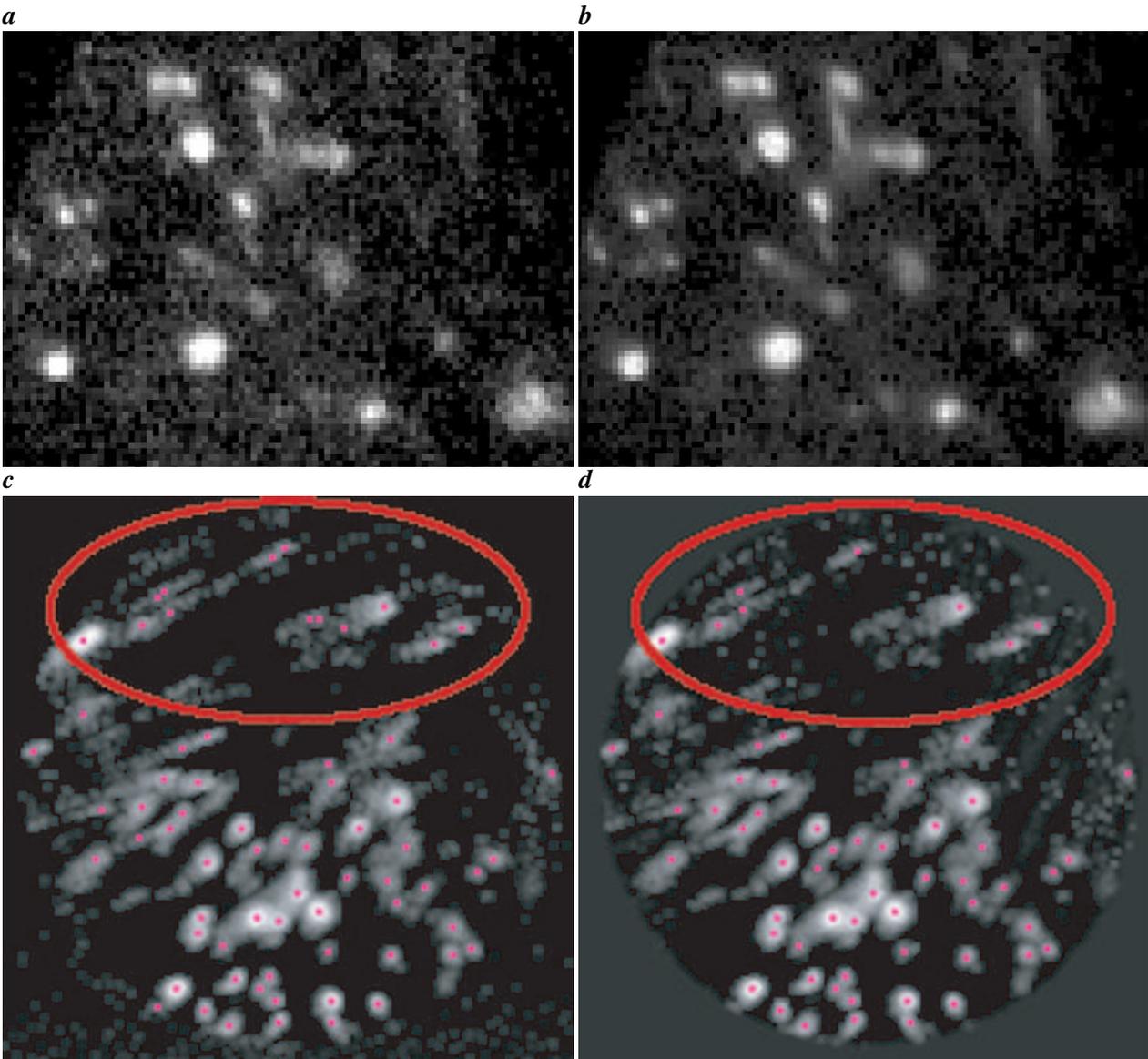
Figure 7.4: *Correction of fixed pattern noise in the particle images. **a** , **c** Original images. **b** , **d** Corrected images. Clearly, the noise level is reduced, and the gray value distributions become more regular. Note that no smoothing is involved in the correction. The diameter of the particles is not changed, and the smallest particles are not blurred out. In **c** and **d** , the result of the segmentaion is indicated by the red dots. The FPN correction significantly reduces the segmentation of spurious particles (false positives).*

7.3.2   Temporal highpass filter

In the first segmentation step, we apply an algorithm proposed by Haussecker [1993] for **motion segmentation** of moving sand grains in submerged subsoil. This method can even detect objects hardly recognizable by the human eye, as demonstrated in [Haussecker, 1993] and confirmed here. Objects are segmented by their feature 'motion'. For details of the algorithm, see [Haussecker, 1993]. We only give a brief sketch of the basic ideas.

The stationary image structures, i.e. the image 'background' (gravel layer, illumination inhomogeneities)
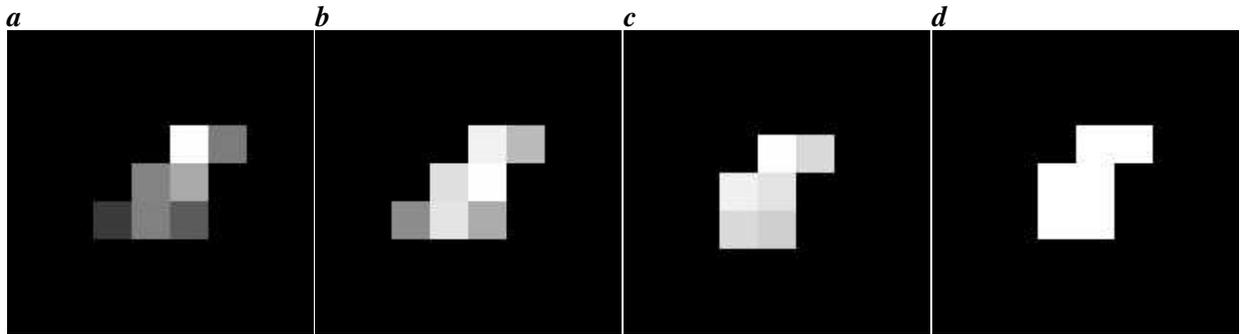
Figure 7.5: *Different preprocessing operations enhance the signal-to-noise ratio of particle images.* **a** *Original image.* **b** *Image after temporal smoothing with a 3-tap binomial filter.* **c** *Image after correction of fixed pattern noise.* **d** *Image after correction of fixed pattern noise and temporal smoothing.*
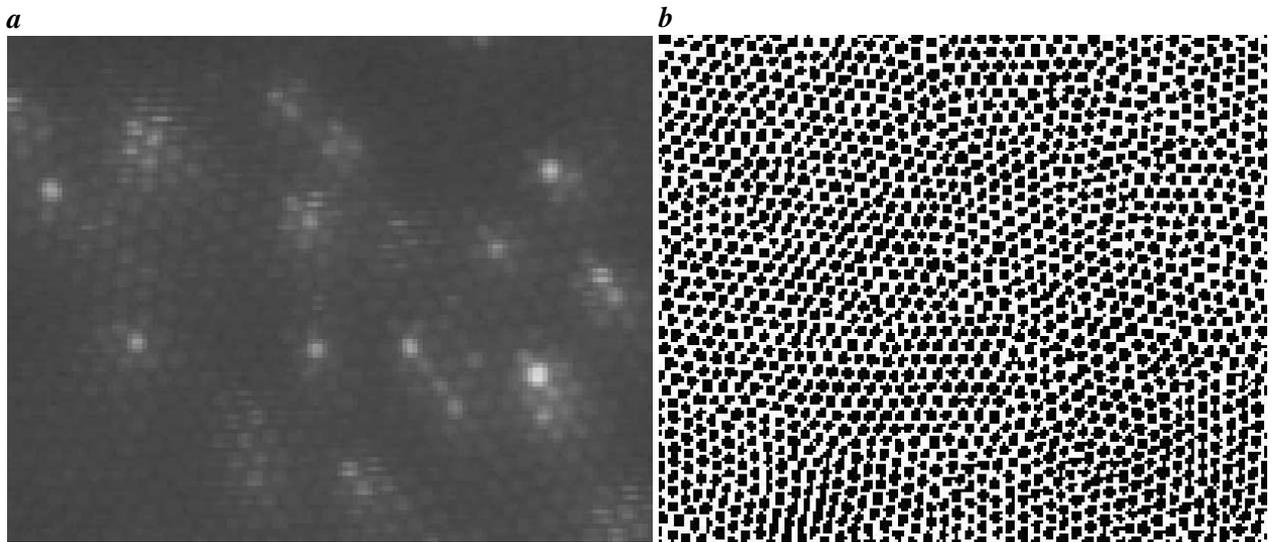

Figure 7.6: a *Enlarged section of a particle image acquired by the Jagst-setup. The endoscopes in this setup have only* 10000 *fibers, which are arranged with a lower density than in the endoscopes of the other two setups. The individual fibers have a relatively large cladding, which does not transmit light. As a result, the inter-fiber distances are larger, and the hexagonal lattice structure of the fiber bundle is visible in the images. Small particle images may temporarily disappear, if they move from one fiber to another and thereby cross the claddings.* **b** *Hexagonal lattice structure obtained by thresholding an endoscope image of a uniform white plane.*

are extracted by a temporal convolution of the image sequence. The convolution is computed using spatial and temporal binomial masks, which can be efficiently implemented by spatial and temporal image pyramids. The kernel size of the required filter masks depends on the size and the velocity of the particles to be segmented and has been derived analytically by Haussecker [1993]. Since in our application the image background is completely static throughout the whole image sequence, we can simply compute it as a pixelwise temporal mean image over the sequence.

A comparison between the image background and the original images yields the moving objects. As a measure of similarity $S(\mathbf{x}, \mathbf{t})$, a spatially smoothed version of the squared difference between background and original image is used:

$$S(\mathbf{x}, t) = \mathcal{B}_{\mathbf{x}}^{n_x}[(\mathcal{I} - \mathcal{B}_t^{n_t})g(\mathbf{x}, t)]^2, \tag{7.2}$$

where n_x resp. n_t is the size of the spatial $\mathcal{B}_\mathbf{x}$ resp. temporal $\mathcal{B}_t$ binomial filter masks and $\mathcal{I}$ is the identity operator. Haussecker [1993] shows that $S(\mathbf{x}, \mathbf{t})$ approximates the energy of an ideal temporal high pass filter. It is equivalent to the local temporal variance in the image sequence. $S(\mathbf{x}, \mathbf{t})$ will be large for pixels of moving objects and low for pixels of stationary structures. Hence, moving objects can be segmented by thresholding $S(\mathbf{x}, \mathbf{t})$, which introduces the problem of finding an appropriate threshold. Equation (7.2) can also be written as

$$S(\mathbf{x}, t) = S_0(\mathbf{x}, t) + \sigma_g^2, \tag{7.3}$$

where $S_0(\mathbf{x}, \mathbf{t})$ is the value of S in a sequence without noise and σ_g^2 is the noise variance. Hence, if the noise variance is known, an appropriate threshold can be chosen, e.g. a two-sigma or three-sigma threshold.

As shown in figure 4.3, the noise variance of a pixel depends on its gray value. Since we know this dependency $\sigma_g^2(g)$ from the analysis in section 4.1.2, we can estimate the noise variance of each pixel and choose a pixel-based threshold on S. Hence, we use a local threshold, which adapts to the image content. The treshold takes into account the statistical significance of the gray value deviation between a moving object and the image background, as compared to the expected noise level. In the practical implementation, we estimate the local noise level as $N(g) = 3\sigma_g(g)$ and segment those pixels with a signal-to-noise ratio S_0/N larger than two. An example of the resulting segmentation is shown in figure 7.7a.

7.3.3   Resolving overlapping particle images

Figure 7.7a,b shows that the motion segmentation presented in the previous section tends to segment several particles as one connected object. Two effects are responsible for this. First, if two particles are close to each other, the light reflected between them may introduce a glow of the water, since the latter is not perfectly clear but contains a certain amount of particulate matter. Second, since a volume illumination is used, there will always be a number of particles in different depths which overlap in the image.

Maas [1992] presents a statistical analysis to derive the probability of overlapping particle images as a function of the imaging parameters (image size, average particle size and average particle density). Bastiaans et al. [2002] carries out a similar investigation and gives the following formula to compute the fractional amount of overlapping particle images for randomly distributed particles:

$$P_{overlap} = 1 - \exp(-4N_s), \tag{7.4}$$

where N_s is the image source density (the fraction of the total image area that is covered by particles). The number of expected particle overlaps for the parameters of the present setups is shown in figure 7.8. The curves are strictly valid only for symmetric particle shapes and a homogeneous particle density. Inhomogeneities in shape and density will increase the average number of overlaps. The plots show that up to $\approx 20\%$ of overlapping particle images have to be expected (depending on the particle density). Hence, this problem cannot be neglected.

Perkins and Hunt [1989] suggest using the features shape and area to separate overlapping particle images. They use a 'sector fitting' algorithm, which successively computes best fits of Gaussians to overlapping particle images. The fitted peak is removed and another fit is computed for the remaining sector of the overlapping particle. Since both particle shape and area may be quite inhomogeneous in the case of endoscopic imaging, this approach is not feasible here.

Maas [1992] presents an analysis of different approaches to separate particle images and introduces an 'anisotropic thresholding operator' to resolve overlapping particle images. This operator is basically a
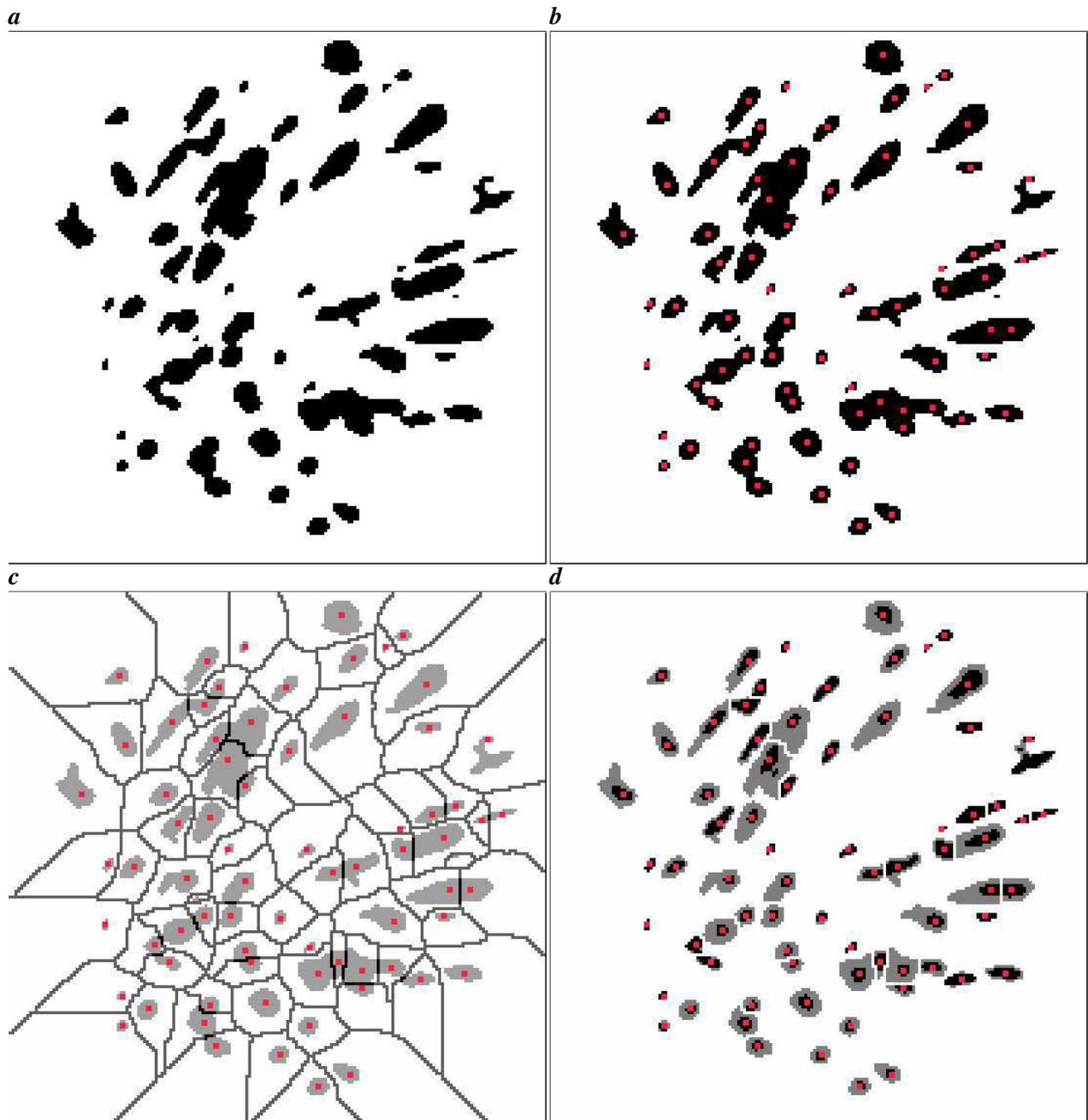
Figure 7.7: *Particle segmentation in an endoscope image (Höllbach-setup).* ***a*** *Result of the segmentation-from-motion described in section 7.3.2. In* ***b*** *, the local gray value maxima as given by the final segmentation results are indicated by the red dots. This shows that the motion segmentation tends to segment several particles as one connected object. To separate the overlapping particle images, we apply a watershed algorithm. The separation lines ('watersheds') are shown in* ***c*** *. The final segmentation result is shown in* ***d*** *(black areas). The size of the segmented particles is reduced by selecting only those pixels with a gray value larger than a fraction of $1/e$ of the particle's maximum gray value.*

region-growing method that performs a connectivity analysis in a local neighbourhood, including a discontinuity criterion to separate two local gray value maxima. The approach is very similar to a well-known and very powerful image segmentation method known as **watershed transformation** in morphological image
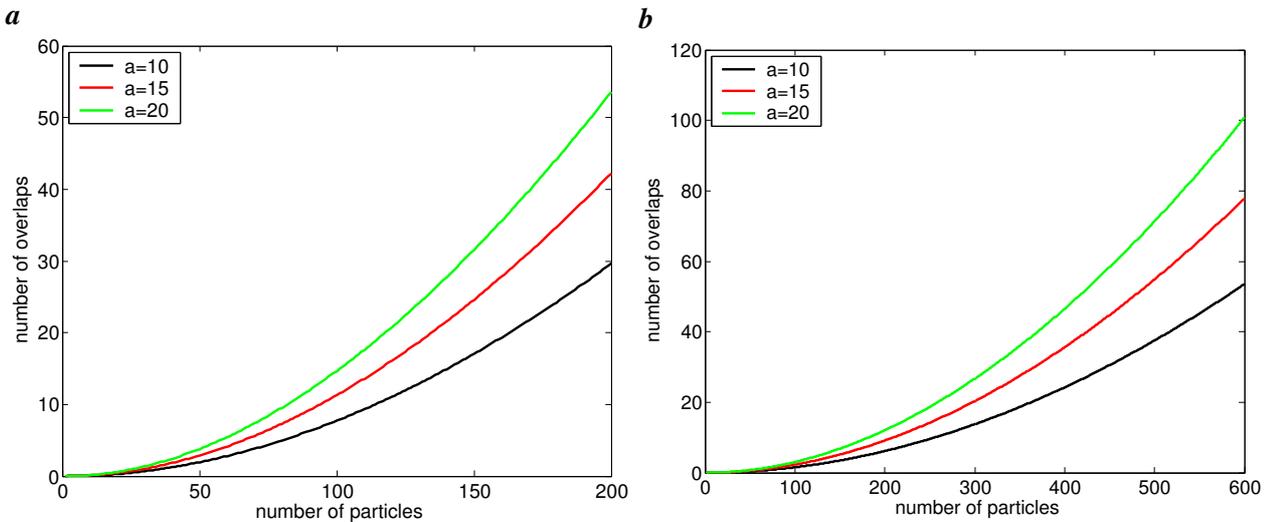
a

b

Figure 7.8: *Statistical estimation of the number of overlapping particles per image as a function of the total number of particles per image, according to [Maas, 1992]. The three different curves correspond to three different average particle sizes of $a = 10, 15, 20$ pixels. Depending on the flow conditions, the average particle size is between* 10 *and* 20 *pixels in all setups.* ***a*** *Höllbach-setup.* ***b*** *Neckar-setup.*

processing [Soille, 1998]. We use a watershed transformation to separate overlapping particles.

The watershed transformation groups the pixels around the local image minima and separates adjacent regions along so-called watersheds. The basic principle of the watershed transformation can be explained using a demonstrative analogon [Dougherty and Lotufo, 2003]. If we consider the inverse of a particle image as a topographic landscape, the local minima in this image (the local maxima in the original image, i.e. the particles) correspond to more or less deep valleys or basins within the plateau of the image background. If rain starts to fall onto this landscape, the water will be collected in the basins, and the water level will start to rise in each particular basin. If two basins are adjacent to each other, there will be a point where the water level in one of them reaches the border between them, and the two basins will start to merge. As soon as this point is reached, we start to build a dam (a watershed) on the border to keep the two basins separated. If the water level finally reaches the plateau, the watersheds represent a segmentation of the image. Overlapping particle images will be separated along the curve of local gray value minima between the two maxima, see figure 7.7c.

Belonging to the group of region-growing methods, the watershed transformation needs seeding points to start the region-growing. In practical applications, image noise makes the computation of the watershed transformation on the original images infeasible. Noise introduces many local gray value minima, which will cause an over-segmentation of the image. Hence, a so-called *marker image* is necessary, that indicates the relevant objects in the image and distinguishes them from the irrelevant noise minima. In our application, we can use the segmentation mask from the motion segmentation (figure 7.7a) as marker image. The input image for the watershed transformation is computed by multiplying the original image with the marker image, applying a further noise reduction by smoothing with a (3×3)-binomial filter and inverting the result.

After the watersheds have been computed, we finally reduce the size of the segmented areas by choosing only those pixels with a gray value larger than a fraction of $1/e$ of the particle's maximum gray value. Only the bright pixels are statistically relevant and carry information on the position and shape of the particles.

Further, we remove all objects with a size smaller than three pixels, since they are probably caused by noise and do not allow for an accurate determination of their position. The final segmentation result is shown by the black areas in figure 7.7d.

7.3.4   Position and shape by gray value moments

The next step after the segmentation is to estimate the parameters describing the shape and position of the particles. A precise determination of the particle position is particularly important both for the particle-tracking and the final 3-D reconstruction of the trajectories. The two most common methods to estimate the particle positions are the computation of the gray value centroid (e.g. [Maas, 1992]) or model-based approaches that fit a Gaussian to the image data (e.g. [Marxen et al., 2000]). [Wernet and Pline, 1993] derive the theoretical (Cramer-Rao-) lower bound for the error in estimating a particle's position. If the intensity distribution is Gaussian and the only noise source is photon shot noise, a theoretical optimum accuracy of ≈ 0.015 pixels may be achieved. [Marxen et al., 2000] find a comparable lower bound of 0.03 pixels from an analysis of synthetic images. Typical accuracies obtained on real image data are of the order of magnitude of 0.1 pixels, both for centroiding and Gaussian fits [Maas, 1992; Wernet and Pline, 1993; Marxen et al., 2000]. We use first and second order gray value moments to describe the shape and position of a particle image. This basically represents a modeling of a particle image as an ellipse.

Let $g_p(x, y)$ be the intensity distribution of a particle image. The **gray value moments** of the particle are defined as

$$m_{pq} = \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} x^p y^q g_p(x, y) dx dy. \tag{7.5}$$

The zero-order moment m_{00} is the sum of the gray values of the particle:

$$m_{00} = \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} g_p(x, y) dx dy. \tag{7.6}$$

For a binary image $b(x, y)$ (e.g. a segmentation mask of a segmented particle image), we have $b(x, y) = 1$ if the pixel (x, y) belongs to a particle and $b(x, y) = 0$ otherwise. In this case, the zero-order moment is simply the number of pixels belonging to the particle, i.e. its area. If we think of $g_p(x, y)$ as the density $\rho_p(x, y)$ of the particle, the zero-order moment becomes the total mass of the particle. Similarly, the first-order moments define the center of mass of the particle:

$$\begin{aligned} m_{10} &= \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} x g_p(x, y) dx dy = \langle x \rangle\, m_{00} \Rightarrow \langle x \rangle = x_p = \frac{m_{10}}{m_{00}}, \\ m_{01} &= \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} y g_p(x, y) dx dy = \langle y \rangle\, m_{00} \Rightarrow \langle y \rangle = y_p = \frac{m_{01}}{m_{00}}. \end{aligned} \tag{7.7}$$

Since the shape description of the particle images should be translation invariant, we use the following *central moments*, which are related to the center of mass:

$$\mu_{pq} = \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} (x - \langle x \rangle)^p (y - \langle y \rangle)^q g_p(x, y) dx dy. \tag{7.8}$$

Since the first-order central moments are zero by definition, $\mu_{10} = \mu_{01} = 0$, the shape description starts with the second-order central moments:

$$\mu_{20} = \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} (x - \langle x \rangle)^2 g_p(x, y) dx dy,$$

$$\mu_{02} = \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} (y - \langle y \rangle)^2 g_p(x, y) dx dy,$$

$$\mu_{11} = \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} (x - \langle x \rangle)(y - \langle y \rangle) g_p(x, y) dx dy. \tag{7.9}$$

Again by analogy to mechanics, these three second-order moments form the components of the inertia tensor for rotation of the object around its center of mass resp. rotation of the particle image around its centroid.

In gray scale particle images, where $g_p(x, y) \in [0, ..., g_{max}]$, we may also regard $g_p(x, y)$ as a discrete 2-D probability density distribution of the particle position over (x, y). In this case, we should have $m_{00} = 1$, which can be obtained by normalizing the original image by m_{00}:

$$g_p'(x, y) = g_p(x, y) / m_{00}. \tag{7.10}$$

Using this normalization, the inertia tensor becomes the covariance matrix of the particle position, see appendix A.

Under ideal circumstances, the image of a tracer particle under Mie-scattering is given by an Airy function, which is well approximated by a Gaussian gray value distribution in the image plane [Adrian, 1991; Raffel et al., 1998]. In this case of perfect rotational symmetry and hence isotropic light scattering, we can equally well compute the particle position by fitting a Gaussian or simply by computing the gray value centroid. However, in practical applications, there are always deviations from this ideal case. Irregularities in the particle surface, influences of the illumination, gloom in dirty water or camera noise limit the accuracy of the position computation.

Following Maas [1992], we estimate the influence of noisy gray values on the computation of the particle centroid position by Gaussian error propagation. For the x-component x_p (the same derivation also holds for the y-component), the discrete version of (7.7) reads

$$x_p = \frac{\sum x_i g_i}{\sum g_i}, \tag{7.11}$$

where the summation is carried out over the segmented particle area. Gaussian error propagation yields

$$\sigma_{x_p}^2 = \frac{1}{(\sum g_i)^2} \sum (x_i - x_p)^2 (\sigma_{g_i})^2. \tag{7.12}$$

To estimate the noise component σ_{g_i}, we can use the results of the radiometric analysis in chapter 4, compare figure 4.3. As a result, typical values for σ_{x_p} are in the range from 0.01 to 0.1 pixels. Maas [1992] also finds a centroid error of roughly 0.1 pixels for realistic camera noise conditions. The most important factor that influences the accuracy is the choice of the segmentation threshold, i.e. the decision which pixels should contribute to the summation in (7.12). This choice is particularly critical if the shape of the particle is asymmetric, e.g. in the case of overlapping particle images. Thus, the accuracy will rather be at the upper bound of the range given above, i.e. of order of magnitude of 0.1 pixels. This is confirmed in the performance analysis in section 8.2.2.
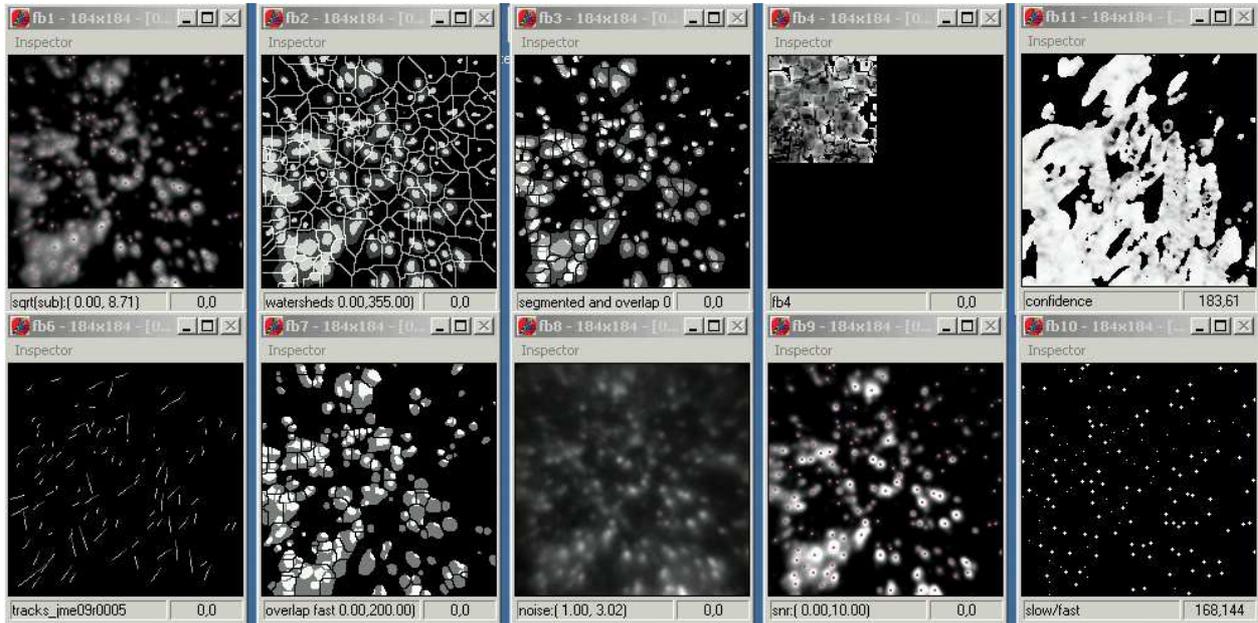
Figure 7.9: *Screenshot of the output windows of the 2-D PTV algorithm during the tracking of a particle sequence. Top row: window one to five, bottom row: window six to ten (from left to right). For explanations, refer to the text in section 7.4.*

7.4   2-D particle tracking

In this section, we discuss the 2-D particle-tracking algorithm that is used to determine the particle trajectories in the image planes of the left resp. right camera. The basic task that has to be solved is the matching of new measurements (here: particles) appearing in frame k to the features (here: trajectories) of the previous frame $k - 1$. This process is called the **motion correspondence problem** in computer vision or the **data association problem** in target tracking applications. A screenshot of the output windows of the PTV algorithm during the tracking is shown in figure 7.9. The various information displayed in these windows is used for the tracking and will be explained in the following subsections. The algorithm is a hybrid approach (see section 3.6), which combines low-level motion estimators (optical flow and PIV) with a Kalman tracker.

7.4.1   Tracking strategy

Temporal neighbourhood.

The tracking is carried out on a **temporal neighbourhood** of nine frames. Within this sequence, the central frame (number five) is the actual frame, i.e. we try to establish the particle links between frame five and frame six, as indicated by the arrow:

$$[1]\ [2]\ [3]\ [4]\ [5] \rightarrow [6]\ [7]\ [8]\ [9]$$

After all particles in frame five have been processed, the linked particles of frame one are appended to a list of trajectories that is created throughout the tracking. Frame one is deleted, a cyclic shift of the sequence is carried out, and the next frame is appended at position nine. The current trajectories are shown in window six of figure 7.9.
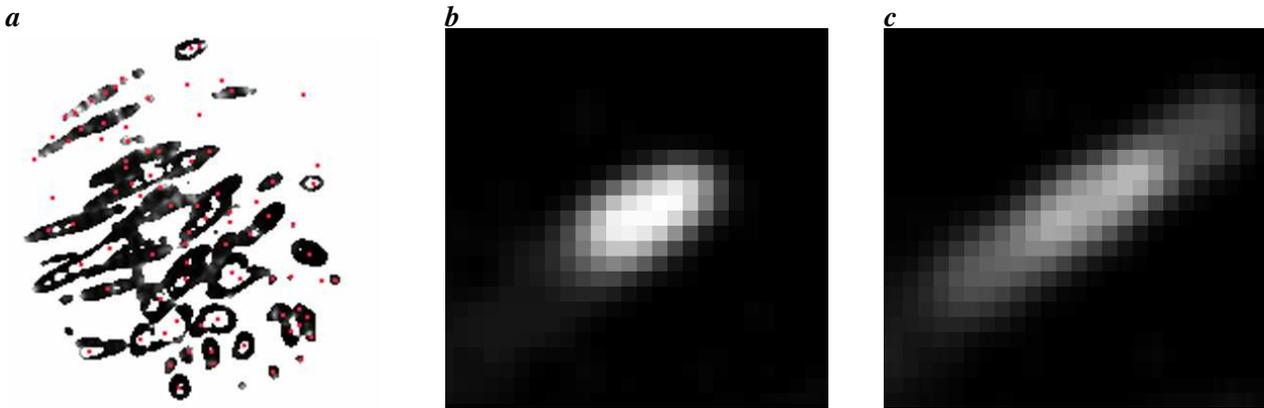
Figure 7.10: *The PTV algorithm takes into account an extended temporal neighbourhood.* **a** *Confidence measure (coherence) of the optical flow computed by the structure tensor method (black areas have a high confidence). The confidence is high along the particle trajectories. Note that large and bright particles only have a high confidence measure at the border, since in the inner particle area the spatial gray value gradients are too small.* **b** *Enlarged image of a particle.* **c** *The same particle in a temporally smoothed image.*

There are three reasons to work on an extended temporal neigbourhood:

1. The tracking can be stabilized by temporal smoothing. In the optical flow computation (section 7.4.3), the structure tensor is computed taking into account a temporal neighbourhood in addition to the usual spatial neighbourhood (temporal smoothing using a 9-tap binomial filter). As we can see in figure 7.10a, the motion of the particles is revealed by the confidence measure [Jähne, 1997] of the optical flow, which is high along the particle trajectories. A second benefit of temporal smoothing is the increased particle size in the direction of motion, see figure 7.10b,c. The latter can be used to identify the corresponding particle in the next frame by the spatial overlap of the two images (if the local particle density is sufficiently low, so that the overlap is unique).

   Temporal smoothing is only helpful, if the particles are 'slow', i.e. if they do not move by more than their diameter between two subsequent frames. Otherwise, temporal smoothing will create spurious particles, since the smoothed particle image will have several local maxima. Hence, we have to discrimate between slow and fast particles and apply the smoothing only to the slow ones. The discrimination is achieved by computing the pixelwise temporal variance on a sequence of three frames around the actual frame (frame four, five, six). Tests have shown that particles may be identified as 'fast', if the temporal standard deviation of the gray values within the particle area is higher than half of the temporal mean gray value. An example is shown in figure 7.9, second row, window ten: slow particles are indicated by small dots, fast particles by large dots.

2. Trajectories are only considered valid if they have a length of at least four frames. Shorter trajectories are rejected.

3. Three previous frames are necessary to compute the velocity and acceleration in the second order kinematic model of the Kalman filter (section 7.4.2).

At the beginning of the tracking, the temporal sequence is initialized by computing the particle segmentation for the first eight images.

Initialization of a tracking step.

The tracking is carried out in frame-by-frame steps. Since the idea of the hybrid approach is to use information from low-level velocity estimators to guide the tracking, we first have to compute this information before the actual tracking is performed. Each tracking step is initialized by the following operations:

- Compute the **particle segmentation** of the new frame appended at position nine.

- Determine the particle displacements by **spatial overlaps** between particles in the actual frame (frame five) and particles in the next frame. If the particle density and velocity is not too high, unique overlaps can be established either directly, by temporal smoothing (see above) or by a morphologic dilation. Further, we can use the image lag introduced by the amplifiers of the CMOS cameras (see section 9.3.4) to establish overlaps of fast particles.

- Compute the displacement vector field of the actual frame by the **structure tensor method** (section 3.4.1). The computation is carried out on a spatial neighbourhood defined by a 5-tap binomial mask and a temporal neighbourhood defined by a 9-tap binomial mask. By choosing a small spatial mask we try to avoid several particles occuring within the neighbourhood, which introduces motion discontinuities.

- Compute the displacement vector field of the actual frame by a **multigrid PIV algorithm** (section 3.2). The size of the basic interrogation window is 16×16, and the first two levels of a Gaussian pyramid are used. Since we are basically interested in the velocity of single particles, the latter should be in the center of the interrogation window. To account for the random position of the particles, we compute the correlation on a *dense* regular grid defined by every other pixel in the original image. Then we choose the displacement vector that is closest to the particle of interest, resp. the mean displacement vector over the area of the particle.

For each particle, a list of parameters is initialized. This list contains the following information about the particle:

- segmentation parameters: particle position and covariance matrix (determined from the gray value moments), area, sum of gray values, maximum gray value, noise variance (spatial mean over the particle area) (windows one, two, seven, eight in figure 7.9),

- correspondence partner in the next frame determined from spatial overlap (if the latter is unique) (windows three and six in figure 7.9),

- velocity determined by the structure tensor method (spatial mean over the particle area),

- confidence measure of the structure tensor (spatial mean over the particle area) (window five in figure 7.9),

- velocity determined by the multigrid PIV method (spatial mean over the particle area),

- confidence measure of the multigrid PIV method (spatial mean over the particle area) (window four in figure 7.9, linear size reduced by a factor two since displacements are estimated for every other pixel in the original image),

- flag indicating slow or fast particle (window ten in figure 7.9),

- normalized signal-to-noise ratio of the particle (window nine in figure 7.9).

The discussion of flow visualization methods in chapter 3 has shown that region-based approaches like the structure tensor method and PIV assume a spatially coherent optical flow field in the image plane. This is the reason why such methods in general are not suitable for images obtained with volume illumination, where particles in different depth ranges are projected simultaneously on the same image. However, if the particle density is not too high, only a small part of the image will suffer from motion discontinuities. In the major part of the image, the motion will be locally coherent, so that region-based methods can be applied. Note that we try to avoid motion discontinuity by computing the optical flow on a small neighbourhood of the order of magnitude of the particle size. Further, the results of the region-based methods are only used to initialize trajectories if the corresponding confidence measure indicates a good estimate. Discontinuities are detected by a low confidence. In this case, the trajectory intialization fails.

Order of trajectory processing.

With each new frame, we get a list of the segmented particles that have to be linked to the already established trajectories. The latter are stored in a list. This rises the question, in which order the list should be processed. To define this order, we assign a **score** S_T to each trajectory, which is computed as follows:

$$S_T = L_T + SNR, \tag{7.13}$$

where L_T is the current length of the trajectory in frames and SNR is the normalized signal-to-noise ratio of the last particle that has been linked to the trajectory ($SNR \in [0, 1]$). The first processing step for a new frame (frame six) is to compute S_T for all particles in frame five. The latter may already be linked to a trajectory ($L_T > 1$) or define a new trajectory ($L_T = 1$). Subsequently, the trajectories are processed by decreasing score. Hence, we first try to continue the tracking of long trajectories, since these are considered to be most reliable and thus have the highest priority. If several trajectories have the same length, the one with the highest signal-to-noise ratio will be processed first. In doing so, we prevent noisy particles from 'stealing measurements' from the more reliable particles.

7.4.2   Kalman filter

As the basic tracking framework, we set up a Kalman filter for each trajectory. Under the usual Gaussian noise assumptions, the Kalman filter represents a recursive version of the maximum-likelihood estimation of the particle trajectories. It estimates the particle positions by an optimal (in a statistical sense) combination of the measurements from the low-level image processing (segmentation and velocity estimation) with the predictions given by a linear kinematic model. Kalman filtering is a very common and powerful method used for tracking and other sequential state estimation problems. Hence, it is the subject of extensive research and application [Maybeck, 1979; Gelb et al., 1992; Brookner, 1998; Blackman and Popoli, 1999; Stone et al., 1999]. For a short introduction, see [Welch and Bishop, 2001].

State vector and process model.

A **state vector** x is assigned to each trajectory, which contains the current position, velocity and acceleration of the corresponding particle:

$$\boldsymbol{x} = [x, y, v_x, v_y, a_x, a_y]^T .\qquad(7.14)$$

The process to be estimated by the Kalman filter is given by a constant acceleration model. The corresponding **state transition equation** reads

$$\boldsymbol{x}_k = \boldsymbol{A}\boldsymbol{x}_{k-1} + \boldsymbol{w}_{k-1},\qquad(7.15)$$

where $\boldsymbol{A}$ is the state transition matrix

$$\boldsymbol{A} = \begin{pmatrix} 1 & 0 & 1 & 0 & 0.5 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0.5 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},\qquad(7.16)$$

and $\boldsymbol{w}_{k-1}$ is the **process noise**, which is assumed to be Gaussian with zero mean and covariance matrix $\boldsymbol{Q}$:

$$\boldsymbol{w}_k \sim N(0, \boldsymbol{Q}).\qquad(7.17)$$

Measurements.

In each tracking step, we obtain new measurements, which have to be assigned to the trajectories. The relation between the state vector $\boldsymbol{x}_k$ and the measurement vector $\boldsymbol{z}_k$ is given by

$$\boldsymbol{z}_k = \boldsymbol{H}\boldsymbol{x}_k + \boldsymbol{v}_k,\qquad(7.18)$$

with $\boldsymbol{H} = diag(1, 1, 1, 1, 1, 1)$. The vector $\boldsymbol{v}_k$ is the **measurement noise**, which is also assumed to be Gaussian with zero mean and covariance matrix $\boldsymbol{R}$:

$$\boldsymbol{v}_k \sim N(0, \boldsymbol{R}).\qquad(7.19)$$

The form of the matrix $\boldsymbol{H}$ implies that measurements are available for all components of the state vector. While the position measurement is always available from the particle segmentation, velocity and acceleration have to be computed by first resp. second order finite differences using the particle positions in the previous frames. If only one previous frame is available, we assume zero acceleration and account for possible errors by larger entries in the covariance matrix $\boldsymbol{R}$ of the measurement noise. If no previous frame is available, i.e. a particle appears for the first time, we have to initialize the trajectory as outlined in section 7.4.3.

Process noise and measurement noise.

To quote Y. Bar-Shalom, one of the leading scientists in the design of Kalman filters for tracking applications, 'tuning of a Kalman filter is an art' [Bar-Shalom, 1987]. The word 'tuning' refers to the choice of the

process noise and measurement noise covariance matrices. The former represent the degree of confidence in the kinematic model resp. the predictability of the particle motion, and the latter quantify the precision of the measurements. In principle, both matrices can be time-dependent. For example, the process noise should depend on the dynamics of the tracked object. Many sophisticated process noise models have been suggested, see e.g. [Blackman and Popoli, 1999].

For our application, we choose a simplified approach, which nevertheless shows good performance. The measurement covariance is given by the confidence ellipses obtained in the particle segmentation, see figure 7.11b, section 7.3.4 and appendix A. We use the same confidence ellipses to assess the precision of position, velocity and acceleration. This accounts for a higher uncertainty along the trajectory of the particle than perpendicular to it, since the major axes of the ellipses are aligned along the trajectory. Another option is to use constant entries for the velocity and acceleration noise, where the acceleration noise is typically higher to account for (turbulent) fluctuations of the particle motion.

The process noise Q is also assumed to be constant. Good results are obtained with the choice $Q = diag(0.01, 0.01, 0.1, 0.1, 1.0, 1.0)$.

Prediction of search area.

The first step of the Kalman filter cycle is the so-called **time update**, i.e. the prediction of a search area or *validation region* in the next frame, where the corresponding particle will be most probably found. The time update equations are

$$\boldsymbol{x}_k^- = \boldsymbol{A}\boldsymbol{x}_{k-1}, \tag{7.20}$$

$$\boldsymbol{P}_k^- = \boldsymbol{A}\boldsymbol{P}_{k-1}\boldsymbol{A}^T + \boldsymbol{Q}. \tag{7.21}$$

An ellipsoidal search region around the predicted position $\boldsymbol{x}_k^-$ is defined by the confidence ellipse given by $\boldsymbol{P}_k^-$. By an appropriate scaling of this ellipse (e.g. corresponding to a three-sigma threshold), we obtain the contours of equal Mahalanobis distance of the measurements around their predicted value. Figure 7.11c shows the predicted search areas, together with the current end points of the trajectories (red dots).
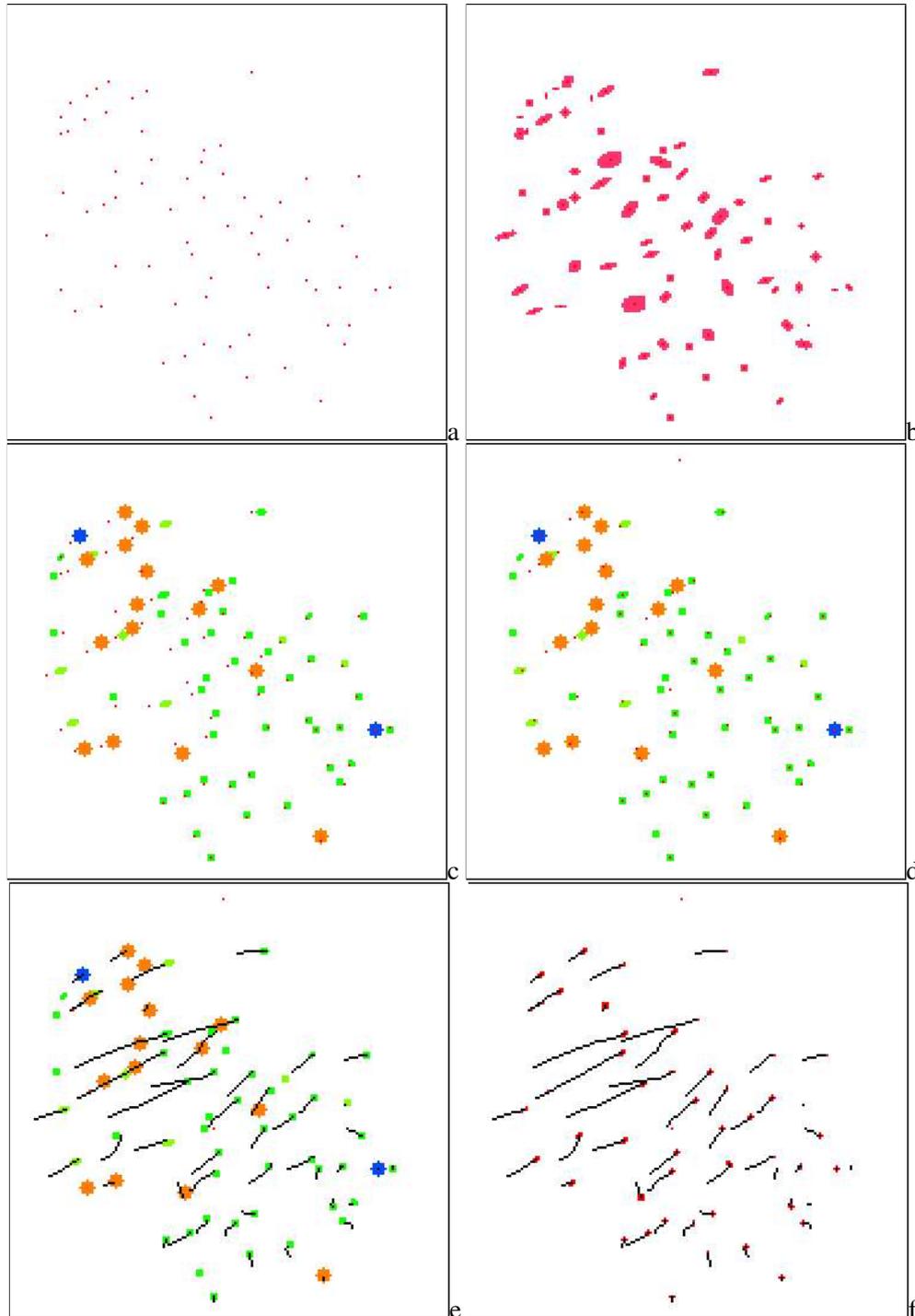
Figure 7.11: *The various stages of one tracking step in the Kalman filter cycle. **a** Particles in the current frame. **b** Uncertainty ellipses of the particles. **c** Predicted search areas and current end points of the trajectories (red dots): orange = predicted by optical flow, green = predicted by the kinematic model of the Kalman filter, blue = predicted by spatial overlap. **d** Predicted search areas and new measurements (red dots). Note that almost all new measurements are uniquely assigned to the search regions of the trajectories. **e** Trajectory update. **f** State update. The red areas indicate the uncertainty ellipses of the updated particle states. Note that the updated ellipses are significantly smaller than both the measurement noise and the size of the search regions.*

Assignment of new particles to trajectories.

After the time update step, we have to solve the **particle matching**, i.e. we have to look for a unique correspondence partner within the predicted search region. If only one particle is found within the search region, it is matched to the corresponding trajectory. However, three cases of ambiguities may also occur:

- several possible candidates are within the search region ('one-to-many'),

- several trajectories have the same search region ('many-to-one'),

- several trajectories have the same search region containing several particles ('many-to-many').

These ambiguities occur if the particle density is high and trajectories are located close to each other or cross each other, see also section 7.4.5. To solve the one-to-many ambiguity, we take again a simplified approach and choose the particle with the **minimum Mahalanobis distance** to the predicted position. While this approach is still simple, it will perform better than the usual nearest-neighbour strategy, which is based on the Euclidean distance. The latter is an isotropic distance measure, while the Mahalanobis distance takes into account the elliptical shape of the predicted search region. The many-to-one ambiguities are resolved in a 'greedy' fashion by the processing order of the trajectories, see section 7.4.1. Note that for typical particle densities occuring in our application, the number of ambiguities is rather small. For example, figure 7.11d shows that almost all new measurements are uniquely assigned to the search regions of the trajectories. Many ambiguities can already be resolved in the segmentation step by the separation of overlapping particle images (section 7.3.3).

Measurement update.

The final step of the Kalman filter cycle is the so-called **measurement update**, which computes the statistically optimal particle position by combining the position predicted by the kinematic model with the measured position of the particle that has been assigned to the trajectory in the matching step. The measurement update equations read

$$\boldsymbol{x}_k = \boldsymbol{x}_k^- + \boldsymbol{K}_k \left(\boldsymbol{z}_k - \boldsymbol{H}\boldsymbol{x}_k^- \right), \tag{7.22}$$
$$\boldsymbol{P}_k = \left(\boldsymbol{I} - \boldsymbol{K}_k \boldsymbol{H} \right) \boldsymbol{P}_k^-, \tag{7.23}$$

with the so-called *gain matrix* or *Kalman gain*

$$\boldsymbol{K}_k = \boldsymbol{P}_k^- \boldsymbol{H}^T \left(\boldsymbol{H}\boldsymbol{P}_k^- \boldsymbol{H}^T + \boldsymbol{R} \right)^{-1}. \tag{7.24}$$

The new particle state is represented by $(\boldsymbol{x}_k, \boldsymbol{P}_k)$. In figure 7.11f, we see that the uncertainty of the particle position (ellipse given by $\boldsymbol{P}_k$) is smaller than both the measurement uncertainty (ellipse given by $\boldsymbol{R}$) and the model uncertainty (size of the search region, ellipse given by $\boldsymbol{P}_k^-$), which is the basic effect of combining measurements and model predictions as done in (7.23).

7.4.3 Trajectory initialization

Trajectory initialization refers to the case when a particle enters the field of view for the first time. In this case, a new trajectory is created, and no previous information is available. We have to find a way to initialize

the state vector (7.14), which basically means that we need some estimate of the particle velocity and acceleration (the position is available from the segmentation). This is the place where the low-level velocity estimators enter our hybrid tracking approach. As already mentioned in section 7.4.1, we estimate the particle velocity by spatial overlaps, the structure tensor method and a PIV method, and use this information to initialize the state vector. First we look for unique image overlaps. If none are found, we use the optical flow estimation. If the latter is also not available because the confidence measure is too low, we use the PIV estimation. The first two methods will fail if the displacements are large. In this case, we also use the PIV result. If all low-level velocity estimators have failed, the trajectory cannot be initialized, and the particle is rejected. If there are no discontinuities in the local displacement field in the image, the multigrid PIV estimator is able to initialize particle trajectories with displacements of up to 20 pixels per frame.

7.4.4   Trajectory termination

A trajectory is terminated if no matching particle is found within the search region of the Kalman filter. Trajectory termination occurs in two cases. The first (and natural one) are particles leaving the field of view, either across the image border or along the depth direction with a decreasing signal-to-noise ratio. The second case is a **tracking failure**, either caused by a segmentation failure (false negative) or a wrong estimation of the search region. To distinguish between the two cases, we check whether the particle is close to the image border or its signal-to-noise ratio is low and has been continuously decreasing for some frames. If none of these conditions are given, the termination is probably the result of a tracking failure. Tracking failures can be resolved by trying to improve either the segmentation or the definition of the search area (or both). We only follow the first approach, see section 7.4.5.

7.4.5   Occlusions and segmentation failures

The most difficult part of a tracking algorithm is to solve the problems related to occlusions, segmentation failures and ambiguities. In simple tracking approaches, the trajectory will be terminated in these cases, since a (unique) matching particle cannot be found. We have already mentioned our simplified approach to resolve ambiguities in section 7.4.2. More sophisticated tracking algorithms try to resolve ambiguities by extending the temporal scope of the matching over several frames. One example is the 'Multiple Hypothesis Tracker' mentioned in section 3.5.2, that takes into account several different possible trajectories ('hypotheses') and tries to find the set of correct ones. However, the resulting combinatorial optimization problems become very complex[1] and computationally intensive.

In the present application, tests have shown that the number of trajectory terminations due to occlusions, i.e. crossing trajectories, is rather low. Typically, less than 15% of the trajectories are terminated due to an unresolved occlusion. In previous implementations, occlusions occured more frequently, since overlapping particles and particles that are very close to each other were not resolved. The application of the watershed segmentation (section 7.3.3) has significantly improved this situation. Hence, no further effort has been spent to resolve the few remaining occlusions. According to the tracking strategy described in section 7.4.1, in case of an occlusion, the longer trajectory will be continued, and the shorter one will be terminated. As an option, we try to link broken trajectories in a postprocessing step. Towards this end, we look for trajectory

---

[1] the simultaneous particle matching for n frames is a multidimensional assignment problem, which is NP-complete for $n \geq 3$ [Nemhauser and Wolsey, 1999]
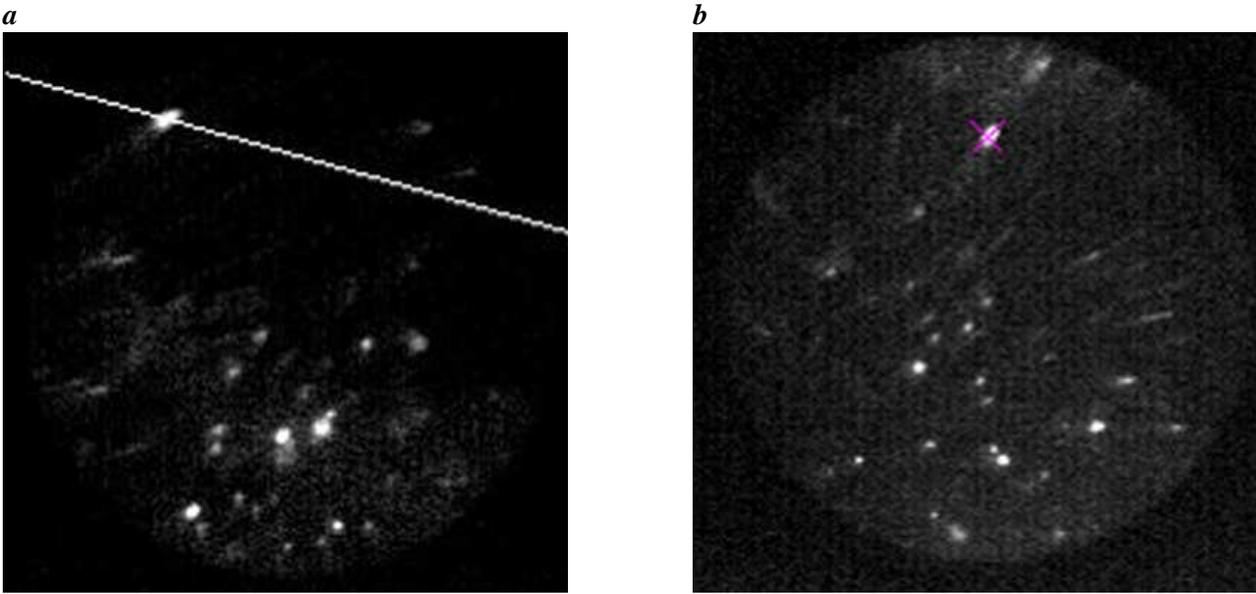
Figure 7.12: *Search for correspondences in stereoscopic particle images, utilizing the epipolar constraint. A particle in the right image has been selected (red cross in **b**). The corresponding epipolar line has been drawn in the left image (**a**). In this case, the unique correspondence partner can be easily identified. However, in general, several possible correspondence partners will be found on the epipolar line in the left image.*

ends and trajectory starts that are close to each other and compare the velocity vectors of the corresponding end resp. start points in magnitude and direction. If the vectors are similar, the two trajectories are connected by linearly interpolating the gap between them.

Trajectories may also be terminated due to segmentation failures, i.e. the corresponding particle is not found in the predicted search area because it has not been segmented ('false negative'). To deal with such cases, a feedback between segmentation and tracking has been introduced. All images are actually segmented twice, using a high and a low threshold on the signal-to-noise ratio. If a tracking failure is detected, we assume that it is caused by a false negative of the segmentation. Hence, we look for a corresponding particle in the second segmentation image that has been obtained using the lower threshold. If a particle is found there, we use it to continue the trajectory. In this way, the tracking becomes more robust to noise without producing many spurious trajectories that are initialized by false positives.

7.5 Stereo correspondence analysis and 3-D reconstruction

7.5.1 Stereo correlation of trajectories

To establish the stereoscopic correspondences between the particle images in the left and the right view, we use the epipolar constraint as described in section 6.1. Figure 7.12 shows a simple example in an image pair of low particle density. In this case, a unique correspondence can be established, since only one particle is located on the epipolar line. However, one can easily imagine that for higher particle densities the correspondence will not be unique because several particles are located on the epipolar line resp. in the epipolar plane in 3-D space. Figure 7.13a illustrates the simplest case of an ambiguity of two particles. Since in general the particles cannot be distinguished by further individual object properties like size, shape
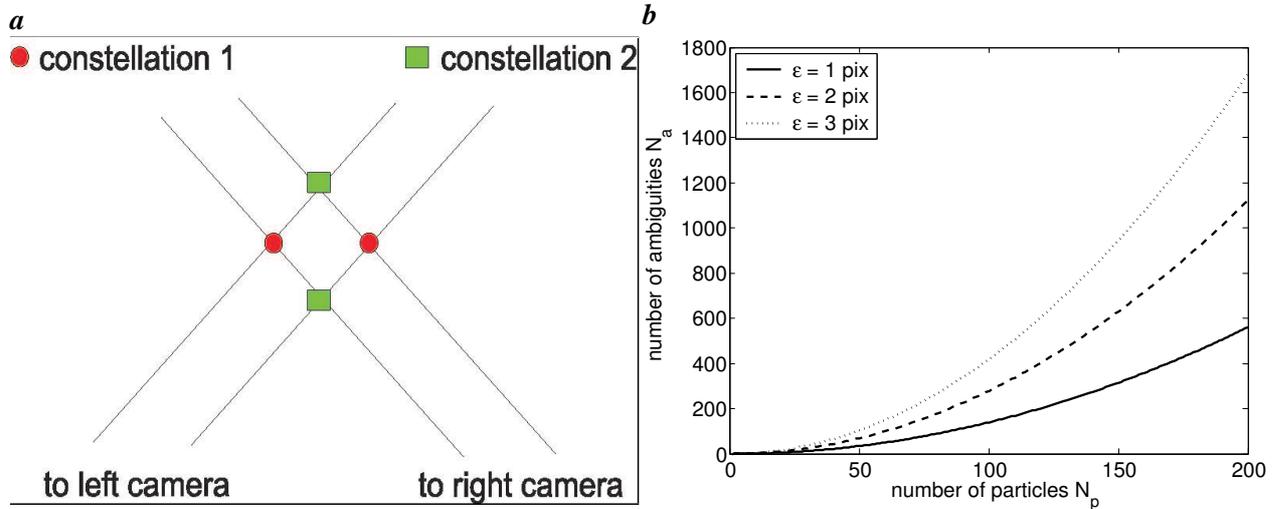
Figure 7.13: a Simple example of ambiguities occuring in the stereo correlation of particle images. The colored markers represent two different possible constellations of two particles, which lie in the same epipolar plane. The resulting ambiguity cannot be resolved without taking into account further information. b Number of expected ambiguities for an endoscopic stereo setup.

or color, the stereoscopic reconstruction of single particles is infeasible. Maas [1992] derived the following formula to compute the expected number of ambiguities N_a in dependence of the number of particles in the image and the geometric parameters of the stereo rig:

$$N_a = (N_p^2 - N_p)\frac{2f\epsilon b(Z_{max} - Z_{min})}{A_i Z_{min} Z_{max}}, \tag{7.25}$$

where N_p is the number of particles in the image, f is the focal length, b is the length of the stereo baseline, A_i is the image area, $[Z_{min}, Z_{max}]$ is the depth range of the observation volume and ϵ is the width of the epipolar search area perpendicular to the epipolar line ('**epipolar window**'). The latter is necessary to account for inaccuracies in the camera parameters and the particle positions. Figure 7.13b shows the resulting curves for the parameters of an endoscopic setup and epipolar windows of up to three pixels. It follows that up to ≈ 10 particles are expected to lie on the same epipolar line, depending on the particle density.

There are two possibilities to resolve the stereoscopic ambiguities. They correspond to the two different approaches to extend 2-D PTV to the third spatial dimension (discussed in section 3.5.3). The most common approach is to use a third camera. The additional view restricts the search area from the epipolar line to the intersection of two epipolar lines. The number of ambiguities is significantly reduced, see e.g. [Maas et al., 1993]. If at least three views are available, a 3-D reconstruction of single particle positions and a subsequent tracking in 3-D space is possible.

Our algorithm is based on the second approach, which does not use an additional camera. The particle-tracking is performed in image space, as described in section 7.4. The particles in the left resp. right image sequence are tracked independently of each other. This results in two lists of 2-D particle trajectories. After the 2-D tracking, the stereo correlation is solved for particle trajectories instead of single particles. This means that the objects which are to be reconstructed into 3-D space are *trajectories*. The latter can be distinguished from each other by their geometrical shape. Especially for fluctuating flow fields, the particle trajectories are expected to differ remarkably from each other. In other words, ambiguities are reduced,

because the epipolar constraint is not imposed on single particles, but rather on all particles forming a trajectory.

To make the above consideration more precise, consider the set of planes in space that contain the baseline of the stereo rig. This set is called the **pencil of epipolar planes**. Any two epipolar planes can be aligned with each other by a rotation around the baseline. The positions of a particle along its trajectory define a discrete subset out of the pencil of epipolar planes. We define a cylindrical coordinate system where the baseline is the symmetry axis. In this coordinate system, the position vector of a particle is given by $\boldsymbol{x}_{cyl} = [z, r, \phi]^T$, where z is the component in the direction of the baseline, r is the component in the radial direction perpendicular to the baseline, and ϕ is the angular component describing the rotation around the baseline. The necessary and sufficient condition for the 3-D reconstruction of the trajectory is the uniqueness of the angular motion of the particle, described by the ϕ-component. The latter is a reasonable assumption for a fluid flow field.

The basic principle of our method is to use the temporal information that is available from the tracking of single particles to resolve spatial ambiguities. A similar approach is proposed by Chang and Aggarwal [1997]. However, in their method they use spatial relations to resolve temporal ambiguities in the tracking of a monoscopic image sequence. Willneff and Gruen [2002] present a 3-D PTV algorithm that performs a combined simultaneous analysis of temporal and spatial correspondences.

The present algorithm for stereoscopic correspondence analysis is based on previous work by Netsch [1995], Garbe [1998] and Engelmann [2000]. An improved version has been presented by Klar [2001]. The latter is used here. The correspondence search starts by checking the epipolar constraint for all possible pairs of left and right trajectories and thereby creating a list of possible correspondence partners. The algorithm is carried out in four passes. In the first pass, a small epipolar window of 0.1 pixels is used, and only trajectories longer than a specified minimum length are considered. If unique correspondences are found, they are removed from the list of remaining correspondence partners and marked as matched. The search continues by requiring lower minimum trajectory lengths up to a minimum length of four frames. As before, all unique correspondences are removed from the list. In a second pass, we recompute the list of remaining correspondence partners using a larger epiplar window of one pixel. The search is carried out in the same fashion by successively lowering the required trajectory length and removing all unique matches. In the final pass, we use an epipolar window of three pixels. In each pass, different heuristic criteria are applied to resolve trajectory ambiguities. For the details of the implementation see [Klar, 2001].

7.5.2   3-D reconstruction of trajectories

Triangulation of 3-D points.

If a unique correspondence between the left and right image of a trajectory has been established, the 3-D trajectory is computed by the midpoint triangulation method described in section 6.2. To obtain a confidence measure for the triangulation, we compute the mean Euclidean distance between the two projection rays over all points of the trajectory. If the mean distance exceeds a threshold, the trajectory is rejected. The threshold is set according to the 3-D reconstruction accuracy that has been determined in section 6.2.3.

Velocity computation.

Given the successive 3-D particle positions $\boldsymbol{X}_{i-1}, \boldsymbol{X}_i, \boldsymbol{X}_{i+1}$ along a trajectory, the particle velocity is approximated by a second order central difference scheme:

$$\boldsymbol{V}_i = \frac{\boldsymbol{X}_{i+1} - \boldsymbol{X}_{i-1}}{2\Delta t} + O(\Delta t^2), \tag{7.26}$$

where Δt is the frame period of the camera in seconds.

Malik et al. [1993] use the second order central scheme and state that the latter yields an accurate estimate of the local Lagrangian mean velocity over the time interval $2\Delta t$. They also test linear and higher order schemes and report that they perform worse, provided that the particle acceleration is approximately constant over the three frames. In their tracking scheme, they use the criterion of minimum change of acceleration to resolve tracking ambiguities. A similar approach is taken here, since we use a constant acceleration model in the Kalman filter.

Lüthi [2002] and Willneff and Lüthi [2003] do not compute velocities by numerical differentiation because of the high risk of large errors due to noise in the tracer particle locations. They fit third-order polynomials in time to 21 successive particle positions along the trajectory (one polynomial for each component). The particle velocity is computed by analytical differentiation of the polynomials. The basic effect of these 'moving cubic splines' is a temporal lowpass filtering of the position data.

In our implementation, we apply a temporal smoothing of the position data along the trajectories before computing the velocities. The smoothing is carried out by a convolution of the trajectories with a 3-tap binomial filter. The half width of the transfer function of this filter is half the Nyquist frequency. In case of the CMOS cameras, this means that noisy fluctuations with a frequency of $100\,Hz$ and more are reduced in amplitude by at least a factor of two. Depending on the noise level, several successive convolutions may be applied, which increases the effective filter size resp. reduces the half width of the corresponding transfer function.

7.6 Summary

The 3-D PTV algorithm developed in this thesis has been presented. The task of this algorithm is to compute the 3-D particle trajectories of flow tracers, given a stereoscopic image sequence of the flow. Towards this end, two correspondence problems have to be solved: the temporal correspondence of particles between subsequent frames ('tracking') and the spatial (stereoscopic) correspondence of particles between the two views.

It is difficult to solve the stereoscopic correspondences of an ensemble of tracer particles. The reason is that the epipolar geometry of a stereo setup only restricts the search area of a corresponding particle to a line in the image. Except for very low particle densities, the probability that several particle images will be found on this line is high. The resulting ambiguities cannot be resolved, since the particle images cannot be distinguished by further features like size, shape or intensity.

In our approach, we start with a 2-D tracking of the particles in each sequence. Hence, we first solve the temporal correspondences in image space. In a second step, we analyze the stereoscopic correspondences of trajectories rather than of single particles. In this case, we can use an additional constraint to resolve ambiguities, namely the geometric shape of the trajectories. Because the epipolar constraint is imposed

on all particles of a trajectory rather than on a single particle, the number of ambiguities is significantly reduced.

The main difficulty of our approach is the particle-tracking in image space, which is more complex than tracking in 3-D space. The principal reason is the volume illumination that is necessary in a 3-D technique. This has two consequences. First, the size and brightness of the particles may vary considerably because of illumination inhomogeneities and out-of-focus imaging. Second, the motion of particles in different depth layers is projected simultaneously onto the sensor. This creates complex optical flow patterns in the image plane. Particles may overlap or occlude each other, which results in asymmetric particle shapes and crossing or merging trajectories. Finally, we have to keep in mind that both the resolution and the signal-to-noise ratio is generally lower in endoscopic images. Hence, the need for a robust tracking algorithm becomes clear.

A powerful segmentation algorithm has been developed, that achieves a reliable determination of the particle positions. The particles are extracted by taking into account both features that distinguish them from the background, namely their motion and their intensity. A watershed algorithm is used to split overlapping particle images, which reduces the number of tracking ambiguities.

The tracking is performed within a Kalman filter framework. Different low-level motion estimators are integrated to initialize the trajectories. The reliability of both the segmentation and the tracking is increased by taking into account an extended temporal neighbourhood around the current frame. To further reduce the number of tracking ambiguities, the simple nearest neighbour search of previous implementations has been improved by using the Mahalanobis distance instead of the Euclidean distance. This enables a more precise definition of search areas, where the correct matching particle is most probably located.

The main limitation of the tracking algorithm is that the remaining tracking ambiguities are not explicitly resolved. Only a greedy approach has been implemented. However, because of the relatively moderate particle densities in the present application, this drawback is not too severe.

Chapter 8

Performance and accuracy assessment

Performance evaluation is an essential step in the design of any algorithm. Such an evaluation should proof the correctness of the algorithm and quantify the accuracy and precision of its results. The performance often depends on certain parameters of the input data. For example, the tracking efficiency of a particle-tracking algorithm depends on the particle density in the images. In such cases, parameter ranges resulting in optimal performance resp. limitations of the algorithm should be indicated. To analyze the correctness and accuracy of an algorithm, the correct results ('ground truth') have to be known and compared to the results of the algorithm. Since the acquisition of image sequences with precisely known ground truth is very elaborate and involves expensive equipment, synthetic images are commonly used instead.

This chapter provides a comprehensive performance analyis of the 3-D PTV algorithm. In section 8.1 we introduce the image parameters and the performance measures that will be used. The analysis is carried out both for synthetic (section 8.2) and real image sequences (section 8.3). Section 8.4 presents the accuracy analysis of the velocity measurements. The results of this chapter are summarized in section 8.5.

8.1   Image parameters and performance measures

We summarize the definitions of the image parameters and the performance measures that are used in the evaluation. The image parameters describe the properties of the particle images, and the performance measures quantify the quality of the tracking results. The quantities described here are commonly used for performance assessment by many authors of PIV and PTV algorithms [Adrian, 1991; Malik et al., 1993; Kieft, 2000; Bastiaans et al., 2002].

Image source density.

The image source density N_s is the fraction of the total image area that is covered by particles:

$$N_s = \frac{\pi}{4} d_p^2 \frac{N_p}{A_i},$$

(8.1)

where d_p is the particle diameter, N_p is the number of particles in the image and A_i is the image area. Note that the source density is different from the particle density. The latter may be defined as the number of particles per image or the number of particles per pixel.

Mean particle spacing.

The mean particle spacing Λ_0 (more precisely: the mean minimum particle distance between nearest neighbours) is given by

$$\Lambda_0 = \left(\frac{A_i}{4N_p} \right)^{1/2}. \tag{8.2}$$

Sampling quality.

The (temporal) sampling quality r_p [Bastiaans et al., 2002] is the ratio of the average particle spacing Λ_0 and the average particle displacement between two successive frames Λ_t:

$$r_p = \frac{\Lambda_0}{\Lambda_t}. \tag{8.3}$$

The same quantity is termed the 'particle spacing displacement ratio' by Malik et al. [1993]. Tracking is easy, if $r_p >> 1$ and becomes difficult for $r_p \approx 1$. Tracking is practically impossible for $r_p << 1$ due to a large number of ambiguities in the temporal correspondence analysis. The mean particle displacement Λ_t depends on the frame rate of the camera and the flow velocity.

Measures of efficiency.

Typically, the **tracking efficiency** of a 3-D PTV algorithm is defined as the ratio of the number of particles tracked in 3-D to the number of reconstructed 3-D positions per time step, e.g. [Willneff and Gruen, 2002]. Since in our algorithm we track the particles in the image plane and perform a 3-D reconstruction of the resulting trajectories, a 3-D particle position always implies that this particle is also tracked. We define the following three efficiency measures.

The **detection efficiency** η_{det} is the ratio of the number N_{seg} of segmented particles to the number N_P of particles that the image actually contains:

$$\eta_{det} = \frac{N_{seg}}{N_P}. \tag{8.4}$$

The detection efficiency can only be computed if ground truth data is available, since only in this case the actual number of particles N_P is known.

The **2-D tracking efficiency** η_{2d} is defined as follows:

$$\eta_{2d} = \frac{N_{track}}{N_{seg}}, \tag{8.5}$$

where N_{track} is the number of tracked particles and N_{seg} is the number of segmented particles in the image. The 2-D tracking efficiency characterizes the performance of the 2-D particle tracking described in section 7.4. Only particles that are tracked in at least four consecutive frames are added to N_{track}.

Finally, to assess the performance of the 3-D reconstruction, we define the **3-D reconstruction efficiency** η_{3d} as

$$\eta_{3d} = \frac{N_{rec}}{N_{track}}, \tag{8.6}$$

where N_{rec} is the number of particles reconstructed in 3-D.
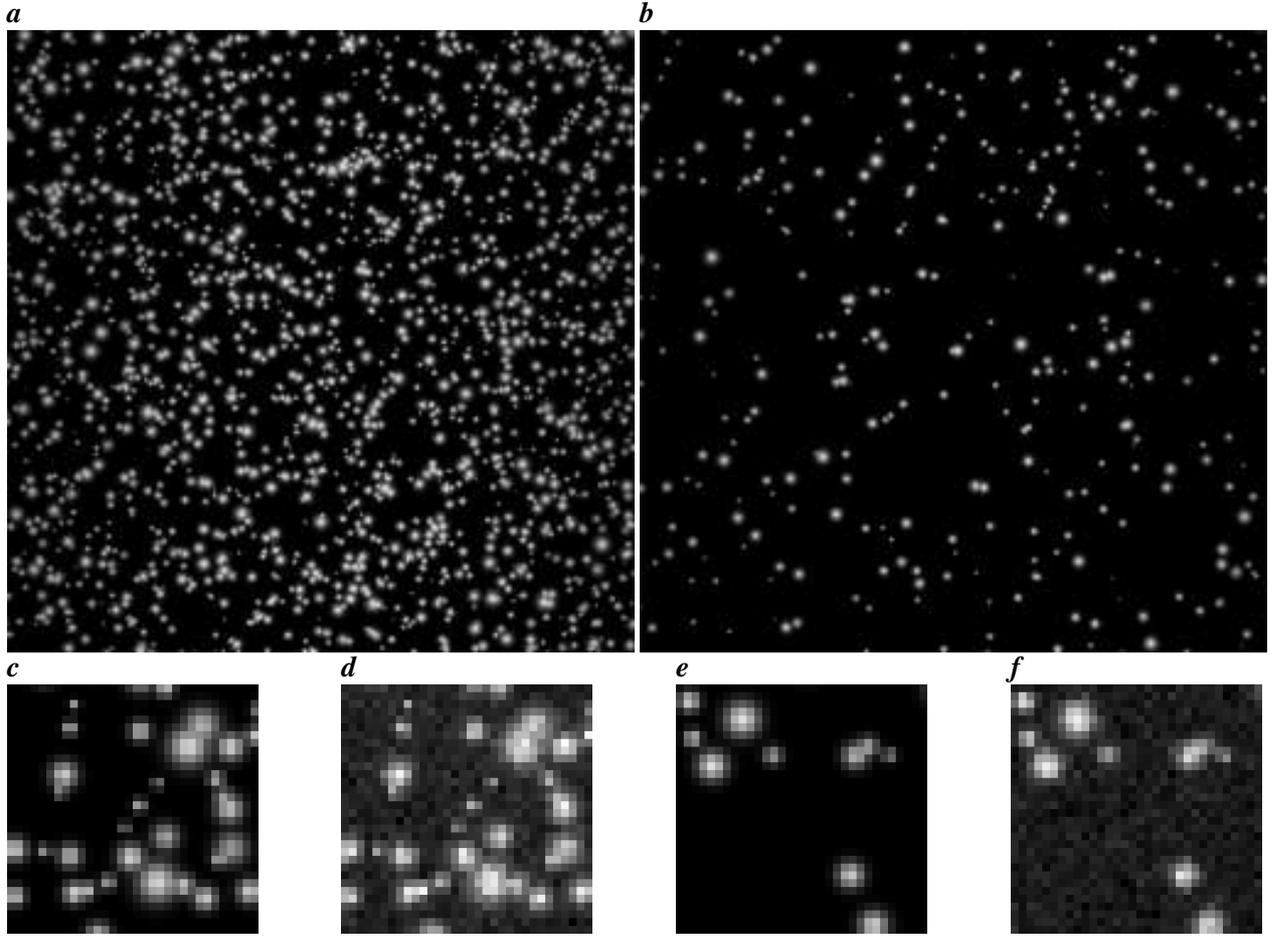
Figure 8.1: *Example test images of the PIV-STD Project provided by the Visualization Society of Japan.* ***a*** *Test case 351.* ***b*** *Test case 352.* ***c*** *Detail from* ***a*** *.* ***d*** *Detail from* ***a*** *with additive Gaussian and Poisson noise.* ***e*** *Detail from* ***b*** *.* ***f*** *Detail from* ***b*** *with additive Gaussian and Poisson noise.*

Combining the above definitions, the **total fractional vector yield** η_{tot} is defined as the number of reconstructed 3-D particle positions to the number of particles in the image:

$$\eta_{tot} = \eta_{det}\,\eta_{2d}\,\eta_{3d} = \frac{N_{seg}}{N_P}\frac{N_{track}}{N_{seg}}\frac{N_{rec}}{N_{track}} = \frac{N_{rec}}{N_P}. \tag{8.7}$$

Note that also in the ideal case of $\eta_{det} = \eta_{2d} = 1$, the total yield η_{tot} may be smaller than one if a fraction of the particles is only visible in one of the two image sequences. This fraction can be minimized by a careful design of the experimental setup. However, it will certainly not be reduced to zero in practical applications.

8.2 Synthetic test images

8.2.1 Properties of the test images

The Visualization Society of Japan has initiated the so-called PIV-STD Project. The objective is to develop PIV standards as a common basis for a comparison of different algorithms. Within this project, many
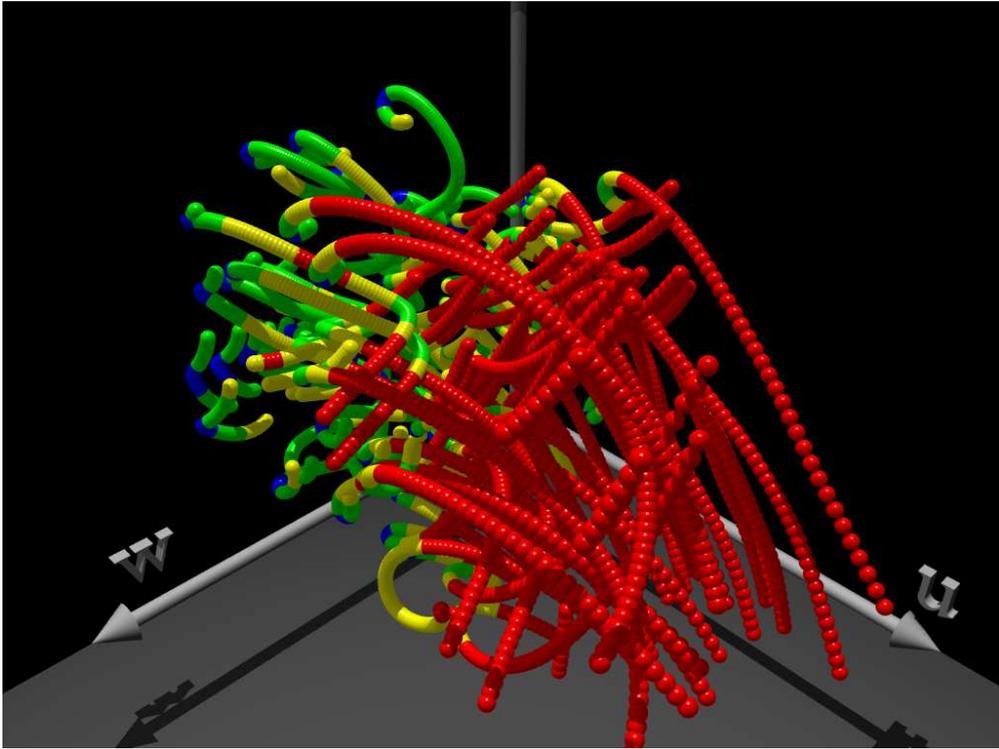
Figure 8.2: *3-D trajectories of the jet shear flow starting in the first frame of test case 352 of the PIV-STD Project. The color code is the same as in figure 3.5.*

synthetic PIV image sequences have been generated along with the corresponding ground truth velocity fields. The sequences are distributed via the internet[1]. Many of the images are designed for PIV evaluation (high particle density, only two to four successive images, light sheet illumination, stationary velocity field). However, the test cases also include some sequences which are suitable for processing by a tracking algorithm.

Here we use the test cases 351 and 352. These images have been generated using the velocity field of a jet shear flow, which has been calculated numerically by a Large Eddy Simulation. A volumetric illumination has been simulated, and image sequences from three different views are provided. The images show a time-dependent turbulent flow field containing rich three-dimensional vortex motion on many different scales, see figure 8.2. The latter is important for an efficient evaluation of a 3-D technique. The data set also includes calibration images for the geometric camera calibration. Thus, a full evaluation of the 3-D PTV is possible. The camera calibration has been carried out using the same method as described in sections 5.3.1–5.3.2.

For each test case, a sequence of 145 images is provided (only 136 are used in the evaluation). The image size is 256×256 pixels. The two test cases 351 and 352 used here only differ in the particle density. Test case 351 has a high particle density, which is suitable for PIV evaluation, but is rather high for PTV. Test case 352 has a particle density typical for PTV. To give a visual impression, figure 8.1 shows example images of the test cases. Since the original images are free of noise, additional tests with additive noise have also been carried out. The noise contains two different components: a Gaussian component that is independent of the gray value and a Poisson component with a variance proportional to the mean gray

---

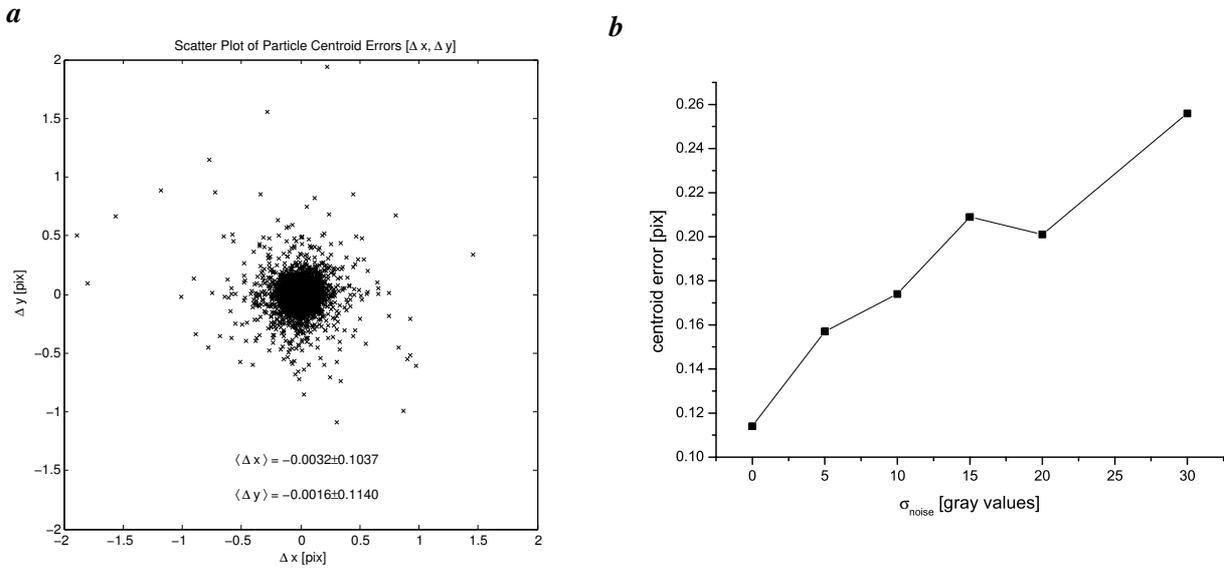[1]The images can be downloaded from the website `http://www.vsj.or.jp/piv`.

a

b

Figure 8.3: a *Scatter plot of the particle centroid errors (test case 352, particle positions from 136 images).* ***b*** *Noise dependence of the particle centroid error.*

value. Figure 8.1d,f shows examples with a medium noise level. More details on the PIV-STD test images are given by Okamoto et al. [2000b] and Okamoto et al. [2000a].

8.2.2 Results

Accuracy of 2-D position determination.

We test the accuracy of the extracted particle positions by computing the rms-values of the difference between the estimated particle positions and the ground truth for all segmented particles in test case 352. The results are shown in figure 8.3. The particle positions are estimated to an accuracy of about 0.1 pixels, which confirms the accuracy estimation carried out in section 7.3.4. The position error increases linearly with the noise level.

Performance of 3-D PTV.

We process the image sequences of test cases 351 and 352 with the 3-D PTV algorithm described previously in chapter 7. The same parameter settings are used as in the processing of real flow measurements (most relevant: epipolar windows of up to three pixels and a minimum trajectory length of four frames). The results of the performance analysis are compiled in table 8.1. We can draw the following conclusions:

- The mean particle diameter and image source density determined by the PTV algorithm are lower than the true values. The reason is that the segmentation only takes into account pixels with a gray value larger than a fraction of $1/e$ of the particle's maximum gray value.

- The performance in test case 352 is much better than in test case 351. The reason is the very high particle density in case 351. Typically, image source densities in PTV applications are well below 10%. Test case 351 has a source density of 27%, which is a great challenge for a tracking algorithm

Table 8.1: *Image parameters and performance measures obtained from the synthetic test sequences. Columns two and three contain the ground truth, columns four and five the results of the 3-D PTV algorithm. Two different reconstruction efficiencies are shown: (A) obtained from the application of the complete 3-D PTV algorithm, (B) obtained using ground truth 2-D trajectories and only performing the 3-D reconstruction. For a discussion, see the text in this section.*

test case	351 truth	352 truth	351 PTV	352 PTV
particles per image N_P resp. N_{seg}	1739 ± 56	293 ± 21	1459	293
particles per pixel $[10^{-3}]$	27 ± 1	4.5 ± 0.3	22	4.5
mean particle diameter [pix]	5 ± 2	5 ± 2	4	3
image source density [%]	-	-	27	3.5
mean particle spacing Λ_0 [pix]	3.1	7.5	3.4	7.5
mean 2-D displacement Λ_t [pix]	3.5 ± 2.6	3.4 ± 2.5	2.9 ± 2.4	3.3 ± 2.7
maximum 2-D displacement [pix]	19	17	18.7	16.9
sampling quality r_p	0.89	2.21	1.17	2.27
mean trajectory length [frames]	36 ± 32	37 ± 33	12	26
X-range of stereo volume	$[-1.41, 1.41]$	$[-1.40, 1.41]$	$[-1.40, 1.41]$	$[-1.40, 1.40]$
Y-range of stereo volume	$[-1.26, 1.26]$	$[-1.26, 1.26]$	$[-1.24, 1.24]$	$[-1.23, 1.23]$
Z-range of stereo volume	$[-1.00, 1.00]$	$[-1.00, 1.00]$	$[-1.02, 1.02]$	$[-1.01, 1.00]$
total number of 3-D vectors	201584	34042	50707 (A)	21072 (A)
			193540 (B)	32987 (B)
wrong correspondences			15%	0.5%
detection efficiency η_{det}			84%	100%
tracking efficiency η_{2d}			75%	87%
reconstruction efficiency η_{3d}			34% (A)	60% (A)
			96% (B)	97% (B)

because the particles are very close to each other. This is also expressed in a sampling quality r_p of only 0.89, which means that the average particle displacements are larger than the mean particle distance. Keeping this in mind, the tracking efficiency of 75% has to be considered as very good. However, as a result of occlusions and unresolved particle overlaps, the recovered trajectories are interrupted. The mean length of the trajectories drops to about a third of the actual length. As outlined in section 7.5, the probability of spatial ambiguities in the stereo correspondence search is higher for shorter trajectories. Hence, the reconstruction efficiency is only 34%. Further, the fraction of wrong correspondences is high. The latter are introduced by wrong matches of short trajectory segments.

- The images in test case 352 have a source density that is typical for PTV applications. The 3-D PTV algorithm performs quite well: practically all particles are segmented, 87% of them are tracked in at least four frames, and 60% of the tracked particles can be reconstructed in 3-D. This corresponds to a total 3-D vector yield of $\approx 52\%$. This value is lower than in typical 3-D PTV systems with three or

more cameras (see table 3.2). But keeping in mind that we use only two cameras, the result is very satisfactory. Compared to the previous implementation of the algorithm by Engelmann et al. [1999], the vector yield is increased by at least a factor of two. Note that the ratio of recovered vectors to the number of possible 3-D vectors (with particle images visible in both image sequences) is 60%. The lower total yield of 52% (according to definition (8.7)) implies that a certain amount of particles is visible in only one of the two sequences.

- To find out the main limiting part of the 3-D PTV algorithm, we carry out the following two tests. In table 8.1 the reconstruction efficiency of the two tests is indicated by (A) and (B). In test (A), the full 3-D PTV algorithm is applied, including both the 2-D tracking and the 3-D reconstruction. In test (B), we only compute the 3-D reconstruction from the ground truth 2-D trajectories, i.e. we do not perform the tracking. A comparison of the resulting reconstruction efficiencies clearly shows that the limiting factor is the 2-D tracking, especially in test case 351. If the correct 2-D trajectories are used, a reconstruction efficiency of almost 100% is obtained in both test cases. Hence, the lower reconstruction efficiencies of the full 3-D PTV are mainly caused by failures of the tracking algorithm. The latter result in interrupted short trajectory fragments, which cannot be reconstructed in 3-D. In the present applications, the particle densities are much lower than in test case 351. Hence, no further attempts have been made to improve the tracking. Some advanced tracking techniques that aim to achieve high tracking efficiencies at high particle densities have been mentioned in section 3.5.2.

Accuracy of 3-D position and velocity measurements.

The synthetic test sequences also provide a possibility to check the accuracy of the measurements by computing the mean errors between the estimated 3-D positions resp. displacements and the known ground truth. In these tests, we use the full 3-D PTV algorithm, including tracking and 3-D reconstruction. In test case 351, we obtain an mean 3-D position error (mean of the Euclidean distance between the estimated position and the true position, computed for all particles in the sequence) of $\Delta_{XYZ} = (60.1 \pm 27.4)\ \mu m$. The mean error of the displacement estimation (difference of two successive particle positions) is $\Delta_V = (10.8 \pm 9.7)\ \mu m$. In test case 351, the corresponding errors are $\Delta_{XYZ} = (46.5 \pm 14.3)\ \mu m$ and $\Delta_V = (5.4 \pm 4.3)\ \mu m$. The lower displacement errors show the high degree of correlation in the position errors. The common systematic error in the particle positions due to miscalibration cancels out if a difference of two nearby positions is computed to estimate the velocity, see also section 8.4.

We also estimate the errors in the geometric camera calibration. The 3-D position rms errors computed in the calibration images are $[\sigma_{X,rms}, \sigma_{Y,rms}, \sigma_{Z,rms}]^T = [14.3\mu m, 7.6\mu m, 22.1\mu m]^T$. The mean distance of the calibration points to their estimated reference planes is $\langle D_\perp \rangle = (7.6 \pm 5.2)\mu m$. For the details of the computation of these estimates, see section 5.4.2. The 3-D errors in the calibration points are lower than those in the 3-D particle positions. This shows the influence of the determination of the particle positions. The latter become inaccurate if the intensity distribution of a particle is not symmetric or if particles are only partly visible due to overlaps.

Note that the calibration data supplied together with the test sequences also takes into account refraction effects due to the multimedia environment (see section 5.2.3). The latter are not included in our camera model. Still we obtain a good accuracy, which is comparable to the typical accuracies obtained in other 3-D PTV implementations (see table 3.2).
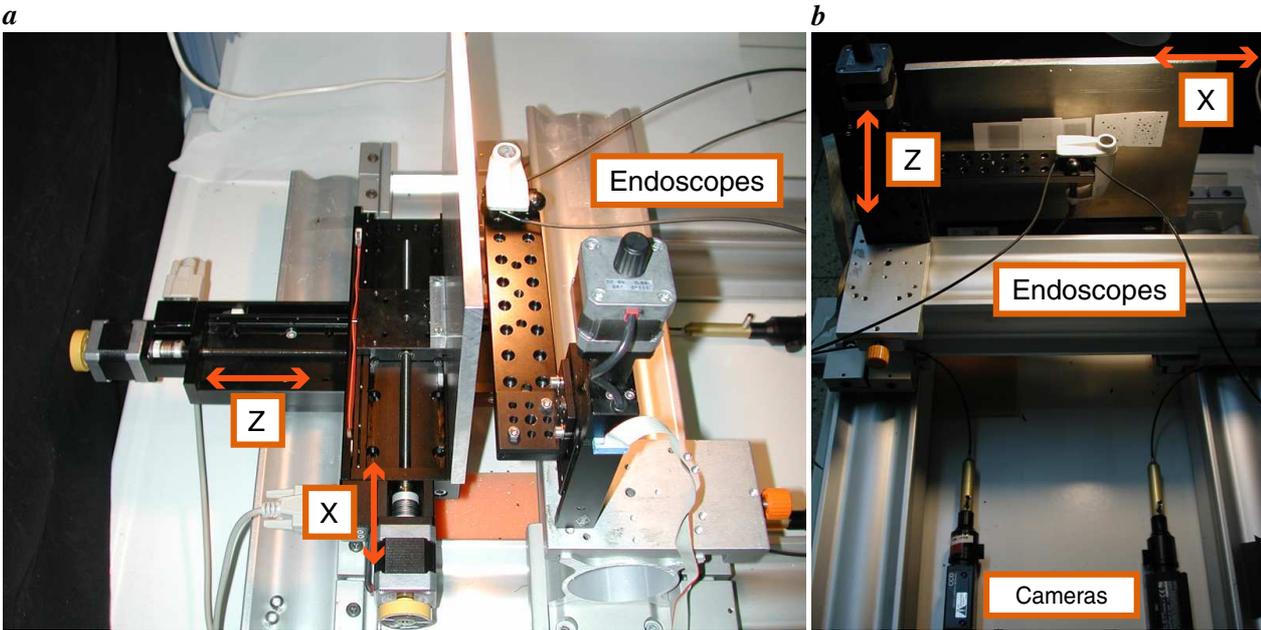
Figure 8.4: *Acquisition of real test sequences with known ground truth (of the displacements). An ensemble of particles is printed on a sheet of paper and moved using two perpendicular translation stages. **a** Side view. **b** Front view.*

Noise influence.

We test the influence of image noise on the 3-D accuracy and the total vector yield. A mixture of Gaussian and Poisson noise is added to the images. The standard deviation of the Gaussian noise is increased from 5 to 30 gray values. The results show an approximately linear increase of the 3-D position errors with the noise level. The triangulation error (computed as the perpendicular distance between the two projection rays that should ideally meet in the 3-D point) increases by a factor of two, from 12 μm at zero noise to 25 μm at the maximum noise level. The vector yield is less sensitive to noise. The tracking efficiency stays approximately constant at 90%, independent of the noise level. The relative decrease of the reconstruction efficiency is only 15% (from $\eta_{3d} = 60\%$ to $\eta_{3d} = 51\%$). This shows that the segmentation and tracking algorithms also perform well for relatively high noise levels. The main effect of the noise is the degraded 3-D accuracy.

8.3 Real test images

8.3.1 Tests using real images with ground truth

Real test sequences with known ground truth are generated using a setup with two perpendicular translation stages, see figure 8.4. An ensemble of particle images is printed on a sheet of paper. The paper is attached to a planar metal surface, which is mounted on the translation stages. An endoscopic stereo rig is installed in front of the surface. This setup allows to record image sequences with particles moving in the (lateral-) X-direction or (depth-) Z-direction (or a combination of both). The position of the translation stages is known to an accuracy of 1 μm. The Z-translation stage is also used to acquire several images of a planar grid target for the geometric camera calibration of the stereo rig. The calibration is carried out as described in

Table 8.2: *Image parameters and performance measures obtained from four representative real pore flow sequences.*

test case	kmb39	kmb33	kmb19	kmb04
velocity	'slow'	'slow'	'medium'	'fast'
particles per image N_{seg}	108	161	115	77
particles per pixel [10^{-3}]	4.5	6.7	4.8	3.2
mean particle diameter [pix]	3.6	3.5	4.1	4.7
image source density [%]	5.2	6.9	5.6	4.9
mean particle spacing Λ_0 [pix]	7.5	6.1	7.2	8.8
mean 2-D displacement Λ_t[pix]	0.15 ± 0.13	0.28 ± 0.2	0.5 ± 0.4	2.5 ± 2.1
maximum 2-D displacement [pix]	4.4	6.4	5.6	21.2
sampling quality r_p	50	22	14	3.5
mean trajectory length [frames]	137	80	51	17
total number of 3-D vectors	105327	148562	113069	66147
tracking efficiency η_{2d}	96%	97%	97%	89%
reconstruction efficiency η_{3d}	52%	53%	52%	50%

section 5.4.2. While the true absolute 3-D coordinates still are not available, the relative 3-D displacements of the particles within these sequences are known to a high precision and can be used as ground truth.

Accuracy of 3-D displacement measurement.

Both camera calibration and 3-D PTV are performed on the test sequences using the same algorithms and parameter settings as in the flow measurements. These tests mainly serve as a check on the accuracy of the velocity measurements. The following results have been obtained.

We track the motion of a set of ten particles in two sequences. In the first sequence, the true displacement between two frames is $0.25 \ mm$ in the X-direction, $\boldsymbol{D}_{true,1} = [\Delta X_1, \Delta Y_1, \Delta Z_1]^T = [0.25, 0, 0]^T$. In the second sequence, the true displacement is $0.25 \ mm$ in the Z-direction, $\boldsymbol{D}_{true,2} = [\Delta X_2, \Delta Y_2, \Delta Z_2]^T = [0, 0, 0.25]^T$. Since the same displacements should be measured for all particles, we compute the mean and standard deviation of the displacements over all reconstructed 3-D trajectories. The following mean displacement vectors have been obtained: $\boldsymbol{D}_{exp,1} = [0.243 \pm 0.009, 0.007 \pm 0.005, 0.001 \pm 0.004]^T$ for the first sequence and $\boldsymbol{D}_{exp,2} = [0.000 \pm 0.002, -0.001 \pm 0.001, 0.270 \pm 0.005]^T$ for the second sequence (all values in mm). The results show that the statistical velocity error in these tests is well below 5% of the mean velocity. The systematic error in the X-direction is also less than 5%, whereas there is a bias of about 8% in the Z-direction. The latter is probably caused by the lower accuracy of the stereo rig in the Z-direction (a stereo rig with a convergence angle of $60°$ has been used).

8.3.2 Tests using real flow sequences

Performance tests on synthetic images provide valuable information because the correct results are known and can be used for comparison. However, there is a number of additional effects (camera noise, properties

of the optical system, illumination inhomogeneities, etc.) that influence the performance of the same algorithm on real images. Hence, tests should also be carried out on real flow sequences. While obviously a comparison with ground truth is not possible in this case, the behaviour of the algorithm and its dependence on the image parameters can be studied.

We choose four representative pore flow sequences from the measurements described in chapter 10. All sequences have been acquired by the Kocher-setup (CMOS cameras, $400\,Hz$ frame rate, see section 9.3). The image parameters and performance measures are compiled in table 8.2. The endoscopic images typically contain between 100 and 150 particles. This corresponds to an image source density of about five percent, which is rather low.

The typical particle displacements are in the range from below one pixel for very slow flows to maximum displacements of more than 20 pixels for fast flows. On average, the sampling quality is much larger than one, which indicates that reasonable tracking results can be obtained without the need for sophisticated global optimization strategies to resolve temporal ambiguities. Indeed, the tracking efficiency almost reaches 100%, except for the case kmb04, which corresponds to fast turbulent pore flow induced by large surface waves in the open-channel flow.

The reconstruction efficiency is around 50% for all test cases. The main reason for the limited reconstruction efficiency is not a failure of the 3-D reconstruction due to unresolved stereo ambiguities. Most of the trajectories that cannot be triangulated do not have a correspondence partner of sufficient length at all. There are several reasons that might be responsible for this. First, particles may be visible in only one image sequence due to a shortcoming of the experimental setup. Second, particles may not be tracked long enough in one of the two sequences due to unresolved occlusions or segmentation failures. Third, two actually different trajectories may be spuriously linked if they cross each other. Such trajectories are filtered out in the 3-D reconstruction, since they do not have a correspondence partner. Further investigations are necessary to clarify this issue.

8.4   Velocity accuracy and dynamic range

Particle velocities are computed by the second order central difference scheme (7.26). Gaussian error propagation yields the following conservative estimates of the absolute and relative errors of one velocity component V:

$$\sigma_V = \frac{1}{\sqrt{2}}\frac{\sigma_X}{\Delta t}, \qquad \frac{\sigma_V}{V} = \sqrt{2}\frac{\sigma_X}{\Delta X}, \tag{8.8}$$

where σ_X is the error in a single particle coordinate X and $\Delta X = X_{i+1} - X_{i-1}$. We assume that the error in the frame period Δt of the cameras can be neglected. Equation (8.8) shows, that the relative velocity error decreases with increasing particle displacements. The velocity resolution is limited by the triangulation accuracy. Due to the high frame rates of the setups Neckar, Höllbach and Kocher, the resulting velocity errors are of the order of magnitude of one to three mm/s (see rows A and B in table 8.3), which would limit the minimum resolvable velocities. However, since we compute velocities by the difference of two particle positions that are close to each other, we can take advantage of the correlation between the position errors. As shown e.g. by Förstner [1996], systematic errors due to miscalibration yield high values of the correlation coefficient ρ. The common calibration bias will cancel out if a difference of two closely spaced positions is computed. The standard deviation of the difference of correlated measurements (as given by σ_V in (8.8)) has to be multiplied by a factor $\sqrt{1-\rho}$, where $\rho \in [0,1]$ is the correlation between the

Table 8.3: *Compilation of different estimations of the 3-D velocity accuracy. Case A: Estimation using the rms-error of the Z-component of 3-D positions as a measure of the 3-D position error. Case B: Estimation using the mean distance of calibration points to their reference planes as a measure of the 3-D position error. Case C: Estimation taking into account the correlation of the errors. All position errors have been obtained from the analysis of the calibration accuracy in section 5.4. For further discussion, see the text.*

		Neckar	Jagst	Höllbach, Kocher
	frame period Δt $[s]$	0.0077	0.02	0.0025
A	$\sigma_X = (\Delta Z)_{rms}$ $[\mu m]$	31	12	10
	σ_V $[mm/s]$	2.8	0.4	2.8
B	$\sigma_X = \langle D_\perp \rangle$ $[\mu m]$	27	7	5
	σ_V $[mm/s], \rho = 0$	2.5	0.25	1.4
C	$\sigma_X = \langle D_\perp \rangle$ $[\mu m]$	27	7	5
	σ_V $[mm/s], \rho = 0.95$	0.55	0.06	0.31
	$V_{min,exp}$ $[mm/s]$	≈ 1	≈ 0.3	≈ 0.3
	$V_{max,exp}$ $[mm/s]$	≈ 1000	≈ 30	≈ 150
	dynamic range DR_{exp}	≈ 1000	≈ 100	≈ 500

measurements. The correlation has been estimated by computing the variances and covariances of the errors of calibration points of two successive calibration planes separated by 0.2 mm along the Z-direction. The known positions of the calibration points have been used as ground truth (see also section 5.4.2). Correlation values of $\rho = 90\% - 95\%$ have been found. Hence, the actual velocity errors are expected to be a factor three to five lower (row C in table 8.3). Similar findings are mentioned by other authors (a factor of two by Virant [1996] and a factor of four by Stüer [1999]). The tests on synthetic images in section 8.2.2 have shown that the 3-D velocity error is almost a factor ten lower than the 3-D position error.

A further check of the velocity accuracy is carried out using a real flow sequence where the flow velocity is practically zero. This sequence has been acquired in the test flume at the BAW. Both the water level and the flow discharge were practically zero, only the sand and gravel layer were submerged in water. Hence, the velocities obtained from this sequence should be close to zero. The mean and standard deviation of the absolute 3-D velocity has been computed for the Kocher and Höllbach measurements, with a resulting mean velocity of $V_{min,exp} = (0.3 \pm 0.2)$ mm/s. This is in good agreement with the accuracy obtained from the above considerations on error propagation.

The maximum velocities that can be measured are limited by the maximum particle image displacements that can be tracked by the 2-D PTV algorithm. The latter depend on the particle density. The main difficulty is the trajectory initialization of fast particles. Initialization of fast particles can only be achieved by the multigrid PIV method, with maximum displacements of about 20 pixels per frame. Note that larger displacements can also be tracked if the corresponding trajectory is already initialized, so that the displacement in the next frame can be predicted by the Kalman filter. The manual inspection of some of the tracking results from the measurements described in chapter 10 has shown that displacements of up to 50 pixels per

frame have been correctly tracked. However, to be on the safe side, we use the maximum displacement (20 pixels per frame) of the PIV method to estimate the maximum detectable velocity of the 3-D PTV. The resulting velocity ranges of the different setups are shown in table 8.3.

The dynamic range DR of the velocity measurement in image space is approximately $1:1000$. This value results from the assumed maximum possible displacement of 20 pixels and the minimum possible displacement. The latter is given by the subpixel accuracy of the particle position determination, which has been shown to be ≈ 0.1 pixels. Assuming a high correlation of the errors of two closely spaced successive particle positions in the image plane (see the discussion above), the resulting accuracy of the displacement is ≈ 0.02 pixels, which yields the dynamic range of $1:1000$. However, some uncertainties are involved in translating this dynamic range in image space to the 3-D object space (calibration errors, varying lateral magnification throughout the depth of the observation volume). Hence, we estimate the dynamic range of the 3-D velocity measurement from the minimum and maximum values obtained in the experiments. The obtained values DR_{exp} are also shown in table 8.3. Except for the Neckar setup, they are lower than the DR in image space, which can basically be traced back to the lower image resolution and accuracy of endoscopic imaging.

8.5   Summary

A performance analysis of the 3-D PTV algorithm in terms of its correctness, accuracy and efficiency has been given. The evaluation has been carried out on synthetic images of a jet shear flow (with known ground truth), on real test image sequences with precisely known 3-D displacements and on real flow sequences obtained in the measurements presented in chapter 10. The results are summarized as follows:

- The two main image parameters that determine the performance are the particle density resp. image source density and the sampling quality. The best results are achieved for particle densities below 0.01 particles per pixel and sampling qualities larger than one. Most of the flow sequences acquired in the experiments comply with these values.

- A total 3-D vector yield of 60% is achieved on the synthetic data. The vector yield on real images is typically 50%, which is a factor two larger than in previous versions of the algorithm. The main limitation is the 2-D particle-tracking (problems related to occlusions and crossing trajectories) and particles that are visible in only one image sequence due to shortcomings of the illumination setup.

- The tests on the synthetic images have shown that the 3-D reconstruction performs very well even for high particle densities. Reconstruction efficiencies of over 95% have been obtained. However, this is only possible if the correct 2-D trajectories are precisely known, without interruptions due to occlusions and overlaps caused by particles crossing each other.

- In the synthetic test case 352 (corresponding to reasonable sampling quality and particle density for PTV), 0.5% of the reconstructed 3-D trajectories were wrong. In this test, the minimum trajectory length has been set to four. The only way to further reduce the probability of wrong correspondences is to increase the minimum required trajectory length.

- For the typical flow conditions investigated in chapter 10 (pore flow velocity larger than $5\ mm/s$), the relative velocity error is well below 10%.

Part III

Experiments and results

Chapter 9

Experimental setup

The following two chapters constitute the third and final part of the thesis, which is devoted to the flow measurements carried out at the BAW in Karlsruhe. This chapter provides a detailed description of the experimental setup that has been designed and implemented. The next chapter describes the extensive systematic measurements that have been performed using the new setup. The obtained results provide new insight into the flow structures within and above a permeable wall.

We start with a description of the experimental flume at the BAW Karlsruhe, where the measurements have been performed (section 9.1). The experimental setup for flow measurement consists of three parts. A separate section is devoted to each part: measurement of sand motion (section 9.2), pore flow (section 9.3) and free surface flow (section 9.4). In section 9.5 we take a brief look at the pressure sensors that have been developed and applied by the Institute for Hydromechanics of Karlsruhe University. In most of the experiments, pressure and flow measurements have been carried out simultaneously. Some aspects of the synchronization and the realtime data storage are discussed in section 9.6. Finally, a summary of the experimental setup is given in section 9.7.

9.1 Experimental flume

All experiments have been carried out in an experimental flume located at the Federal Waterways Engineering and Research Institute (Bundesanstalt für Wasserbau, BAW) in Karlsruhe. All important parameters of this facility are summarized in section 9.1.1. Section 9.1.2 shows the organization of the observation area and introduces the particular subsystems of the experimental setup.

9.1.1 Flume parameters

A sketch of the experimental facility is shown in figure 9.1a. The flume is $L = 40\ m$ long and $B = 0.9\ m$ wide. In the flume, a sand layer of height $H_S = 0.5\ m$ is covered by a permeable gravel layer of height $H_P = 0.04 - 0.2\ m$. The sediment layers start approximately $10\ m$ downstream of the water inlet and have a length of $30\ m$. The mean grain diameter of the gravel is $d_{mD} = 10.2\ mm$, with a degree of non-uniformity of $C_c = d_{60}/d_{10} = 1.25$, where d_{60} is the mean diameter of the grains at 60% sieve fraction. Hence, the gravel material is rather uniform, as compared to natural river beds. Further parameters of the gravel and the sand, as reported by Detert et al. [2004b], are compiled in Table 9.1.

a

b   *c*

Figure 9.1: *a Sketch of the experimental flume at the BAW Karlsruhe (from [Detert et al., 2004b]). All distances are given in m. The sketch is not to scale. b View along the experimental flume, downstream of the observation area. c Movable weir at the flume outlet (top) and wave generator (bottom).*

The flume is connected to the central water supply system for all the experimental models at the BAW. Water is delivered from large reservoir tanks located on the roof of a dedicated building, at a maximum flow rate of $0.275 \ m^3/s$. At the water inlet of the flume, the flow discharge is measured by an inductive flow-meter. Wave dampers are used to reduce fluctuations of the water level upstream of the observation area. The water level in the flume is adjusted by a movable weir at the water outlet, see figure 9.1b. The

Table 9.1: *Physical parameters of the experimental flume: dimensions, fluid properties and hyraulic parameters, parameters of the sediment layers.*

parameter	symbol	value
flume length	L	$40\ m$
flume width	B	$0.9\ m$
height of sand layer	H_S	$0.5\ m$
height of gravel layer	H_P	$0.04\ m, 0.1\ m, 0.2\ m$
gravitational acceleration	g	$10\ m/s^2$
density of water	ρ	$1000\ kg/m^3$
specific weight of water	$\gamma_w = g\rho$	$10^4\ N/m^3$
dynamic viscosity	μ	$10^{-3}\ Ns/m^2$
kinematic viscosity	$\nu = \mu/\rho$	$10^{-6}\ m^2/s$
maximum flow discharge	Q_{max}	$0.275\ m^3/s$
water level	h	$0.1 - 0.4\ m$
maximum flow velocity	v_{max}	$\approx 1\ m/s$
mean grain diameter of gravel	d_{mD}	$10.2\ mm$
mean grain diameter of sand	d_s	$1\ mm$
absolute density of gravel	ρ_{abs}	$2.464\ g/cm^3$
bulk density of gravel layer	ρ_{bulk}	$1.538\ g/cm^3$
gravel porosity (void fraction)	$\phi = 1 - \rho_{bulk}/\rho_{abs}$	0.38
solid fraction of gravel	ρ_{bulk}/ρ_{abs}	0.62
hydraulic conductivity of gravel	k_f	$\approx 0.9\ m/s$
geometric permeability of gravel	$K = (\mu/\gamma_w)k_f$	$\approx 10^{-7}\ m^2$
critical shear stress of gravel	τ_{0c}	$8.8\ N/m^2$

water level is controlled by an ultrasonic gauge and two mechanical pointers upstream and downstream of the observation area. At the outlet of the flume, the water is collected in a large tank in the basement of the building before it is pumped back to the reservoirs. Additional monitoring of the mean velocity profiles in the channel is carried out using a 1D Acoustic Doppler Current Profiler (ADCP), see [Detert et al., 2005]. Steady flow conditions for each experiment are provided by controlling the gate valve at the inlet and the movable weir at the outlet automatically.

To enable the investigation of unsteady flow conditions, the flume is equipped with a **wave generator** $9.25\ m$ downstream of the observation area (figure 9.1c). Waves are created by periodically dipping a displacement body into the water. In this way, surface waves with amplitudes of up to $15\ cm$ and periods between one and three seconds can be generated. During the wave experiments, the water level is recorded by an ultrasonic wave gauge.

At the observation area, optical access to the flume is given by glass windows inserted into the channel walls on a length of $4\ m$. Further it is possible to access the obervation area from a pit below the flume to insert experimental equipment through the bottom of the flume.
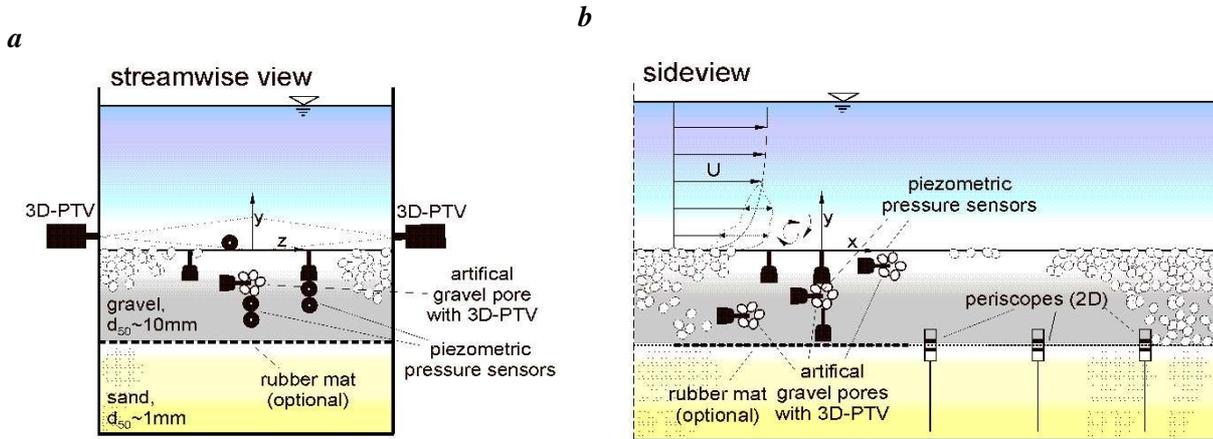
Figure 9.2: *Streamwise view (**a**) and sideview (**b**) of the obervation area (from [Detert et al., 2004b]). The experimental setup consists of several subsystems: 3-D PTV to observe the free surface flow near the gravel bed (indicated by the cameras at the flume walls), 3-D PTV to observe the pore flow in three 'artifical' gravel pores, observation of the sand-gravel-interface by three rigid endoscopes ('periscopes') and several piezometric pressure sensors at different locations.*

Table 9.2: *Positions of the setups for image acquisition. The coordinates given in the second and third columns indicate the center of the observation volume. The spanwise position of all setups is at the centerline of the flume ($z = 0$). Max($\Delta x/y/z$) is the maximum inner diameter of the artificial pores resp. the maximum extension of the observation volume in the corresponding directions.*

setup	x	y	Max(Δx)	Max(Δy)	Max(Δz)
	[cm]	[cm]	[mm]	[mm]	[mm]
Jagst	-18	-18/-6.5/-3	13	13	18
Kocher	1.5	-1.5/-5.5/-2	13	11	10
Höllbach	26	-1	12	11	12
Neckar	46	2.5	100	50	100
	26	2.5	100	50	100
Elbe-u	80	-10/-4	15	15	1
Elbe-m	130	-10/-4	15	15	1
Elbe-d	180	-10/-4	15	15	1

9.1.2   Observation area

The observation area is located in the middle of the flume. Influences of the water inlet and outlet are considered negligible, and a fully developed turbulent wall boundary layer is assumed in the observation area. Different experimental instrumentation is installed, compare the sketches in figure 9.2. The data acquisition is carried out simultaneously by all these subsystems, which are described in more detail in the following sections. The synchronization is explained in section 9.6. For an easy identification of the subsystems, we refer to them by the name of the corresponding workstation that has been used for the data

acquisition[1]. In particular, the setup consists of the following parts:

- a stereo camera setup for the investigation of the free surface flow directly above the gravel layer ('**Neckar**'),

- three endoscopic stereo setups to record the flow fields inside specially prepared 'artificial' gravel pores within the gravel layer: '**Jagst**' (upstream part of the observation area), '**Kocher**' (middle part of the observation area) and '**Höllbach**' (downstream part of the observation area),

- three rigid endoscopes inserted from below the channel to observe the sand-gravel-interface: '**Elbe-d**' (downstream), '**Elbe-m**' (middle) and '**Elbe-u**' (upstream),

- up to ten pressure sensors[2] at arbitrary locations within the gravel layer (three of the sensors are attached to the artificial gravel pores used for the flow measurement).

To indicate the locations of the different subsystems in the flume and relative to each other, we define the following right-handed coordinate system:

- x: coordinate in streamwise direction, increasing in flow direction, origin $1.5\ cm$ upstream of the middle artificial pore ('Kocher'),

- y: vertical coordinate, origin at the top of the gravel layer,

- z: spanwise coordinate, origin at the centerline of the flume, increasing to the right when viewing downstream.

The locations of the observation points, defined in this coordinate system, are compiled in table 9.2. In all experiments, the spanwise position of the artificial pores is at the centerline of the flume ($z = 0$). The table also shows the size of each obervation volume. Note that the pore volume of the artificial pores extends approximately $10\ mm$ in each direction. The stereoscopic observation volume of the endoscopes, however, is smaller, see figure 9.7. The different y-values in table 9.2 correspond to the three different heights of the gravel layer that have been investigated.

9.2 Measurement of sand motion

Subsoil instabilities may occur at the interface between the gravel filter and the base layer of fine sand, depending on the hydromechanic load conditions. This section presents the experimental setup used to detect and quantify the motion of sand grains within a 2-D cross section perpendicular to the horizontal interface layer. Both the experimental approach and the image processing techniques (see section 3.4) have been successfully applied before in a number of projects to study hydromechanic motion in submerged subsoil [Spies et al., 1999, 2000; Köhler and Koenders, 2003]. The setup presented here is an improved version of the setup used in these previous studies. The improved version provides a larger field of view and a better illumination. It has also been used in additional experiments in a wave-experiment test pit at the BAW [Jehle et al., 2004; Klar et al., in preparation].

---

[1]All the names are actually names of German rivers or streams, which are however not directly related to the measurements performed in this thesis.

[2]The pressure measurements are carried out by the Institute for Hydromechanics of the University of Karlsruhe (e.g. [Detert et al., 2004a, 2005]) and are not the subject of this thesis.
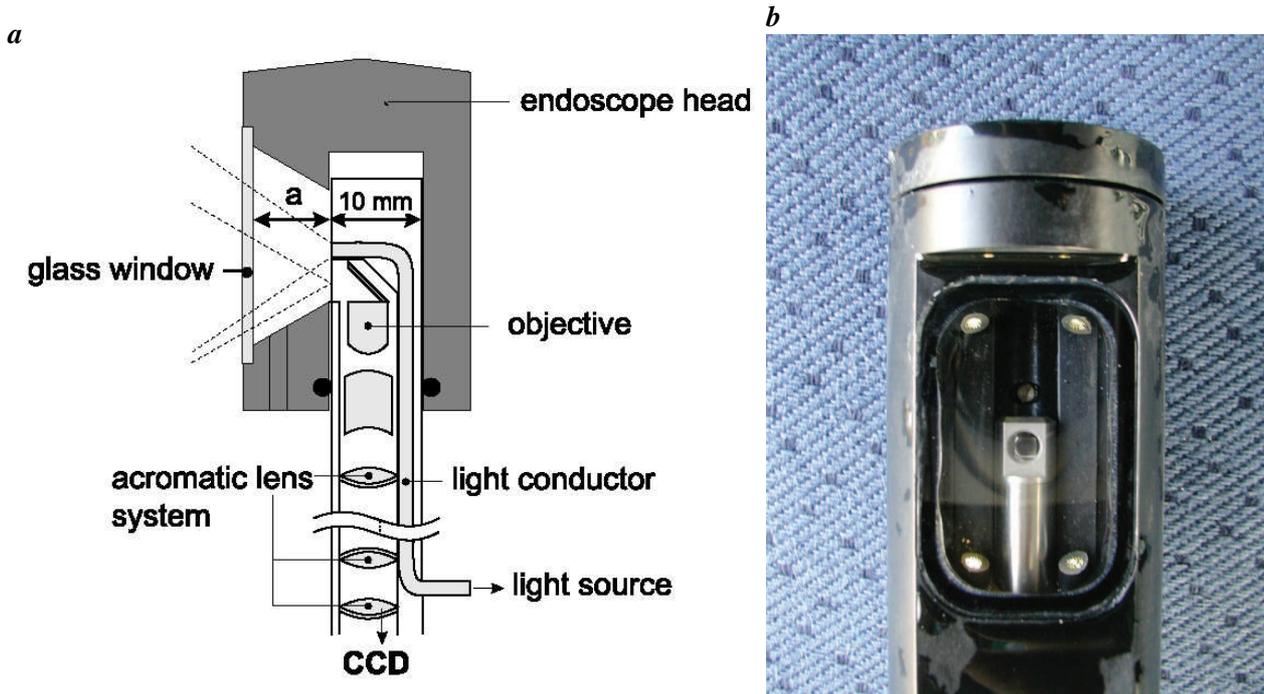
Figure 9.3: *a* *Sketch of a rigid endoscope. The observation direction is at right angles to the endoscope axis, hence the term 'periscope' is also used.* ***b*** *Endoscope protection head with illumination fibers in the corners of the observation window.*

9.2.1   Rigid endoscopes ('periscopes')

Sediment movement is detected by three **rigid endoscopes** (figure 9.3), inserted into the flume from below (figure 9.4). They are 110 cm long and have a diameter of 1 cm. The viewing direction of the endoscopes is perpendicular to the endoscope axis, hence, they are also referred to as '**periscopes**'. Within the periscope, the image is transmitted through a rod lens system which consists of several tens of lens components. With an aperture angle of 90 degrees and a viewing distance of approximately 7 mm, a circular area of about 15 mm in diameter can be observed. The wide-angle lens introduces a significant amount of radial distortion to the images, see section 5.4.1. CCD cameras attached to the periscopes enable the recording of image sequences of the sand-gravel boundary, which can be analyzed by the optical flow techniques described in section 3.4. The result is a time-resolved 2-D/2-C vector field within the observation plane.

The requirements of the endoscopic imaging are the following:

- The observation area should be as large as possible.

- A working distance of 7 mm has to be kept.

- Sufficient and homogeneous illumination should be provided.

- The sensitive optical setup has to be protected in the rough environment of the flume (sand, gravel, water).

In order to meet these demands, special protection heads with a glass window have been designed (figure 9.3b, figure 9.4b). To prevent the sediment from blocking the endoscope view and to keep the

Figure 9.4: a *Three rigid endoscopes used to observe the sand-gravel boundary ('periscopes'). The periscopes are 110 cm long and have a diameter of 1 cm.* **b** *One of the periscopes inserted into the sand layer from below the flume. For protection purposes, the periscope is covered by a special illumination head and a waterproof steel tube. The direction of observation is at right angles to the periscope axis.* **c** *Top view of the observation area during the installation of the three periscopes.* **d** *View from below the flume. A CCD camera and an illumination fiber is attached to each periscope.*

working distance of $7\ mm$, the endoscopes are inserted into these protection heads. Further protection of the endoscopes is achieved by covering them with waterproof steel tubes.

9.2.2 Illumination and flow visualization

The illumination system has been integrated into the protection heads. The light from an external **halogen cold light source** is coupled to the endoscope heads via a glass fiber light conductor, which separates into four fibers within the head. The sediment is illuminated by these four fibers at the corners of the glass window, see figure 9.3b. In this way, a small angle of incidence of the light is achieved, avoiding specular reflections on the glass window, which have drastically reduced the image quality in previous applications [Spies et al., 1999]. The new setup produces a relatively homogeneous illumination without any over- or underexposed regions.

No tracer particles are added to the flow at the sand-gravel-boundary. Hence, the water flow is not visible in the image sequences. Only motion of sand or gravel grains or motion of gas bubbles in the subsoil will be detected.

9.2.3   Image acquisition

Standard **CCD-cameras** (SONY XC-73CE) have been used for image acquisition. The analog images delivered by the cameras are digitized by PCI-bus-framegrabbers. Two of the cameras are connected to an ELTEC PCEye4 stereo framegrabber with two 8 bit input channels, which resides in a standard desktop PC. The third camera is connected to one of the color channels of an ELTEC PCEye2plus color framegrabber. The two remaining channels of the latter are used to acquire the images of one of the artifical gravel pores.

Image sequences are acquired at a resolution of 512×512 pixels and a frame rate of $25\ Hz$ (full frames). The data thruput is approximately $20\ MB/s$ for all three periscopes. The data can be written to a RAID system in real-time during the acquisition. The radiometric properties of the cameras are analyzed in section 4.2, the geometric calibration is discussed in section 5.4.1.

All image acquisition systems are running on standard desktop PCs under WindowsNT or Windows2000.

9.3   Measurement of pore flow

Flow measurements inside three single pores of the gravel layer are carried out using three **miniaturized endoscope stereo setups**. The basic principle of these setups is to acquire stereoscopic image sequences of the flow field inside the pore volume by viewing it from two different directions. Towards this end, two flexible fiberoptic endoscopes of 2.4 mm diameter (section 9.3.1) are attached to an adapted artificial gravel pore made of grains fixed to each other (section 9.3.2). The artificial pores are installed in the gravel layer, illuminated by a further fiber bundle connected to a cold light source (section 9.3.3), and image sequences are acquired and stored in real-time (section 9.3.4).

Two of the pore flow setups ('Höllbach' and 'Kocher') have been re-designed using superior equipment, on the basis of past experience by Stybalkowski [2001] and Klar et al. [2002]. The third setup ('Jagst') is a replication of the setup applied by the latter, using the same hardware.

9.3.1   Flexible endoscopes

Optical access to the flow within the gravel layer is given by **flexible fiberoptic endoscopes**[3], see figure 9.5. Endoscopes of this type are typically used for quality assurance in technical applications or for medical examinations. The image is transmitted by 30000 glass fibers which are arranged in a fixed, regular array (a so-called *coherent fiber bundle*). Inside every fiber, the light propagates forward by total reflection at the fiber cladding. The fibers have a diameter of $\approx 10\ \mu m$, which is very close to the pixel size of the cameras (see section 9.3.4). At the endoscope entrance, an objective lens system produces an image of the observed object and projects it onto one end of the fiber bundle. At the endoscope exit at the other end of the fiber bundle, a zoom lens system projects the transmitted image onto the sensor of a camera. The zoom is used to adjust the number of sensor pixels covered by the image to the resolution of the fiber bundle.

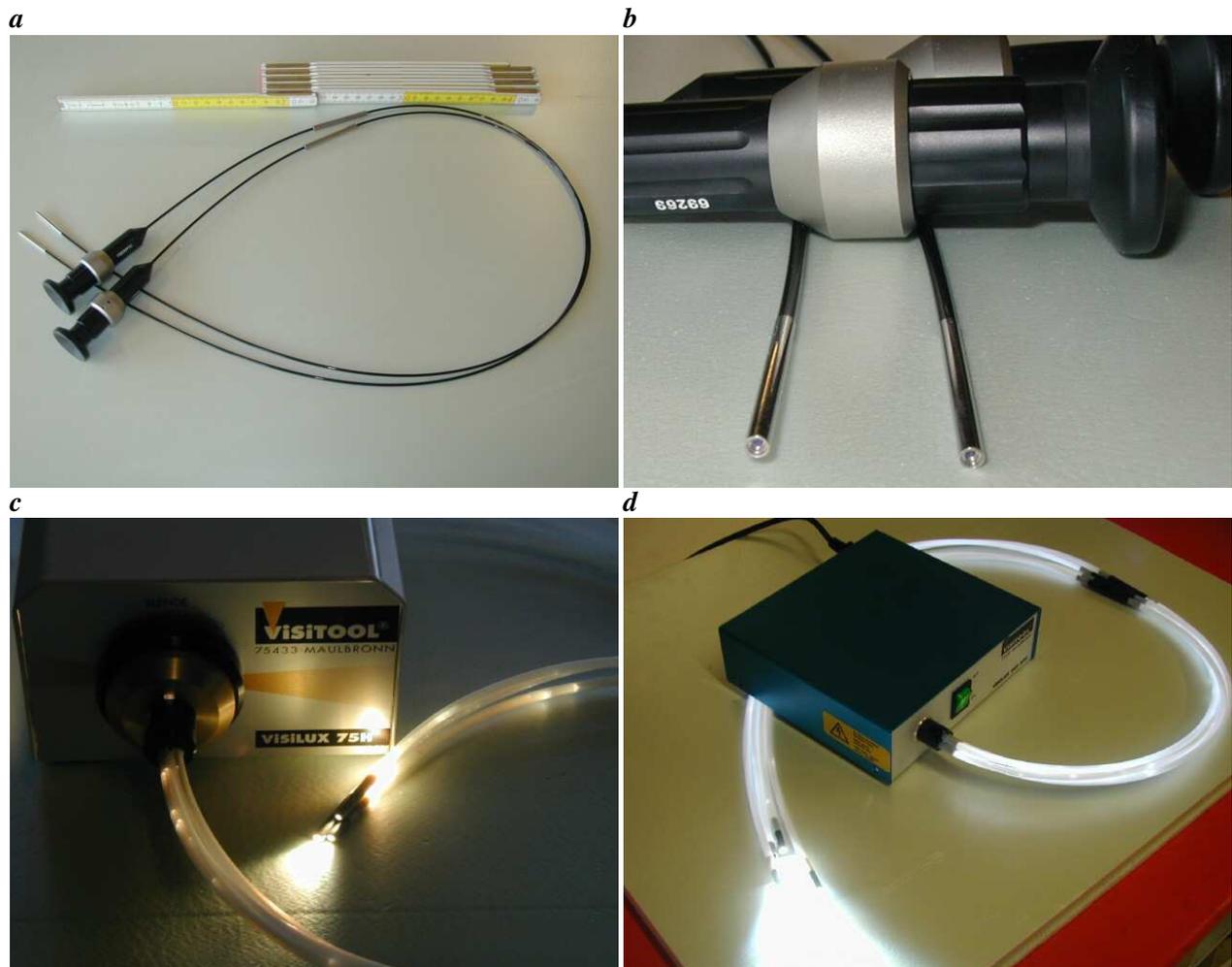[3]manufactured by the Eltrotec Sensor GmbH, `http://www.eltrotec.com`

Figure 9.5: **a** *Flexible fiberoptic endoscopes used to acquire image sequences from within the gravel layer. The endoscopes are* 80 *cm long and have a diameter of* 2.4 *mm. The image is transmitted through a regular array of* 30000 *single fibers.* **b** *Enlarged view of the endoscope tips with the objective lenses (bottom of image) and the rear ends of the endoscopes with the zoom optics (top of image).* **c** *Halogen cold light source (color temperature* 3200 *K).* **d** *Short arc lamp (color temperature* 5600 *K).*

9.3.2 Artificial gravel pores

The core of the experimental setup for flow measurement within the gravel layer is a so-called **artificial gravel pore**, made of pebbles fixed to each other, see figure 9.6. The purpose of the artificial pore is to hold the endoscopes and the illumination fiber at a fixed relative position and to keep surrounding grains in the gravel layer from blocking the endoscopes' views. The endoscopes are fixed in a stereo rig consisting of two PVC spheres with drilled holes holding the endoscopes at a fixed position relative to each other. This stereo rig in turn is attached to the artificial pore. Before and after the flow measurements, it can be taken off for calibration purposes without changing the relative position of the endoscopes. The observation area is defined by the interior volume of the artificial pore. The average linear dimension of this volume is approximately 1 cm, see table 9.2. It is much smaller than the volume of the pore used by Stybalkowski [2001] and Klar et al. [2002] (which is ≈ 8 cm^3, see figure 7.2a), but still larger than the natural pore spaces within the gravel layer. Further size reduction is not possible without major modifications of the
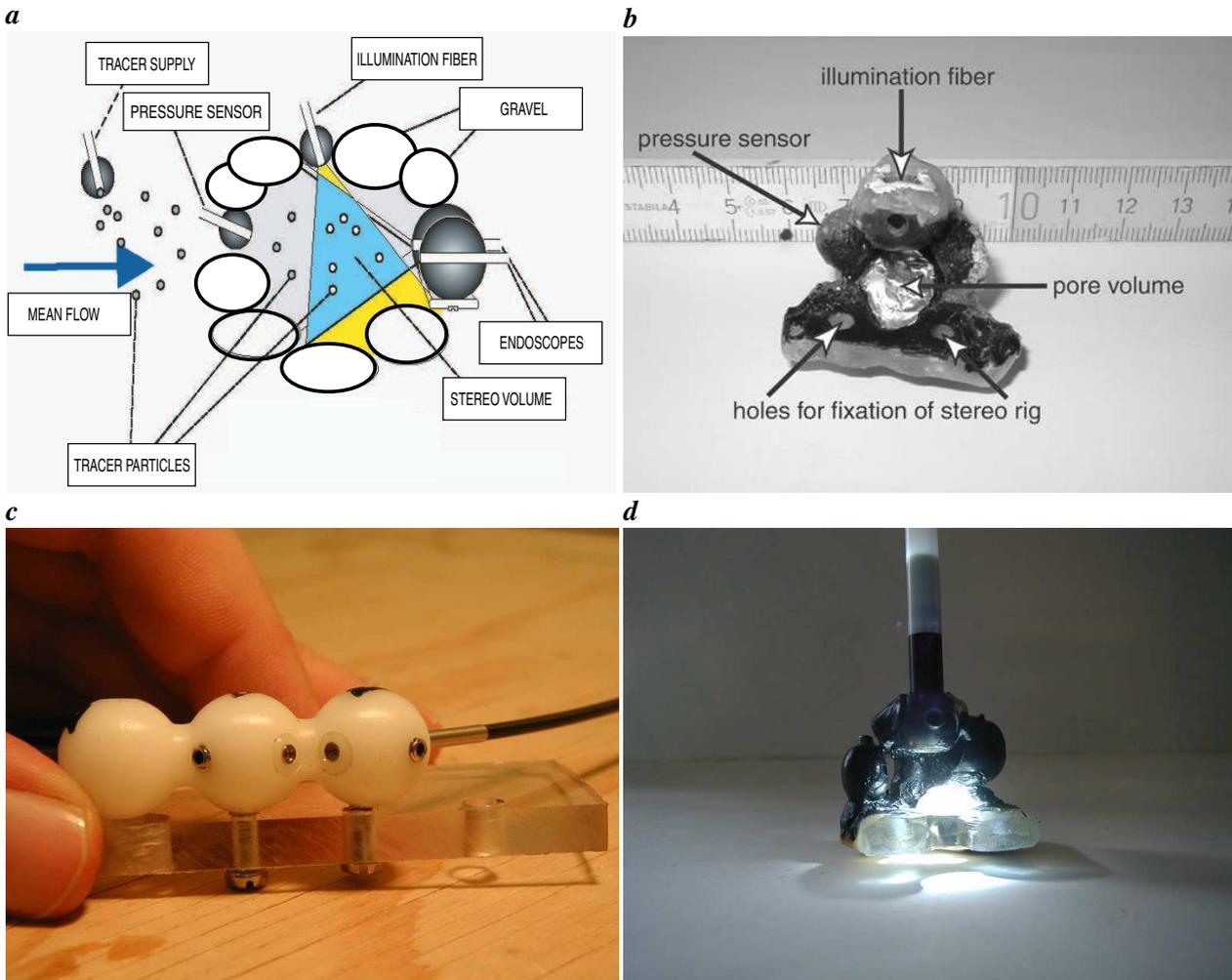
Figure 9.6: **a** *Sketch of an artificial gravel pore.* **b** *Front view of an artificial gravel pore. The endoscope stereo rig has been removed.* **c** *Endoscope stereo rig. To perform velocity measurements in the gravel layer, this stereo rig is attached to the artificial pore, viewing the pore volume inside.* **d** *Artificial pore with illumination fiber.*

hardware (smaller endoscopes and illumination fibers, higher frame rates). Thus, an influence of the larger pore volume has to be considered in the analysis of the measurement results.

The size reduction of the artificial pores was possible due to a change of the convergence angle of the stereoscopic geometry from $\alpha = 60°$ (as used by Stybalkowski [2001] and Klar et al. [2002]) to $\alpha = 90°$. Figure 9.7 shows that the larger convergence angle reduces the size of the stereo volume and moves it closer to the objective lenses of the endoscopes. Since tracer particles may cover such a small volume very rapidly, cameras with high frame rates are necessary, see section 9.3.4. Otherwise, the displacements of the particle images between two successive frames become too large. Further, the depth of field of the endoscopes has to start very close to the objective lenses. The endoscopes shown in figure 9.5 have been specially adapted to meet this demand. The depth of field of the previous endoscopes used by Klar et al. [2002] starts approximately $5 \, mm$ in front of the objective lenses, which is one of the reasons why a larger pore volume with a smaller convergence angle was necessary.

To perform flow measurements, the three artificial pores are embedded in the gravel layer at different

α = 90˚,
z_0 = 1.5 mm,
z_1 = 6.5 mm,
Δx = 4 mm @ z = 2.25 mm

α = 60˚,
z_0 = 2.5 mm,
z_1 = 18 mm,
Δx = 4.75 mm @ z = 4.5 mm
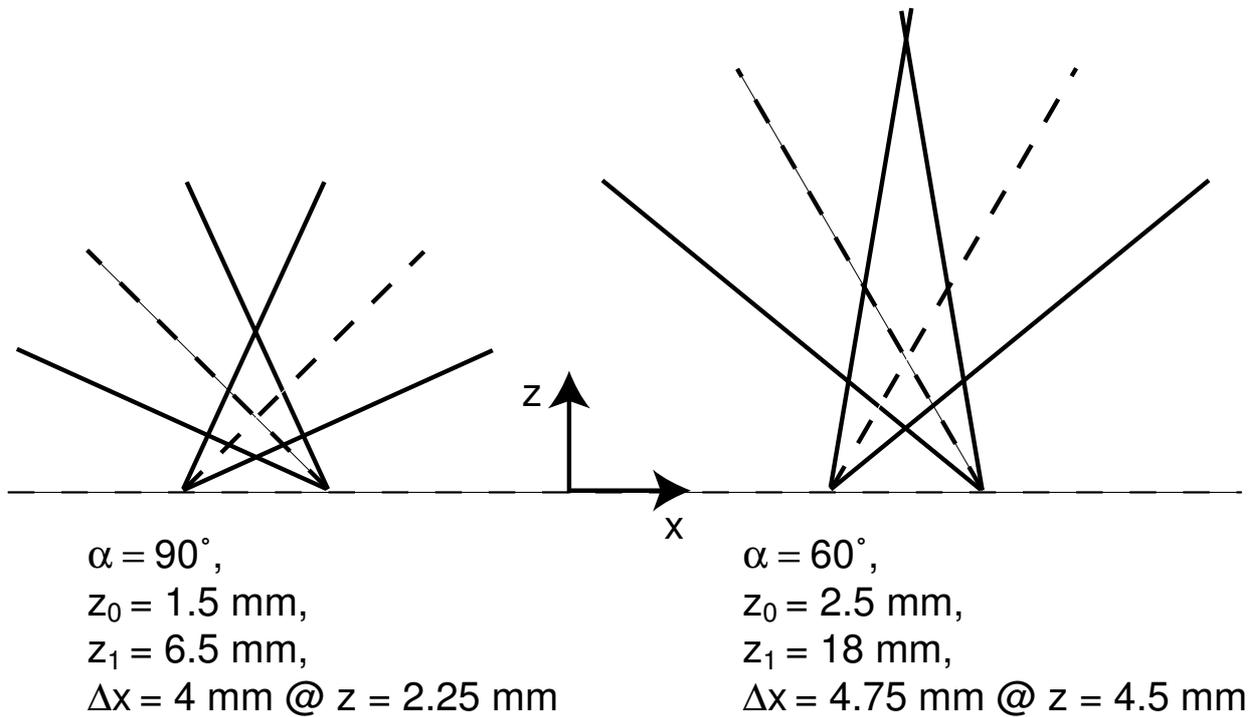
Figure 9.7: *Top-view of a horizontal cross section of the stereoscopic observation volumes for two different conver-gence angles: $\alpha = 90°$. (left) and $\alpha = 60°$ (right). In both cases, the stereo baseline has a length of 6 mm and the angle of view of the endoscopes (in water) is $\theta = 40°$. z_0 is the point of the observation volume that is closest to the endoscopes, z_1 is the farthest point (along the axis of symmetry of the stereo rig). Δx is the maximum extension of the stereo volume in the x-direction perpendicular to the axis of symmetry. The geometry with $\alpha = 90°$ is implemented in the 'Kocher'- and 'Höllbach'-setups, in the 'Jagst'-setup the convergence angle is $\alpha = 60°$.*

positions, see figure 9.8. A grid of steel rods fixed to the flume enables the precise determination of the pore positions. The pressure sensors are also fixed to this grid. A solution of tracer particles is added locally to the flow upstream of the pores, and particle image sequences of the two different endoscope views are recorded simultaneously.

9.3.3 Illumination and flow visualization

Illumination of the pore volume is provided by an optical fiber-bundle guiding the light from a cold light source into the pore. Achieving a sufficient illumination for endoscopic flow measurements is difficult for two reasons:

1. Intensity losses due to the transmission of the images through the fibers reduce the brightness of the images.

2. Due to the high frame rates of the cameras ($400\ Hz$), the exposure time is only $1\ ms$.

Janßen [2000] implemented the first prototype of an artificial gravel pore for stereo imaging. He used a halogen cold light source and standard CCD cameras acquiring full frames at $25\ Hz$. Hence, the exposure
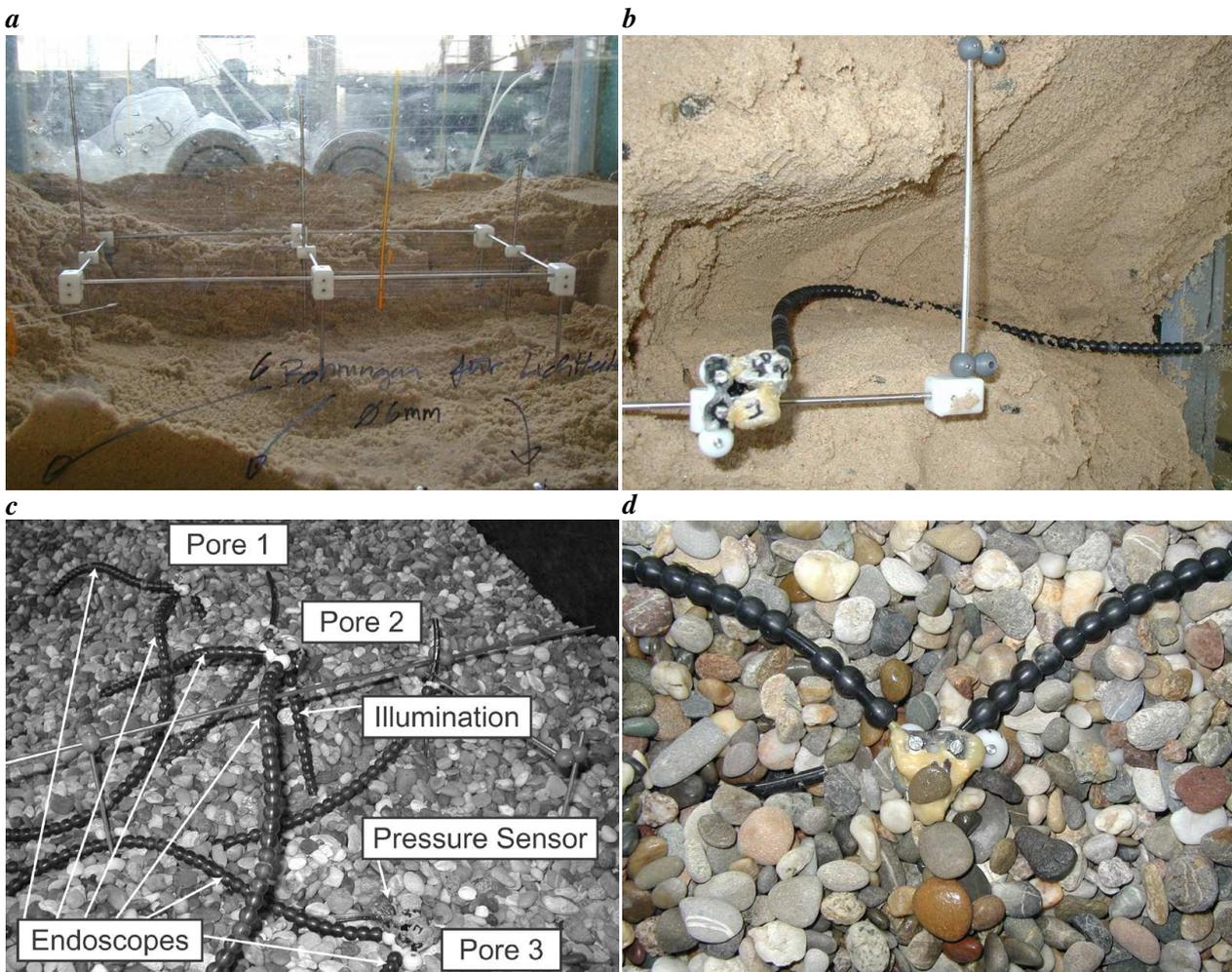
Figure 9.8: *a A grid of steel rods is fixed in the flume. All artificial pores and pressure sensors are attached to this grid, so that the relative position of the measurement equipment is exactly known. **b** Top view of an artificial pore during installation in the flume. The pores are mounted upside down, so that the illumination fiber is laid in the sand layer and enters the pore volume from below. **c** Installation of the three artificial pores in the gravel layer. **d** Top view of one of the pores. For protection purposes, the endoscopes are covered by flexible tubes. Plastic beads of 1 cm diameter are threaded along these tubes to make the geometry more similar to gravel.*

time was much higher. Nevertheless, he reports severe illumination problems. Due to the very low signal-to-noise ratio in his images, he was only able to apply a tracking algorithm after a computationally very intensive denoising procedure (anisotropic diffusion [Spies and Scharr, 2001]). Klar et al. [2002] used the same halogen cold light source to illuminate the pore volume. To achieve a higher signal-to-noise ratio and avoid the denoising, they used an additional illumination fiber (see figure 7.2), i.e. the pore volume was illuminated by two fibers, aligned with the two optical axes of the endoscopes. Since in the experiments carried out in this work, three artifical pores are used simultaneously, the complexity of the experimental setup had to be reduced as far as possible. Attaching two illumination fibers to each pore was not feasible due to spatial limitations in the measurement area and the very laborious preparation of the experiments. Thus, a more powerful illumination setup has been designed[4].

---

[4]manufactured by the ViSiTool GmbH, `http://www.visitool.de`

Improvements have been achieved in two ways. First, the active diameter of the illumination fibers has been increased, without increasing the outer diameter of the bundle. This was possible by applying a different protection coating. Second, the halogen light source has been replaced by a **short arc lamp** with an elliptical reflector. The color temperature of the short arc lamp is $\approx 5600\ K$, which is almost twice as large as that of the halogen lamp ($\approx 3200\ K$), see figure 9.5c,d. As a consequence, the maximum emission of the halogen lamp is at $\lambda \approx 900\ nm$, whereas that of the short arc lamp is at $\approx 500\ nm$, which is much closer to the maximum spectral response of the CMOS sensors at $\approx 600\ nm$ [Photonfocus, 2003].

The flow is visualized by adding a solution of tracer particles to the pore flow upstream of each artificial pore. Hence, the flow was only seeded *locally*. With the local seeding, mixing of the highly concentrated tracer suspension with the water in the flume obviously starts not before the suspension leaves the supply tube some grain diameters upstream of the artificial pore. Additional stirring is not possible, since the tracer injection has to be done without significantly disturbing the pore flow. The only mixing mechanisms are those of the pore flow itself, which depend on the flow conditions and the local pore geometry of the gravel layer. Therefore it was difficult to achieve a homogeneous tracer density in some of the experiments. However, it was not possible to seed the whole water volume to achieve a homogeneous distribution of tracer particles because of its very large size and the connection of several different experiments to the same water supply.

Concerning the tracer particles, there are two basic requirements. The first is that they should closely follow the fluid flow, since it is the particles' velocity which is actually measured and assumed to be equal to the fluid velocity. Thus, the density of the particles should be as close as possible to the density of the fluid. The second requirement is that the particles should have a large light scattering cross section in order to get a sufficient image contrast. The two requirements are conflicting, since the first favours small particles while the second favours larger ones. The tracer particles used to visualize the pore flow are polystyrene particles (Optimage), with a mean diameter of $30\ \mu m \pm 20\%$, a density of $1.020\ g/cm^3$, and a refraction index of 1.602. They have been specially designed for the use as PIV tracers and used in many previous studies. For more information on the particles, e.g. an investigation of the scattering cross section, see [Leue et al., 1996; Hering, 1996; Garbe, 1998; Engelmann, 2000]. The theory of light scattering at small particles is developed in [van de Hulst, 1981], and an analysis of the motion of small particles in a turbulent flow field can be found in [Hinze, 1975].

9.3.4 Image acquisition

As mentioned in section 9.3.2, high frame rates are necessary to keep the displacements of particle images between two successive frames within reasonable limits. Hence, two of the three artificial pores ('Kocher' and 'Höllbach') are equipped with **digital high-speed Megapixel-CMOS-cameras** (Photonfocus MV-D1024-28). These cameras were among the first of this type that were commercially available. The new CMOS sensor technology has several advantages over the more conventional CCD technology, e.g. blooming resistance, adjustable characteristics, selectable read-out windows (Region Of Interest, ROI), global shutter permitting high-speed applications and low energy consumption at high data rates. The versatility of the MV-D1024-28 allows for the optimization of the camera settings for a given image processing application by the user. The camera settings are stored in an EEPROM and are loaded automatically at bootup. All parameters can be set and stored in the EEPROM by using a command line tool [Photonfocus, 2003].
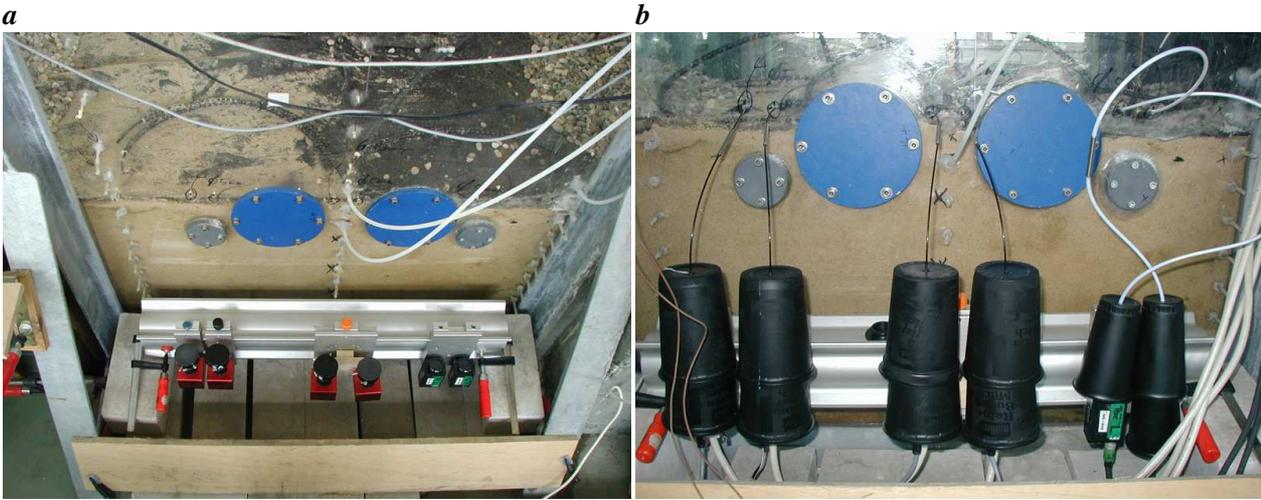
a   *b*

Figure 9.9: *a* *Installation of the three stereo camera setups for the artificial pores at the flume.* ***b*** *The endoscopes are attached to the cameras and inserted into the flume wall. To protect the cameras from leaking water, they are covered by water-proof cases.*

The selection of an arbitrarily sized ROI is of particular importance for the present application. Since only a ROI of size 184×184 pixels (which approximately corresponds to the resolution of the fiber bundles in the endoscopes) is read out, it is possible to increase the frame rate to $400\ Hz$. It is only this high frame rate that makes endoscopic particle tracking feasible at high Reynolds numbers in the open-channel flow.

The MV-D1024-28 camera provides several possibilities for amplifying the video signal on the sensor before A/D conversion and thus increasing the sensitivity of the camera. For applications having short exposure times and low illumination intensity (like the present one), the CMOS sensor can be operated in skim mode, in which a nonlinear amplification, similar to a gamma correction, occurs in the pixel so that small signals are amplified significantly more than large signals. Preceding tests have shown that this mode yields particle images that can be segmented much easier and much more reliably than those acquired without the skim mode. Hence, the skim mode has been used in the measurements. The disadvantage of this mode is the nonlinearity, which deteriorates the subpixel-accurate determination of the particle positions for asymmetric gray value distributions. However, priority has been given to a robust and reliable segmentation. A further effect related to the skim mode is image lag in areas of high contrast, resulting in smearing of the image. In the case of particle sequences, the image lag introduces motion blur, i.e. a blurring of the particle images along their direction of motion. Motion blur also makes the exact determination of the particle position more difficult. On the other hand, it introduces a spatial overlap of successive images of fast particles, which can be used to guide the tracking, see section 7.4.

Among the many benefits of the CMOS technology, there is also one major drawback: CMOS cameras typically show a much larger fixed pattern noise than CCD cameras. A detailed investigation of the radiometric properties and the noise structure of the CMOS cameras is carried out in section 4.1.

The CMOS cameras have a digital output and are connected to Silicon Software microenable high speed framegrabbers via the digital CameraLink interface. The framegrabbers create the timing signals to achieve a pixel-synchronous acquisition of both cameras.

The third artificial pore ('Jagst') is working with standard CCD cameras (Sony XC-73CE, see section 9.2.3) running at $25\ Hz$. To increase the frame rate to $50\ Hz$, the interlaced images are separated into

the even and odd fields and the missing lines are interpolated. The separation into the image fields also reduces blur of fast moving objects due to the time delay between the acquisition of even and odd fields. On the other hand, it deteriorates the vertical resolution due to the necessary interpolation. However, in the present application, the resolution is limited by the number of fibers in the endoscope (≈ 10000), which is much smaller than the number of pixels on the sensor, that are covered by the endoscope image (≈ 200000). No zoom optics were available in the Jagst-setup to reduce the image size on the sensor. Hence, the images in this setup are oversampled and therefore reduced to the first level of a Gaussian image pyramid before processing, see section 7.3.1. Due to the low frame rate of $50\ Hz$, the Jagst-setup can only be used in the lowermost position within the gravel layer, where the flow velocities are expected to be lowest.

The image data of the three setups are written to three RAID arrays (RAID level 0, providing the fastest access rates) in real-time during the acquisition, see section 9.6. Thus, the duration of the sequence acquisition is only limited by the RAID capacity. For a single measurement, a sequence duration of $60\ s$ has been chosen.

The installation of the stereo camera setups at the flume is shown in figure 9.9. To protect the cameras from dirt and water, they are covered with waterproof cases.

Before conducting flow measurements, a geometric camera calibration has to be carried out. The calibration of the endoscopes is described in section 5.4.2.

9.4 Measurement of free surface flow

To study the hydromechanic interaction of the free surface flow and the subsurface pore flow, it is necessary to measure the flow field in the near-wall region of the open-channel flow. Towards this end, a further 3-D PTV setup is used, which is described in this section. This setup has been designed, implemented and applied in preliminary experiments by Klar [2001]. Some modifications were necessary, since Klar [2001] reports the following shortcomings of his setup:

- The density of the tracer particles (air bubbles) in the flume is too high, especially for high Reynolds numbers. The reason is the production of air bubbles at the water inlet. The density and size distribution of the bubbles depends on the flow discharge and cannot be controlled independently.

- Due to strong light scattering at the gravel layer, measurements are only possible up to a minimal distance to the gravel layer of about $0.5\ cm$.

- A frame rate of $60\ Hz$ is too low to study large Reynolds number flows ($Re > \approx 50000$).

- More sophisticated segmentation and tracking techniques are necessary.

Concerning the last two points, the frame rate has been increased to $130\ Hz$ (see section 9.4.3) and the improved algorithms are described in detail in chapter 7. Concerning the first two points, to prevent the production of air bubbles, the flume inlet has been re-designed. But also with the new inlet, it turned out that the amount of air bubbles in the water is still very high, especially at flow rates approaching the low mobility conditions. Further, it was not possible to reduce the light scattering at the gravel surface, see e.g. figure 10.4. Hence, the experimental conditions for the measurement of the near-wall channel flow are still very difficult, and optimal performance cannot be expected at high Reynolds numbers. However, no further measures could be taken to improve the setup since the focus of this work was on the pore flow.
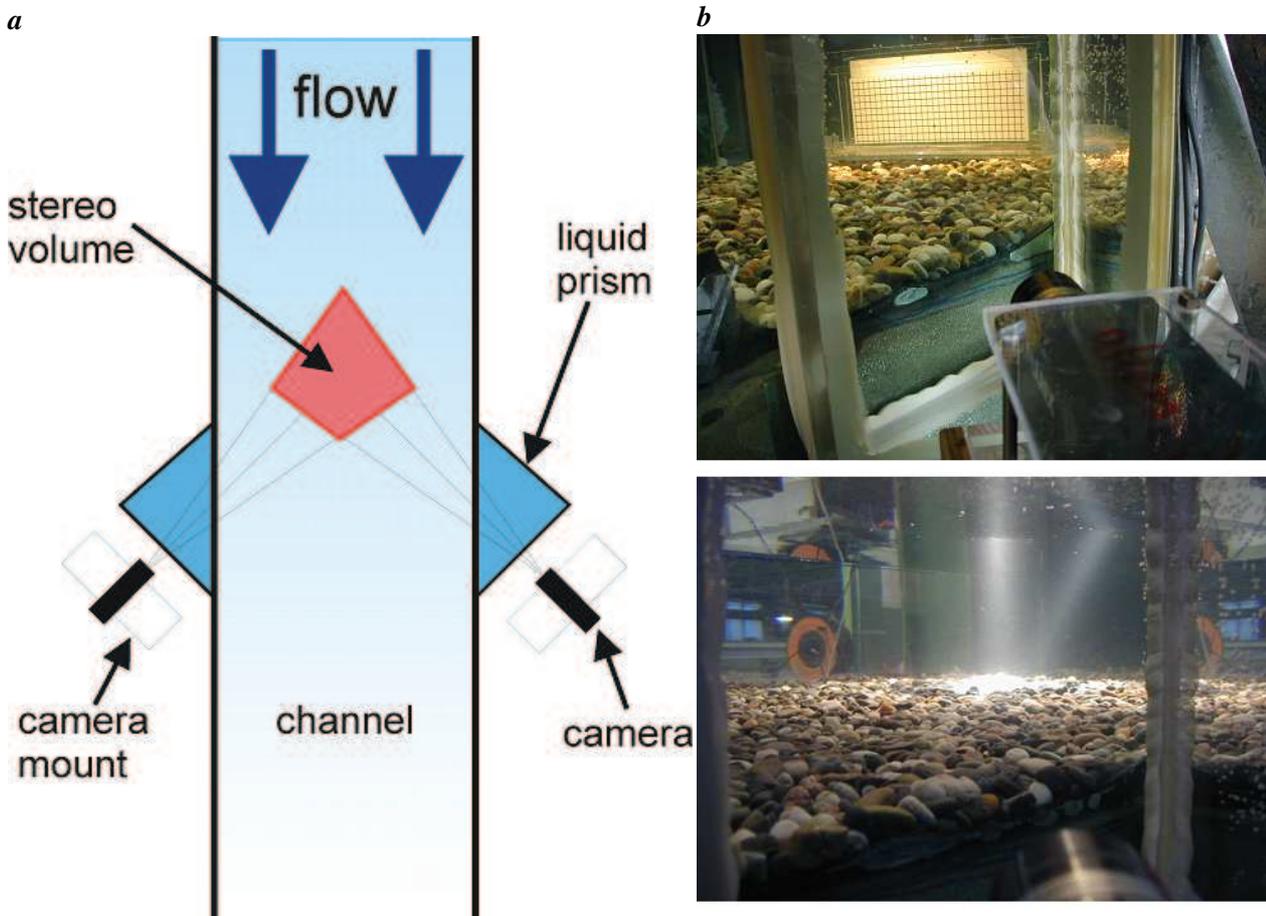
Figure 9.10: a Top view of the stereo camera setup for 3-D PTV above the gravel layer, with a convergence angle of $\alpha = 90°$. The cameras view the channel in upstream direction through water-filled glass prisms attached to the channel side walls. The purpose of the prisms is to reduce refraction and dispersion effects. b Top: View of the right camera through the liquid prism into the flume. The photo was taken during the calibration of the setup, a calibration grid has been placed in the observation volume. Bottom: Same view during flow measurement. The light cones of the halogen lamps illuminating the observation volume from above are inclined towards the cameras to achieve a higher signal-to-noise ratio.

An important advantage of the free surface flow setup is its simplicity as compared to a typical setup for stereoscopic PIV or a 3-component LDA system. The major benefit is that no laser is used. Hence, laser safety issues do not have to be considered. Furthermore, no special high-precision positioning devices or similar expensive optical and mechanical components are necessary, as typically used in laser setups. Most of the components of the setup consist of standard off-the-shelf equipment, which is also very cost efficient.

9.4.1   Optical setup

In order to acquire stereoscopic image sequences of the free surface flow, two CCD cameras (Pulnix TM6701AN) are mounted on the left and right side of the flume, viewing the center of the channel in upstream direction through the glass windows (figure 9.10). The optical axes of the two cameras enclose an angle of $\alpha = 90°$. The resulting size of the stereoscopic observation volume is approximately $5 \, cm$ in all

a

b
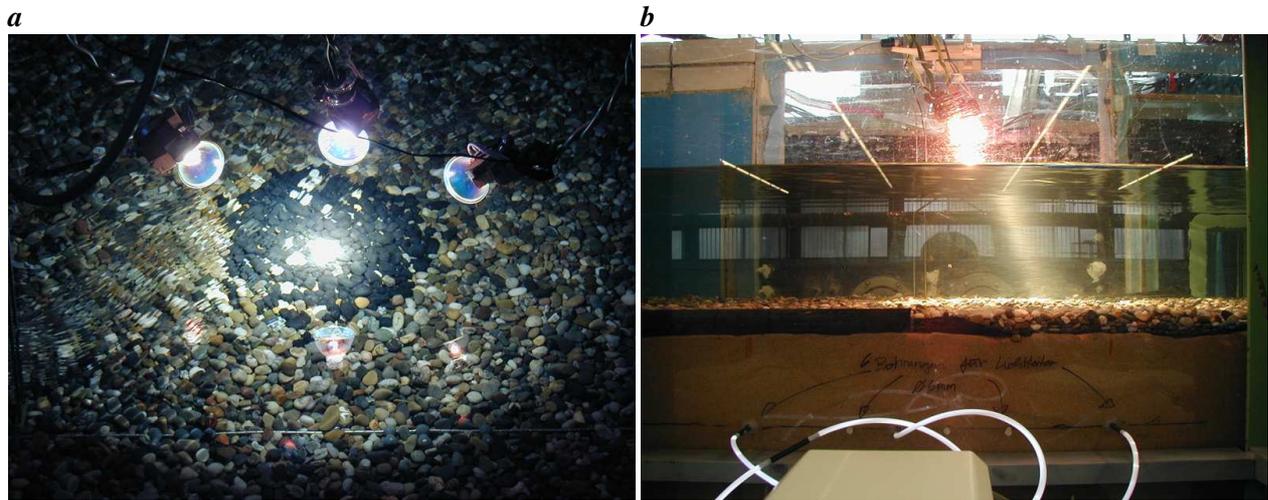
Figure 9.11: ***a*** *Top view of the three halogen lamps used to illuminate the free surface flow. To reduce light scattering at the channel bed, the gravel in the observation area has been colored black. The white spot within the black gravel is one of the artificial pores, mounted in the uppermost position within the gravel layer.* ***b*** *Side view of the observation area during an experiment. The mean flow is from left to right. The halogen lamps are inclined towards the cameras to reduce the scattering angle to a value below* $90°$ *and thereby achieve a higher signal-to-noise ratio in the images (the cameras are beyond the right border of the image and thus not visible). At the bottom of the image, the illumination fibers of the artificial pores can be seen.*

directions[5]. Both cameras are equipped with Schneider Optics **C-mount compact lenses** with an effective focal length of $23\ mm$. The lenses have a locking mechanism for both iris and focus, which is quite important for stereo measurements, because after the system has been calibrated, the camera position, focus and iris settings must not change.

Because the optical axes of the cameras enclose an angle of about $45°$ with the channel walls, the optical rays are refracted at the water-glass and glass-air boundaries of the glass windows. The refraction causes a systematic shift of the image points as compared to optical rays without such multiple media transients. One way to reduce the systematic error introduced by the **multimedia geometry** is to extend the camera model by a model of the imaging geometry. Additional parameters describing the refraction at the interfaces are introduced, e.g. by a ray-tracing approach [Maas, 1992], see section 5.2.3.

In the present setup, the modeling of refraction effects has been avoided by attaching '**liquid prisms**' (glass prisms filled with water) to the channel windows. With the liquid prisms, the cameras continue to have an orthogonal orientation with respect to the air-water interface. Hence, large off-axis angles are avoided and refraction effects are minimized. In this case, the common paraxial approximation is valid and the setup can be modeled by the usual pinhole model. Liquid prisms are also frequently used in stereoscopic PIV applications and have been shown to be extremely efficient at reducing radial distortions arising from a water-air interface [Prasad, 2000; van Doorne et al., 2003].

9.4.2   Illumination and flow visualization

In the design of the illumination setup, three important aspects have to be considered: the light source, the scattering angle between the light source and the camera, and the type of tracer particles. Three standard off-the-shelf **halogen lamps** (Osram, $50\ Watt$) have been used to illuminate the observation volume from above the water surface, see figure 9.11. The lamps are attached to flexible mounts that can be freely moved in all directions. This is important to enable an **optimization of the scattering angle**. Light scattering at small particles with a diameter that is comparable to or slightly larger than the wavelength of the incident light is governed by *Mie-scattering*, which is highly anisotropic [van de Hulst, 1981]. The intensity of the scattered light has a minimum for a scattering angle of $90°$ and a maximum for forward scattering, see e.g. [Leue et al., 1996; Garbe, 1998] for measurements of the scattering cross section. To keep the scattering angle smaller than $90°$, the halogen lamps are inclined towards the cameras, which increases the scattering intensity significantly, see figure 9.11b and figure 9.10b. By placing the two cameras on either side of the flume, it is possible to operate both cameras in forward scatter and thus achieve higher and equal signal-to-noise ratios in both views.

Typically, small polystyrene or glass particles added to the flow serve as tracer particles for flow measurements, see section 9.3.3. The generation of hydrogen and oxygen bubbles by electrolysis is also a common method, see [Engelmann, 2000]. In the present measurements, small **air bubbles** produced in the flume inlet have been used as tracers. The problem of the high bubble density has been mentioned above. The re-design of the flume inlet did not help to overcome this limitation. The bubble concentration and the distribution of bubble sizes cannot be controlled independently but depend on the flow rate. For large flow rates, the bubble density is higher, and the bubble sizes are larger than for small flow rates because of more turbulent inlet conditions. For most of the experiments with Reynolds number $Re >\approx 100000$, the bubble density was too high to enable an optimal performance. Further, the larger bubbles have significant rise velocities which do not represent the fluid flow field. The trajectories of such particles (having large upward velocities over the complete trajectory length) have been removed by a threshold.

9.4.3   Image acquisition

Two **CCD cameras** (Pulnix TM6701AN) have been used for the imaging. The TM6701AN is a progressive scan camera and operates in full resolution of 631×483 pixels at a frame rate of $60\ Hz$. The pixel size is $9 \times 9 \mu m^2$, resulting in a size of the CCD sensor of $5.83 \times 4.36\ mm^2$.

The frame rate is an important parameter with respect to the temporal resolution of the measurements. The higher the frame rate, the higher is the maximum flow velocity that can be measured. Klar [2001] was able to measure maximum flow velocities of about $0.3\ m/s$ using the $60\ Hz$ mode of the cameras. The Pulnix TM6701AN also offers a partial scan mode, reducing the image size to 631×200 pixels, but increasing the frame rate to $130\ Hz$. To enable particle-tracking at higher velocities, this mode has been used for the measurements. However, due to the partial scan of only 200 image lines, the height of the observation volume is reduced to $\approx 5\ cm$.

The two cameras deliver analog image signals which are synchronized and digitized by a color framegrabber (Mikrotron Inspecta 2). This framegrabber has three $8\ bit$ input channels (RGB), that are typically used

---

[5]Actually the stereo volume is larger. Since it is the cross section of two cones in space, it has an irregular shape. The maximum cube that can be inscriped into the volume has a side length of $\approx 5\ cm$.
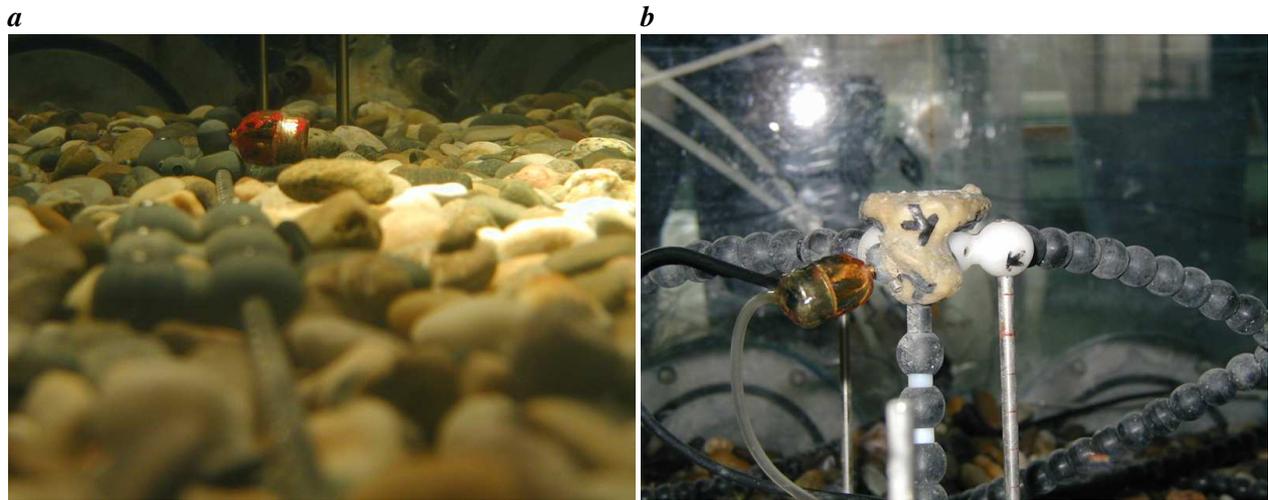
Figure 9.12: a *Pressure sensor fixed at $y = 10\ mm$ above the gravel layer, facing upstream.* **b** *Pressure sensor fixed to one of the artificial pores, facing the pore volume inside.*

for the acquisition of color images. Two of the color channels have been used to grab the monochrome stereo images of the left resp. right camera. The acquired image sequences have to be stored in the computer memory before they are written to a harddisk. The data rate of one camera for the current settings is about $16\ MB/s$. The memory size was $1\ GB$, however, only about $560\ MB$ could be accessed by the framegrabber. Thus, the maximum duration of one stereo image sequence of the free surface flow is about $18.8\ s$, limited by the PC memory.

A standard WindowsNT desktop PC is used to acquire and process the image data.

Before conducting flow measurements, a geometric camera calibration has to be carried out. The calibration of the Pulnix cameras is described in section 5.4.3.

9.5   Pressure measurements

Pressure measurements have been carried out simultaneously with the velocity measurements by the Institute for Hydromechanics of Karlsruhe University. Towards this end, Detert et al. [2004a] developed a new experimental setup to measure pressure fluctuations within and on top of the gravel layer (figure 9.12): **MPPS (Miniaturized Piezoelectric Pressure Sensors)**. The principle of the MPPS is based on the piezoresistive effect. Pressure fluctuations with frequencies of up to $100\ Hz$ can be captured, with an absolute amplitude range of $0 - 6\ kPa$ (corresponding to $0 - 587\ mmWC$ at a temperature of $20°C$) and a resolution of $12\ bit$. The MPPS technique has proven to be a very robust and reliable tool to determine pressure fluctuations down to dissipative scales. For more details, see [Detert et al., 2004a].

9.6   Frame rates, real-time data storage and synchronization

Before installing the equipment at the experimental flume, some preceding tests concerning the frame rates, real-time data storage and synchronization of the seven subsystems have been carried out. The results of these tests are presented in this section.
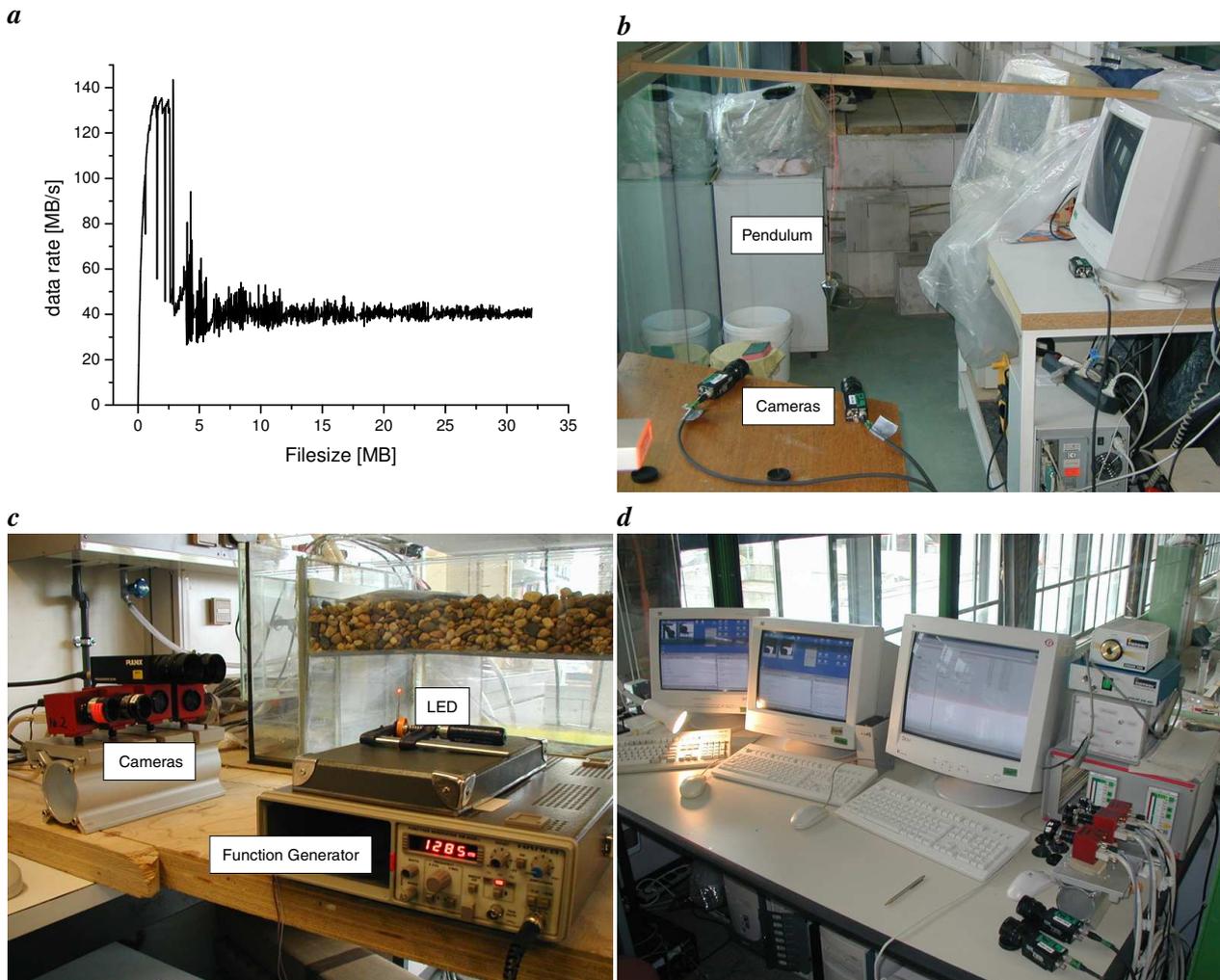
Figure 9.13: *Preceding tests of the RAIDs and the cameras.* ***a*** *Dependence of the sustained data rate of an Infortrend RAID-system on the filesize. For sufficiently large file sizes, a sustained data rate of over* 30 *MB/s can be achieved.* ***b*** *Check if there are missing frames in the sequence acquisition by recording the motion of a pendulum.* ***c*** *Check of the frame rates of the cameras using an LED powered with a sine wave of known frequency. The sine wave is created by a function generator.* ***d*** *Check of the synchronicity of the image acquisition. A lamp (left image border) is recorded by the cameras of two subsystems (right image border) while being switched on and off.*

9.6.1 Frame rates

To check the nominal frame rates of all the cameras, image sequences of an LED connected to a function generator have been acquired, see figure 9.13c. The function generator powers the LED with a sine wave of a precise frequency given by the user. Hence, it is possible to determine the frame rate of the camera using the beat frequency of the LED brightness in the image sequence. The frequency of the sine wave is adjusted until the brigthness of the LED is stationary. Then the frequency of the sine wave is equal to the frame rate of the camera. The results of this test are compiled in Table 9.3. No significant deviations from the nominal frame rates have been found.

Table 9.3: *Frame rates of the cameras.*

setup	camera	frame rate f [Hz]	frame period $T = 1/f$ [s]
Elbe	Sony CCD	25 (full frames)	0.04
Jagst	Sony CCD	50 (fields)	0.02
Kocher	Photonfocus CMOS	400 (ROI)	0.0025
Höllbach	Photonfocus CMOS	400 (ROI)	0.0025
Neckar	Pulnix CCD	130 (partial scan)	0.0077

9.6.2 Real-time storage

To ensure a continuous operation during the measurements the sustained data rates (write rates) of the RAID systems have been determined in dependency of the size of the data packets written to the disks, see figure 9.13a. The plot shows that a sustained data rate of over 30 MB/s can be achieved for sufficiently large file sizes. Thus, the data rate of about 25 MB/s of one stereo camera pair can be easily written to a RAID in real-time during the acquisition. Further tests have been carried out to see if the image sequences are really stored continuously without missing frames. Towards this end, image sequences of an oscillating pendulum have been acquired (see figure 9.13b) and observed manually. No missing frames could be identified.

9.6.3 Synchronization

To ensure simultaneous acquisition of all subsystems, the image acquisition of all cameras should start at the same time. If there is a time lag between the systems, it has to be constant and known in order to correct for it afterwards. Synchronization of the subsystems has been implemented using a hardware trigger module. At the beginning of a measurement, all PCs are sent into a waiting state. To start the image acquisition, a single TTL pulse is created in the trigger module and sent to the waiting PCs. The trigger signal is detected via the parallel port. To get direct access to the IO-ports, the `WinIO` library for Windows is used[6]. As soon as the TTL pulse is detected, the image acquisition is started. Hence, we only ensure a simultaneous acquisition start of all subsystems. After the acquisition start, all systems are running freely without any frame-based or even pixel-based synchronization. However, the acquistion of a stereo pair is always frame-synchronous due to the synchronization by the framegrabber.

The trigger module has been checked by acquiring image sequences of a lamp, which is switched on and off several times. Such a sequence is acquired by two subsystems simultaneously, using the trigger mechanism to start the acquisition, see figure 9.13d. Afterwards, the frames, in which the lamp turns from off to on are detected and the corresponding frame numbers of the two subsystems are plotted against each other. From the axis intercepts of the resulting straight lines, the temporal offset between the two subsystems is determined. The result is that the subsystems Kocher, Höllbach, Elbe and Neckar start the acquisition simultaneously to an accuracy of $\approx 1\ ms$. For the subsystem Jagst, there is a systematic temporal offset of $68\ ms$, i.e. the acquisition starts $68\ ms$ later.

---

[6]`http://www.internals.com`

Table 9.4: *Summary of the seven subsystems (Elbe actually consists of three systems at three different locations) of the experimental setup for flow measurement.*

setup	Pore 1	Pore 2	Pore 3	Sand Motion	Free Surface Flow
	Jagst	Kocher	Höllbach	Elbe	Neckar
frame rate	$50\,Hz$	$400\,Hz$	$400\,Hz$	$25\,Hz$	$130\,Hz$
image size	512×512	184×184	184×184	640×480	631×200
RAM	$1\,GB$	$1\,GB$	$1\,GB$	$512\,MB$	$1\,GB$
speed	$1.7\,GHz$	$1.7\,GHz$	$1.7\,GHz$	$500\,MHz$	$400\,MHz$
OS	W2K	W2K	W2K	W2K	NT
Framegrabber	Eltec PCEye 2plus	SiliconSoftware microenable	SiliconSoftware microenable	Eltec PCEye 4	Mikrotron Inspecta 2
cameras	Sony XC-73CE CCD analog	Photonfocus MV-D1024-28 CMOS digital	Photonfocus MV-D1024-28 CMOS digital	Sony XC-73CE CCD analog	Pulnix TM6701AN CCD analog
RAID	$344\,GB$	$344\,GB$	$344\,GB$	$142\,GB$	-
Tape	-	Sony AIT2	Sony AIT2	$4\,mm$ DDS2	-
seq. length	no restriction	no restriction	no restriction	no restriction	$18.8\,s$
endoscopes/ lenses	Volpi	Eltrotec	Eltrotec	Visitool	Schneider Kreuznach $f = 23\,mm$
illumination	Visitool short arc	Visitool short arc	Visitool short arc	Visitool Halogen	Osram Halogen
data rate	$22\,MB/s$	$25\,MB/s$	$25\,MB/s$	$15\,MB/s$	$31\,MB/s$

9.7   Summary

We summarize the most important facts of the experimental setup, that has been described in this chapter (see also table 9.4). The setup consists of seven subsystems with altogether eleven cameras operating at the same time: three stereo setups to measure the 3-D pore flow within specially prepared artificial gravel pores (setups Kocher, Höllbach and Jagst), one stereo setup to measure the 3-D near-bed free surface flow (setup Neckar) and three monoscopic setups to observe the 2-D motion of sand grains at the sand-gravel-interface (setup Elbe). Optical access to the subsurface flows is provided by endoscopic imaging. Rigid endoscopes ('periscopes') inserted from below the flume are used to observe the sand motion. Flexible endoscopes inserted from the side walls provide image sequences of the pore flow.

The previous work by Janßen [2000], Stybalkowski [2001] and Klar et al. [2002] has shown that 3-D particle-tracking in pore flow sequences is a difficult task in many respects, e.g. the tracking of fast fluctuations on small spatial scales. However, the main reason is the quite limited quality of their endoscopic images (low resolution, geometric distortions, low signal-to-noise ratio). Sophisticated and computationally

intensive image processing had to be applied to achieve good results.

As a general rule, the first thing to do to improve an image processing method is to try to improve the quality of the imagery in the first place, since this may save a lot of algorithmic efforts or even make the application of certain algorithms feasible at all. Therefore, one of the first goals of this work was a re-design of the experimental setup in order to improve its imaging performance. Significant advances have been achieved:

- **Optical quality of the endoscopes:**

  New flexible endoscopes have been introduced that produce images of superior quality. The endoscopes have a higher resolution (30000 instead of 10000 fibers) and show negligible geometric distortion (see also section 5.4.2). The inter-fiber distances are very small and the cladding of the individual fibers is sufficiently thin, so that no hexagonal lattice structure is visible in the images. The endoscopes are equipped with zoom optics, providing the possibility to adjust the image size to the size of the ROI on the image sensor. The depth of field of the endoscopes has been specifically adapted to the application, which enabled a significant reduction of the pore size.

- **CMOS cameras:**

  New digital CMOS cameras have been introduced. The flexibility provided by the latest CMOS technology enables to read out a small ROI (184×184 pixels) on the sensor at a frame rate of $400\,Hz$. A high frame rate is a major demand for endoscopic particle-tracking.

- **Illumination:**

  In order to achieve a sufficient illumination using only one fiber bundle, the effective diameter of the illumination fibers has been increased. The halogen cold light source has been replaced by a short arc lamp, providing a higher color temperature. The resulting spectrum matches the spectral sensitivity of the sensors. As a result, under laminar and moderately fluctuating flow conditions, particle images with a high SNR are obtained that can be segmented easily. For high turbulence levels, motion blur reduces the SNR.

- **Real-time storage:**

  The experiments presented in chapter 10 require the continuous acquisition of long image sequences to gather a sufficient data basis for a statistical evaluation. The need to record long sequences has been mentioned in many previous works [Spies, 1998; Ehrbächer, 1999; Janßen, 2000; Engelmann, 2000; Klar, 2001; Stybalkowski, 2001]. A reliable continuous real-time storage of the sequences has been achieved for the first time. The images are written to a RAID system during the acquistion, and no intermediate storage in the PC memory is necessary. Sequences of arbitrary length can be acquired, with the only restriction being the RAID capacity.

A great advantage of the experimental setup is the exclusive use of cold light sources resp. a short arc lamp for illumination. No lasers are necessary. Hence, laser safety issues do not have to be considered. Further, no high precision optical and mechanical components are necessary, and the setup becomes quite cost efficient. Note that the price of a typical commercial time-resolved 3-D PIV system[7] is at least three

---

[7]as offered e.g. by LaVision or DantecDynamics, including the software

times the price of the hardware for a complete pore flow stereo setup (light source, two endoscopes, two CMOS cameras, PC including framegrabber, RAID and backup system, and a calibration unit). The main expense factor of the commercial systems is the powerful laser that is necessary for a sufficient illumination.

Together with the pressure sensors developed by Detert et al. [2004a], the experimental setup enables temporally and spatially high-resolved measurements of pressure and 3-D flow fields. For the first time, a synoptic view of pressure and velocity fields within and above a porous gravel layer becomes possible.

Chapter 10

Flow measurements and results

This chapter presents the application of the developed technique to study the flow within and above a gravel layer. The recorded image sequences are processed by the 3-D PTV algorithm developed in chapter 7 to extract the 3-D velocities of flow tracers. The simultaneous acquisition of the flow at different locations within and above the gravel layer using the extended experimental setup described in chapter 9 has provided a wealth of information that will be analyzed in a systematic evaluation in future. Here we only present some of the possible approaches and illustrate the results with some examples. Nevertheless, these examples provide new experimental insight into the flow within a porous wall that has been obtained in this form for the first time.

We start with an overview of the experimental program in section 10.1 and a description of the experimental procedure in a typical experiment in section 10.2. Section 10.3 presents the results, and section 10.4 gives a summary of the chapter.

10.1   Overview of the experimental program

The experiments are partitioned into seven major test cases. They are distinguished by the height of the gravel layer H_P, see table 10.1. In some test cases, a rubber mat has been put between the sand and the gravel, see figure 10.1. In these cases, the periscopes for observation of sand motion are not used and the flow in the gravel layer can be studied without hydraulic contact to the sand layer.

The seven test cases are identified by the letters 'm', 'mb', 'mc', 'md', 'me', 'mf', 'mg'. In test case 'me', the gravel in the experimental test section has been replaced by glass beads of $1\ cm$ diameter. In test case 'mg', an obstacle has been put on top of the gravel layer to enable a study of the influence of extra turbulence on the pore flow.

In all cases except 'me' and 'mg', both stationary flow conditions and surface waves are investigated at two different water levels of $h = 0.2\ m$ and $h = 0.4\ m$. In the stationary flows, the mean flow velocity in the free surface flow is typically varied between $0.31\ m/s$ and $0.86\ m/s$. The highest velocity corresponds to low mobility conditions in the gravel bed, i.e. transport of single grains occurs.

An **instability criteria** $\tau_0/\tau_{0,c}$ is defined as the ratio of the bed shear stress to the critical bed shear stress (where motion of gravel grains occurs). The different mean flow velocities correspond to instability criteria between $\tau_0/\tau_{0,c} = 0.09$ and $\tau_0/\tau_{0,c} = 0.59$. The flow conditions in a typical experimental test case with

Table 10.1: *The seven test cases that are studied in the experiments.*

test case	H_P [m]	
m	0.2	rubber mat, incl. waves
mb	0.1	rubber mat, incl. waves
mc	0.1	no rubber mat, incl. waves
md	0.04	rubber mat, incl. waves
me	0.04	rubber mat, glass beads
mf	0.04	no rubber mat, incl. waves
mg	0.04	no rubber mat, obstacle
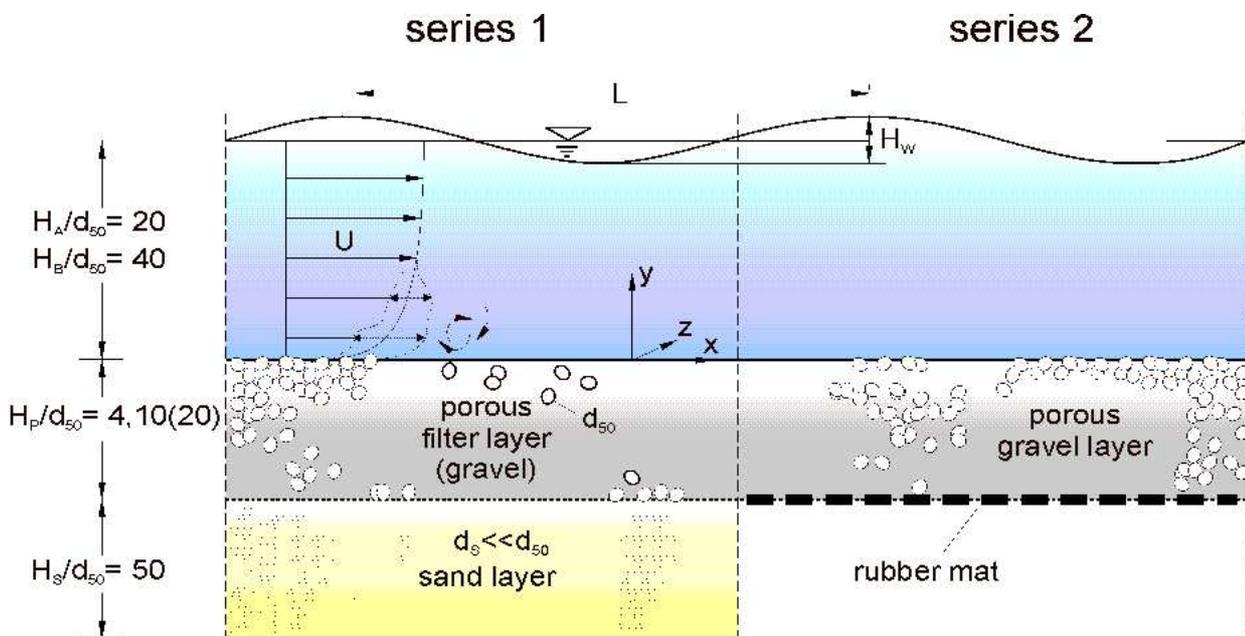
Figure 10.1: *Sketch of the different experimental boundary conditions. Two major series are distinguished. The first corresponds to the normal configuration (sand-gravel-water). In the second, hydraulic contact between sand and gravel is prevented by putting a rubber mat inbetween. Sketch from [Detert et al., 2004b].*

increasing instability criteria at $H_P = 0.1\ m$ are shown in table 10.2. The values given in the table have been determined by Detert et al. [2004b]. The shear velocities u_* have been determined from roughness parameters that were identified in calibration tests. The error in the given values of u_* is estimated to $\pm 5\%$.

All measurements have been carried out in cooperation with the Institute for Hydromechanics of the University of Karlsruhe (IfH). Simultaneously with the velocity measurements, pressure measurements have been carried out by the IfH using the specially developed equipment described in section 9.5 [Detert et al., 2004b, 2005].

In the test cases 'm', 'mb' and 'mc', velocity profiles as a function of the vertical position in the gravel layer have been measured. Towards this end, the Höllbach-pore has been installed in different vertical positions throughout the gravel layer, in steps of 2 cm ('m') resp. 1 cm ('mb','mc'). Flow measurements

Table 10.2: *Experimental conditions in the test cases 'mb' and 'mc' ($H_P = 0.1\ m$, $h = 0.2\ m$).*

$\tau_0/\tau_{0,c}$	$[-]$	0.09	0.18	0.36	0.48	0.55	0.59
Q	m^3/s	56.0	81.8	120.5	149.8	173.0	193.4
h	m	0.201	0.203	0.207	0.219	0.234	0.249
U	m/s	0.31	0.45	0.65	0.76	0.82	0.86
u_*	m/s	0.026	0.040	0.063	0.078	0.073	0.085
Re_*	$[-]$	260	410	640	800	740	870

under different flow conditions have been conducted sequentially.

The experimental data of all the test cases has been acquired during a time period of six months (not including the preliminary tests and the basic preparation of the flume). Altogether, more than one Terabyte of image data has been recorded.

10.2 Experimental procedure

In this section, we describe the typical experimental procedure of one of the test cases shown in table 10.1.

10.2.1 Geometric camera calibration

The first step of the experimental procedure is the geometric camera calibration. All setups have been re-calibrated before and after each test case to account for possible changes in the relative position of the stereo rig, which may occur during the preparation of the experiments or evolve gradually over time. All setups are calibrated in the flume, with the calibration targets submerged in water. The calibration of the artificial pore setups is shown in figure 10.2. The details of the calibration have been described in sections 5.4.1–5.4.3.

10.2.2 Preparation of the experiment

After the geometric calibration, the setups are installed in the flume, see figure 10.3. All artificial pores and pressure sensors are fixed to the grid of steel rods at their corresponding positions. The coordinate system in the flume is defined relative to this grid. After all setups have been installed, a test run is carried out with a low flow discharge of about $10\ l/s$. The test run is performed before the setups are covered with gravel to ensure that the illumination and the tracer supply work properly. The test run is also used to saturate the sediment layers with water *slowly*. The increase of the water level has to be slow, since otherwise large air bubbles rise up through the gravel layer, which may transport sand grains into the gravel pores.

After the test run, the experimental setups are covered with gravel, the gravel layer is leveled, and the desired flow conditions are adjusted.

10.2.3 Image acquisition

When stationary flow conditions have been obtained, the measurements are performed, typically starting at low flow velocities and increasing them up to low mobility conditions, compare table 10.2. Figure 10.4

a   *b*

c   *d*
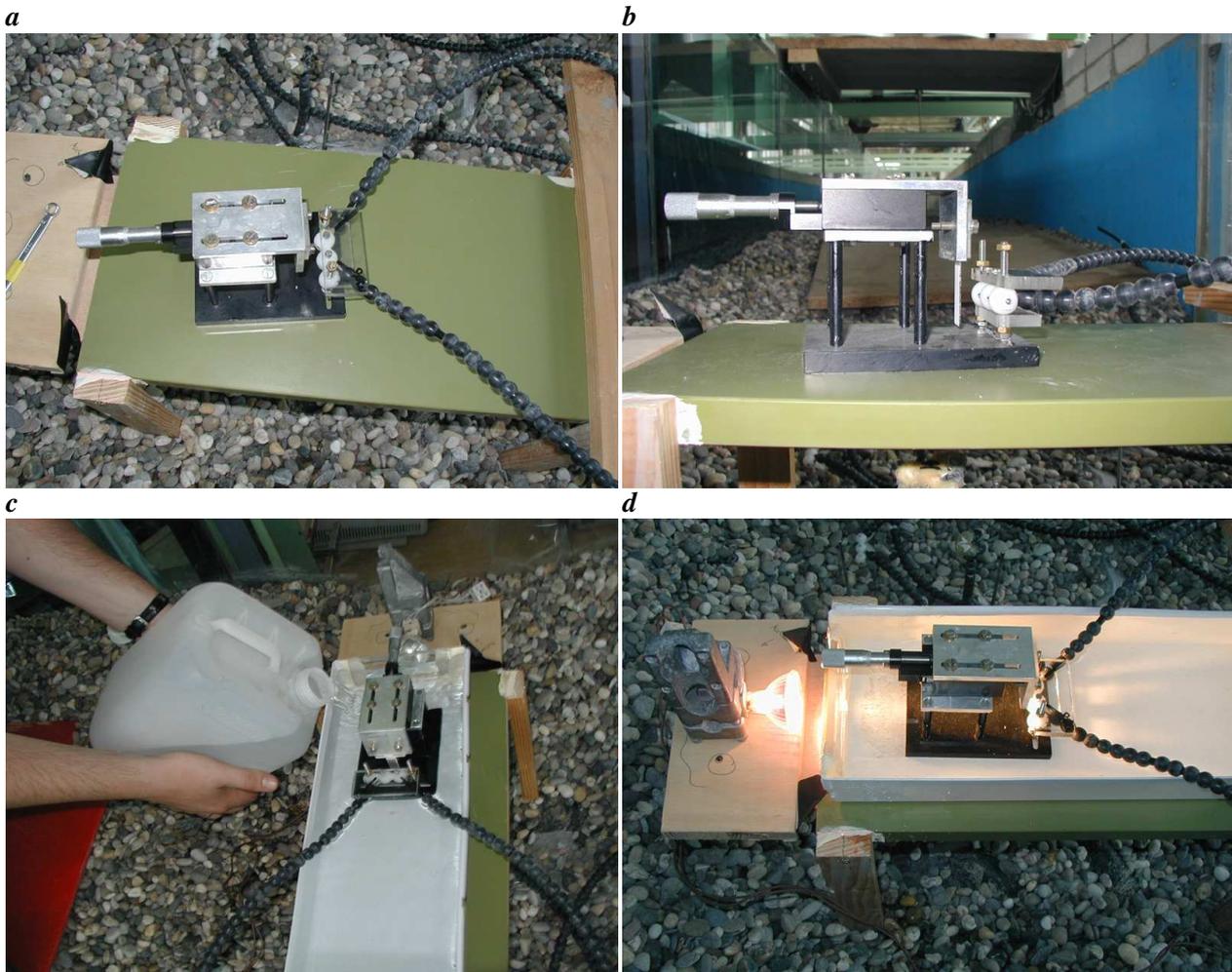
Figure 10.2: *Calibration of the flexible endoscopes in the flume.* **a** *The endoscope stereo rig is mounted on the calibration device (i.e. the linear stage creating the simulated 3-D target, see section 5.4.2).* **b** *Side view of the calibration device.* **c** *Since the calibration images have to be acquired with the endoscopes submerged in water, the calibration device is put in a vessel that is filled with water.* **d** *Illumination of the calibration target from the backside with a halogen lamp.*

gives an impression of an experiment with large surface waves.

10.3   Results

10.3.1   Preliminary notes

Some important aspects have to be kept in mind when analyzing the obtained measurement data. They are summarized here.

a

b

c

d
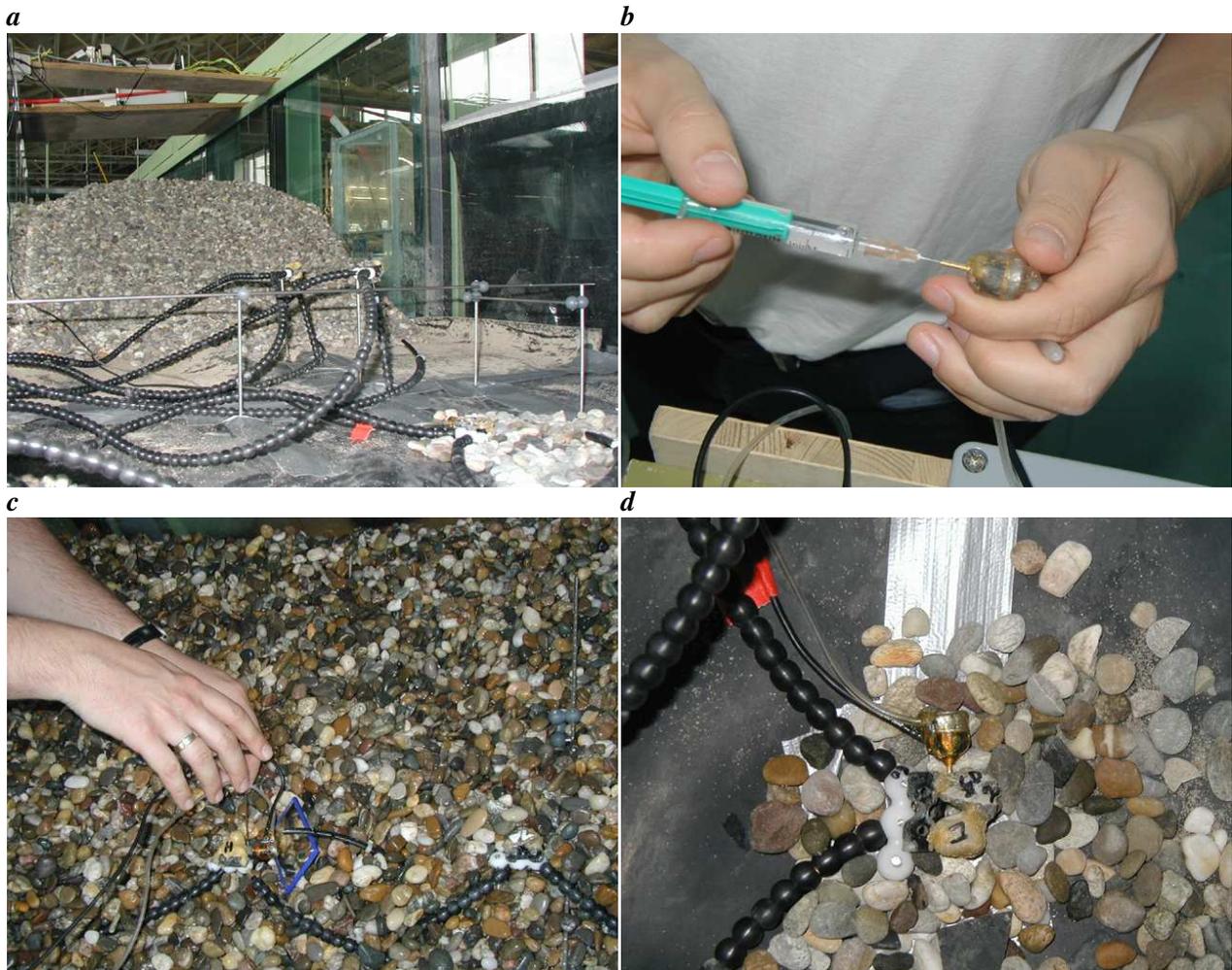
Figure 10.3: *Installation of the experimental setups in the flume. **a** The artificial pores are fixed to the steel grid. **b** The pressure sensors are filled with water using a syringe to remove the air inside the sensor, which would otherwise disturb the measurements. **c** Mounting of a pressure sensor to an artificial pore. **d** Prepared pore with pressure sensor at the bottom of the gravel layer. The photos have been taken during test case 'm'. The rubber mat at the bottom of the gravel layer can be seen.*

Character of the 3-D PTV data.

In the 3-D PTV method, the underlying real velocity field is sampled at **random** points both in space and time. Velocity information is available only at those positions and time instants where tracer particles could be found and successfully tracked. If the flow is seeded homogeneously, this is not a severe limitation. Note that in PIV (section 3.2) the distribution of tracers is also random, but the density is high and homogeneous enough to enable the determination of velocity vectors on a regular grid.

In the experiments presented here, it was not possible to seed the whole water volume with tracer particles, as mentioned in section 9.3.3. Thus, tracers had to be added to the pore flow locally upstream of the artifical pores. With this seeding method, a homogeneous tracer distribution could not always be obtained. The seeding density in the artificial pore depends on the hydromechanic dispersion of the tracers in the gravel layer. The latter represents a highly stochastic geometric system of channels, indirections and

a
b

c
d

Figure 10.4: *Wave experiments.* ***a*** *Surface waves propagating along the channel.* ***b*** *,* ***c*** *,* ***d*** *Three successive images of a surface wave passing the observation area. These images also show a basic limitation of the free surface flow setup: the bright light scattering at the gravel surface. Further, the instationary water surface is responsible for large illumination inhomogeneities due to refraction.*

dead-end pores. The geometry of this system is changed between the different test cases because of the adjustment of the experimental setups. This obviously requires to remove the gravel and replace it after the adjustment. Depending on the flow conditions and the current gravel geometry, in some experimental runs a rather homogeneous particle density was observed in the artificial pores, while in others this was not the case.

As a result, the number of velocity vectors per frame is not constant. In extreme cases it may also drop to zero, i.e. there may be time periods where no velocity information is available at all since no tracer particles reached the observation volume. Two further effects are contributing to this fluctuating information density. First, it was observed that during some experimental runs dirt particles were temporarily deposited in the artificial pores. Sometimes these dirt particles completely blocked one of the endoscopes' view or reduced the intensity of the illumination and thus the signal-to-noise ratio. The second effect is related to limitations of the image processing. Under low mobility conditions, the turbulence intensity in the upper grain layers becomes large. In these cases, the maximum pore flow velocities may reach values beyond the limits of
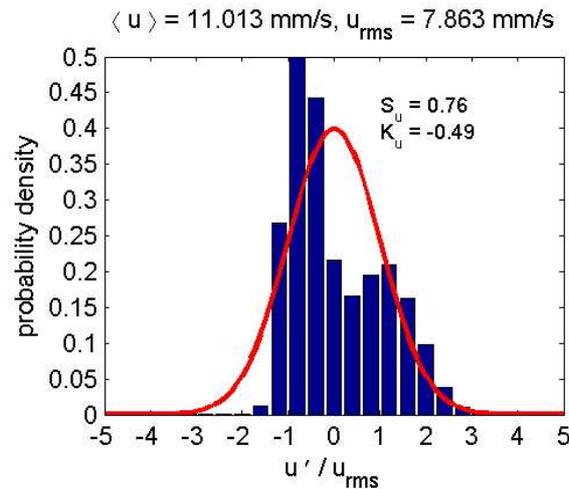
Figure 10.5: *Bimodal histogram of velocity fluctuations in the main flow direction along the flume, obtained within a gravel pore (test case 'mb', Höllbach-setup, $\tau_0/\tau_{0,c} = 0.35$, measurement hmb22). The red curve shows a standard normal distribution for comparison.*

the current endoscopic 3-D PTV. The interframe particle displacements in the image sequences become too large and cannot be tracked any more. Again, the number of recovered velocity vectors drops and the velocity statistics get biased towards lower velocities.

To summarize, the endoscopic 3-D PTV yields information that is **distributed randomly in space and time** and with a fluctuating information density that may cause statistical biases. Thus, a straight-forward analysis of this kind of data using standard statistical tools is not really suitable. More sophisticated statistical analyses are necessary. The results presented in the following sections have been obtained by applying only simple approaches to look at the data. Hence, we have to keep in mind the above considerations.

Spatial structure of the flow within the pores.

Probability density histograms and profiles of mean velocities are computed by double averaging over space and time, compare section 2.4.2. Time series of velocity are computed by frame-wise spatial averaging of all vectors within the artificial pores. For the free surface flow, time series are computed by spatial averaging all vectors within a small volume close to the gravel bed.

The spatial averaging over the pore volume to obtain a time series of mean velocity neglects the spatial structure in the flow field and considers the obtained data as a point measurement of the flow within the pore. This introduces problems if the flow field contains eddy motions. An example is shown in figure 10.5. In this measurement, the flow field inside the Höllbach-pore shows a swirling motion imposed on the main flow, with particles in the upper part of the observation volume moving in the main flow direction and particles in the lower part moving in opposite direction. The motion in the main flow direction along the channel dominates, hence the resulting mean velocity component $\langle u \rangle$ is positive. However, the histogram is bimodal, which shows the presence of the counter-rotating motion. If vectors of this flow are spatially averaged at an instant of time, the u-components with opposite sign may cancel each other. The resulting instantaneous 'point measurement' of u will indicate a mean velocity close to zero.

The natural way to interpret the data obtained by 3-D PTV would be the analysis in a Lagrangian

framework, see [Pope, 1994] for the theory or e.g. [Virant, 1996; Voth, 2000] for Lagrangian analyses of experimental flow data. Note that even if we do not evaluate the data in a Lagrangian way here, this does not question the application of a 3-D PTV approach. The latter is the only feasible method to obtain quantitative measurements of the strongly three-dimensional pore flow at all.

Superficial, intrinsic and local velocity.

As outlined in section 2.6.1, different velocities have to be distinguished in porous media. We note here that the velocity obtained by the 3-D PTV method is a **local** velocity (resp. velocity field) within a single pore. This velocity is different both from the superficial and the intrinsic velocity, since the latter are macroscopically averaged quantities.

We can only obtain a 'local intrinsic' velocity by averaging over a single pore (more precise: averaging over the fraction of the pore volume that is covered by the stereoscopic observation volume of the endoscopes). To obtain superficial or intrinsic velocities, the averaging has to be carried out over at least $15 \times 15 \times 15$ pores, which is the minimum size of a macroscopic sample according to [Dullien, 1992]. Averages over submacroscopic samples may fluctuate. This has to be kept in mind when comparing the measurement results to macroscopic phenomenological flow models.

Size of the pore volume.

The average inner diameter of the artificial pores is $\approx 1\ cm$, which is larger than the natural pore spaces. Smaller diameters are not possible due to limitations of the experimental setup. Hence, a possible influence of the larger pore size on the measurement data has to be considered.

10.3.2   Classification of flow regimes

We give a general classification of the flow regimes that have been studied, according to some Reynolds numbers (for the definitions, see section 2.9).

Free surface flow.

The Reynolds number in the free surface flow is always above 60000 and reaches values of up to ≈ 300000 for the fastest flow conditions. The minimum roughness Reynolds number is 260 and reaches values of up to ≈ 1000. Hence, the flow above the gravel layer is **fully turbulent** and in the **completely rough** regime. Typical viscous wall units are $y_0 \approx 10^{-5}\ m$. The viscous sublayer can be neglected, and near-wall turbulence will be dominated by roughness effects. The permeability Reynolds number is in the range of ≈ 8 to ≈ 30. Hence, there is an influence of permeability. However, it is supposed to be much smaller than the influence of roughness.

Figure 10.6 shows histograms of the near-wall free surface flow velocities. All components are normalized by their rms-values. The lower right plot shows the histogram of the normalized 3-D speed, where the 3-D speed is defined as $s = \sqrt{u^2 + v^2 + w^2}$. The red curves show standard Gaussian distributions for comparison, resp. a transformed χ^2-distribution in case of the speed (the square of the speed follows a χ^2-distribution with three degrees of freedom, if the single components are Gaussian distributed and assumed independent).
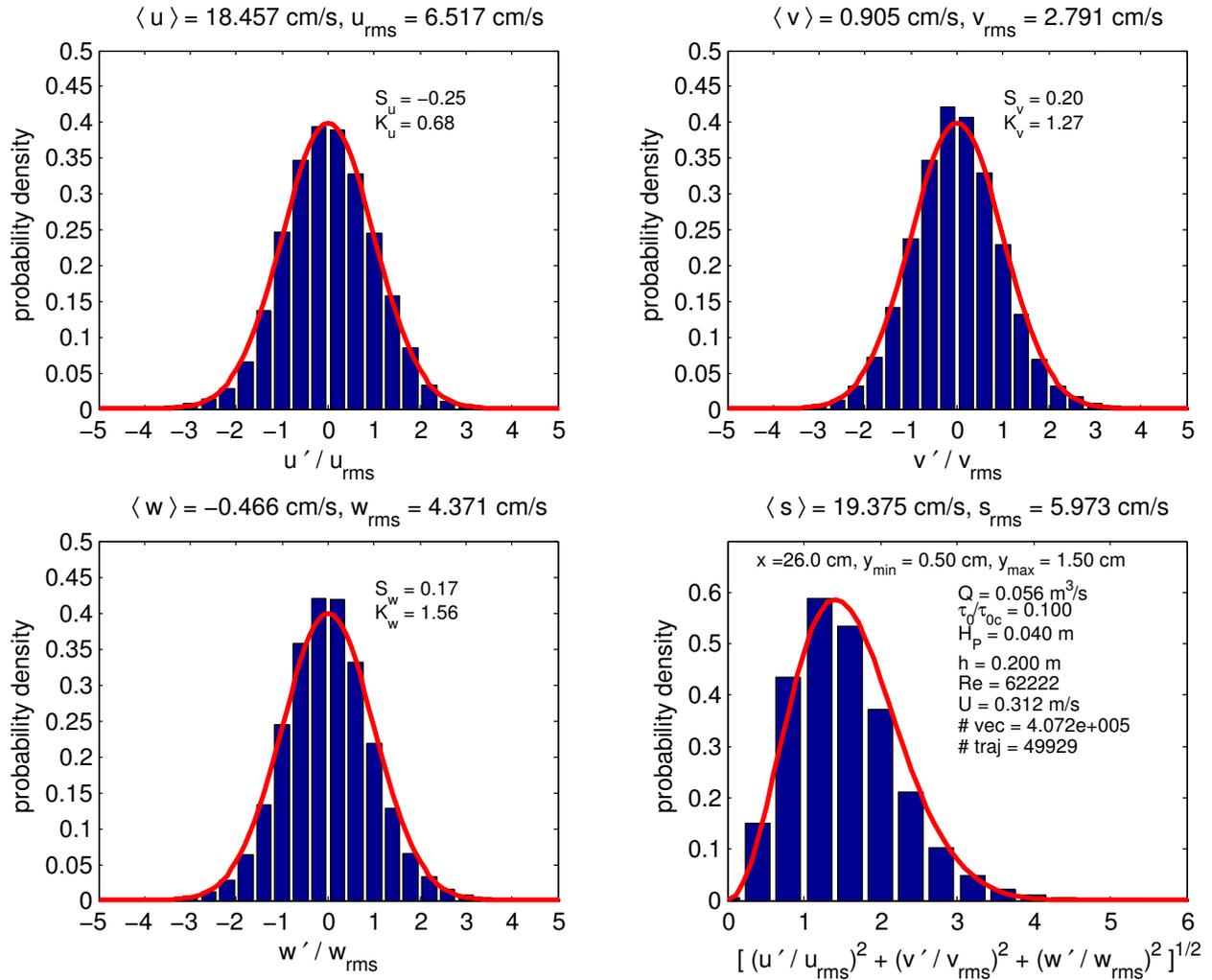
Figure 10.6: *Histograms of velocity fluctuations in the near-wall free surface flow ($\tau_0/\tau_{0,c} = 0.09$, $H_P = 0.04\,m$, $h = 0.2\,m$).*

The plots show that the near-wall fluctuations can be assumed to be Gaussian to a good approximation. This confirms the results of Hofland [2005], who assumed Gaussian near-wall velocity fluctuations and derived a pdf for near-wall pressure fluctuations. The latter has been confirmed in experimental measurements, both by Hofland [2005] and in the measurements presented here [Detert et al., 2005].

The histogram of the speed fluctuations appears to be slightly shifted to the left as compared to the theoretical curve. This indicates that the effective number of degrees of freedom is smaller than three. The latter has to be expected, since clearly the fluctuations within a turbulent channel flow are not independent of each other.

Pore flow.

Typical (local) mean flow speeds s measured within the artificial pores are in the range of well below $1\,mm/s$ in the deepest grain layers and for the lowest free surface flow velocities, up to more than $150\,mm/s$
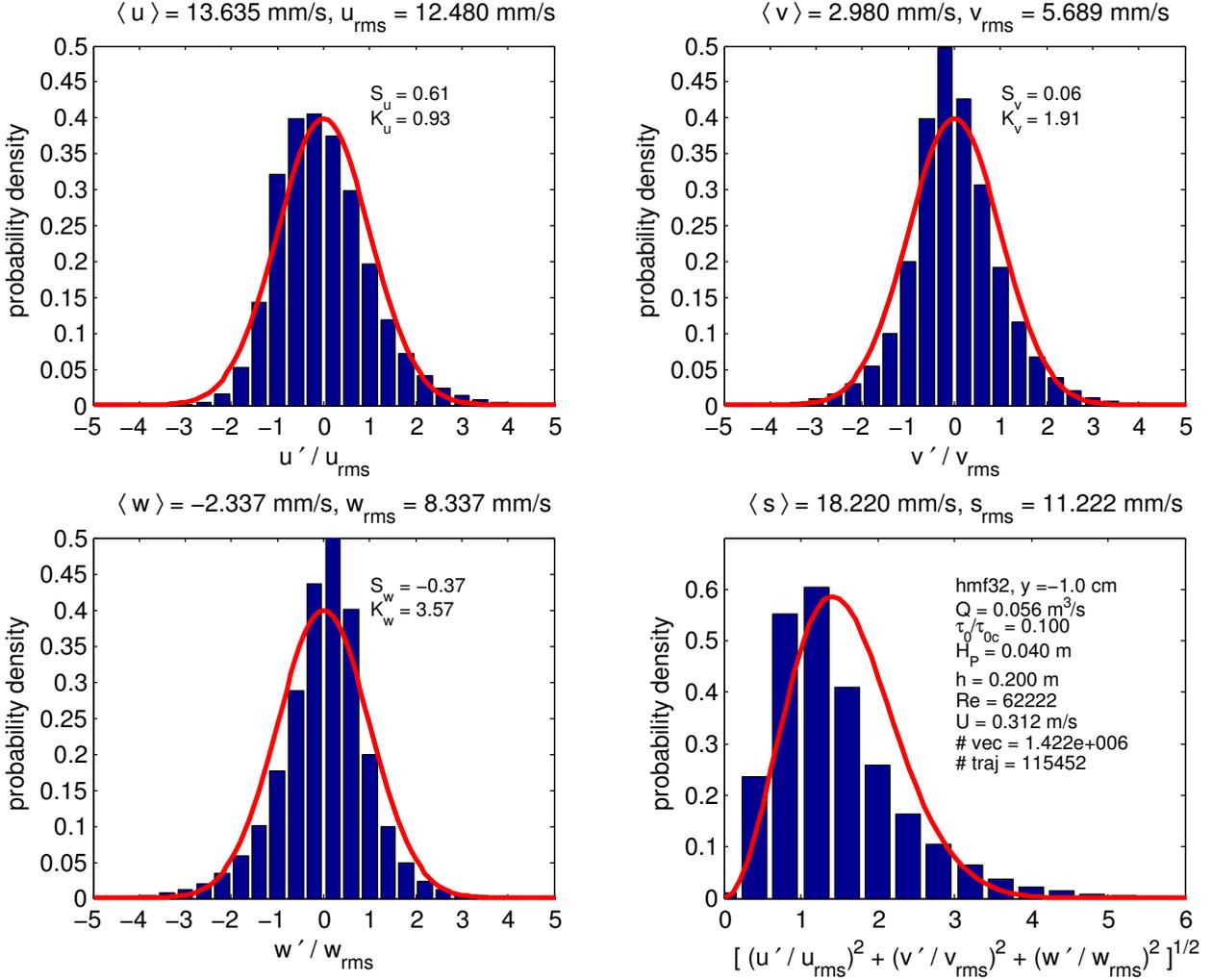
Figure 10.7: *Histograms of velocity fluctuations in the pore flow at $y = -1\,cm$. ($\tau_0/\tau_{0,c} = 0.09$, $H_P = 0.04\,m$, $h = 0.2\,m$).*

in the experiments with surface waves. Typical values in the upper grain layers are in the range of $\approx 1\,mm/s$ up to $\approx 40\,mm/s$. Note that the flow in the pores is strongly fluctuating. The standard deviations of the mean velocities are typically of the order of magnitude of the mean values. Using a mean inner diameter of the artificial gravel pores of $1\,cm$, the resulting pore Reynolds numbers are between ≈ 5 for slow flow in deeper grain layers and ≈ 1500 for flow induced by surface waves. Clearly, the latter is turbulent, which is also observed from the image sequences. The dominant flow regime in the pore flow for stationary conditions in the free surface flow seems to be a transitional one, with periods of quasi-laminar flow interrupted by turbulent spots.

The intermittency of the pore flow is also shown in the histograms of figure 10.7 and figure 10.8. These histograms have a much higher kurtosis and do not agree with the standard Gaussian distributions. The reason is the intermittency of the pore flow, which is not fully turbulent, but rather transitional. A comparison of the two plots shows that the kurtosis is higher for $y = -2\,cm$ than for $y = -1\,cm$. This tendency of increasing kurtosis - and hence increasing intermittency and decreasing degree of turbulence - with increasing
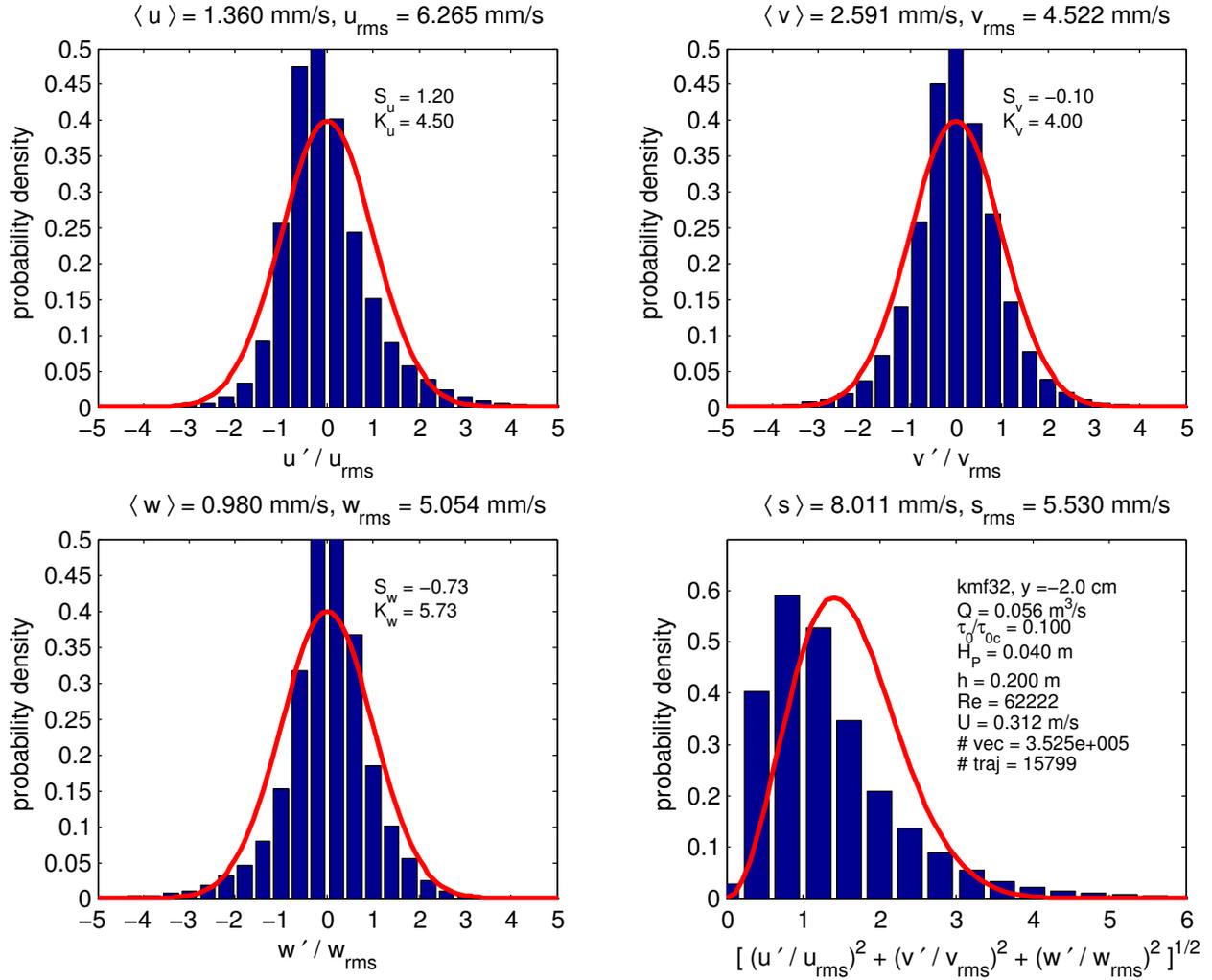
Figure 10.8: *Histograms of velocity fluctuations in the pore flow at $y = -2\,cm$. ($\tau_0/\tau_{0,c} = 0.09$, $H_P = 0.04\,m$, $h = 0.2\,m$).*

depth in the gravel layer is confirmed by many other measurements.

10.3.3 Free surface flow

We determine the parameters u_*, k_s and the vertical shift Δy of the zero-level of the logarithmic wall law by curve fits. The results are shown in figure 10.9 and figure 10.10. We use three approaches.

First, we assume a constant vertical shift of $\Delta y = 0.25k$, as also done by [Detert et al., 2005] and commonly done in engineering practice [Dittrich, 1997]. The fitted parameters are u_* and k_s. All fits are performed using a robust method (Least Median of Squares, see also section 5.3.2), which automatically removes outliers. The results are shown by the yellow curve. The regression coefficient is close to one, however the obtained values of u_* and k_s are larger than those given by [Detert et al., 2005] ($u_* = 0.026\,m/s$, $k_s = 1.68 d_{mD}$). The latter have been obtained by independently determining u_* from roughness parameters and subsequently fitting only k_s to the logarithmic profile.
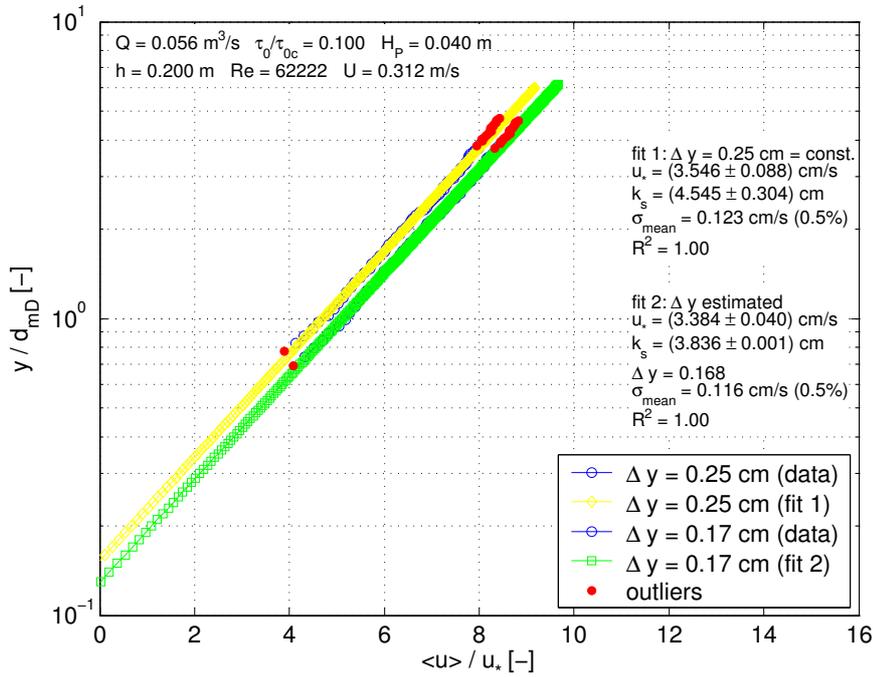
Figure 10.9: *Determination of the parameters u_*, k_s and Δy by fitting the logarithmic wall law to the experimental data.*
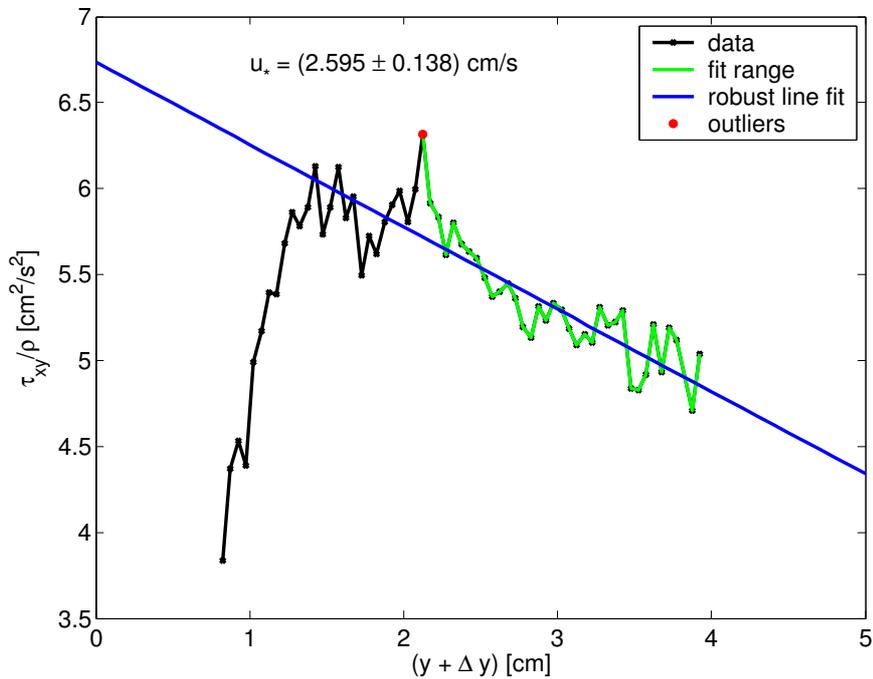
Figure 10.10: *Determination of the shear velocity u_* from the ordinate intercept of the vertical Reynolds stress profile.*

In a second approach, we use the method described in section 2.5.2. In this method, the vertical shift Δy is not assumed constant but also estimated from the data. A value of $\Delta y = 0.17\ cm$ is obtained. The results are shown by the green curve. The roughness parameters are again larger than those of Detert et al.

[2005].

The reason of the larger values of both u_* and k_s are probably correlations of these two parameters, which are fitted simultaneously. Koll [2002] mentions the shortcomings of these fitting procedures and introduces a new method to determine the roughness parameters and u_*. However, this method has not been used here.

Instead, we follow a third approach. The shear velocity u_* can also be determined from the vertical profile of the Reynolds stress term $\overline{u'v'}$, see section 2.5. The result is shown in figure 10.10. The obtained value of u_* agrees very well with that of [Detert et al., 2005]. Subsequently, we use the obtained u_* and $\Delta y = 0.25k$ to determine only k_s from the logarithmic profile. The result is $k_s = (1.71 \pm 0.11)\ cm$, which is also in excellent agreement with the value of Detert et al. [2005].

Figure 10.10 shows that the experimental data is rather fluctuating. This shows that statistical convergence has not really been reached. The data basis of this plot consists of ten image sequences, with a total length of $188\ s$ (ten successive recordings under the same flow conditions). Averaging is carried out over horizontal slices with a height of $0.5\ mm$. These slices typically contain only ≈ 20000 vectors, which seems to be too less to arrive at a converged estimate of the higher order moments.

Nevertheless the results shown in this section have proven that reasonable results are obtained by the 3-D PTV. The roughness parameters determined by 3-D PTV agree with those determined by [Detert et al., 2005], who used a different, independent measurement technique (1-D Acoustic Doppler Current Profiler). The deviations of u_* and k_s obtained in the simultaneous fits are rather related to shortcomings of the fit method than to shortcomings of the data.
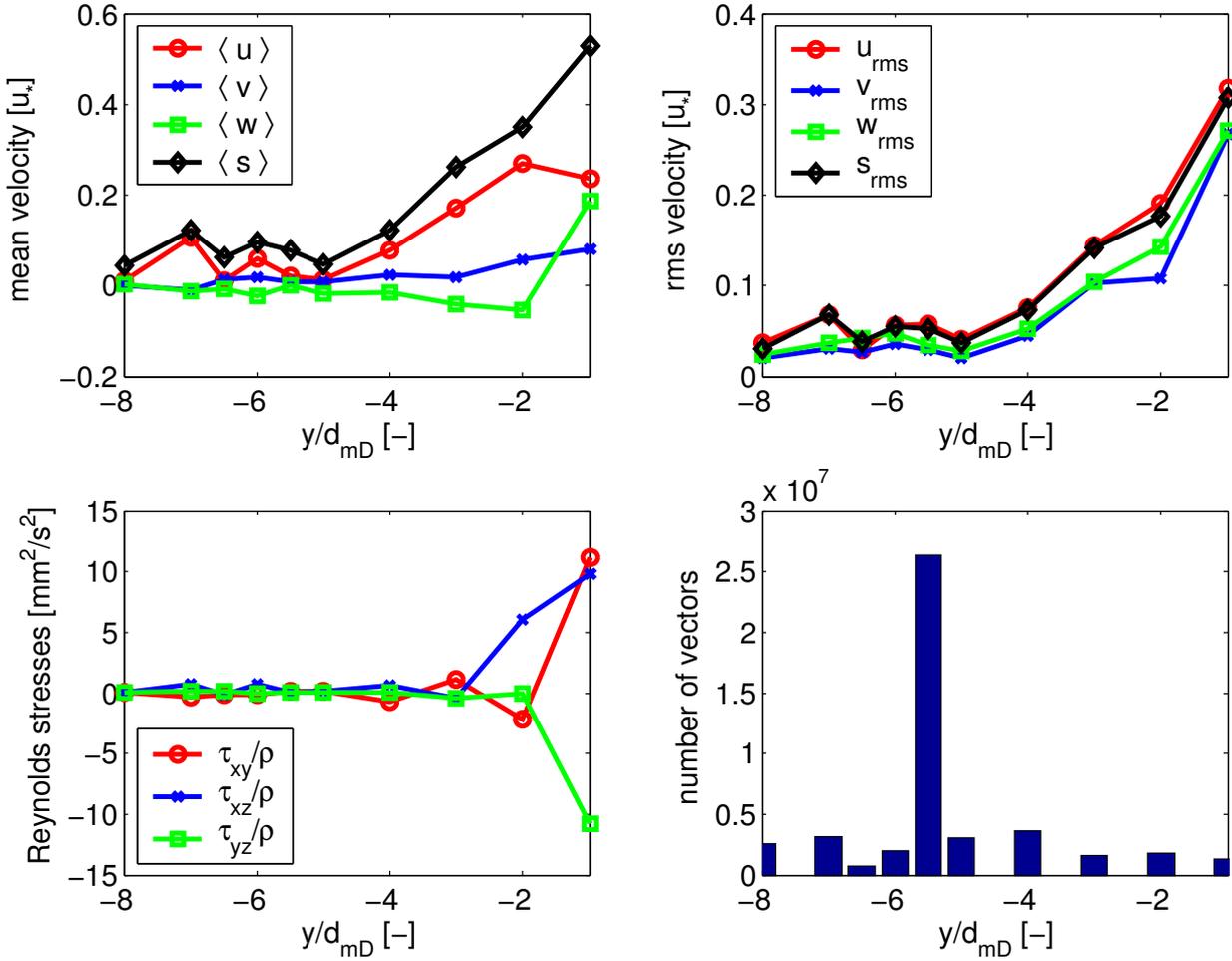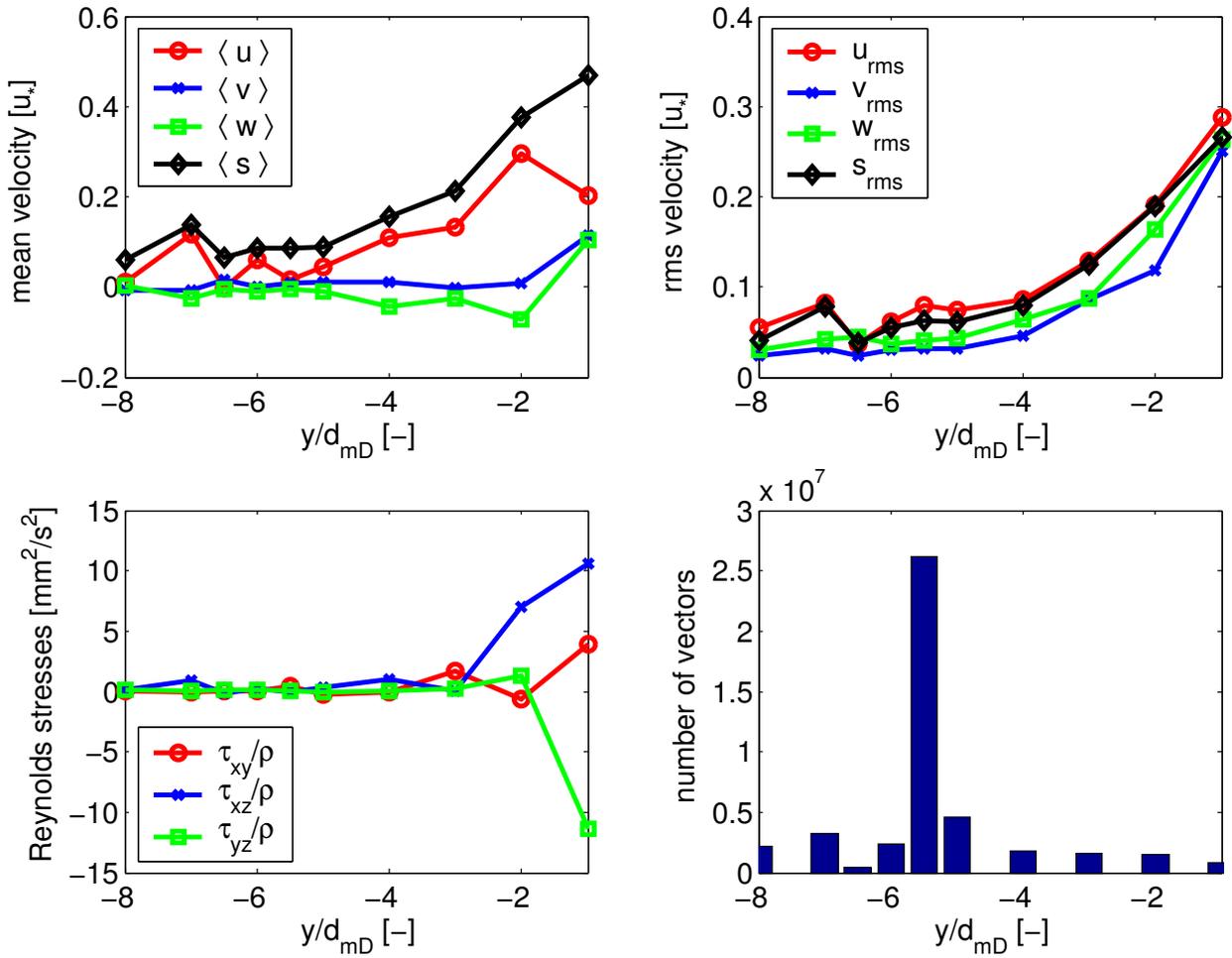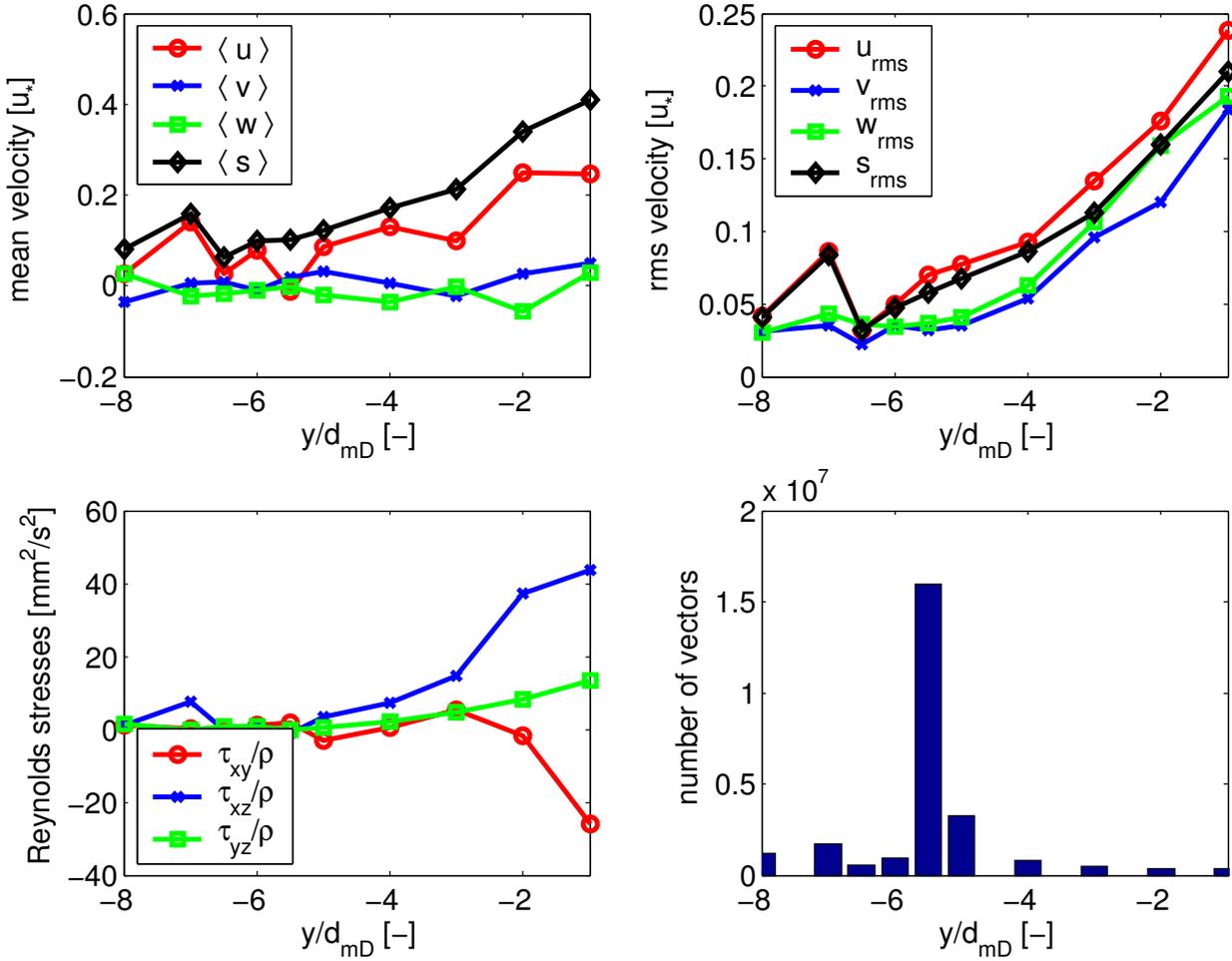
Figure 10.11: *Vertical profiles of local intrinsic velocity, rms-velocity and Reynolds stress throughout the gravel layer. The data is taken from test cases 'mb' and 'mc', $\tau_0/\tau_{0,c} = 0.09$, $H_P = 0.1\ m$, $h = 0.2\ m$.*

10.3.4   Pore flow

Profiles of intrinsic velocity, rms-velocity and Reynolds stress.

To illustrate the pore flow measurements, we show vertical profiles of the local intrinsic velociy, rms-velociy and Reynolds stress throughout the filter layer in Figs. 10.11–10.13. Velocities are normalized by u_*, and the vertical coordinate is normalized by d_{mD}. We show both the three velocity components and the speed. The lower right plot shows the number of vectors that have been averaged to obtain the corresponding data point. We average all measurements that were available for the corresponding flow conditions. The reason why the data point at $y = -5.5\ cm$ has much more vectors is that at this position the Kocher-setup was located. The other points of the profiles have been obtained by shifting the Höllbach-setup sequentially throughout the gravel layer. For each measurement, a sequence of the Kocher-setup (which always resides at the same position) has also been acquired. Hence, the data basis at this point is much larger.

In agreement with the observations of other researchers that have been reviewed in sections 2.7–2.8 there is a strong velocity gradient in the upper grain layers. The decrease of the intrinsic speed seems to be

Figure 10.12: *Vertical profiles of local intrinsic velocity, rms-velocity and Reynolds stress throughout the gravel layer. The data is taken from test cases 'mb' and 'mc', $\tau_0/\tau_{0,c} = 0.09$, $H_P = 0.1 \, m$, $h = 0.4 \, m$.*

linear, whereas the rms-velocity and the Reynolds stress follow an exponential decay. Below $y/d_{mD} < -4$, all profiles become uniform. This is in good agreement with the results of [Detert et al., 2005], who has shown that the high-frequency pressure fluctuations in the gravel layer (obtained simultaneously in the same experiments) also follow an exponential decay, compare Fig. 12 in [Detert et al., 2005]. Similar observations are reported by [Vollmer, 2005].

The data of figure 10.11 has been obtained at a water level of $h = 0.2 \, m$. In figure 10.12, we show the profiles that have been obtained at the same instability criteria but at a water level of $h = 0.4 \, m$. The results are very similar. A comparison of the two plots does not suggest that the fluctuations penetrate deeper into the gravel layer at the larger water level. Hence, the dominant source of fluctuations appears to be the near-wall turbulence, which scales with the inner variables u_* and k_s, which are the same for both plots. Note that the data shown here has been obtained at a rather low Reynolds number of the outer flow. At larger Reynolds numbers, the influence of large-scale vortical motions in the outer flow may also be of importance.

Figure 10.13: *Vertical profiles of local intrinsic velocity, rms-velocity and Reynolds stress throughout the gravel layer. The data is taken from test cases 'mb' and 'mc', $\tau_0/\tau_{0,c} = 0.35$, $H_P = 0.1\,m$, $h = 0.2\,m$.*

In figure 10.13, we show the velocity profiles again at a water level of $h = 0.2\,m$, but at a larger instability criteria of $\tau_0/\tau_{0,c} = 0.35$. Here, all profiles penetrate deeper into the gravel layer up to $y/d_{mD} \approx -6$. All profiles show a rather linear decrease. The deeper penetration indicates again that the dominant fluctuations originate from the near-wall turbulence. The corresponding shear velocity for this flow condition is $u_* = 0.063\,m/s$, which is roughly a factor of two larger than in the plots of figure 10.11.

Figure 10.14 presents the dependency of intrinsic velocity, rms-velocity and Reynolds stress on the instability criteria $\tau_0/\tau_{0,c}$. The data has been obtained from the Höllbach-setup, which was located at $y/d_{mD} = -1$. The water level was $h = 0.2\,m$, and the height of the gravel layer was $H_P = 0.1\,m$. The velocities are normalized by the 'critical shear velocity' $u_{*,c}$, which is derived from the critical shear stress of the gravel $\tau_{0,c} = 8.8\,Pa$ as $u_{*,c} = \sqrt{\tau_{0,c}/\rho}$. The increase of the rms-velocities with $\tau_0/\tau_{0,c}$ appears to be stronger than linear. The maximum rms-value of the streamwise velocity reaches about $u_{rms}/u_{*,c} \approx 2.7$ at $\tau_0/\tau_{0,c} = 0.48$. Hence, turbulent velocity fluctuations may be much larger than the mean values not only in the near-wall region in the channel, but also in the upper grain layers within the gravel.
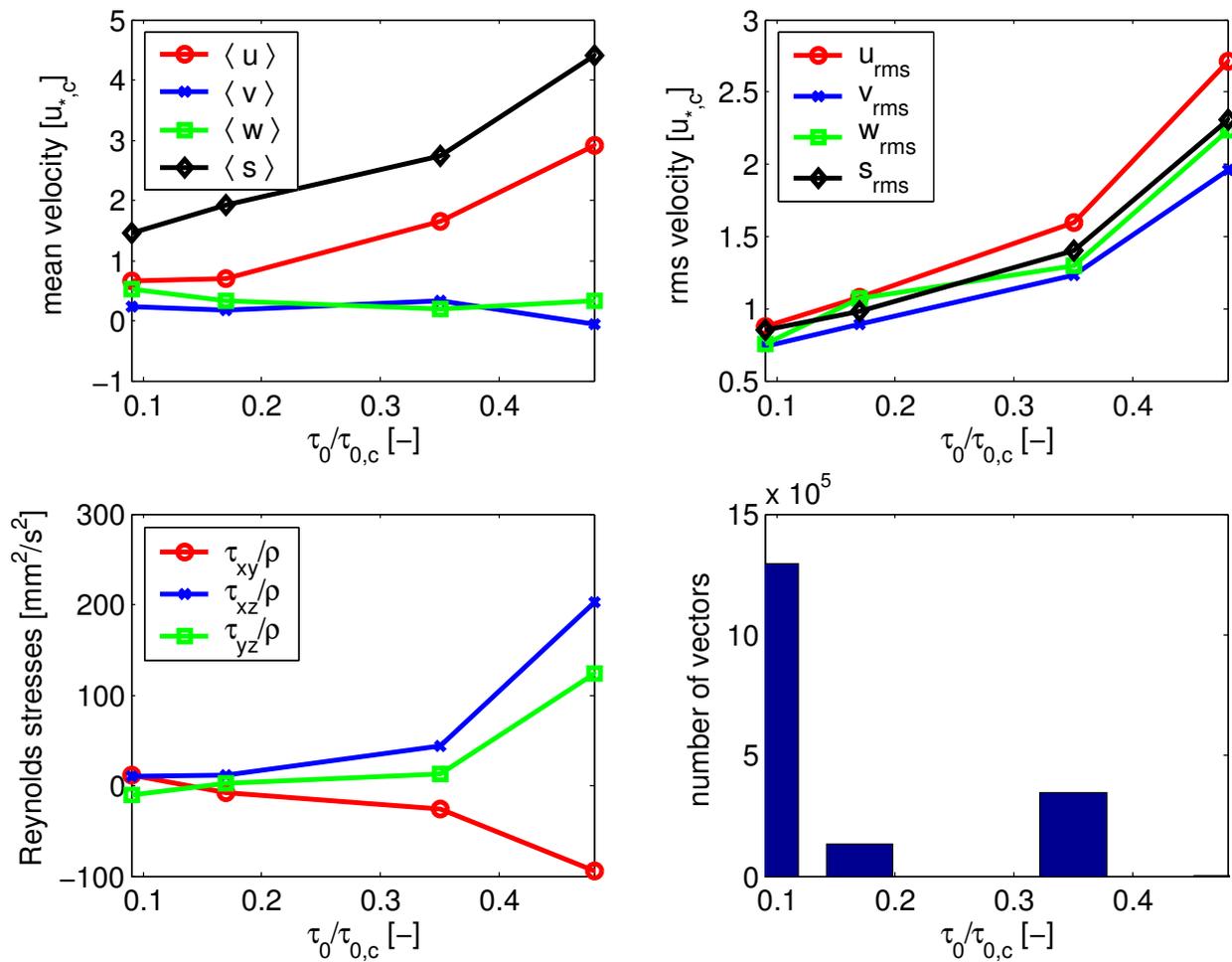
Figure 10.14: *Dependency of local intrinsic velocity, rms-velocity and Reynolds stress on the instability criteria* $\tau_0/\tau_{0,c}$. *The data has been obtained from the Höllbach-setup located at* $y/d_{mD} = -1$, $H_P = 0.1\ m$, $h = 0.2\ m$.
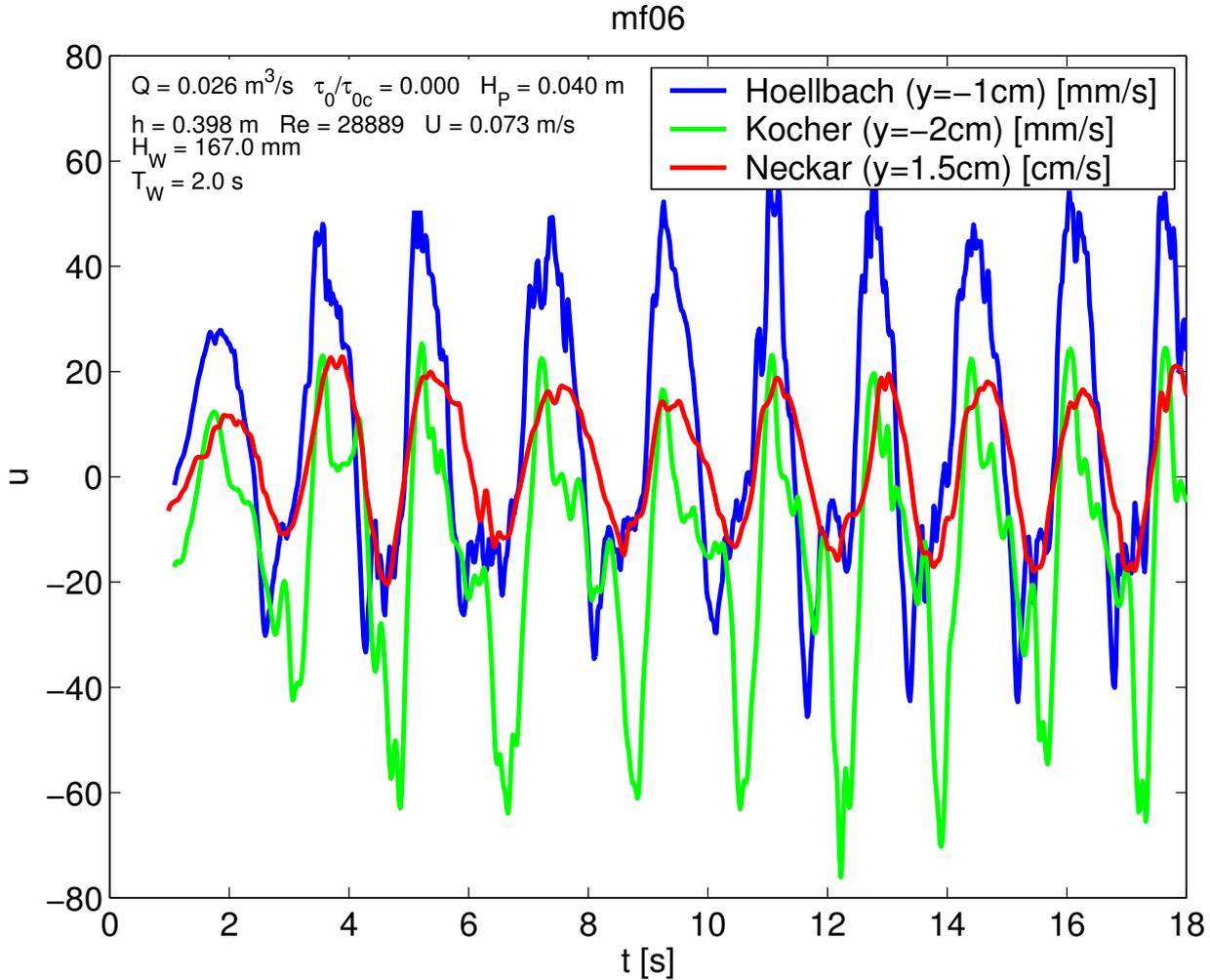
Figure 10.15: *Simultaneous time series of the free surface flow and pore flow, obtained under instationary flow conditions in the open-channel flow due to surface waves.*

Interaction of free surface flow and pore flow.

Figure 10.15 shows simultaneous time series of free surface flow and pore flow. The time series of the free surface flow has been obtained by frame-wise spatial averaging of all vectors within a cube of $2\ cm$ side length. The center of this cube is located at $y/d_{md} = 1.5$, immediately above the Höllbach-pore. The latter is located at $x/d_{md} = 26, y/d_{md} = -1$, and the Kocher-setup is located at $x/d_{md} = 1.5, y/d_{md} = -2$. Surface waves are propagating along the flume, with an amplitude of $\approx 8\ cm$ and a period of $\approx 2\ s$. Note that the velocity of the free surface flow is shown in cm/s (red curve, Neckar-setup) and the pore velocities are shown in mm/s on the same y-axis. The plot clearly indicates the instantaneous penetration of the wave-induced pressure field into the gravel layer, which in turn drives the flow fields within the pores. Similar observations are reported for wave-induced pressure variations within a gravel layer by Vollmer [2005] and, on a larger spatial scale, by Schwab and Köhler [2003]. They show that wave-induced pressure fluctuations are damped exponentially with increasing depth in the gravel layer. The amplitudes of the velocity fluctuations in the gravel shown here are approximately equal, because both pores are located in
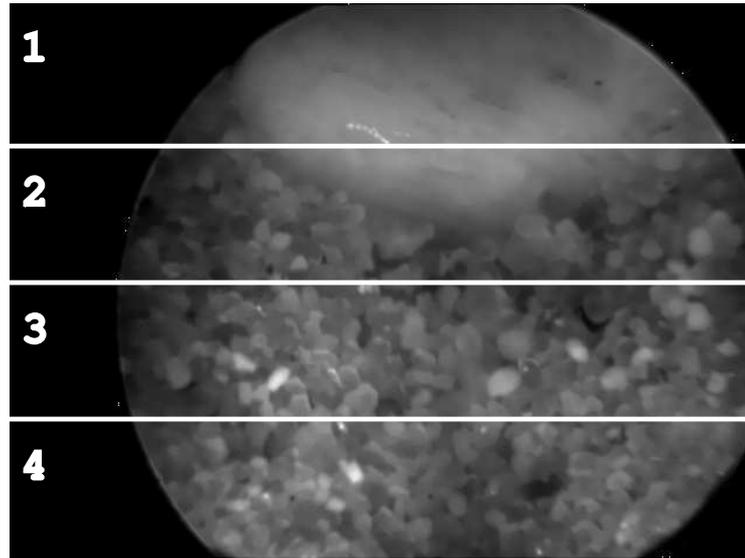
Figure 10.16: *Partition of the periscopic observation area into four horizontal slices.*

the upper grain layers. Interestingly, the mean streamwise velocity of the Höllbach-setup at $y/d_{md} = -1$ is positive, whereas that of the Kocher-setup at $y/d_{md} = -2$ is negative.

10.3.5 Sand motion

In this final section, we show an exemplary result of the detection of sand motion using the rigid endoscopes ('periscopes', Elbe-setup, see section 9.2). A comprehensive analysis of the subsoil observation in the flume experiments has been compiled in a technical report [Klar, 2004b].

In a first step, all image sequences are processed by the same motion detection algorithm that has been used to segment the tracer particles in the 3-D PTV algorithm, see section 7.3.2. This analysis has shown that the only flow conditions where significant motion of sand grains is observed are those with large surface waves. In previous applications, image sequences showing hydromechanically induced subsoil motion have been processed by an optical flow algorithm (reviewed in section 3.4), e.g. [Spies et al., 1999]. Similar small-scale motions with slow velocities have not been observed in the present application. The entrainment of sand grains by the wave-induced pore flow leads to fast particle motions. The resulting displacements between two successive images are too large to be processed by an optical flow algorithm. Hence, we only present the results of the motion detection algorithm.

Figure 10.16 shows a typical periscope image. In this image, a partition of the observation area into four horizontal slices is indicated. In the upper part of the image, a gravel grain can be seen. The interface between the gravel layer and the sand layer is located in slice number two. Figure 10.17 shows the results of the motion detection. In the upper left plot, a time trace of the sand motion is shown. The plot has been obtained by simply counting all the pixels in the image where motion has been detected and plotting the result as a function of time. A quasi-periodic motion can be seen. This means that in some wave cycles, sand grains have been transported, but not in all. The upper right plot also shows a time series of the motion counts. Here, the counts have been accumulated separately for each of the horizontal slices indicated in figure 10.16. The plot shows that the maximum motion occurs in slice one and two, where the gravel layer
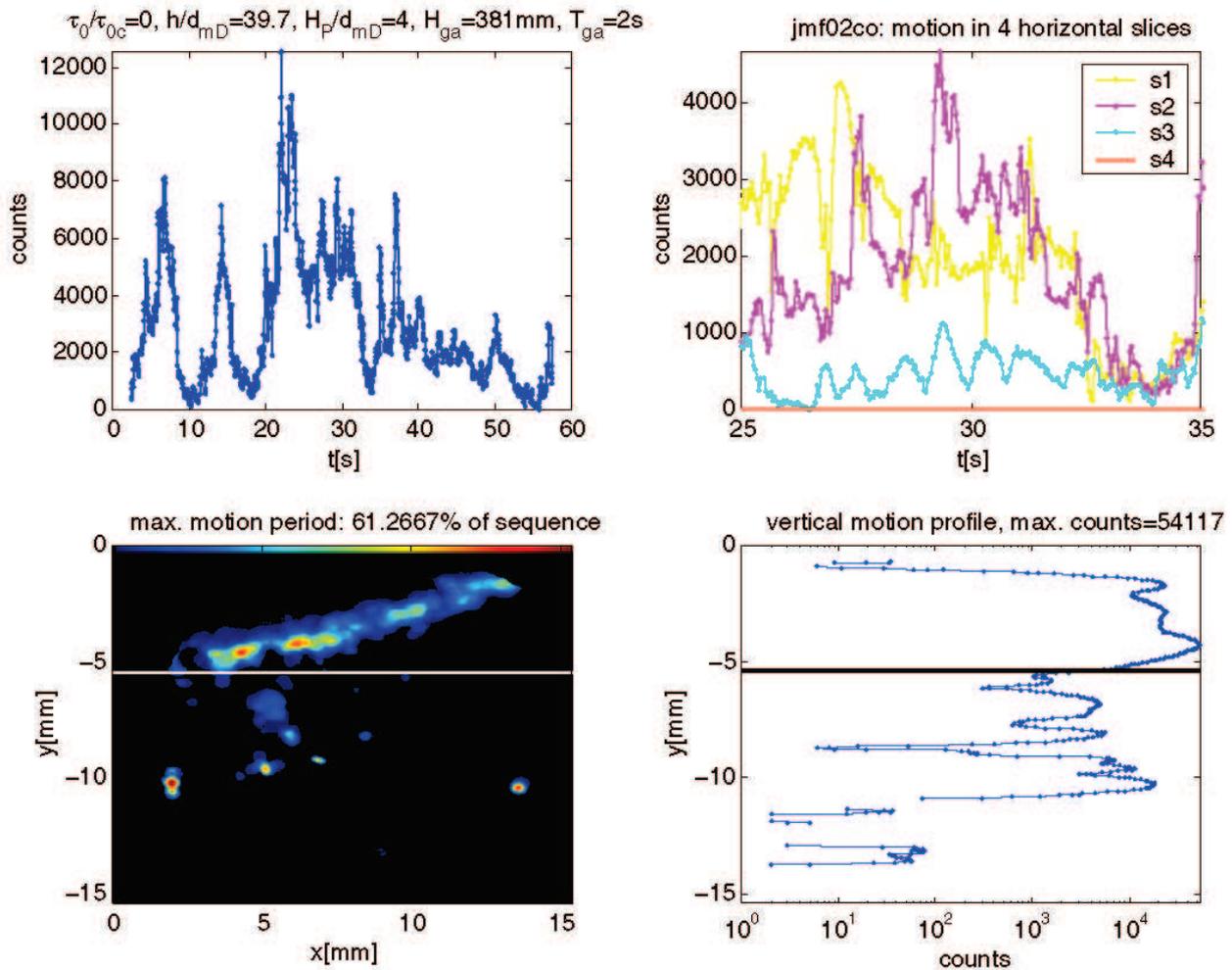
Figure 10.17: *Results of the sand motion detection. For explanations, see text.*

resp. the interface of sand and gravel is located.

In the lower left plot, the motion counts are shown as a function of space. The white horizontal line indicates the approximate location of the sand-gravel-interface. The counts are color-coded. The colormap is shown at the top of the image. Blue corresponds to 'no motion', and red corresponds to 'maximum motion'. The plot clearly shows that motion of sand grains occurs along the interface between the sand and gravel layer. Grain motion in the deeper sand layers (slice three and four) only occurs for single grains, which move periodically with a small amplitude, but are not transported away, e.g. because of a fluidization of the sand layer. A corresponding vertical profile of the motion counts is shown in the lower right plot.

10.4   Summary

In this chapter, exemplary analyses of the 3-D flow data to study the influence of turbulent free surface flow on the pore flow within the gravel layer have been presented. Statistical evaluations based on spatial averaging resp. double averaging over space and time have been carried out. The difficulties in applying such approaches to Lagrangian 3-D PTV data have been outlined.

The observed flow regimes have been characterized by their Reynolds numbers. It has been shown that the dominating flow regime in the pore flow is a time-dependent transitional one, with periods of quasi-stationary flow interrupted by turbulent bursts. In the upper grain layers and for pore flows induced by large surface waves, turbulence is also observed in the pore spaces.

The 3-D PTV technique has been verified by estimating the shear velocity u_* and the equivalent sand roughness k_s from curve fits of the logarithmic wall law and the Reynolds stress profile. The obtained values agree very well with those determined independently from roughness parameters and ADCP data.

Profiles of intrinsic velocity, rms-velocity and Reynolds stress in the pore flow have been computed as a function of the vertical coordinate through the gravel layer. These profiles show the strong interaction of the free surface flow and the pore flow in the upper grain layers. Velocity fluctuations originating from near-wall turbulence of the open-channel flow penetrate into the gravel layer and thereby induce exchange of mass and momentum. As a result, turbulent velocity fluctuations much larger than the mean values occur not only in the near-wall region of the open-channel flow, but also in the upper gravel grain layers. These results are in good agreement with the pressure measurements acquired simultaneously by Detert et al. [2005] and with results obtained in different experiments by Vollmer [2005].

Chapter 11

Conclusion and outlook

In this thesis a novel technique for three-dimensional flow measurements within a permeable gravel layer has been presented. This includes both the experimental realisation and the development of an algorithmic framework for 3-D Particle-Tracking Velocimetry. For the first time, three-dimensional measurements of the pore flow within a gravel layer, driven by a turbulent open-channel flow on top, have been accomplished with a high temporal and spatial resolution. This final chapter summarizes the achievements and discusses perspectives of future work.

11.1   Summary

The following major achievements have been presented in this thesis.

First, it has been shown that precise quantitative 3-D measurements of highly dynamic processes can be obtained from endoscopic stereo imagery, in spite of limitations concerning resolution and image quality. This has been made possible by a careful choice of the experimental approach and a thorough design of both the experimental setup and the image processing algorithms. A detailed survey of current state-of-the art techniques for quantitative flow visualization (chapter 3) has shown that 3-D PTV is the method that is best suited for the presented application, namely flow measurements within a gravel layer.

A 3-D PTV algorithm (chapter 7) has been adapted to the particular challenges of endoscopic imaging. Image preprocessing methods have been implemented that take into account the particular noise structure of the applied CMOS cameras (chapter 4). As a result, systematic noise patterns could be significantly reduced. This improves both subpixel accuracy and the performance of optical flow algorithms. Nevertheless, the possibility of blunders in the feature extraction from endoscopic images cannot be ruled out. This problem is of particular significance for the geometric camera calibration. Hence, robust camera calibration methods have been introduced (chapter 5), which can tolerate up to 50% of corrupted data and still obtain the correct results.

The most difficult part of the 3-D PTV is the particle-tracking in image space. To identify the particle images, a segmentation-from-motion approach has been combined with a watershed transformation to separate overlapping particle images. The tracking is performed within a Kalman filter framework that integrates different low-level motion estimators to initialize particle trajectories. A subsequent 3-D reconstruction is based on a stereoscopic correspondence analysis of the particle trajectories, constrained by the epipolar geometry of the stereo rig (chapter 6).

A performance analysis on synthetic and real test sequences (chapter 8) has shown that typically 50% of the segmented particles can be reconstructed in 3-D with a relative accuracy in object space of $\approx 1 : 300$. Due to the high correlations of systematic calibration errors, the accuracy of the velocity measurements is much higher than the accuracy of single 3-D coordinates. The relative velocity error for the typical velocity ranges studied in the experiments is below 10%.

Second, an extended experimental setup for flow measurements within a gravel layer and at its interfaces with an open-channel flow on top and a sand layer below has been presented. All important issues of this unique setup have been addressed in chapter 9. For the first time, a synoptic investigation of the flow fields within and above a gravel layer becomes possible.

Third, the applicability of the setup has been demonstrated in an extensive series of systematic flow measurements. These experiments have been carried out in cooperation with the Institute for Hydromechanics of the University of Karlsruhe. They are part of an international research project initiated by the Federal Waterways Engineering and Research Institute (Bundesanstalt für Wasserbau) in Karlsruhe. The main issue of this project is to study the interaction of turbulent open-channel flow with the pore flow in a gravel bed, in particular with respect to processes that initiate the motion of sand and gravel grains. The long-term goal is to improve hydraulic engineering design criteria for bed protection of waterways and to validate flow and sediment transport models.

First results of the measurements have been presented, which clearly show the penetration of velocity fluctuations originating in the near-wall turbulence into the upper grain layers, with fluctuation amplitudes significantly exceeding the mean values. Current approaches to describe the destabilization of single grains only take into account the forces induced by the flow above the grain. The results obtained here show that the flow field below the grain is also relevant.

11.2   Outlook

The data basis obtained in the measurements provides a wealth of information for further analyses. A wide range of different flow conditions, including surface waves and flow around an obstacle, has been studied. The main subject of further work will be the comprehensive hydromechanic analysis of this data. This may include correlations with the simultaneously obtained pressure signals, Lagrangian analyses of the flow fields or conditional sampling techniques to identify coherent structures. Further, a detailed comparison of the data with numerical simulations and theoretical flow models should be made.

Apart from the further evaluation of the measurement results, the following improvements of the algorithms are suggested:

- In the current implementation, linear methods are used both for camera calibration and triangulation. The accuracy is sufficient to obtain reasonable results, but may be improved upon, especially for the endoscope setups. The best known approach to compute 3-D coordinates is a photogrammetric bundle adjustment. The bundle adjustment estimates an optimal solution of the set of camera parameters and the 3-D coordinates of object points simultaneously. It provides the flexibility to incorporate arbitrary further correction models to account for systematic deviations. For example, finite-element approaches are used in photogrammetry to correct for sensor unflatness [Hastedt et al., 2002]. Such an approach might also be used to model the systematic deviations between the entry and exit sides of an endoscopic fiber bundle.

- It has been shown that the main factor limiting the total vector yield of the 3-D PTV is the particle-tracking in image space. The latter is very difficult for complex motions at high particle densities. The most powerful approaches to find the particle correspondences are those that take into account an extended temporal scope. The resulting multidimensional assignment problems are NP-complete, but can be solved approximately by heuristic techniques of combinatorial optimization. Within such a framework, temporal and spatial correspondences may also be solved simultaneously, similar to the approach proposed by Willneff [2003].

- The randomly distributed Lagrangian data may be regularized in a postprocessing step. For example, if the time-dependence of the pore flow is moderate, a physically-based interpolation may be carried out, which enforces the Stokes equation to fill in the gaps between data points.

Finally, we note that there are many other application areas for quantitative flow visualization in confined geometries, for example in chemical engineering and combustion diagnostics. Currently, the combustion behaviour of spark-ignition IC engines is studied by endoscopic 2-D PIV approaches [Dierksheide et al., 2001]. The endoscope tips of the setup presented in this thesis are resistant to both high pressures and temperatures. Hence, they could possibly also be applied to obtain 3-D flow information of in-cylinder flow fields in IC engines without major modifications of the setup. The main limiting factor in the investigation of these highly turbulent reacting flows might only be the frame rate of the cameras.

Appendix

Appendix A

Multivariate Gaussian distribution

An important probability distribution of an n-dimensional random variable $\boldsymbol{x} = [x_1, x_2, ..., x_n]^T$ with mean $\boldsymbol{\mu} = [\mu_1, \mu_2, ..., \mu_n]^T$ and covariance matrix $\boldsymbol{\Sigma}$ is a multivariate normal or Gaussian distribution:

$$p(x_1, x_2, ..., x_n) = \frac{1}{\sqrt{(2\pi)^n \det(\boldsymbol{\Sigma})}} \exp\left(-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{x} - \boldsymbol{\mu})\right) \tag{A.1}$$

with

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_{x_1}^2 & \sigma_{x_1 x_2} & \cdots & \sigma_{x_1 x_n} \\ \sigma_{x_2 x_1} & \sigma_{x_2}^2 & \cdots & \sigma_{x_2 x_n} \\ . & . & \cdots & . \\ \sigma_{x_n x_1} & \sigma_{x_n x_2}^2 & \cdots & \sigma_{x_n}^2 \end{pmatrix}. \tag{A.2}$$

In the one-dimensional case, (A.1) reduces to

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right). \tag{A.3}$$

The interval $[\mu - \sigma, \mu + \sigma]$ is the *standard confidence interval*. The probability that x falls into the standard confidence interval is about 68.27%. If $\mu = 0$ and $\sigma = 1$, the distribution is called the *standard Gaussian* (or *normal*) *distribution*.

In the following, we consider the case of the position of a 2-D point in an image: $\boldsymbol{p} = [p_x, p_y]^T$. The precision of the point coordinates is given by the covariance matrix

$$\boldsymbol{\Sigma}_p = \begin{pmatrix} \sigma_{p_x}^2 & \sigma_{p_x p_y} \\ \sigma_{p_y p_x} & \sigma_{p_y}^2 \end{pmatrix}. \tag{A.4}$$

Such pairs $(\boldsymbol{p}, \Sigma_p)$ are typically the result of feature extraction algorithms, as for example those described in section 5.1 and section 7.3. The non-diagonal element, the *covariance* $\sigma_{p_x p_y} = \sigma_{p_y p_x}$, measures to which extent the fluctuations of p_x and p_y are related to each other. The covariance is related to the corresponding variances by the correlation coefficient $r_{p_x p_y}$:

$$r_{p_x p_y} = \frac{\sigma_{p_x p_y}}{\sigma_{p_x} \sigma_{p_y}}, \tag{A.5}$$

a

b

c

$\sigma_x = 0.0342, \sigma_y = 0.0503, r_{xy} = -0.00, \sigma'_x = 0.0342, \sigma'_y = 0.0503, \theta = 0.1$

d

$\sigma_x = 0.0330, \sigma_y = 0.0439, r_{xy} = 0.20, \sigma'_x = 0.0316, \sigma'_y = 0.0449, \theta = -17.1$
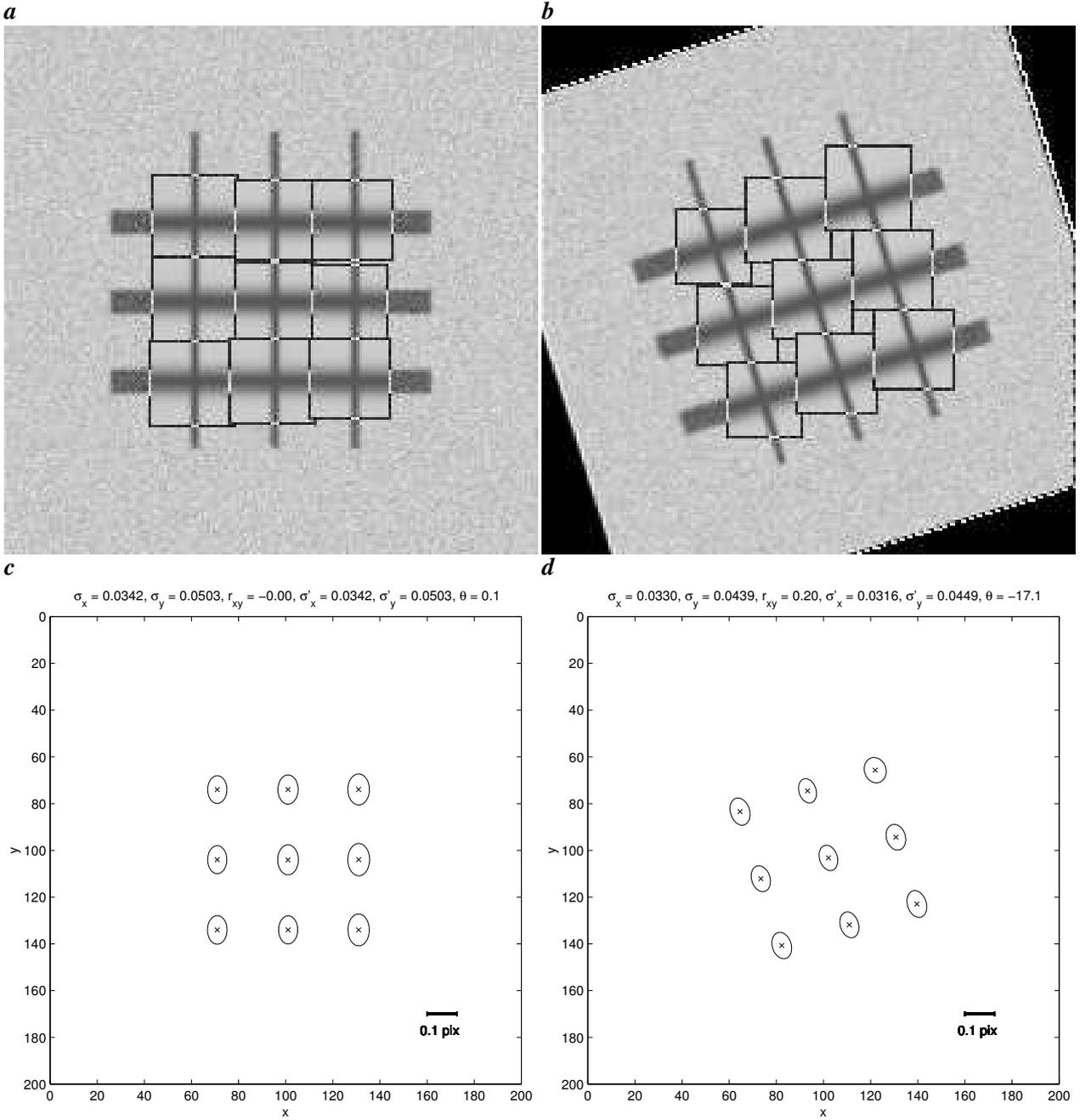

*Figure A.1: Feature extraction using synthetic calibration grid images: **a** and **b** show the results of the subpixel-precise determination of the grid crossings, using the method described in Section 5.1. Image **b** has been obtained by rotating image **a** by an angle $\theta \approx 17°$. **c** and **d** show the corresponding standard error ellipses with the standard deviations σ_x, σ_y and the correlation coefficient r_{xy} in the image coordinate frame (x, y), as well as the standard deviations σ'_x, σ'_y in coordinate frame (x', y') rotated by the angle θ to align the major axis of the ellipses with the x'-axis. In the (x', y')-frame, the correlation between the errors in the coordinates is zero.*

with $|r_{p_x p_y}| \leq 1$. Assuming that the probability distribution of $\boldsymbol{p}$ is Gaussian, the probability density is constant on the ellipses defined by

$$(\boldsymbol{x} - \boldsymbol{p})^T \boldsymbol{\Sigma}_p^{-1} (\boldsymbol{x} - \boldsymbol{p}) = c \tag{A.6}$$

for some positive constant c. This is the equation of an ellipse centered around $\boldsymbol{p}$. From an eigenvector analysis of the positive definite matrix $\boldsymbol{\Sigma}_p$, we can determine the orientation of the ellipse as well as the minor and major axes lengths (diagonalization of $\boldsymbol{\Sigma}_p$):

$$\boldsymbol{\Sigma'}_p = \boldsymbol{U}^T \boldsymbol{\Sigma}_p \boldsymbol{U}, \tag{A.7}$$

where the diagonal elements of

$$\boldsymbol{\Sigma'}_p = \begin{pmatrix} \sigma'^2_{p_x} & 0 \\ 0 & \sigma'^2_{p_y} \end{pmatrix} \tag{A.8}$$

are the eigenvalues and the columns of the rotation matrix $\boldsymbol{U}$ are the eigenvectors of of $\boldsymbol{\Sigma}_p$. In the rotated coordinate system, (A.6) is of the form

$$\frac{(x' - p_x)^2}{\sigma'^2_{p_x}} + \frac{(y' - p_y)^2}{\sigma'^2_{p_y}} = c', \tag{A.9}$$

where

$$\sigma'^2_{p_x} = \frac{\sigma^2_{p_x} + \sigma^2_{p_y} + \sqrt{(\sigma^2_{p_x} - \sigma^2_{p_y})^2 + 4\sigma^2_{p_x p_y}}}{2}, \tag{A.10}$$

$$\sigma'^2_{p_y} = \frac{\sigma^2_{p_x} + \sigma^2_{p_y} - \sqrt{(\sigma^2_{p_x} - \sigma^2_{p_y})^2 + 4\sigma^2_{p_x p_y}}}{2}, \tag{A.11}$$

and

$$\tan 2\theta = \frac{2\sigma_{p_x p_y}}{\sigma^2_{p_x} - \sigma^2_{p_y}}. \tag{A.12}$$

The length of the major axis of the ellipse is given by $\sqrt{c'}\sigma'_{p_x}$, the length of the minor axis is given by $\sqrt{c'}\sigma'_{p_y}$, and θ is the angle between the $x-$axis and the $x'-$axis. Statistically, there is a chance of 38% that the true position of the point will fall within the error ellipse defined by $c' = 1$. This ellipse is called the *standard ellipse* or *standard confidence region*. Similarly, there is a 90% and 99% probability that the true position would fall within the ellipses defined by $c' = 4.6$ and $c' = 9.2$ respectively.

Figure A.1a,b shows the results of the feature extraction algorithm of Section 5.1, applied to synthetic images of a calibration grid. The extracted grid positions and their standard error ellipses are shown in Figure A.1c,d. The width of the grid horizontal grid lines is very large, which results in a larger uncertainty in the $y-$direction than in the $x-$direction,as can be seen from the error ellipses, which are elongated along the $y-$axis. In Figure A.1b, the grid has been rotated by an angle of $\approx 17°$. This rotation introduces a correlation ($r_{xy} = 0.2$) between the errors in the $x-$ and $y-$coordinates, thus the standard error ellipses are rotated by the same angle.

Bibliography

Y. I. Abdel-Aziz and H. M. Karara. Direct linear transformation into object space coordinates in close-range photogrammetry. In *Proc. Symp. Close-Range Photogrammetry*, pages 1–18, Urbana, 1971. Univ. of Illinois at Urbana-Champaign.

A. A. Adamczyk and L. Rimai. Reconstruction of a 3-dimensional flow field from orthogonal views of seed track video images. *Experiments in Fluids*, 6:380–386, 1988.

R. J. Adrian. Particle-imaging techniques for experimental fluid mechanics. *Annu. Rev. Fluid Mech.*, 23: 261–304, 1991.

R. J. Adrian, C. D. Meinhart, and C. D. Tomkins. Vortex organization in the outer region of the turbulent boundary layer. *J. Fluid Mech.*, 422:1–54, 2000.

J. Agui and J. Jimenez. On the performance of particle tracking. *J. Fluid Mech.*, 185:447–468, 1987.

Y. Altunbasak, R. M. Mersereau, and A. J. Patti. A fast parametric motion estimation algorithm with illumination and lens distortion correction. *IEEE Trans. Image Processing*, 12(4):395–408, 2003.

P. Anandan. A computational framework and an algorithm for the measurement of visual motion. *Intern. J. Comput. Vis.*, 2:283–310, 1989.

X. Armangue, J. Salvi, and J. Batlle. A comparative review of camera calibrating methods with accuracy evaluation. *Pattern Recognition*, 35(7):1617–1635, 2002.

K. B. Atkinson, editor. *Close Range Photogrammetry and Machine Vision*. Whittles Publishing, Caithness, Scotland, UK, 1996.

A. Bab-Hadiashar and D. Suter. Robust optic flow computation. *International Journal of Computer Vision*, 29(1):59–77, 1998.

S. J. Baek and S. J. Lee. A new two-frame particle tracking algorithm using match probability. *Experiments in Fluids*, 22:23–32, 1996.

E. P. Baltsavias. *Multiphoto Geometrically Constrained Matching*. PhD thesis, University of Zürich, 1991.

Y. Bar-Shalom. *Tracking and data association*, volume 179 of *Mathematics in Science and Engineering*. Academic Press Professional, Inc., San Diego, CA, USA, 1987.

J. L. Barron, D. J. Fleet, and S. S. Beauchemin. Performance of optical flow techniques. *Intern. J. Comput. Vis.*, 12(1):43–77, 1994.

A. Bartoli. A unified framework for quasi-linear projective, affine and metric bundle adjustment and pose estimation. In *Proc. 16th ICPR*, volume 2, pages 560–563, Quebec City, Canada, 2002.

R. J. M. Bastiaans, G. A. J. van der Plas, and R. N. Kieft. The performance of a new PTV algorithm applied in super-resolution PIV. *Experiments in Fluids*, 32:346–356, 2002.

A. Bayer. *X-ray attenuation techniques to explore the dynamics of water in porous media*. Phd thesis, University of Heidelberg, 2005.

J. Bear. *Dynamics of Fluids in Porous Media*. Dover Publications, New York, 1988.

S. S. Beauchemin and J. L. Barron. The computation of optical flow. *ACM Computing Surveys*, 27(3): 433–467, 1995.

H. Beyer. *Geometric and radiometric analysis of a CCD-camera based photogrammetric close-range system*. PhD thesis, University of Zürich, 1992.

R. B. Bird, W. E. Stewart, and E. N. Lightfoot. *Transport phenomena*. John Wiley and Sons, New York, 1960.

S. Blackman and R. Popoli. *Design and Analysis of Modern Tracking Systems*. Artech House, Boston, London, 1999.

G. Bollrich. *Technische Hydromechanik*. Verlag für Bauwesen, Berlin, 1996.

H. Bopp and H. Krauss. An orientation and calibration method for non-topographic applications. *Photogrammetric Engineering and Remote Sensing*, 44(9):1191–1196, 1978.

O. J. Borchers. Bestimmung von Grösse und Geschwindigkeit der dispersen Phase in Gas-Flüssig-Strömungen mittels digitaler Bildverarbeitung. Diploma thesis, University of Heidelberg, 1997.

B. S. Boufama and R. Mohr. A stable and accurate algorithm for computing epipolar geometry. *Int. Journal of Pattern Recognition and Artificial Intelligence*, 12(6):817–840, 1998.

W.-P. Breugem. *The influence of wall permeability on laminar and turbulent flows*. Phd thesis, Delft University of Technology, 2004.

H. C. Brinkman. A calculation of the viscous force exerted by a flowing fluid on a dense swarm of particles. *Appl. Sci. Res.*, 1:27–34, 1947.

E. Brookner. *Tracking and Kalman Filtering made easy*. John Wiley and Sons, Inc., New York, 1998.

D. C. Brown. Decentering distortion of lenses. *Photogrammetric Engineering*, 32(3):444–462, 1966.

D. C. Brown. Close-range camera calibration. *Photogrammetric Engineering*, 37(8):855–866, 1971.

M. Z. Brown, D. Burschka, and G. D. Hager. Advances in computational stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(8):993–1008, 2003.

F. Carosone, A. Cenedese, and G. Querzoli. Recognition of partially overlapped particle images using the Kohonen neural network. *Experiments in Fluids*, 19:225–232, 1995.

T. P. Chang, N. A. Wilcox, and G. B. Tatterson. Application of image processing to the analysis of three-dimensional flow fields. *Opt. Eng.*, 23(3):283–287, 1984.

T. P. K. Chang, A. T. Watson, and G. B. Tatterson. Image processing of tracer particle motions as applied to mixing and turbulent flow. I. The technique. *Chem. Eng. Sci.*, 40(2):269–275, 1985a.

T. P. K. Chang, A. T. Watson, and G. B. Tatterson. Image processing of tracer particle motions as applied to mixing and turbulent flow. II. Results and discussion. *Chem. Eng. Sci.*, 40(2):277–285, 1985b.

Y.-L. Chang and J. K. Aggarwal. Line correspondences from cooperating spatial and temporal grouping processes for a sequence of images. *Computer Vision and Image Understanding*, 67(2):186–201, 1997.

S. Chen and G. D. Doolen. Lattice Boltzmann method for fluid flows. *Ann. Rev. Fluid Mech.*, 30:329–364, 1998.

D. Chetverikov and J. Verestoy. Feature point tracking for incomplete trajectories. *Computing*, 62(4):321–338, 1999.

T. A. Clarke and J. G. Fryer. The development of camera calibration methods and models. *Photogrammetric Record*, 16(91):51–66, 1998.

T. A. Clarke and X. Wang. The control of a robot end-effector using photogrammetry. In *IAPRS*, volume 23, Amsterdam, 2000.

T. A. Clarke, S. Robson, and J. Chen. A comparison of three techniques for the 3-d measurement of turbine blades. In *ISMTII*, volume SPIE 2101, pages 1–12, 1993.

T. A. Clarke, J. F. Fryer, and X. Wang. The principal point and CCD cameras. *Photogrammetric Record*, 16(92):293–312, 1998.

N. J. Clifford, J. R. French, and J. Hardisty, editors. *Turbulence. Perspectives on flow and sediment transport.* John Wiley and Sons, Chichester, England, 1993.

W. J. Cook, W. H. Cunningham, W. R. Pulleyblank, and A. Schrijver. *Combinatorial Optimization.* John Wiley and Sons, Inc., New York, 1998.

E. R. Corino and R. S. Brodkey. A visual investigation of the wall region in turbulent flow. *J. Fluid Mech.*, 37:1–30, 1969.

K. Cornelis, M. Pollefeys, and L. Van Gool. Lens distortion recovery for accurate sequential structure and motion recovery. In *Proc. ECCV*, volume 2351 of *Springer LNCS*, pages 186–200, 2002.

E. Cowen and S. Monismith. A hybrid digital particle tracking velocimetry technique. *Experiments in Fluids*, 22:199–211, 1997.

I. J. Cox. A review of statistical data association techniques for motion correspondence. *International Journal of Computer Vision*, 10(1):53–66, 1993.

I. J. Cox and S. L. Hingorani. An efficient implementation of Reid's multiple hypothesis tracking algorithm and its evaluation for the purpose of visual tracking. *PAMI*, 18(2):138–150, 1996.

M. M. Cui and R. J. Adrian. Refractive index matching and marking methods for highly concentrated solid-liquid flows. *Experiments in Fluids*, 22:261–264, 1997.

J. H. Cushman, editor. *Dynamics of Fluids in Hierarchical Porous Media*. Academic Press, London, 1990.

H. Daebel. Einfluss einer turbulenten Gerinneströmung auf den momentanen Porenwasserdruck in einer Kiessohle. Experimentelle Vertieferarbeit, University of Karlsruhe, 2001.

S. B. Dalziel. Decay of rotating turbulence: some particle tracking experiments. In F. T. M. Nieuwstadt, editor, *Flow Visualization and Image Analysis*. Kluwer Academic Publishers, 1993.

C. L. Dancey, P. Diplas, A. Papanicolaou, and M. Bala. Probability of Individual Grain Movement and Threshold Condition. *Journal of Hydraulic Engineering*, 128(12):1069–1075, 2002.

C. Q. Davis, Z. Z. Karu, and D. M. Freeman. Equivalence of subpixel motion estimators based on optical flow and block matching. In *International Symposium on Computer Vision*, Coral Gables, Florida, 1995.

M. Davis, H.-J. Köhler, and M. A. Koenders. Unsaturated subsoil erosion protection in turbulent flow conditions. *J. Hydr. Res.*, submitted, 2003.

M. Davis, M. A. Koenders, and H.-J. Köhler. Risk analysis of unsaturated, non-cohesive subsoil due to turbulent flow fluctuations. In *Proc. Coastal Engineering 2004*, volume 1, pages 1861–1873, Lisbon, Portugal, 2004.

M. Detert, G. H. Jirka, M. Klar, M. Jehle, B. Jähne, H.-J. Köhler, and T. Wenka. Pressure fluctuations within subsurface gravel bed caused by turbulent open-channel flow. In M. Greco, A. Carravetta, and R. Della Morte, editors, *Riverflow 2004: 2nd International Conference on Fluvial Hydraulics*, Naples, Italy, 2004a. Balkema, Rotterdam.

M. Detert, C. Lang, G. H. Jirka, and V. Weitbrecht. Bericht Nr. 808: Druckfluktuationen im Interstitial bei Belastung durch die Hauptströmung bis zur einsetzenden Destabilisierung. Technical report, Institute for Hydromechanics, University of Karlsruhe, 2004b.

M. Detert, M. Klar, T. Wenka, and G. H. Jirka. Pressure- and velocity-measurements above and within a porous gravel bed at the threshold of stability. In *Proc. Gravel Bed Rivers 2005*, 2005.

S. Deusch, H. Merava, T. Dracos, and P. Rys. Measurement of velocity and velocity derivatives based on pattern tracking in 3D LIF images. *Experiments in Fluids*, 29:388–401, 2000.

S. Devasenathipathy, J. G. Santiago, S. T. Wereley, C. D. Meinhart, and K. Takehara. Particle imaging techniques for microfabricated fluidic systems. *Exp. Fluids*, 34:504–514, 2003.

F. Devernay and O. Faugeras. Straight lines have to be straight. *Machine Vision and Applications*, 13:14–24, 2001.

S. Dey, H. Sarker, and K. Debnath. Sediment threshold under stream flow on horizontal and sloping beds. *Journal of Engineering Mechanics*, 125 (5):545–554, 1999.

T. Dierig. *Gewinnung von Tiefenkarten aus Fokusserien*. PhD thesis, University of Heidelberg, 2002.

U. Dierksheide, P. Meyer, T. Hovestadt, and W. Hentschel. Endoscopic 2D-PIV flow field measurements in IC engines. In *Proc. 4th International Symposium on Particle Image Velocimetry*, 2001.

U. Dierksheide, P. Meyer, T. Hovestadt, and W. Hentschel. Endoscopic 2D particle image velocimetry (PIV) flow field measurements in IC engines. *Experiments in Fluids*, 33:794–800, 2002.

J. Dieter. Messung von Strömungen in der viskosen Grenzschicht an der Wasseroberfläche. Diploma thesis, University of Heidelberg, 1994.

A. Dittrich. *Wechselwirkung Morphologie/Strömung naturnaher Fliessgewässer*. Habilitationsschrift, Universität Karlsruhe, 1997.

A. Dittrich, F. Nestmann, and P. Ergenzinger. Ratio of lift and shear forces over rough surfaces. In *Coherent flow structures in open channels*. John Wiley and Sons Ltd., 1996.

D.-H. Doh, D.-H. Kim, S.-H. Choi, S.-D. Hong, T. Saga, and T. Kobayashi. Single-frame (two-field image) 3-D PTV for high speed flows. *Experiments in Fluids*, 29:S85–S98, 2000.

E. R. Dougherty and R. A. Lotufo. *Hands-on Morphological Image Processing*. SPIE Press, Bellingham, Washington, USA, 2003.

T. Dracos, editor. *Three-Dimensional Velocity and Vorticity Measuring and Image Analysis Techniques*. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1996.

F. A. L. Dullien. *Porous Media - Fluid Transport and Pore Structure*. Academic, New York, 2nd edition, 1992.

U. Ehrbächer. Voruntersuchungen für ein messsystem zur analyse von strömungen in sedimentporen mittels digitaler bildfolgenanalyse. Diploma thesis, University of Heidelberg, 1999.

M. T. El-Melegy and A. A. Farag. Statistically robust approach to lens distortion calibration with model selection. In *Proc. CVPR Workshop*, volume 8, 2003.

D. Engelmann. *3D-Flow Measurement by stereo imaging*. PhD thesis, University of Heidelberg, 2000.

D. Engelmann, M. Stöhr, C. Garbe, and F. Hering. Particle-tracking velocimetry. In B. Jähne, H. Haußecker, and P. Geißler, editors, *Handbook of Computer Vision and Applications*, volume 3, chapter 31, pages 663–697. Academic Press, 1999.

T. Etoh, K. Takehara, and K. Okamoto. The particle mask correlation method. In *Proc. Flow Visualization 8*, pages 283.1–283.11, 1998.

T. E. Faber. *Fluid dynamics for physicists*. Cambridge University Press, Cambridge, 1995.

W. Faig. Calibration of Close-Range Photogrammetric Systems: Mathematical Formulation. *Photogrammetric Engineering and Remote Sensing*, 41(12):1479–1486, 1975.

O. Faugeras. *Three-Dimensional Computer Vision: A Geometric Viewpoint*. MIT Press, Cambridge, MA, 1993.

O. Faugeras, Q.-T. Luong, and T. Papadopoulo. *The Geometry of Multiple Images*. MIT Press, Cambridge, MA, 2001.

O. D. Faugeras, Q.-T. Luong, and S. J. Maybank. Camera self-calibration: Theory and experiments. In *Proc. ECCV*, pages 321–334, 1992.

J. Feng and S. Weinbaum. Lubrication theory in highly compressible porous media: the mechanics of skiing, from red cells to humans. *J. Fluid Mech.*, 422:281–317, 2000.

A. M. Fincham and G. R. Spedding. Low cost, high resolution DPIV for measurement of turbulent fluid flow. *Experiments in Fluids*, 23:449–462, 1997.

J. Finnigan. Turbulence in plant canopies. *Annu. Rev. Fluid Mech.*, 32:519–571, 2000.

G. Florou and R. Mohr. What accuracy for 3D measurements with cameras? In *Proc. ICPR*, pages 354–358, 1996.

W. Förstner. 10 pros and cons against performance characterization of vision algorithms. In *Proc. ECCV Workshop on Performance Characteristics of Vision Algorithms*, Cambridge, UK, 1996.

W. Förstner. Image preprocessing for feature extraction in digital intensity, color and range images. In *Proc. Int. Summer School on Data Analysis and the Statistical Foundations of Geomatics*, Springer Lecture Notes on Earth Sciences, Chania, Crete, Greece, 1998.

B. Fowler, A. E. Gamal, D. Yang, and H. Tian. A Method for Estimating Quantum Efficiency for CMOS Image Sensors. *Proc. SPIE*, 3301:178–185, 1998.

U. Frisch. *Turbulence*. Cambridge University Press, Cambridge, 1995.

R. Frischholz. Motion tracking. In B. Jähne, H. Haußecker, and P. Geißler, editors, *Handbook of Computer Vision and Applications*, volume 3, chapter 15. Academic Press, 1999.

J. G. Fryer, T. A. Clarke, and J. Chen. Lens distortion for simple C-mount lenses. *International Archives of Photogrammetry and Remote Sensing*, 30(5):97–101, 1994.

D. Fuß. *Kombinierte Höhen- und Neigungsmessung von winderzeugten Wasserwellen am Heidelberger Aeolotron*. Phd thesis, University of Heidelberg, 2004.

A. E. Gamal, B. Fowler, H. Min, and X. Liu. Modeling and Estimation of FPN Components in CMOS Image Sensors. *Proc. SPIE*, 3301:168–177, 1998.

C. Garbe. Entwicklung eines Systems zur dreidimensionalen Particle Tracking Velocimetry mit Genauigkeitsuntersuchungen und Anwendung bei Messungen in einem Wind-Wellen Kanal. Diploma thesis, University of Heidelberg, 1998.

C. S. Garbe. *Measuring heat exchange processes at the air-water interface from thermographic image sequence analysis*. Phd thesis, University of Heidelberg, 2001.

M. Garcia, Y. Nino, and F. Lopez. Laboratory observations of particle entrainment into suspension by turbulent bursting. In P. J. Ashworth, S. L. Bennett, J. L. Best, and S. J. McLelland, editors, *Coherent Flow Structures in Open Channels*. John Wiley & Sons Ltd., 1996.

T. Geis, G. Rottenkolber, M. Dittmann, B. Richter, K. Dullenkopf, and S. Wittig. Endoscopic PIV-Measurements in an Enclosed Rotor-Stator System with Pre-Swirled Cooling Air. In *Proc. 11th Int. Symp. on Application of Laser Techniques to Fluid Mechanics*, Lisbon, Portugal, 2002.

A. Gelb, J. F. Kasper, R. A. Nash, C. F. Price, and A. A. Sutherland. *Applied Optimal Estimation*. The M.I.T. Press, Cambridge, Massachusetts, 1992.

O. v. Genabeek. *Velocity fluctuations in slow flow through porous media*. Phd thesis, Department of Earth, Atmospheric and Planetary Sciences, MIT, 1992.

P. Gill, W. Murray, and M. Wright. *Practical Optimization*. Academic Press, 1981.

J. Gindele and U. Spicher. Investigation of in-cylinder flow inside IC engines using PIV with endoscopic optics. In *Proc. 9th Int. Symp. on Application of Laser Techniques to Fluid Mechanics*, Lisbon, Portugal, 1998.

R. Godding. Geometric Calibration and Orientation of Digital Imaging Systems. Technical report, AICON 3D Systems GmbH, 1998.

S. Gold, A. Rangarajan, C.-P. Lu, and S. Pappu. New algorithms for 2D and 3D point matching: pose estimation and correspondence. *Pattern Recognition*, 31:1019–1031, 1998.

G. H. Golub and C. F. van Loan. *Matrix Computations*. The Johns Hopkins University Press, Baltimore and London, 3rd edition, 1996.

G. H. Granlund and H. Knutsson. *Signal processing for Computer Vision*. Kluwer Academic Publishers, Norwell, MA, 1995.

A. J. Grass. Structural features of turbulent flow over smooth and rough boundaries. *J. Fluid Mech.*, 50: 233–255, 1971.

A. J. Grass, R. J. Stuart, and M. Mansour-Tehrani. Vortical structures and coherent motion in turbulent flow over smooth and rough boundaries. *Phil. Trans. R. Soc. Lond. A*, 336:35–65, 1991.

H. Gröning. *Radiometrische Kalibrierung und Charakterisierung von CCD und CMOS Bildsensoren und monokulares 3D-Tracking in Echtzeit*. PhD thesis, University of Heidelberg, 2003.

A. Gruen. Fundamentals of videogrammetry - a review. *Human Movement Science*, 16:155–187, 1997.

A. Gruen and T. S. Huang, editors. *Calibration and Orientation of Cameras in Computer Vision*. Springer, Berlin, Heidelberg, 2001.

A. W. Gruen. Adaptive least squares correlation: a powerful image matching technique. *S Afr J of Photogrammetry, Remote Sensing and Cartography*, 14(3):175–187, 1985.

Y. G. Guezennec and N. Kiritsis. Statistical investigation of errors in particle image velocimetry. *Experiments in Fluids*, 10:138–146, 1990.

Y. G. Guezennec, R. S. Brodkey, N. Trigui, and J. C. Kent. Algorithms for fully automated three-dimensional particle tracking velocimetry. *Experiments in Fluids*, 17:209–219, 1994.

L. Gui and W. Merzkirch. A comparative study of the MQD method and several correlation-based PIV evaluation algorithms. *Experiments in Fluids*, 28:36–44, 2000.

L. Gui and S. T. Wereley. A correlation-based continuous window-shift technique to reduce the peak-locking effect in digital PIV evaluation. *Experiments in Fluids*, 32:506–517, 2002.

L. Gui, S. T. Wereley, and S. Y. Lee. Digital filters for reducing background noise in micro-PIV measurements. In *Proc. 11th Int. Symp. on Application of Laser Techniques to Fluid Mechanics*, Lisbon, Portugal, 2002.

L. C. Gui and W. Merzkirch. A method for tracking ensembles of particle images. *Experiments in Fluids*, 21:465–468, 1996.

S. Hahn, J. Je, and H. Choi. Direct numerical simulation of turbulent channel flow with permeable walls. *J. Fluid Mech.*, 450:259–285, 2002.

Y. Hardalupas, K. Pantelides, and J. H. Whitelaw. Particle tracking velocimetry in swirl-stabilised burners. *Experiments in Fluids*, 29:S220–S226, 2000.

D. P. Hart. Piv error correction. *Experiments in Fluids*, 29:13–22, 2000a.

D. P. Hart. Super-Resolution PIV by Recursive Local-Correlation. *Journal of Visualization*, 3(2), 2000b.

R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, Cambridge, UK, 2000.

R. I. Hartley. In defence of the 8-point algorithm. *PAMI*, 19(6), 1997a.

R. I. Hartley. Minimizing algebraic error. In *Proc. DARPA Image Understanding Workshop*, pages 631–637, 1997b.

R. I. Hartley. Minimizing algebraic error in geometric estimation problems. In *Proc. ICCV*, 1998.

R. I. Hartley and R. Kaucic. Sensitivity of calibration to principal point position. In *Proc. 7th ECCV*. Springer LNCS, 2002.

R. I. Hartley and T. Saxena. The cubic rational polynomial camera model. In *Proc. DARPA Image Understanding Workshop*, pages 649–654, 1997.

R. I. Hartley and P. Sturm. Triangulation. *Computer vision and image understanding*, 68(2):146–157, 1997.

Y. A. Hassan and R. E. Canaan. Full-field bubbly flow velocity measurements using a multiframe particle tracking technique. *Experiments in Fluids*, 12:49–60, 1991.

H. Hastedt, T. Luhmann, and W. Tecklenburg. Image-variant interior orientation and sensor modelling of high-quality digital cameras. In *ISPRS Comm. V*, Corfu, 2002.

H. Haussecker. Mehrgitter-Bewegungssegmentierung in Bildfolgen mit Anwendung zur Detektion von Sedimentverlagerungen. Diploma thesis, University of Heidelberg, 1993.

H. Haußecker and H. Spies. Motion. In B. Jähne, H. Haußecker, and P. Geißler, editors, *Handbook of Computer Vision and Applications*, volume 2, chapter 13. Academic Press, 1999.

H. Haußecker, H. Spies, and B. Jähne. Tensor-based image sequence processing techniques for the study of dynamical processes. In *Proc. Intern. Symp. On Real-time Imaging and Dynamic Analysis*, pages 704–711, Hakodate, Japan, 1998. International Society of Photogrammetry and Remote Sensing, ISPRS, Commision V.

H. W. Haussecker and D. J. Fleet. Computing optical flow with physical models of brightness variation. *PAMI*, 23(6):661–673, 2001.

G. E. Healey and R. Kondepudy. Radiometric CCD Camera Calibration and Noise Estimation. *PAMI*, 16 (3):267–276, 1994.

J. Heikkilä. Geometric camera calibration using circular control points. *PAMI*, 22(10):1066–1077, 2000.

J. Heikkilä and O. Silven. Calibration procedure for short focal length off-the-shelf ccd cameras. *ICPR*, 1996.

J. Heikkilä and O. Silven. A four-step camera calibration procedure with implicit image correction. In *Proc. CVPR*, pages 1106–1112, San Juan, Puerto Rico, 1997.

F. Hering. *Lagrangesche Untersuchungen des Strömungsfeldes unterhalb der wellenbewegten Wasseroberfläche mittels Bildfolgenanalyse*. PhD thesis, University of Heidelberg, 1996.

F. Hering, M. Merle, D. Wierzimok, and B. Jähne. Particle tracking in space time sequences. In *Springer Lecture Notes in Computer Science*, volume 970, 1995a.

F. Hering, M. Merle, D. Wierzimok, and B. Jähne. A robust technique for tracking particles over long image sequences. In *Proc. of ISPRS Intercommission Workshop 'From Pixels to Sequences', Int'l Arch. of Photog. and Rem. Sens.*, 1995b.

F. Hering, D. Wierzimok, C. Leue, and B. Jähne. Particle tracking velocimetry beneath water waves. Part I: visualization and tracking algorithms. *Experiments in Fluids*, 23(6):472–482, 1997.

F. Hering, D. Wierzimok, C. Leue, and B. Jähne. Particle tracking velocimetry beneath water waves. Part II : Water waves. *Experiments in Fluids*, 24(1):10–16, 1998.

S. Heuel. *Statistical Reasoning in Uncertain Projective Geometry for Polyhedral Object Reconstruction*. PhD thesis, University of Bonn, 2002.

V. Hilsenstein. *Design and implementation of a passive stereo-infrared imaging system for the surface reconstruction of water waves*. Phd thesis, University of Heidelberg, 2004.

K. D. Hinsch. Holographic particle image velocimetry. *Meas. Sci. Technol.*, 13:R61–R72, 2002.

J. O. Hinze. *Turbulence*. McGraw-Hill, New York, 2nd edition, 1975.

G. J. C. M. Hoffmans, H. den Adel, and H. J. Verheij. Shear stress concept in granular filters. In *Intern. Symposium on Scour of Foundations*, Melbourne, Australia, 2000.

B. Hofland. Measuring the flow structures that initiate stone movement. In M. Greco, A. Carravetta, and R. Della Morte, editors, *Riverflow 2004: 2nd International Conference on Fluvial Hydraulics*, Naples, Italy, 2004. Balkema, Rotterdam.

B. Hofland. The probability density function for instantaneous drag forces and shear stresses on a bed. *Journal of Hydraulic Engineering*, submitted, 2005.

B. K. P. Horn and B. G. Schunk. Determining optical flow. *Artificial Intelligence*, 17:185–204, 1981.

H. T. Huang, H. F. Fielder, and J. J. Wang. Limitation and improvement of PIV, part I. Limitation of conventional techniques due to deformation of particle image patterns. *Experiments in Fluids*, 15:168–174, 1993a.

H. T. Huang, H. F. Fielder, and J. J. Wang. Limitation and improvement of PIV, part II. Particle image distortion, a novel technique. *Experiments in Fluids*, 15:263–273, 1993b.

P. J. Huber. *Robust Statistics*. John Wiley and Sons, New York, 1981.

V. Hwang. Tracking feature points in time-varying images using an opportunistic selection approach. *Pattern Recognition*, 22:247–256, 1989.

H. M. Jaeger and S. R. Nagel. Physics of granular states. *Science*, 255:1524, 1992.

B. Jähne. *Digitale Bildverarbeitung*. Springer-Verlag, Berlin, Heidelberg, 1997.

B. Jähne. Continuous and digital signals. In Jähne, B., Haußecker, H., and Geißler, P., editors, *Handbook of Computer Vision and Applications*, volume 2, chapter 2. Academic Press, 1999a.

B. Jähne. Multiresolutional signal representation. In Jähne, B., Haußecker, H., and Geißler, P., editors, *Handbook of Computer Vision and Applications*, volume 2, chapter 4. Academic Press, 1999b.

B. Jähne. Interpolation. In Jähne, B., Haußecker, H., and Geißler, P., editors, *Handbook of Computer Vision and Applications*, volume 2, chapter 8. Academic Press, 1999c.

B. Jähne. Neighborhood operators. In Jähne, B., Haußecker, H., and Geißler, P., editors, *Handbook of Computer Vision and Applications*, volume 2, chapter 5. Academic Press, 1999d.

B. Jähne. *Digital Image Processing*. Springer-Verlag, Berlin, Heidelberg, 5th edition, 2002.

B. Jähne and H. Herrmann. Vermessung der Kenngrössen digitaler Kameras. Technical report, IWR Heidelberg, AEON Bildverarbeitung Hanau, 2003.

B. Jähne, H. Haußecker, H. Scharr, H. Spies, D. Schmundt, and U. Schurr. Study of dynamical processes with tensor-based spatiotemporal image processing techniques. In *ECCV*, pages 322–336. Springer, 1998.

D. F. James and A. M. J. Davis. Flow at the interface of a model fibrous porous medium. *J. Fluid Mech.*, 426:47–72, 2001.

J. R. Janesick. *Scientific charge-coupled devices*. The International Society for Optical Engineering, SPIE Press, Bellingham, Washington, USA, 2001.

J. R. Janesick, K. P. Klaasen, and T. Elliott. Charge-coupled-device charge-collection efficiency and the photon-transfer technique. *Optical Engineering*, 26(10):972–980, 1987.

C. Janßen. Ein miniaturisiertes Endoskop-Stereomesssystem zur Strömungsvisualisierung in Kiesbetten. Diploma thesis, University of Heidelberg, 2000.

M. Jehle, M. Klar, H.-J. Köhler, and M. Heibaum. Bewegungsdetektion und Geschwindigkeitsanalyse in Bildfolgen zur Untersuchung von Sedimentverlagerungen. In *Mitteilungen des Instituts für Grundbau und Bodenmechanik*, volume 77. TU Braunschweig, 2004.

J. Jimenez, M. Uhlmann, A. Pinelli, and G. Kawahara. Turbulent shear flow over active and passive porous surfaces. *J. Fluid Mech.*, 442:89–117, 2001.

G. H. Jirka, D. Lehmann, and C. Lang, editors. *Workshop 'Filter- und Interstitialforschung - Strömung und Turbulenz'*. Institut für Hydromechanik, Universität Karlsruhe, Karlsruhe, 2001.

D. Joseph and S. Collins. Modeling, Calibration, and Correction of Nonlinear Illumination-Dependent Fixed Pattern Noise in Logarithmic CMOS Image Sensors. *IEEE Transactions on Instrumentation and measurement*, 51(5):996–1001, 2002.

C. J. Kähler and J. Kompenhans. Fundamentals of multiple plane stereo particle image velocimetry. *Experiments in Fluids*, 29:S70–S77, 2000.

G. Kamberova. Understanding the Systematic and Random Errors in Video Sensor Data. Technical report, GRASP Laboratory, Department of Computer and Information Science, University of Pennsylvania, 1996.

G. Kamberova. The Effect of Radiometric Correction on Multicamera Algorithms. Technical report, GRASP Laboratory, Department of Computer and Information Science, University of Pennsylvania, 1997.

G. Kamberova and R. Bajcsy. Sensor errors and the uncertainties in stereo reconstruction. In *IEEE Workshop on Empirical Evaluation Techniques in Computer Vision, in conjunction with CVPR 1998*, 1998.

K. Kanatani. *Statistical Optimization for Geometric Computation: Theory and Practice*, volume 18 of *Machine Intelligence and Pattern Recognition*. Elsevier, Amsterdam, 1996.

N. Kasagi and K. Nishino. Probing turbulence with three-dimensional particle tracking velocimetry. In *International Symposium on Engineering Turbulence Modelling and Measurements*, pages 299–314, Dubrovnik, 1990.

R. Keane, R. J. Adrian, and Y. Zhang. Super-resolution particle imaging velocimetry. *Meas. Sci. Technol.*, 6:754–768, 1995.

R. D. Keane and R. J. Adrian. Optimization of particle image velocimeters. Part I: Double pulsed systems. *Meas. Sci. Technol.*, 1:1202–1215, 1990.

R. D. Keane and R. J. Adrian. Optimization of particle image velocimeters. Part II: Multiple pulsed systems. *Meas. Sci. Technol.*, 2:963–974, 1991.

R. D. Keane and R. J. Adrian. Theory of cross-correlation analysis of PIV images. *Applied Scientific Research*, 49:191–215, 1992.

R. Kieft. *Mixed convection behind a heated cylinder*. Phd thesis, Technische Universiteit Eindhoven, 2000.

R. N. Kieft, K. R. A. M. Schreel, G. A. J. van der Plas, and C. C. M. Rindt. The application of a 3D PTV algorithm to a mixed convection flow. *Experiments in Fluids*, 33:603–611, 2002.

H.-B. Kim and S.-J. Lee. Performance improvement of two-frame particle tracking velocimetry using a hybrid adaptive scheme. *Meas. Sci. Technol.*, 13:573–582, 2002.

M. Klar. 3-D Particle-Tracking Velocimetry applied to Turbulent Open-Channel Flow over a Gravel Layer. Diploma thesis, University of Heidelberg, 2001.

M. Klar. Untersuchung des Rauschens von CMOS- und CCD-Kameras. Technical report, IWR, University of Heidelberg, 2004a.

M. Klar. BAW-Messungen Sohlstabilität, Teil 2: Auswertung der Sedimentbewegungen. Technical report, IWR, University of Heidelberg, 2004b.

M. Klar, P. Stybalkowski, H. Spies, and B. Jähne. A miniaturised 3-D Particle-Tracking Velocimetry System to measure the Pore Flow within a Gravel Layer. In *11th International Symposium Applications of Laser Techniques to Fluid Mechanics*, Lisbon, Portugal, 2002.

M. Klar, M. Jehle, H. Spies, and B. Jähne. Investigation of structure changes in unsaturated subsoil. In B. Jähne, editor, *Springer LNCS: Image Sequence Analysis to Investigate Dynamic Processes*, chapter 16. Springer-Verlag, in preparation.

S. J. Kline, W. C. Reynolds, F. A. Schraub, and P. W. Runstadler. The structure of turbulent boundary layers. *J. Fluid Mech.*, 30:741–773, 1967.

M. A. Koenders, S. Reymann, and H.-J. Köhler. Theoretical and experimental analysis of flow in a turbulent filter layer. In W. Wolski and J. Mlynarek, editors, *Filters and Drainage in geotechnical and environmental engineering*, Rotterdam, Netherlands, 2000. A. A. Balkema.

H.-J. Köhler. Filterbemessung aus geotechnischer Sicht. In G. H. Jirka, D. Lehmann, and C. Lang, editors, *Workshop 'Filter- und Interstitialforschung - Strömung und Turbulenz'*, Karlsruhe, 2001. Institut für Hydromechanik, Universität Karlsruhe.

H. J. Köhler and M. A. Koenders. Direct visualisation of underwater phenomena in soil-fluid interaction and analysis of the effects of an ambient pressure drop on unsaturated media. *J. Hydr. Res.*, 41(1):69–78, 2003.

K. Koll. *Feststofftransport und Geschwindigkeitsverteilung in Raugerinnen*. Phd thesis, University of Karlsruhe, 2002.

R. Krepki, Y. Pu, H. Meng, and K. Obermayer. A new algorithm for the interrogation of 3D holographic PTV data based on deterministic annealing and expectation minimization optimization. *Experiments in Fluids*, 29:S99–S107, 2000.

P. K. Kundu. *Fluid Mechanics*. Academic Press, San Diego, USA, 1990.

A. La Porta, G. A. Voth, A. M. Crawford, J. Alexander, and E. Bodenschatz. Fluid particle accelerations in fully developed turbulence. *Nature*, 409:1017–1019, 2001.

L. Lading, G. Wigley, and P. Buchhave. *Optical Diagnostics for Flow Processes*. Plenum Press, New York, 1994.

L. D. Landau and E. M. Lifschitz. *Hydrodynamik*. Akademie-Verlag, Berlin, Germany, 1991.

J. M. Lavest, M. Viala, and M. Dhome. Do we really need an accurate calibration pattern to achieve a reliable camera calibration? In *Proc. of the 5th ECCV*, volume 1, pages 158–174, Freiburg, 1998. Springer.

J. M. Lavest, G. Rives, and J. T. Laprest. Underwater camera calibration. In *Proc. 6th ECCV*, Springer LNCS, pages 654–668, 2000.

R. K. Lenz. Linsenfehlerkorrigierte Eichung von Halbleiterkameras mit Standardobjektiven für hochgenaue 3D-Messungen in Echtzeit. In *Proc. 9. DAGM Symposium Mustererkennung*, volume 149 of *Informatik Fachberichte*, pages 212–216, Braunschweig, 1987.

R. K. Lenz and R. Y. Tsai. Techniques for calibration of the scale factor and image center for high accuracy 3-d machine vision metrology. *PAMI*, 10(5):713–720, 1988.

R. K. Lenz and R. Y. Tsai. Calibrating a cartesian robot with eye-on-hand configuration independent of eye-to-hand relationship. *PAMI*, 11(9):916–928, 1989.

F. Lesage, N. Midoux, and M. A. Latifi. New local measurements of hydrodynamics in porous media. *Experiments in Fluids*, 37:257–262, 2004.

C. Leue, P. Geissler, F. Hering, and B. Jähne. Segmentierung von Partikelbildern in der Strömungsvisualisierung. In *Proceedings of 18th DAGM-Symposium Mustererkennung, Informatik Aktuell*, pages 118–129, 1996.

M. Li and J. M. Lavest. Some aspects of zoom lens camera calibration. *PAMI*, 18(11):1105–1110, 1996.

A. Liberzon, R. Gurka, and G. Hetsroni. XPIV-multi-plane stereoscopic particle image velocimetry. *Experiments in Fluids*, 36:355–362, 2004.

H. J. Lin and M. Perlin. Improved methods for thin, surface boundary layer investigations. *Experiments in Fluids*, 25:431–444, 1998.

B. Liu and R. Männer. A linear iterative least-squares method for estimating the fundamental matrix. In *Proc. 7th Int. Symp. on Signal Processing and its applications (IEEE-ISSPA 2003)*, Paris, France, 2003.

Z. Liu, R. J. Adrian, and T. J. Hanratty. Large-scale modes of turbulent channel flow: transport and structure. *J. Fluid Mech.*, 448:53–80, 2001.

M. Loose, K. Meier, and J. Schemmel. CMOS image sensor with logarithmic response and self calibrating fixed pattern noise correction. *Proc. SPIE*, 3410:168–177, 1998.

B. D. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *Proc. Imaging Understanding Workshop*, pages 121–130, 1981.

T. Luhmann. *Nahbereichsphotogrammetrie*. Wichmann Verlag, Heidelberg, 2000.

B. Lüthi. *Some Aspects of Strain, Vorticity and Material Element Dynamics as Measured with 3D Particle Tracking Velocimetry in a Turbulent Flow*. Phd thesis, Swiss Federal Institute of Technology Zürich, 2002.

H.-G. Maas. New developments in multimedia photogrammetry. In A. Grün and H. Kahmen, editors, *Optical 3-D Measurement Techniques*, Karlsruhe, 1995. Wichmann Verlag.

H. G. Maas, A. Gruen, and D. Papantoniou. Particle tracking velocimetry in three-dimensional flows, Part I: Photogrammetric determination of particle coordinates. *Experiments in Fluids*, 15:133–146, 1993.

H.-G. Maas, A. Stefanidis, and A. Gruen. From pixels to voxels: Tracking volume elements in sequences of 3-d digital images. *International Archives of Photogrammetry and Remote Sensing*, 30, 1994.

H.-J. Maas. *Digitale Photogrammetrie in der dreidimensionalen Strömungsmeßtechnik*. PhD thesis, University of Zürich, 1992.

N. A. Malik and T. Dracos. Interpolation schemes for three-dimensional velocity fields from scattered data using Taylor expansions. *J. Comput. Phys.*, 119:231–243, 1995.

N. A. Malik, T. Dracos, and D. Papantoniou. Particle tracking velocimetry in three-dimensional flows, Part II: Particle tracking. *Experiments in Fluids*, 15:279–294, 1993.

N. S. Martys and J. G. Hagedorn. Multiscale modeling of fluid transport in heterogeneous materials using discrete Boltzmann methods. *Materials and Structures*, 35:650–659, 2002.

M. Marxen, P. E. Sullivan, M. R. Loewen, and B. Jähne. Comparison of Gaussian particle center estimators and the achievable measurement density for particle tracking velocimetry. *Experiments in Fluids*, 29: 145–153, 2000.

P. S. Maybeck. *Stochastic Models, Estimation and Control*. Academic, New York, 1979.

S. P. McKenna and W. R. McGillis. Performance of digital image velocimetry processing techniques. *Experiments in Fluids*, 32:106–115, 2002.

P. Meer, D. Mintz, and A. Rosenfeld. Robust regression methods for computer vision: A review. *International Journal of Computer Vision*, 6(1):59–70, 1991.

C. D. Meinhart, S. T. Wereley, and J. G. Santiago. PIV measurements of a microchannel flow. *Exp. Fluids*, 27:414–419, 1999.

E. Memin and P. Perez. Dense estimation and object-based segmentation of the optical flow with robust techniques. *IEEE Transactions on Image Processing*, 7(5):703–719, 1998.

W. Merzkirch. *Flow Visualization, Second Edition*. Academic Press, London, 1987.

L. Mortara and A. Fowler. Evaluations of charge-coupled device (CCD) performance for astronomical use. *Proc. SPIE*, 290:28–33, 1981.

M. Mühlich and R. Mester. The Role of Total Least Squares in Motion Analysis. In *Proc. ECCV*, pages 305–321. Springer, 1998.

M. Mühlich and R. Mester. Subspace methods and equilibration in computer vision. In *Scandinavian Conference on Image Analysis (SCIA 2001)*, 2001.

M. Mühlich, D. Feiden, and R. Mester. A note on error metrics and optimization criteria in 3d vision. In *Workshop Vision, Modeling and Visualization*, pages 149–156. Infix-Verlag, 1999.

K. Mühlmann. *Design und Implementierung eines Systems zur schnellen Rekonstruktion dreidimensionaler Modelle aus Stereobildern*. Phd thesis, University of Mannheim, 2002.

B. F. Murray and D. W. Buxton. *Experiments in the machine interpretation of visual motion*. MIT Press, Cambridge, MA, 1990.

M. Nakagawa, S. A. Altobelli, A. Caprihan, E. Fukushima, and E. K. Jeong. Noninvasive measurements of granular flows by magnetic-resonance-imaging. *Experiments in Fluids*, 16:54–60, 1993.

G. Nemhauser and L. Wolsey. *Integer and Combinatorial Optimization*. John Wiley and Sons, Inc., New York, 1999.

T. Netsch. *Dreidimensionale Particle Tracking Velocimetry*. PhD thesis, University of Heidelberg, 1995.

I. Nezu and H. Nakagawa. *Turbulence in open-channel flows*. A.A.Balkema, Rotterdam, Netherlands, 1993.

F. T. M. Nieuwstadt, editor. *Flow Visualization and Image Analysis*. Kluwer Academic Publishers, 1992.

J. Nikuradse. Strömungsgesetze in rauhen Rohren. *Forsch. Arb. Ing.-Wes.*, 361, 1933.

K. Nishino, N. Kasagi, and M. Hirata. Three-Dimensional Particle Tracking Velocimetry Based on Automated Digital Image Processing. *J. Fluid Eng.*, 111:384–391, 1989.

J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer-Verlag, 1999.

J. Nogueira, A. Lecuona, and P. A. Rodriguez. Local field correction PIV: on the increase of accuracy of digital PIV systems. *Experiments in Fluids*, 27:107–116, 1999.

K. Ogawa, T. Matsuka, S. Hirai, and K. Okazaki. Three-dimensional velocity measurement of complex interstitial flows through water-saturated porous media by the tagging method in the mri technique. *Meas. Sci. Technol.*, 12:172–180, 2001.

K. Ohba, T. Ishihara, and H. Inooka. Model-based velocity estimation using the kalman filter. In *Proc. Flow Visualization 6*, pages 843–847, 1992.

K. Ohmi and H.-Y. Li. Particle tracking velocimetry with new algorithms. *Meas. Sci. Technol.*, 11(6):603–616, 2000.

K. Okamoto, Y. A. Hassan, and W. D. Schmidl. New tracking algorithm for particle image velocimetry. *Experiments in Fluids*, 19:342–347, 1995.

K. Okamoto, S. Nishio, T. Kobayashi, T. Saga, and K. Takehara. Evaluation of the 3D-PIV Standard Images (PIV-STD Project). *J. Visualization*, 3(2):115–124, 2000a.

K. Okamoto, S. Nishio, T. Saga, and T. Kobayashi. Standard images for particle-image velocimetry. *Meas. Sci. Technol.*, 11:685–691, 2000b.

M. G. Olsen and R. J. Adrian. Out-of-focus effects on particle image visibility and correlation in microscopic particle image velocimetry. *Experiments in Fluids*, 29:S166–S174, 2000.

J. Ortiz-Villafuerte, W. D. Schmidl, and Y. A. Hassan. Three-dimensional ptv study of the surrounding flow and wake of a bubble rising in a stagnant liquid. *Experiments in Fluids*, 29:S202–S210, 2000.

D. Papantoniou and T. Dracos. Analyzing 3-d turbulent motions in open channel flow by use of stereoscopy and particle tracking. In *Advances in Turbulence 2*, pages 278–285. Springer, 1989.

F. Pereirra, M. Gharib, D. Dabiri, and M. Modarress. Defocusing PIV: a three component 3-D DPIV measurement technique. Application to bubbly flows. *Experiments in Fluids*, 29:S78–S84, 2000.

R. J. Perkins and J. C. R. Hunt. Particle tracking in turbulent flows. In *Advances in Turbulence 2*, pages 286–291. Springer, 1989.

B. Peuchot. Camera virtual equivalent model: 0.01 pixel detector. *Computerized Medical Imaging and Graphics*, 17(4/5):289–294, 1993.

Photonfocus. *MV-D1024-28 CMOS area scan camera, Camera User's Manual*, 2003. URL `www.photonfocus.com`.

S. B. Pope. Lagrangian pdf methods for turbulent flows. *Annu. Rev. Fluid Mech.*, 26:23–63, 1994.

A. K. Prasad. Stereoscopic particle image velocimetry. *Experiments in Fluids*, 29:103–116, 2000.

W. H. Press, S. A. Teukolsky, W. Vetterling, and B. Flannery. *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, New York, 1992.

R. G. Racca and J. M. Dewey. A method for automatic particle tracking in a three-dimensional flow field. *Experiments in Fluids*, 6:25–32, 1988.

M. Raffel, C. Willert, and J. Kompenhans. *Particle Image Velocimetry. A Practical Guide.* Springer, Berlin, Heidelberg, New York, 1998.

M. Reeves and N. J. Lawson. Evaluation and correction of perspective errors in endoscopic piv. *Experiments in Fluids*, 36:701–705, 2004.

J. E. Rehm and N. T. Clemens. An improved method for enhancing the resolution of conventional double-exposure single-frame particle image velocimetry. *Experiments in Fluids*, 26:497–504, 1999.

K. Robinson. Coherent motions in the turbulent boundary layer. *Ann. Rev. Fluid Mech.*, 23:601–639, 1991.

S. Robson, T. A. Clarke, and J. Chen. The suitability of the Pulnix TM6CN CCD camera for photogrammetric measurement. In *Optical tools for manufacturing and advanced automation*, volume SPIE 2067, pages 66–77, 1993.

W. Rodi. *Turbulence Models and their Application in Hydraulics*. IAHR, Delft, 3rd edition, 1993.

N. Roussell, H.-J. Köhler, and M. A. Koenders. Analysis of erosion protection measures in partially saturated subsoils. In W. Wolski and J. Mlynarek, editors, *Filters and Drainage in geotechnical and environmental engineering*, Rotterdam, Netherlands, 2000. A. A. Balkema.

A. G. Roy, T. Buffin-Belanger, H. Lamarre, and A. D. Kirkbride. Size, shape and dynamics of large scale turbulent flow structures in a gravel bed river. *J. Fluid Mech.*, 500:1–27, 2004.

A. Ruiz, P. E. Lopez-de Teruel, and G. Garcia-Mateos. A note on principal point estimability. In *Proc. 16th ICPR*, 2002.

M. Sahimi. *Flow and Transport in Porous Media and Fractured Rock*. VCH Verlagsgesellschaft mbH, Weinheim, Germany, 1995.

V. Salari and I. K. Sethi. Feature point correspondence in the presence of occlusion. *PAMI*, 12:87–91, 1990.

S. Saleh, J. F. Thovert, and P. M. Adler. Measurement of two-dimensional velocity fields in porous media by particle image displacement velocimetry. *Experiments in Fluids*, 12:210–212, 1992.

F. Scarano and M. L. Riethmuller. Iterative multigrid approach in PIV image processing with discrete window offset. *Experiments in Fluids*, 26:513–523, 1999.

F. Scarano and M. L. Riethmuller. Advances in iterative multigrid PIV image processing. *Experiments in Fluids*, 29:S51–S60, 2000.

H. Scharr. *Optimal Operators in Digital Image Processing*. PhD thesis, University of Heidelberg, 2000.

M. Schaum. Subpixelgenaue Positionsbestimmungen in digitalen Bildern. Diploma thesis, University of Heidelberg, 1993.

A. E. Scheidegger. *Physics of flow through porous media*. University of Toronto Press, Toronto, Canada, 1974.

A. Schimpf, S. Kallweit, and J. B. Richon. Photogrammetric Particle Image Velocimetry. In *5th International Symposium on Particle Image Velocimetry*, 2003.

H. Schlichting and K. Gersten. *Grenzschicht-Theorie*. Springer-Verlag, Heidelberg, Germany, 9th edition, 1997.

M. Schultz. Geometrische Kalibration von CCD-Kameras. Diploma thesis, University of Heidelberg, 1997.

R. Schwab and H.-J. Köhler. Behaviour of near-saturated soils under cyclic wave loading. In *International Symposium on Deformation Characteristics of Geomaterials*, Lyon, France, 2003.

P. Sechet and L. Guennec. Bursting phenomenon and incipient motion of solid particles in bed-load transport. *J. Hydr. Res.*, 37(5):683–696, 1999.

I. K. Sethi and R. Jain. Finding trajectories of feature points in a monocular image sequence. *PAMI*, 9: 56–73, 1987.

J. Shi and C. Tomasi. Good features to track. In *Proc. CVPR*, 1994.

A. Shields. Anwendung der Ähnlichkeitsmechanik und der Turbulenzforschung auf die Geschiebebewegung. In *Mitteilungen der Preuss. Versuchsanst. f. Wasserbau u. Schiffbau*, volume 26, Berlin, 1936.

S.-W. Shih, Y.-P. Hung, and W.-S. Lin. When should we consider lens distortion in camera calibration. *Pattern Recognition*, 28(3):447–461, 1995.

Y. Shimizu, T. Tsujimoto, and H. Nakagawa. Experiment and macroscopic modelling of flow in highly permeable porous medium under free-surface flow. *Journal of Hydroscience and Hydraulic Engineering*, 8(1):69–78, 1990.

B. Shvidchenko and G. Pender. Macroturbulent structure of open-channel flow over gravel beds. *Water Resources Research*, 37(3):709–719, 2001.

A. Singh. *Optic Flow Computation: A Unified Perspective*. IEEE Computer Society Press, Los Alamitos, California, USA, 1991.

C. S. Slama, editor. *Manual of Photogrammetry, Fourth Edition*. American Society of Photogrammetry, Falls Church, USA, 1980.

A. J. Smits and T. T. Lim. *Flow Visualization: Techniques and Examples*. World Scientific Publishing Company, 2000.

P. Soille. *Morphologische Bildverarbeitung*. Springer-Verlag, Berlin, Heidelberg, 1998.

G. R. Spedding and E. J. M. Rignot. Performance analysis and application of grid interpolation techniques for fluid flows. *Experiments in Fluids*, 15:417–430, 1993.

H. Spies. Bewegungsdetektion und Geschwindigkeitsanalyse in Bildfolgen zur Untersuchung von Sedimentverlagerungen und Porenströmungen. Diploma thesis, University of Heidelberg, 1998.

H. Spies. *Range Flow Estimation: The Movement of Deformable Surfaces*. PhD thesis, University of Heidelberg, 2001.

H. Spies and H. Scharr. Accurate optical flow in noisy image sequences. In *ICCV*, pages 587–592, Vancouver, Canada, July 2001.

H. Spies, O. Beringer, H. Gröning, and H. Haussecker. Analyzing Particle Movements at Soil Interfaces. In B. Jähne, H. Haussecker, and P. Geissler, editors, *Handbook of Computer Vision and Applications*, volume 3, pages 699–718. Academic Press, New York, 1999.

H. Spies, H. Haußecker, and H. J. Köhler. Material transport and structure changes at soil-water interfaces. In W. Wolski and J. Mlynarek, editors, *Filters and Drainage in geotechnical and environmental engineering*, Rotterdam, Netherlands, 2000. A. A. Balkema.

C. Steger. Analytical and empirical performance evaluation of subpixel line and edge detection. In K. J. Bowyer and P. J. Phillips, editors, *Empirical evaluation methods in computer vision*, pages 188–210, 1998.

G. P. Stein. Lens distortion calibration using point correspondences. In *Proc. CVPR*, pages 602–609, 1997.

M. Stellmacher and K. Obermayer. A new particle tracking algorithm based on deterministic annealing and alternative distance measures. *Experiments in Fluids*, 28:506–518, 2000.

C. V. Stewart. Robust Parameter Estimation in Computer Vision. *SIAM Review*, 41(3):513–537, 1999.

T. Stoesser, J. Fröhlich, and W. Rodi. Identification of coherent flow structures in open-channel flow over rough bed using Large Eddy Simulation. In *Proc. 30th IAHR*, Thessaloniki, Greece, 2003.

T. Stoesser, W. Rodi, and G. Jirka. Large eddy simulation of flow over rough channel beds. In M. Greco, A. Carravetta, and R. Della Morte, editors, *Riverflow 2004: 2nd International Conference on Fluvial Hydraulics*, Naples, Italy, 2004. Balkema, Rotterdam.

M. Stöhr. Entwicklung dreidimensionaler Particle Tracking Velocimetry zur Messung der Zweiphasenströmung in Gas-Flüssig-Reaktoren. Diploma thesis, University of Heidelberg, 1998.

M. Stöhr. *Analysis of Flow and Transport in Refractive Index Matched Porous Media*. PhD thesis, University of Heidelberg, 2003.

L. D. Stone, C. A. Barlow, and T. L. Corwin. *Bayesian Multiple Target Tracking*. Artech House, Boston, London, 1999.

T. M. Strat. Recovering the camera parameters from a transformation matrix. In *Proc. DARPA Image Understanding Workshop*, pages 264–271, 1984.

H. Stüer. *Investigation of separation on a forward facing step*. Phd thesis, Swiss Federal Institute of Technology Zürich, 1999.

P. Sturm and B. Triggs. A factorization based algorithm for multi-image projective structure and motion. In *Proc. 4th European Conference on Computer Vision*, pages 709–720, Cambridge, England, 1996.

P. F. Sturm and S. J. Maybank. On plane-based camera calibration: A general algorithm, singularities, applications. In *IEEE Conf. Computer Vision & Pattern Recognition*, pages 432–437, 1999.

P. Stybalkowski. Strömungsmessung in Sedimentporen mittel 3-D Particle Tracking Velocimetry. Diploma thesis, University of Heidelberg, 2001.

Y. Suzuki, M. Ikenoya, and N. Kasagi. Simultaneous measurement of fluid and dispersed phases in a particle-laden turbulent channel flow with the aid of 3-D PTV. *Experiments in Fluids*, 29:S185–S193, 2000.

R. Swaminathan and S. K. Nayar. Non-metric calibration of wide-angle lenses and polycameras. *PAMI*, 22 (10), 2000.

K. Takehara, R. J. Adrian, G. T. Etoh, and K. T. Christensen. A Kalman tracker for super-resolution PIV. *Experiments in Fluids*, 29:S34–S41, 2000.

T. Tamaki, T. Yamamura, and N. Ohnishi. Unified approach to image distortion. In *Proc. ICPR2002*, pages 584–587, 2002.

H. Tani, H. Kondo, M. Mori, K. Hishida, and M. Maeda. Development of fiberscope PIV system by controlling diode laser illumination. *Experiments in Fluids*, 33:752–758, 2002.

R. Theunissen, A. Stitou, and M. L. Riethmuller. A novel approach to improve the accuracy of PTV methods. In *Proceedings of the 12th International Symposium on Applications of Laser Techniques to Fluid Mechanics*, 2004.

T. Thormählen, H. Broszio, and I. Wassermann. Robust line-based calibration of lens distortion from a single view. In *Proc. Mirage 2003*, INRIA Rocquencourt, France, 2003.

C. R. Thorne, J. C. Bathurst, and R. D. Hey, editors. *Sediment Transport in Gravel-Bed Rivers*. John Wiley & Sons, Chichester, 1987.

H. Tian and A. E. Gamal. Analysis of Temporal Noise in CMOS Photodiode Active Pixel Sensor. *IEEE Journal of Solid-State Circuits*, 36(1):92–101, 2001.

P. T. Tokumaru and P. E. Dimotakis. Image correlation velocimetry. *Experiments in Fluids*, 19:1–15, 1995.

P. Torr and D. W. Murray. The development and comparison of robust methods for estimating the fundamental matrix. *Int J. Computer Vision*, 24(3):271–300, 1997.

A. A. Townsend. *The structure of turbulent shear flow*. Cambridge University Press, Cambridge, 2nd edition, 1976.

B. Triggs, P. McLauchlan, R. Hartley, and A. Fitzgibbon. Bundle Adjustment - A modern Synthesis. In B. Triggs, A. Zisserman, and R. Szeliski, editors, *Proceedings of the International Workshop on Vision Algorithms: Theory and Practice*, volume LNCS 1883, pages 298–372, Corfu, Greece, 2000. Springer.

D. J. Tritton. *Physical Fluid Dynamics*. Van Nostrand Reinhold Company, Wokingham, England, 1977.

E. Trucco and A. Verri. *Introductory Techniques for 3-D Computer Vision*. Prentice Hall, Upper Saddle River, New Jersey, 1998.

R. Y. Tsai. A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses. *IEEE Journal of Robotics and Automation*, 3(4):323–344, 1987.

H. C. van de Hulst, editor. *Light scattering by small particles*. Dover Publications, New York, 1981.

C. W. H. van Doorne, B. Hof, R. H. Lindken, J. Westerweel, and U. Dierksheide. Time Resolved Stereoscopic PIV in Pipe Flow. Visualizing 3D Flow Structures. In *Proc. 5th International Symposium on Particle Image Velocimetry*, Busan, Korea, 2003.

M. Van Dyke. *An Album of Fluid Motion*. Parabolic Press, Stanford, California, 1982.

S. van Huffel and J. Vandewalle. *The Total Least Squares problem: Computational aspects and analysis*. SIAM (Society for Industrial and Applied Mathematics), Philadelphia, 1991.

C. J. Veenman, M. J. T. Reinders, and E. Backer. Establishing motion correspondence using extended temporal scope. *Artificial Intelligence*, 145(1-2):227–243, 2003.

M. Virant. *Anwendung des dreidimensionalen Particle-Tracking Velocimetry auf die Untersuchung von Dispersionsvorgängen in Kanalströmungen*. Phd thesis, Swiss Federal Institute of Technology Zürich, 1996.

M. Virant and T. Dracos. 3D PTV and its application on Lagrangian motion. *Meas. Sci. Technol.*, 8: 1539–1552, 1997.

S. Vollmer. *Einfluß der Oberflächenströmung auf die permeable Gewässersohle*. Phd thesis, University of Karlsruhe, 2005.

S. Vollmer, F. de los Santos Ramos, H. Daebel, and G. Kühn. Micro scale exchange processes between surface and subsurface water. *J. Hydrol. Res.*, 269:3–10, 2002.

L.-A. von Busse. Untersuchungen zur Messung des Gasflusses durch Blasen an der Wasser-Luft-Grenzfläche unter Einsatz bildverarbeitender Methoden. Diploma thesis, University of Heidelberg, 1997.

G. A. Voth. *Lagrangian acceleration measurements in turbulence at large Reynolds numbers*. Phd thesis, Cornell University, 2000.

I. Wahyudi, A. Montillet, and A. O. A. Khalifa. Darcy and post-Darcy flows within different sands. *Journal of Hydraulic Research*, 40(4):519–525, 2002.

X. Wang and T. A. Clarke. An algorithm for real-time 3-D measurement. In *ISPRS*, volume 31, pages 587–592, 1996.

G.-Q. Wei and S. D. Ma. Implicit and explicit camera calibration: theory and experiments. *PAMI*, 16(5): 469–480, 1994.

J. Weickert. *Anisotropic diffusion in image processing*. Teubner, Stuttgart, 1998.

G. Welch and G. Bishop. An introduction to the Kalman filter. Technical report TR 95-041, Department of Computer Science, University of North Carolina at Chapel Hill, 2001.

J. Weng, P. Cohen, and M. Herniou. Camera calibration with distortion models and accuracy evaluation. *PAMI*, 14(10):965–980, 1992.

J. Weng, N. Ahuja, and T. S. Huang. Optimal Motion and Structure Estimation. *PAMI*, 15(9):864–884, 1993a.

J. Weng, T. S. Huang, and N. Ahuja. *Motion and Structure from Image Sequences*, volume 29 of *Springer Series in Information Sciences*. Springer, Berlin, Heidelberg, New York, 1993b.

M. P. Wernet and A. Pline. Particle displacement tracking technique and Cramer-Rao lower bound error in centroid estimates from CCD imagery. *Experiments in Fluids*, 15:295–307, 1993.

G. A. W. West and T. A. Clarke. A survey and examination of subpixel measurement techniques. In *ISPRS Int. Conf. on Close Range Photogrammetry and Machine Vision*, volume SPIE 1395, pages 456–463, 1990.

J. Westerweel. *Digital Particle Image Velocimetry - Theory and Application*. Delft University Press, Delft, 1993.

J. Westerweel. Efficient detection of spurious vectors in particle image velocimetry data. *Experiments in Fluids*, 16:236–247, 1994.

J. Westerweel. Fundamentals of digital particle image velocimetry. *Meas. Sci. Technol.*, 8:1379–1392, 1997.

J. Westerweel. Theoretical analysis of the measurement precision in particle image velocimetry. *Experiments in Fluids*, 29:S3–S12, 2000.

J. Westerweel, D. Dabiri, and M. Gharib. The effect of a discrete window offset on the accuracy of cross-correlation analysis of digital PIV recordings. *Experiments in Fluids*, 23:20–28, 1997.

D. J. White, W. A. Take, and M. D. Bolton. Measuring soil deformation in geotechnical models using digital images and PIV analysis. In *Proc. 10th International Conference on Computer Methods and Advances in Geomechanics*, pages 997–1002. Balkema, 2001.

B. Wieneke. Stereo-PIV using self-calibration on particle images. In *Proc. 5th International Symposium on Particle Image Velocimetry*, Busan, Korea, 2003.

B. Wieneke. Application of Self-Calibration-Stereo-PIV in enclosed measurement volumes. In *Proc. 12th International Symposium Applications of Laser Techniques to Fluid Mechanics*, Lisbon, Portugal, 2004.

D. Wierzimok. *Messung turbulenter Strömungen unterhalb der windwellenbewegten Wasseroberfläche mittels digitaler Bildverarbeitung*. Diss., Univ. Heidelberg, 1990.

C. E. Willert and M. Gharib. Digital particle image velocimetry. *Experiments in Fluids*, 10:181–193, 1991.

C. E. Willert and M. Gharib. Three-dimensional particle imaging with a single camera. *Experiments in Fluids*, 12:353–358, 1992.

J. Willneff. *A Spatio-Temporal Matching Algorithm for 3D Particle Tracking Velocimetry*. Phd thesis, Swiss Federal Institute of Technology Zürich, 2003.

J. Willneff and A. Gruen. A new spatio-temporal matching algortihm for 3D-Particle Tracking Velocimetry. In *9th International Symposium on Transport Phenomena and Dynamics of Rotating Machinery*, Honolulu, Hawaii, 2002.

J. Willneff and B. Lüthi. Particle Tracking Velocimetry Measurements for Lagrangian Analysis of Turbulent Flows. In *6th Conference on Optical 3-D Measurement Techniques*, volume 2, pages 191–198, Zurich, Switzerland, 2003.

R. G. Willson. *Modeling and calibration of automated zoom lenses*. Ph.d. dissertation, Carnegie Mellon University, Pittsburgh, Pa., 1994.

R. G. Willson and S. A. Shafer. What is the center of the image? *J. Opt. Soc. Am. A*, 11(11), 1994.

P. Wunsch, G. Koegel, K. Arbter, and G. Hirzinger. Kalibrierung eines nichtlinearen, binokularen Hand-Auge Systems. Technischer Bericht IB-515-96-27, Institut für Robotik und Systemdynamik, DLR, 1996.

G. Xu and Z. Zhang. *Epipolar Geometry in Stereo, Motion and Object Recognition*. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1996.

K. Yagoh, K. Ogawara, and S. Iida. The particle tracking method using the kalman filter. In *Proc. Flow Visualization 6*, pages 838–842, 1992.

M. S. Yalin. *Mechanics of sediment transport*. Pergamon Press, Oxford, 1977.

M. S. Yalin. *River Mechanics*. Pergamon Press, Oxford, 1992.

F. Yamamoto, T. Uemura, H. Z. Tian, and K. Ohmi. Three-dimensional PTV based on binary cross-correlation method. *JSME International Journal, Series B*, 36(2):279–284, 1993.

W.-J. Yang, editor. *Handbook of Flow Visualization*. Hemisphere Publishing Corporation, New York, USA, 1989.

A. P. Yarlagadda and A. P. Yoganathan. Experimental studies of model porous media fluid dynamics. *Experiments in Fluids*, 8:59–71, 1989.

U. Zanke. Zum Einfluß der Turbulenz auf den Beginn der Sedimentbewegung. In *Mitteilungen des Instituts für Wasserbau und Wasserwirtschaft*, volume 120. Universität Darmstadt, 2001.

Z. Zhang. Parameter estimation techniques: A tutorial with application to conic fitting. Technical Report 2676, INRIA, 1995.

Z. Zhang. On the epipolar geometry between two images with lens distortion. In *Proc. ICPR*, volume 1, pages 407–411, 1996.

Z. Zhang. Determining the Epipolar Geometry and its Uncertainty: A Review. *Int. Journal of Computer Vision*, 27(2):161–198, 1998.

Z. Zhang. A flexible new technique for camera calibration. *PAMI*, 22(11):1330–1334, 2000.

Z. Zhang and O. Faugeras. *3D Dynamic Scene Analysis*, volume 27 of *Springer Series in Information Sciences*. Springer, Berlin, Heidelberg, New York, 1992.

Z. Zhang, R. Deriche, O. Faugeras, and Q.-T. Luong. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. *Artificial Intelligence*, 78(1-2):87–119, 1995.

Acknowledgements