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
Dipl. Phys. MARC GEESE
born in Bad Soden

Day of oral examination: 8.2.2013

Image Sensor Nonuniformity Correction by a Scene-Based Maximum Likelihood Approach

Referees:
Prof. Dr. BERND JÄHNE
Prof. Dr. KARLHEINZ MEIER

Abstract: Image sensors come with a spatial inhomogeneity, known as Fixed Pattern Noise or image sensor nonuniformity, which degrades the image quality. These nonuniformities are regarded as the systematic errors of the image sensor, however, they change with the sensor temperature and with time. This makes laboratory calibrations unsatisfying. Scene based nonuniformity correction methods are therefore necessary to correct for these sensor errors.

In this thesis, a new maximum likelihood estimation method is developed that estimates a sensor's nonuniformities from a given set of input images. The method follows a rigorous mathematical derivation that exploits the available sensor statistics and uses only well-motivated assumptions. While previous methods need to optimize a free parameter, the new method's parameters are defined by the statistics of the input data. Furthermore, the new method reaches a better performance than the previous methods. Specialized developments that include a row- or column-wise and a combined estimation of the nonuniformity parameters are introduced as well and are of relevance for typical industrial applications.

Finally it is shown that the previous methods can be regarded as simplifications of the newly developed method. This deliberation gives a new view onto the problem of scene based nonuniformity estimation and allows to select the best method for a given application.

Zusammenfassung: Bildsensoren besizten räumliche Inhomogenitäten, auch Fixed Pattern Noise genannt, welche die Bildqualität herabsetzten. Diese Inhomogenitäten können als die (meta-stabilen) systematischen Fehler des Bildensors identifiziert werden und verändern sich im Laufe der Zeit, was laborgestützte Kalibrierungen unzureichend macht. Deswegen sind szenenbasierte Verfahren notwendig um diese Fehler erfolgreich zu korrigieren.

In dieser Arbeit wird eine neue szenenbasierte Maximum Likelihood Schätzung vorgestellt, welche die Inhomogenitäten eines Bildsensors aus einer Anzahl an Eingangsbildern schätzt. Die mathematisch detailliert hergeleitete Methode nutzt die statistischen Informationen des Bildsensors aus und fußt auf trivialen und gut motivierten Annahmen. Während bei Referenzmethoden üblicherweise mindestens ein freier Parameter optimiert werden muss, können die Parameter der neuen Methode direkt aus der Statistik der Eingangsdaten geschätzt werden. Des Weiteren erreicht die neue Methode eine bessere Korrekturrate als die Vergleichsmethoden. Spezialisierte Erweiterungen der Methode enthalten eine zeilen- oder spaltenweise und eine kombinierte Schätzung der Parameter und haben eine große Relevanz für die typischen industriellen Anwendungen.

Die Referenzmethoden lassen sich letztlich als Vereinfachungen der neu entwickelten Methode darstellen. Dies eröffnet einen neuen Blickwinkel auf das Problem der szenenbasierten Schätzung und wird in Zukunft weitere Verbesserungen und die Wahl der richtigen Korrekturmethode für ein gegebenes Problem erleichtern.

Erklärung gemäß § 8(3)b) und c) der Promotionsordnung

Hiermit erkläre ich, dass ich in die vorliegende Dissertation selbstständig verfasst und mich dabei keiner anderen als der von mir ausdrücklich bezeichneten Quellen und/oder Hilfen bedient habe. Des Weiteren bestätige ich hiermit, dass ich an keiner anderen Stelle ein Prüfungsverfahren beantragt, bzw. diese Dissertation in der vorliegenden oder anderer Form bereits anderweitig verwendet oder einer anderen Fakultät als Dissertation vorgelegt habe.

Acknowlegements

I want to thank all the people in my life that supported me while doing this thesis. Especially my supervisors Prof. Bernd Jähne and Paul Ruhnau for their constant support and encouragement during the past years. I also strongly appreciated the help and support from my colleagues and friends at the HCI and at Bosch.

I also want to thank the Robert Bosch GmbH for the funding of this work. Without these fundings the work would not have been possible.

On a personal note, I want to thank my parents Hartmut and Birgit and my girlfriend Angela, whose encouragement helped me a lot.

Contents

1. Introduction

1.1. Motivation

Image sensors come with a spatial inhomogeneity which degrades the image quality and which is known as image sensor nonuniformity or fixed pattern noise (FPN). This thesis presents a new maximum likelihood estimation method that estimates the sensor's nonuniformities from a given set of input images. Most nonuniformity correction methods consider a linear model of the single sensors in compliance with the EMVA1288 standard for camera characterization [1]. The nonuniformities therefore decompose into a photo response nonuniformity (PRNU) and a dark signal nonuniformity (DSNU). These nonuniformity components, especially the DSNU, drift with time and depend on temperature and exposure time, which makes laboratory calibrations unsatisfying or impossible.

A classification of the sensor errors into stochastic and systematic errors identifies the nonuniformities as the (meta-stable) systematic errors of the image sensor. Nonuniformity correction methods (NUC-methods) compensate these systematic errors while the more general denoising techniques differ from this approach as they correct stochastic errors as well. A detailed model of the image itself is often necessary for these general denoising approaches. NUC methods that estimate the nonuniformities from a set of given input images are known as scene based NUC methods and become necessary because of the mentioned changes of the systematic image errors. For example, the video based driver assistance systems have to consider operations at high temperatures over a period of more than 10 years without a possibility for recalibration in a defined environmental setup. Slight drifts in the nonuniformity correction parameters may lead to an unsatisfying image quality after some years of the product being in field.

Nonuniformity degradation may also arise suddenly because of other effects than the temporal drifts, like crosstalk in the analogue signal processing chain. Those effects can also be handled by smart scene based NUC methods to improve the image quality in the given scenarios.

1.2. Related Work

NUC related literature usually concentrates on sensor models [1, 2] but in the early 90s first scene based correction methods were proposed by Scribner [3]. His methods mainly exploit the knowledge of the vertebrate retina at the time. The main

structure of his algorithm has been nearly unchanged until today and the further developments exploit the *retinomorphic* hypothesis of IR-cameras with help of the least mean square algorithm [4]. Recent improvements include a signal gating technique [5] and a generalization by a discrete total variation approach [6]. Another strain of development uses the constant statistics assumption which leads to a simple and effective algorithm that exploits the temporal information of the pixels [7]. The further research directions include covariance and Kalman filters [8, 9], but their basic algorithmic principles are already covered by the constant statistics and least mean square algorithms. The Kalman filters are mainly used to obtain a model for the drifts of the sensor nonuniformities.

While all these methods correct the nonuniformities successfully, their performance can still be improved. Furthermore, a motivation of some of the basic assumptions is missing for these methods (e.g. the way of estimating the original image in [5, 4, 6]). The statistical information of the camera's FPN (e.g. mean and standard deviations) has also not yet been considered. The reference methods are discussed in detail in chapter 5.

For performance analysis, a huge variety of quality metrics is used in literature, which are usually based on the actual image material. This of course leads to different correction performances depending on the used sequences, which is unsatisfying because the nonuniformities are a sensor property and not an image property.

1.3. Contribution of this Work

This thesis derives a new maximum likelihood method for the scene based estimation of image sensor nonuniformities. Thereby the statistical models are motivated by evaluations of ground truth data, which leads to reasonable and understandable assumptions. The method's input finally results as the averaged sensor responses for the DSNU correction and as the corresponding pairwise sensor covariances for the PRNU correction. A rigorous mathematical derivation exploits the continuous light signal's sampling and reconstruction as well as the available statistical data of the sensor parameters. For the DSNU-only variants, the derivation as a maximum likelihood approach allows to estimate all the parameters of the method from the input statistics. This is a huge advantage over the reference methods, which leave free parameters to optimize by the user. Further the derivation demands uncorrelated frames to achieve best correction performances, which is confirmed by the experimental results.

The resulting nonuniformity correction algorithm is subject to further modifications to improve its performance and to minimize its computational complexity. Namely these improvements are done as a motivated weighted averaging of the sensor inputs for the DSNU estimation and by approximating the method by a Chua-Young full range cellular neural network (CNN). A further improvement was achieved by extending the method to a combined row- or column-wise and

a row-, column- and pixel-wise combined correction method. Most of these improvements are especially investigated for the DSNU, as this component of the nonuniformities is in general of more interests in the actual applications.

The derived methods are investigated with three different types of image data: Uncorrelated frames that simulate a slow capture rate, high speed correlated frames and frames with a typical capture rate but low valued nonuniformities. If the demanded assumptions for uncorrelated frames are met, the newly developed methods outperform the related methods in a comparison based on ground truth correction patterns. In general the best correction rates are achieved with the uncorrelated type of input and in cases where the assumptions are not met a general performance decrease can be observed for all methods. However, these scenarios are studied as the limiting cases for the new methods and the correction rates of the reference methods cannot be exceeded in all cases.

To avoid a comparison that bases on the actual image content, a new quality metric is introduced that bases on the estimated nonuniformity correction patterns and the quantization definitions given by the EMVA1288 standard. In the end, a theoretical link between all the presented reference methods from literature against the new maximum likelihood method is formed. This link states assumptions under which the reference methods can be regarded as simplifications of the newly developed approach.

The contents of this thesis have been partly published during the research [10, 11, 12, 13] and the method has a patent pending [14].

1.4. Outline of this Thesis

This thesis is written in a way that the chapters build up all the knowledge needed to understand the follow-up chapters, including the needed physical principles. Efforts were made that should allow to read the individual chapters separately if enough general knowledge of the content of the previous chapters is present.

The thesis starts by providing the very basic background information in chapter 2, explaining mainly the physical backgrounds and the working principles of the common light measurement techniques and image sensors. Next the nonuniformities and their characterization and performance measures are introduced in chapter 3. Here the EMVA1288 nonuniformity characterization is explained as well and forms the basis for a new image data independent performance measure.

Chapter 4 introduces the used image material that is used for the experiments conducted in this thesis. This includes an explanation of the used cameras and the nonuniformity characteristics of their image sensors. With the specified image data and the defined performance characterizations chapter 5 then gives a detailed overview of the related work from literature. The reference methods are explained in detail, including a new matrix-wise notation. First experimental results of these

methods are presented as well in chapter 5 and form the motivation for the development of the new method.

The new maximum likelihood method is derived in its very basic, DSNU-only correction variant in chapter 6. This chapter also shows first experimental results, where the method's performance in its basic form shows potential for improvements. Improving modifications are then introduced in chapter 7, which further includes the CNN approach to solve the maximization task in an efficient way and the combined row-, column- and pixel wise DSNU correction. First experimental results for these improvements are discussed and the new methods outperform the reference methods, even if the statistical estimated parameters are used. To exploit all the theoretical possibilities, the maximum likelihood approach is extended in chapter 8 with a sensor model that allows to correct for the PRNU nonuniformities as well. The experimental results are discussed and the new method performs again superior compared to the reference methods.

To gain further insight into the method's working principles chapter 9 provides a theoretical link between the reference methods towards the new maximum likelihood methods. The necessary assumptions to build this link allow to understand the individual methods' disadvantages and advantages. The last chapter then analyzes the new methods against the reference methods in the limiting case of correlated input data and data with low valued nonuniformities. A summary and an outlook into the future research on the presented topics finishes this thesis in chapter 11.

2. General Background

As the thesis concentrates on the estimation of image nonuniformities, knowledge of the acquisition process of a digital image is needed. This chapter contains a review of the most important physical principles on which this work is based on. The chapter starts by introducing a physical model for light and the possibilities for the measurement of the light intensity. Then the digital image sensors are explained and last a mathematical description of the process of image acquisition and signal reconstruction is presented.

2.1. Images, Photographs, Pixels and Lightsensors

Most modern image processing tasks rely on *the picture* or *the image* as the very basic information. While *images* in their basic definition are artifacts that have a similar appearance to a physical object (e.g. a person), the terms *photograph* or *photo* are defined as an image which is created by light falling onto a light-sensitive area. Usually a lens is used to focus the scene's light into a reproduction that is similar to the perception of what the human eye would see. However, in this thesis we will use the terms images and photographs as synonyms but mean the photographs, that represent the perception of the human eye.

With the appearance of digital image sensors, the light signal is sampled at discrete points. Millions of individual light sensors measure the light intensity and each light sensor then contributes a discrete element to the image. As images are also named pictures, the picture elements that are created by a single light sensor's information are called *pixels*. If the image is stored in a memory on a computer the term pixel then usually refers to the part of memory that stores the information of that light sensor. Thus the terms light sensor and pixel are usually used as synonyms in some parts of the image processing community.

2.2. Physical Background

2.2.1. Definition and Physical Model of Light

Light is usually defined as electromagnetic radiation that is visible to the human eye [15]. The kind of electromagnetic radiation that is responsible for the sense of sight has a wavelength in a range from about $380\,\mathrm{nm}$ to about $740\,\mathrm{nm}$. This means

that the visible light range is bordering the invisible infrared, which is found at longer wavelengths than $740\,\mathrm{nm}$ and the invisible ultraviolet, which is found at shorter wavelengths than $380\,\mathrm{nm}$.

In physics, the term light sometimes refers to electromagnetic radiation of any wavelength, whether visible or not. As the context of this thesis deals with modern digital cameras, the definition of light is extended to electromagnetic radiation that is measurable by these sensors. In general this is an extension to the human eye's visibility into the infrared radiation up to $900\,\mathrm{nm}$.

Electromagnetic radiation, and especially light, obeys the wave-particle duality. This means that light is emitted and absorbed in tiny packets called photons that exhibit properties of particles, while the wave properties are exhibited as well. The study of light is also known as optics and is an important research area in modern physics. A wave property description of electromagnetic radiation can be found directly from an analysis of the Maxwell's equations [15] and the basic quantum mechanics give an overview of the additional particle properties of light.

The connection between the wave model of the Maxwell's equations and the mentioned photons from quantum mechanics are given by the following well known equations:

$$\begin{aligned} W = hf = \frac{hc}{\lambda} \quad \text{and} \quad p = \frac{W}{c} = \frac{hf}{c} = \frac{h}{\lambda} \\ f = \frac{c}{\lambda} \quad \text{and} \quad f = \frac{W}{h} \quad \text{and} \quad W = \frac{hc}{\lambda} \end{aligned} \tag{2.1}$$

The introduced variables are: c as the speed of light, f as the wave frequency and λ as the wavelength. Further were introduced: W as Energy[1], h as the Planck constant and p as the momentum, with W and p referring to the measures for a single photon.

The Light Intensity: The most common measurement of light is its intensity I, which is more correctly called *Irradiance* in the photometric and radiometric area, especially if the disposed energy per area is considered. The intensity is defined via the transported energy and in fact, anything that can carry energy may have an intensity associated with it. For an electromagnetic wave with a complex amplitude of the electric field E, the time-averaged energy density is given as [15]:

$$I\left[\frac{\mathrm{W}}{\mathrm{m}^2}\right] = \frac{cn\epsilon_0}{2}|\mathbf{E}|^2\,, \tag{2.2}$$

where n is the refractive index in which the wave propagates and ϵ_0 is the vacuum permittivity. Thus the light intensity is proportional to the squared amplitude of the electric field E.

[1] W is energy, not to confuse with E, which is the electric field

Considering the particle nature of light, the energy for a single light quantum can be calculated. For the visible light the energy ranges from: $1.6\,\mathrm{eV}$ for red light of $\approx 740\,\mathrm{nm}$ to $3.26\,\mathrm{eV}$ for blue light of $\approx 380\,\mathrm{nm}$. Other quantities of electromagnetic waves like its polarization can be found in the literature [15] and are not of concern for typical imaging systems.

2.2.2. Scene Projection and Optics

To be able to capture the light signal in a way similar to the human eye, a lens (or a lens-system) is used to project the scene onto the 2D image sensor. The basic optics are explained in the physical literature [15, 16] and an ideal lens or an ideal lens system maps all points from a given object plane onto points in an image plane (also called focal plane). Points with a distance to the given object plane are mapped in a blurred way onto the image plane.

The wave character of light explains degradations by diffraction effects by the lenses outer limits and an ideal lens needs to be infinitely large to avoid these degradations. However, besides the diffractive limits, real lenses in general tend to create image degradations also for light that enters with a certain distance to the optical axis of the system. The mainly known optical errors of the lens systems are chromatic and spherical aberration. To avoid these errors apertures can be used to block out the light that causes image degradation, but of course a smaller opening in the aperture leads to more diffraction effects.

A common assumption is that the optics reduce the quality of the focal signal by spatial constant effects [16]. Therefore a point spread function (PSF) that models the optics O_{PSF} can be used to model the degradations of the light signal's transfer onto the image plane. A wavelength dependency for the intended color channels should be considered: $O_{\mathrm{PSF}}(\lambda)$. The mathematical operation that describes the degradation effects is a convolution of the intensity signal I with the point spread function of the optics O_{PSF}. This can be expressed also by a multiplication of the Fourier transformed components in the spatial dimensions:

$$F_{x,y} \circ\!\!-\!\!- \hat{F}_{\hat{x},\hat{y}} \tag{2.3}$$

$$I_{\mathrm{World}} * O_{\mathrm{PSF}} = I_{\mathrm{Focal}} \quad \circ\!\!-\!\!- \hat{I}_{\mathrm{World}} \cdot \hat{O}_{\mathrm{PSF}} = \hat{I}_{\mathrm{Focal}}\,. \tag{2.4}$$

Fig. 2.1 shows a sketch of the optical mapping of a generalized optical system. To describe the scaling and the positions of the mapping, the optical system is defined by the four parameters: Object- and image-wise focal length ($f_{\mathrm{Obj.}}$, $f_{\mathrm{Img.}}$) and principal planes ($H_{\mathrm{Obj.}}$, $H_{\mathrm{Img.}}$) [16, 15]. A first order PSF can be given as the Fourier transform of the size and shape of the used aperture (e.g. an airy disc for a circular aperture). But other spatial homogeneous errors of the optics can also be modeled into it's PSF O_{PSF}, e.g. an out-of-focus adjustment. Due to the limited diameter of the PSF, it acts in general as a low pass filter as well as the effects from an out-of-focus adjustment. As an effect of the convolution of the

world intensity signal with the PSF, the high spatial frequencies are damped by any form of projection. This leads to a blurred signal in the focal plane that misses the highest frequencies and this damping effect is sketched in fig. 2.1 as well.

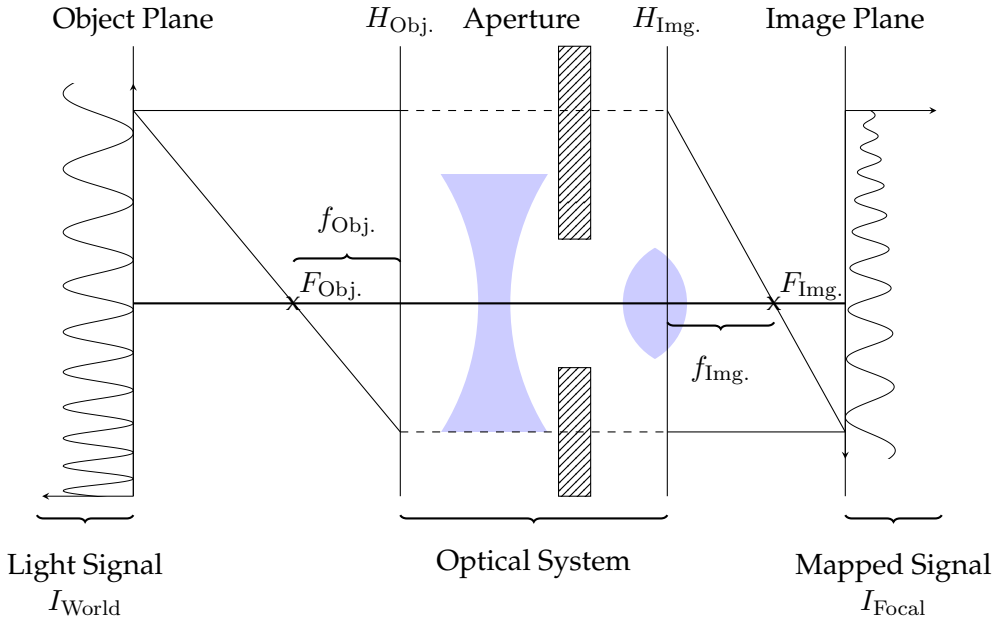


Figure 2.1.: Signal Mapping by an abstract optical system

2.2.3. Measurement Techniques for Light

Given the knowledge of the scene's projection, the measurements for electromagnetic radiation is the next step to be considered. Light measurement techniques exploit the lights interaction with matter and the two most important principles are the inner and outer photoelectric effect. The photoelectric effect in general states that electrons are excited in or emitted from matter as a consequence of the electron's absorption of energy from the electromagnetic radiation (Heinrich Hertz in 1887) [15]. The photons of light have a characteristic energy that is proportional to the frequency of light (see eq. 2.1). Electrons can absorb energy from photons when irradiated and they usually follow an *all or nothing* principle, which means that all the energy of a given photon is transferred to one electron.

The outer photoelectric effect happens if a photon is absorbed by an electron that resides in a material with an electron binding energy that is lower than the photon's energy. Consequently the electron is ejected from the matter, with a kinetic energy that equals the remaining energy after overcoming the binding forces. This effect is used for example in Photomultipliers which are extremely sensitive detectors in the ultraviolet, visible, and near-infrared ranges of the electromagnetic spectrum [15]. The photomultiplier detectors multiply the produced current which is generated by the incident light by a factor 10^8 (i.e. $160\,\text{dB}$) in multiple

dynode stages. This technique allows the detection of individual photons if the incident flux of light is very low.

The inner photoelectric effect on the other hand describes the case in which the electron's acquired energy is not high enough to escape the matter. But the energy may still be high enough to change the electron's energy level inside the matter. In solid state physics it is therefore possible that electrons change their energy band (e.g. from the valence to the conduction band). This realization of the inner photoelectric effect is used in CMOS and CCD sensors as visible light allows the electrons to enter the conduction band. Consequently these electrons contribute to a measurable *photo*-current. The inner and outer photoelectric effects are sketched in fig. 2.2.

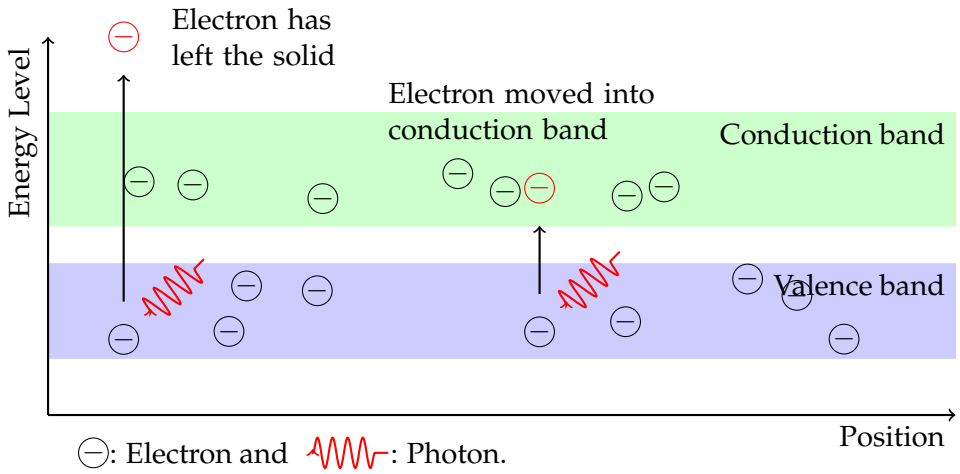


Figure 2.2.: Inner and outer photoelectric effect. A photon hits an electron which is then elevated into a higher energy level (inner effect) or leaves the matter (outer effect).

2.3. Soild State Physics of Silicon

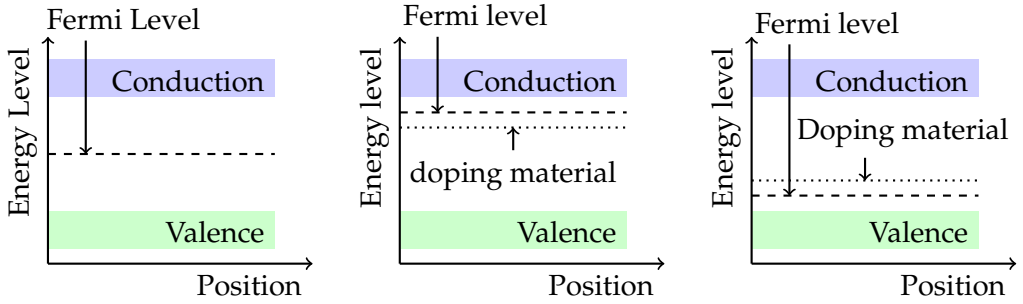
As mentioned, the light intensity is usually measured by exploiting the inner photoelectric effect. This is achieved with help of doped silicon in devices called CMOS or CCD photo diodes. A basic understanding of the properties of silicon and doped silicon, as presented in detail in the physical literature [15, 17], is necessary to understand this light measurement process.

To give a short summary, the solid silicon crystal is a semiconductor which has an energy gap of $\approx 1.1\,\mathrm{eV}$ between its valence and conduction band. The thermal energy per electron at room temperature is just $\approx 25\,\mathrm{meV}$, which is too low to overcome the $\approx 1.1\,\mathrm{eV}$ gap. As only electrons in the conduction band contribute to the conductivity, the silicon has no noteworthy conduction at room temperature. But for the use in electronics the silicon should be conductive which can be achieved

by doping the silicon with impurities, which then result in a shift of the Fermi level. As the Fermi level is the *lowest* allowed energy level, an increase in its value will allow electrons to reach the conduction band at much lower temperatures.

There are two forms of doped silicon: n-type and p-type. For n-type doped silicon, the regular lattice grid of the silicon is disturbed by atoms with one valence electron more than silicon, thus atoms of the 15th period (e.g. Phosphorus) are used. The additional phosphorus brings a Fermi level that lies right below the energy level of the conduction band of silicon and the electrons of the phosphorus can easily enter the conduction band at low temperatures and contribute to the conductivity. Thus the additional electron acts as a negatively loaded charge carrier. In the second variant, the so-called p-type silicon, the lattice grid gets disturbed by atoms of one valence electron less than silicon. These are atoms of the 13th period (e.g. Aluminum) and concerning the crystal structure, the missing electron creates a hole. In this case the Fermi level of the aluminum lies right above the valence band and makes it possible for the valence electrons of the silicon to jump up into this additional bands at very low temperatures. As a consequence the missing electron hole is not fixed and acts as a carrier of a positive charge that contributes to the electric conductivity. Fig. 2.3 shows the energy band structures and the shifted Fermi levels for p- and n-type silicon.

To avoid a destruction in the lattice structure of the crystal, the doped atoms are usually present in the ratio of $1{:}10^6$ only. Summarizing, the doted silicon is conductive in a wide range around room temperature ($300-500\mathrm{K}$). The n-type silicon conducts with negatively charged electrons as charge carriers and the p-type silicon conducts with positively charged holes as charge carriers. However, each kind of silicon is still electrically neutral, as no additional electric charges were added.



(a) pure Silicon (b) n-doped Silicon (c) p-doped Silicon

Figure 2.3.: Shift of the Fermi levels as result of doped silicon.

2.3.1. The P-N-Junction

The most interesting and useful effects in silicon electronics occur if a junction between p-type and n-type silicon is formed, a so-called p-n-junction. This junction

acts as a diode, which is a device that allows an electric current to flow only in one direction . Furthermore, the p-n-junction has the properties to detect light and measure its intensity.

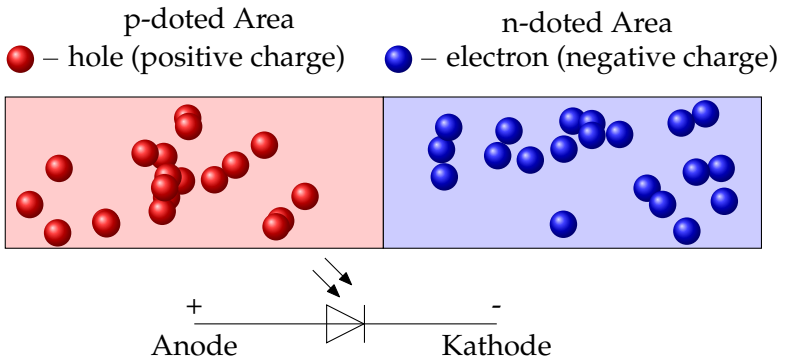


Figure 2.4.: Diode symbol and a p-n silicon junction immediately after junction connection

A simple and limited model of the p-n-junction is given now, while more details can be found in the literature [15, 17]. Fig. 2.4 shows the electric symbol of a diode and the p- and n-type areas (in the very moment the junction is connected). To understand the p-n-junction, one has to consider that the charge carriers are created by the doted materials which has different types of charge carriers in the areas of the p- and n-type silicon. At the junction border, a diffusion of the different charge carriers takes place and electrons diffuse into the p-doted area, while holes move into the n-doted area. As the electrons reach the p-doted area, they result in a negative total charge of the p-doted area, as they violate the electric neutrality of this part of the crystal. The same occurs for the holes that diffuse into the n-doted area. Consequently an electric field builds up between the p- and n-doted areas that pushes the charge carriers against their diffusion direction. In the equilibrium state, the current from the electric field is as large as the current of the diffusion and the voltage between the p- and n-doted crystals is about $0.6V$ at room temperature for common doping parameters.

With regard to conductivity, another effect, called recombination has to be considered. The electrons that moved over the border are attracted to fill the free places of the p-type doping atoms and while the holes take the extra electron from the n-doping atoms. As a result, the diffusing charge carriers are no longer free and don't contribute to the conductivity, meaning that inside the diffusion region the free charge carriers vanish. This zone is called the depletion zone of the p-n-junction. With no available charge carriers in the depletion zone, the conductivity of the p-n-junction is completely interrupted. Fig. 2.5 shows this equilibrium state of the p-n-junction. The application of an external electric field allows now to shrink and finally collapse the depletion zone making the diode conductive. However, if the electric field is applied with reversed polarity, the depletion zone extends and the diode forms an isolator.

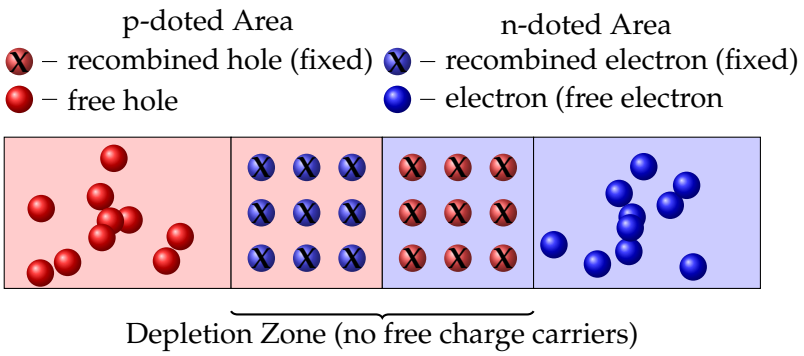


Figure 2.5.: p-n-junction in its equilibrium state

2.3.2. Light Measurement in a CMOS Pixel

This chapter now concentrates on the light measurement ability of the p-n-junction. Most modern image sensors are built upon complementary metal oxide semiconductors (CMOS). The CMOS light sensors and their functionality is now explained in their basic functionality as described in various literature [15, 18]. The other widely spread type of image sensors is based on charge coupled devices (CCD), which have a similar way of measuring the light with help of a p-n-junction. A basic CMOS light sensor has a circuit layout as shown in fig. 2.6.

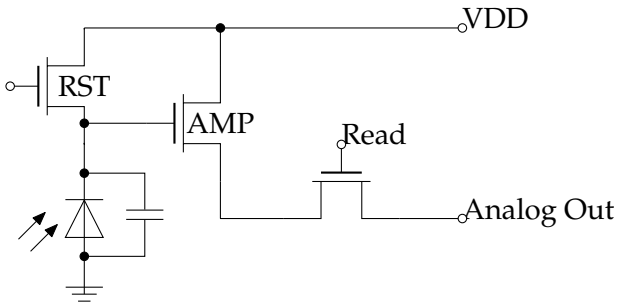


Figure 2.6.: Basic circuit of a CMOS light sensor

To start the light measurement, the *Read* transistor is closed and the *RST* transistor is opened. The photodiode is then in a so-called *reverse-biased* state, which means that the applied voltage extends the depletion zone. This happens as the applied voltage leads to a force on the free charge carriers which pulls them out of the diode system, leading to a decreased charge carrier concentration which allows the diffusion current to extend the depletion zone.

If the *RST* transistor is closed, the diode is isolated on both sides and the depletion zone stays constant. Generally spoken, the depletion zone now builds an isolating layer between two differently charged areas and thus forms a capacitance. Fig. 2.7 shows the extended depletion zone after the RST transistor has been closed.

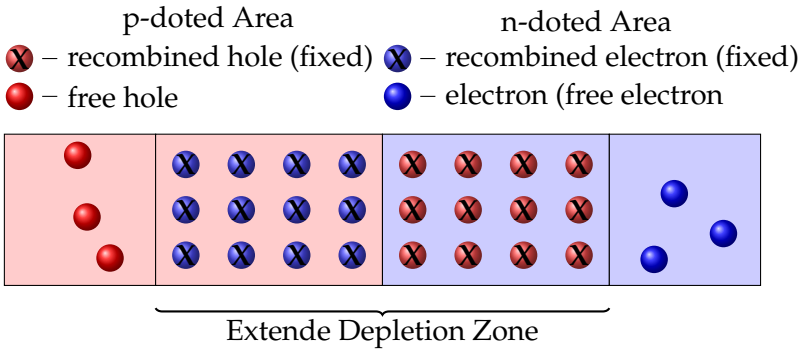


Figure 2.7.: Extendet deplation zone after charging

In this charged stage, the detection of light takes place. Inside the depletion zone, there are no free charge carriers available from the doping material. But it is still possible that electrons from the valence band go into the conduction band if enough energy is transferred. The band gap between valence and conduction band for silicon is $\approx 1.17\,\mathrm{eV}$ at $300\,\mathrm{K}$. An incident photon of visible light has an energy of $1.6\,\mathrm{eV}$ to $3.2\,\mathrm{eV}$ which is enough energy to push an electron into the conduction band. Of course for each electron that leaves the valence band, a new hole is created as well.

If this photon-electron interaction takes place inside the depletion zone, the newly generated free charge carriers are exposed to the present electric field. This electric field will guide the electrons into the n-doted area and the holes into the p-doted area. For a constant intensity, the created charges lead to a constant current that results in a linear voltage decrease of the diode's capacitance. Or in other words, the voltage at the diode is proportional to the number of photons that did hit the diode's depletion zone. Fig. 2.8 illustrates this process.

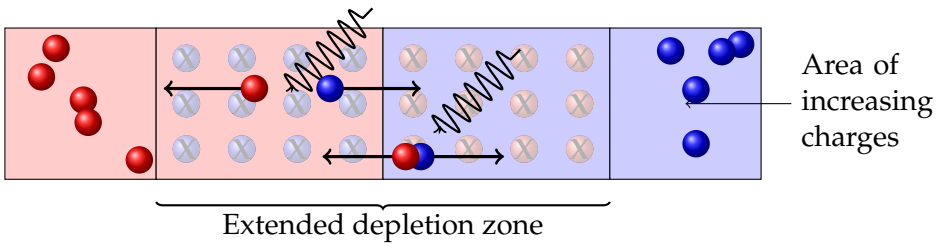


Figure 2.8.: Electric charge generation in the field of the depletion zone

Going back to the circuit design of fig. 2.6, the voltage that is present at the photo diode is amplified with the source following transistor *AMP*. The *Read* transistor then selects the signal of the photo diode and moves it to the *Analog Out* signal from where it will be further amplified and digitalized.

For CCD sensors the physical mechanism of generating charges inside the depletion zone is identical, but instead of a direct signal amplification inside the light

sensor, the charge is shifted outside the light sensing area of the image sensor. Not until then it is amplified, however more details and information can be found in the mentioned literature.

Unwanted effects: Besides this ideal functionality, there are some common non-ideal effects that should be considered. Not every photon of the right energy will create a free electron-hole-pair that accumulates in the capacitance. The photon matter interaction in general obeys to the Possion statistics which has a parameter that depends on the penetration depth of the photon into the matter as well as on the photon's energy.

Considering the whole system, each material has its specific spectral sensitivity depending on its physical parameters, like its reflectance, that hinders the electrons to enter the silicon. Other parameters that influence the sensitivity are the geometry of the photo diode with respect to a sensitive area and the depth and width of the depletion zone. The total number of usable electrons with respect to the number of available photons per light sensor is described as quantum efficacy η and is usually strongly dependent on the energy of the photon, and thus the wavelength ($\eta(\lambda)$).

For longer wavelength photons, the interaction rate is usually lower and thus the diode depth may not allow a high quantum efficacy. Above $900\,\mathrm{nm}$, the silicon is said to appear transparent to photons [15]. Increasing the depth of the diode to achieve a higher quantum efficacy for infrared-like radiation is limited, as skewed photons may then also reach the doped area of the neighboring light sensors, leading to crosstalk effects. Besides the crosstalk effect, the doped silicon usually is very expensive in its fabrication.

Besides the photon electron interactions, electron-hole-pairs are also created randomly from the temperature energy (with respect to their Fermi-Dirac statistics). This effect is called dark current and increases exponentially with temperature as the Fermi-distribution goes from a sharp edge towards a smooth curve with increasing temperature. As each crystal structure has its defects that allow slight changes in the energy levels, the dark current will have individual values for each light sensor.

Another problem occurs from the sensitivity of the analog circuity. The present transistors consist all of p-n-junctions which leads to different amplification rates for each signal. Further, a crosstalk in the analog domain can result in a sudden appearance of a changed set of nonuniformities in certain operating modes (e.g. a crosstalk with the power supply elements). Such problems may also affect the build-in automatic nonuniformity correction circuity of the sensor.

Unwanted Long Term Effects: Besides the above mentioned *unwanted effects* there are also several effects that appear on longer time scales and that changes the light sensor's behavior. Those effects mainly result as a change in the doped

silicon material. Ultra violet (UV) radiation for example is known to degrade the silicon structures [19] which leads to changes in the dark current and amplification behavior of the transistors.

Further to the UV-radiation, other unwanted impurities are present in the material which cause unwanted recombination centers that may be inactive in the first few years [20]. In general, diffusion processes of the doping materials may also lead to a change in the sensor characteristics. All these effects concern the p-n-junctions and their depletion zones, not only in the actual photo-diode, but also in the amplifying transistors as mentioned above are affected. Therefore the conductivity or charge generation ability of the devices changes, which leads to the overall offset and gain characteristics. And for the whole image sensor this results in changes of the nonuniformities.

2.4. The EMVA1288 Sensor Model

In the last section, the basic physical effects of the photon to voltage transformation have been explained methodically without any mathematical modeling (see sec. 2.3.2). The EMVA1288 standard for camera calibration describes a linear signal model that allows to characterize the sensitivity, linearity and temporal noise of a light sensor [1], which is shortly explained now. The standard models CMOS and CCD sensors with the same principle that is compliant with the above described physics:

1. A number of photons hits the photosensitive area A of a light sensor.
2. A number of electrons is created, given a quantum efficacy η.
3. The generated charge is then converted by the light diode's capacitance into a voltage.
4. The voltage is then amplified and converted into a digital number DN.

To calculate the expectation value of the number of photons μ_p for a certain light sensor, its photo sensitive area A, a given irradiance E and the exposure time $t_{\text{exp.}}$ are needed. Furthermore it is assumed that the photons possess an *average* wavelength λ, which allows to calcuate their energy according to eq. 2.1. This number of photons can be further transferred into the number of created electrons with help of the wavelength dependent quantum efficacy $\eta(\lambda)$:

$$\mu_p = \frac{AEt_{\text{exp.}}\lambda}{hc} \tag{2.5}$$

$$\mu_e = \eta(\lambda)\mu_{\text{p}} \tag{2.6}$$

The temperature and exposure time dependent dark signal is considered as an additive signal μ_d and the amplification of the signal is then described by an overall

system gain K, leading to the expected measurement quantity μ_y in digital number DN:

$$\mu_\mathrm{y}[\mathrm{DN}] = K(\mu_\mathrm{e} + \mu_\mathrm{d}) \tag{2.7}$$

As mentioned in sec. 2.3.2, the dark signal μ_d is generated by the temporal integration of the dark current of the randomly generated electron-hole-pairs μ_I that depend on the temperature energy. As the Fermi Dirac distribution can be approximated with an exponential distribution for high temperatures, an exponential model for the dark current μ_I is chosen by EMVA1288:

$$\mu_d = \mu_{d.0} + \mu_\mathrm{I} t_\mathrm{exp.} \tag{2.8}$$

$$\mu_\mathrm{I} = \mu_\mathrm{I,ref} \cdot 2^{\frac{T - T_\mathrm{ref.}}{T_\mathrm{d}}}, \tag{2.9}$$

where T_d is the temperature range that doubles the dark current μ_I and the combined sensor model for the signal results as:

$$\mu_\mathrm{y}[\mathrm{DN}] = K\eta\frac{AEt_\mathrm{exp.}\lambda}{hc} + K\mu_\mathrm{d.0} + K\mu_\mathrm{I}t_\mathrm{exp.}\,. \tag{2.10}$$

Tab. 2.1 gives an overview of the variables used in the EMVA1288 standard.

Parameter	Unit	Description
K	$\frac{\mathrm{DN}}{\mathrm{e}^-}$	overall system gain
$\mu_\mathrm{d,0}$	e^-	constant dark current
$\mu_\mathrm{I}(T)$	e^-	temperature depended dark current
η	$-$	sensor's quantum efficiency
h	$\mathrm{J\,s}$	Planck's constant
c	$\frac{\mathrm{m}}{\mathrm{s}}$	speed of light
$t_\mathrm{exp.}$	ms	exposure time
A	$\mu\mathrm{m}^2$	photo sensitive area of a single sensor
λ	nm	wavelength of the light
E	$\frac{\mu\mathrm{W}}{\mathrm{cm}^2}$	averaged irradiance over the area A

Table 2.1.: Explanation of the EMVA1288 parameters that contribute to the DSNU and PRNU

While this sensor model is very complex in a general case, a constant temperature and a constant exposure time leads to a linearisation.

$$\mu\mathrm{y}[\mathrm{DN}] = \underbrace{K\eta\frac{At_\mathrm{exp.}\lambda}{hc}}_{a} E + \underbrace{K\mu_{d.0} + K\mu_\mathrm{I}(T)t_\mathrm{exp.}}_{b} \tag{2.11}$$

$$\mu_\mathrm{y}[\mathrm{DN}] = aE + b \tag{2.12}$$

This model is valid for each individual light sensor, meaning that each light sensor will have its own offset $b_{i,j}$ and gain $a_{i,j}$, caused by the effects mentioned above:

$$\mu_{\mathrm{y},i,j}[\mathrm{DN}] = a_{i,j}E + b_{i,j} \tag{2.13}$$

These individual different offset and gain factors produce the nonuniformities in the measured image signals and will be further explained in chapter 3.

2.5. Digital Image Sensors

Most modern digital photographs of light signals are acquired by CMOS or CCD light sensors with a functionality as explained in sec. 2.3.2. The described light sensors measure the averaged light irradiation over a given area A in a given time $t_{\mathrm{exp.}}$. To be able to capture the light signal in a way that it can be reconstructed, many light sensors are arranged close to each other. Such an assembly of light sensors then forms an image sensor and typical modern image sensors contain several million light sensors. Many quantities with concern to the light sensors are given in the common data sheets. E.g. the actual size of one light sensor element, however it is not said that the whole area is photo sensitive. The fill-factor gives an estimate on how much of the actual area belongs to the photo sensitive part of the light sensor.

In order to measure the light signal, the photo sensitivity for each light sensor and over the whole sensing area of the image sensor should be known. For CMOS imagers, each pixel contains the photo diode and the readout and pre-amplification transistors. For CCD imagers the readout circuity is much smaller but also present (Especially if dark-rows are embedded in between the light sensitive rows). While in recent years, metal layers were build on top of the light sensitive silicon, new back-illumination techniques can avoid these obstructions, leading to an increased sensitivity. Another method to improve the sensitivity is to bput micro lenses on top of each light sensor.

Given a 2D top view of a light sensor, the photo sensitivity is not distributed constantly and fig. 2.9 shows, the measured sensitivity of a CMOS light sensor (with thanks to Daniele Passeri [21]). It is clearly visible that the obstructions from the readout circuity lead to severe variations in the photo responsibility of the individual pixels. Measurements like this vary from manufacturer to manufacturer and are not part of the common data sheets.

Additionally to the sensitivity of each light sensor, the positions of the individual sensors are of interest. Usually the light sensors are positioned onto a regular grid as shown in figure 2.10. The supply and readout electronics for the light sensors are then usually connected row and column wise, with the signal readout path connected on the shorter routes as the signal noise increases with line length.

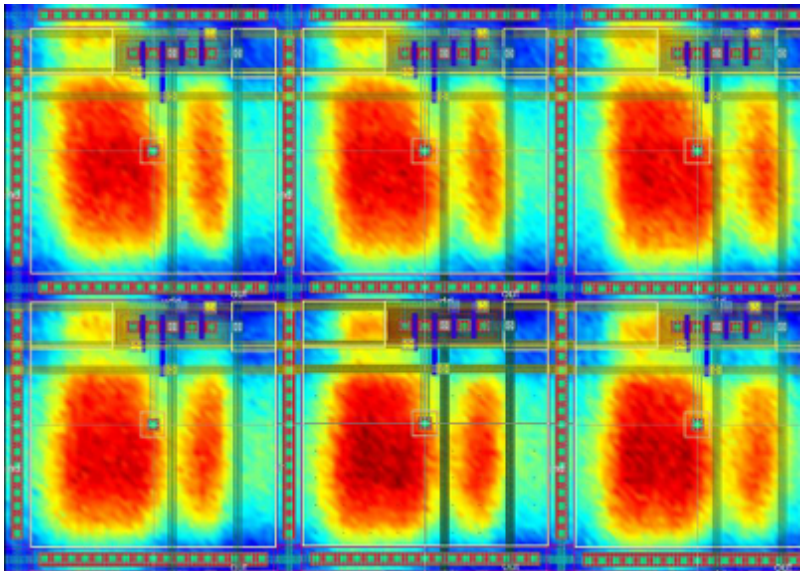


Figure 2.9.: Exemplary photo sensitivity of a given pixel, superimposed with the actual pixel layout (red stays for a high quantum efficacy)

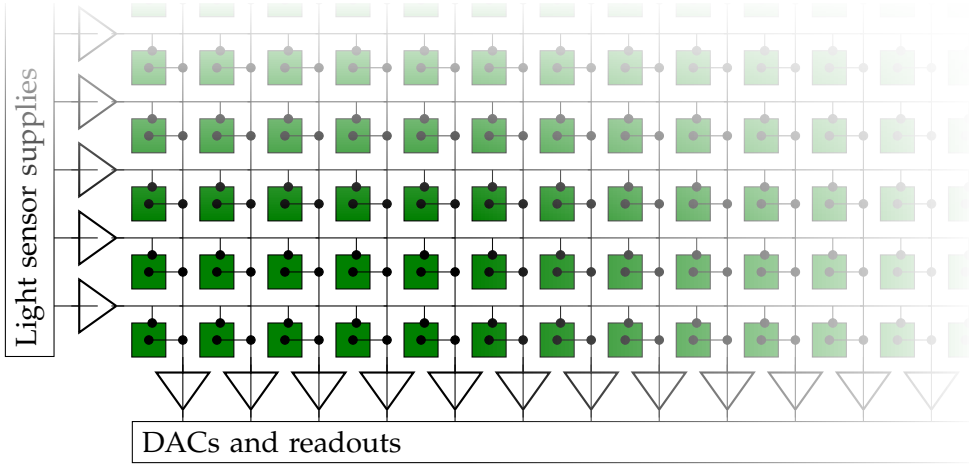


Figure 2.10.: Imager layout with many lightsensors and their row-wise power supply and column wise read-out lines.

It has been shown in several works that the nonuniformities of image sensors have a row and column wise dependency [2, 22], which results from the row and column wise connections. With a more refined sensor model, nonuniformity effects that result from row and column wise gains and offsets can be considered, which is done in chapter 7.

2.5.1. Colored Images

An image sensor as described above is not able to capture color information. An exception is the Foveon Chip [23] that exploits the wavelength dependent photon interaction with respect to the penetration depth of the photons into the silicon material. However, to be able to extract color information with a typical image sensor as described above, the sensors have to be equipped with individual color filters. Typically the sensors are considered in groups of four and the light sensors are then equipped with a regular color filter array (CFA) as shown in figure 2.11. With the knowledge of the CFA structure a reconstruction of the color information is possible to some extend and further details are described below.

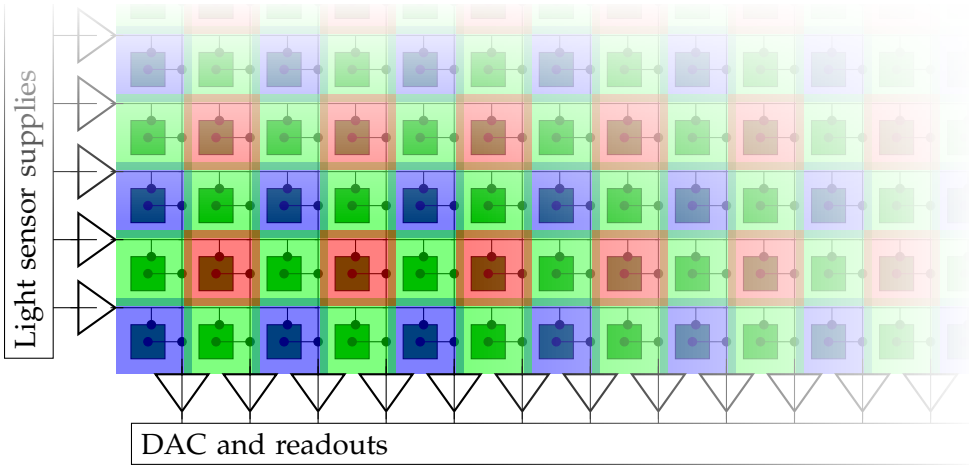


Figure 2.11.: A typical color filter arrays (CFA) on an image sensor

2.6. Image Acquisition and Reconstruction

With the knowledge of light sensors and their combination to image sensors, the description of the physical processes and their mathematical models are completed. Now a mathematical description of the continuous light signal and its reconstruction from the light sensor measurements is given, according to the description found in various literature [24, 16].

2.6.1. Signal Sampling with the Light Sensors

The optically low-pass filtered and mapped signal I_{Focal} according to eq. 2.4 is measured by the light sensors. Regarding the physics, the light sensors average this signal temporally over the exposure time and spatially over the sensitive area of the individual light sensors. As shown in fig. 2.9 and 2.10 in sec. 2.5, we assume that all sensors have an identical layout of their sensitive area and are ordered in a rectangular grid.

While the temporal averaging is described by the sensor models (eq. 2.13), the spatial averaging of the light signal by the sensors can be expressed by a convolution of the focal plane light signal I_{Focal} with the spatial sensor sensitivity. The evaluation of this signal at the sensor positions $(i, j) \in S$ then gives the coefficients of the sampled intensity signal I_{Sampled}. This mathematical calculation is possible because the light sensor's sensitivity can be considered as the light sensor's PSF L_{PSF} [16]. The evaluation at the sensor position is expressed by a sampling with Dirac impulses at the positions of the sensors, and leads to the sampled intensity signal I_{Sampled} as:

$$\begin{aligned} I_{\mathrm{Sampled}} &= (I_{\mathrm{Focal}} * L_{PSF}) \cdot \Delta(x, y) \\ \text{with} \quad \Delta(x, y) &= \sum_{(i,j)} \delta(x - iD_x)\delta(y - D_y)\,. \end{aligned} \tag{2.14}$$

With $\Delta(x, y)$ as a two dimensional Dirac comb and D_x and D_y the sensor's grid constants. The signal I_{Sampled} is completely defined by the sample coefficients of the Diarc pulses $C_{\mathrm{S},i,j}$:

$$C_{\mathrm{S},i,j} = (I_{\mathrm{Focal}} * L_{\mathrm{PSF}})|_{(i,j)\in S}\,. \tag{2.15}$$

Given the definition of I_{Focal} from eq. 2.4 and the commutativity of convolutions, then the above equations can be simplified using a combined PSF for the whole system S_{PSF}:

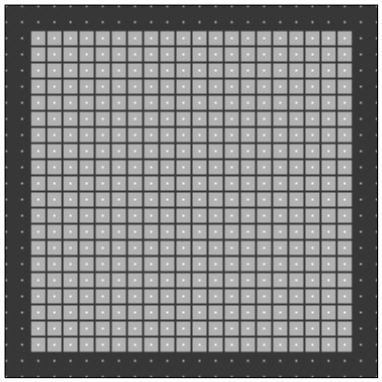
$$I_{\mathrm{Sampled}} = (I_{\mathrm{World}} * \underbrace{O_{\mathrm{PSF}} * L_{\mathrm{PSF}}}_{S_{\mathrm{PSF}}}) \cdot \Delta(x, y) \tag{2.16}$$

$$C_{\mathrm{S},i,j} = (I_{\mathrm{World}} * S_{\mathrm{PSF}})|_{(i,j)\in S} \tag{2.17}$$

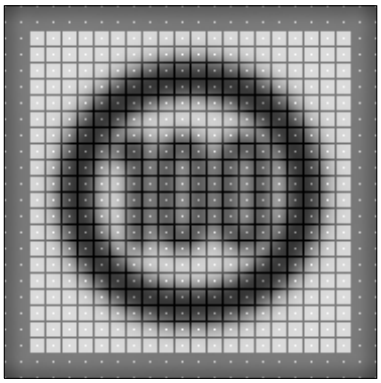
The samples $C_{\mathrm{S},i,j}$ can be used as input values of a light sensor modeling, as they correspond to the averaged intensity that (ideal) light sensors would measure.

Fig. 2.12 shows the schematics of the light signal sampling as described above. The continuous signal (fig. 2.12a) is projected onto the array of light sensors (2.12b). Fig. 2.12c shows the projected signal onto the sensor areas, where the projection includes the low pass filtering of the optics. Further, the averaging over the sensitive areas gives the coefficients of the sampled signal that are shown in fig. 2.12d in a block representation for each coefficient.

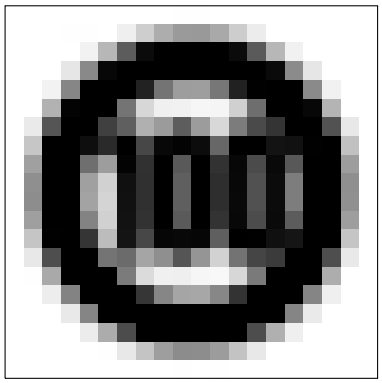
(a) Continuous light signal



(b) Imager with photosensitive Areas



(c) Signal onto the imager



(d) Pixel representation of the signal's sampling coefficients

Figure 2.12.: Schematics of the sampling of the continuous light signal

2.6.2. Signal Reconstruction

After having the signal converted into a set of discrete coefficients, the reconstruction of the continuous signal needs to be considered. The Fourier transform of the sampling process writes as:

$$I_{\text{Sampled}} = (I_{\text{World}} * S_{\text{PSF}}) \cdot \Delta(x,y) \quad \circ\!\!-\!\!- \quad \hat{I}_{\text{Sampled}} = (\hat{I}_{\text{World}} \cdot \hat{S}_{\text{PSF}}) * \hat{\Delta}(\hat{x},\hat{y})\,, \tag{2.18}$$

and the Fourier transform of the Dirac comb Δ is a Dirac comb in the Fourier space again

$$\Delta(x,y) = \sum_{(i,j)} \delta(x - iD_x)\delta(y - D_y) \quad \circ\!\!-\!\!- \quad \hat{\Delta}(\hat{x},\hat{y}) = \sum_{(m,n)} \delta(\hat{x} - \frac{m}{D_x})\delta(\hat{y} - \frac{n}{D_y})\,. \tag{2.19}$$

Therefore in the Fourier space, the convolution with the Dirac comb gives a summation of the replications of the signal $(\hat{I}_{\text{World}} \cdot \hat{S}_{\text{PSF}})$ with the frequency of $\frac{1}{D_x}$ in $\hat{x}$ direction and $\frac{1}{D_y}$ in $\hat{y}$ direction. Fig. 2.13a shows the absolute value of the complex Fourier transform of the sampled signal from fig. 2.12 and the repetitions of the signal are clearly visible. Concluding, a reconstruction is only possible by cutting the additional replications out. This is done for example by multiplying the Fourier transform of the sampled signal with a reconstruction function $\hat{\Phi}$. The ideal reconstruction function will be a two dimensional rectangular function $\hat{\Phi}() = \text{rect}() = \sqcap()$. The reconstruction of the signal then expresses as

$$I_{\text{Reco}} = I_{\text{Sampled}} * \Phi(x,y) \quad \circ\!\!-\!\!- \quad \hat{I}_{\text{Reco}} = \hat{I}_{\text{Sampled}} \cdot \hat{\Phi}\,. \tag{2.20}$$

As the repetitions have to be cut out in $\hat{x}$ and $\hat{y}$ direction, the reconstruction function $\hat{\Phi}$ can be separated into its $\hat{x}$ and $\hat{y}$ components. This further allows to find its inverse Fourier transform as the product of the Fourier transforms of the separable functions [25]:

$$\hat{\Phi}(\hat{x},\hat{y}) = \hat{\Phi}_x(\hat{x}) \cdot \hat{\Phi}_y(\hat{y}) \quad \circ\!\!-\!\!- \quad \Phi(x,y) = \Phi_x(x) \cdot \Phi_y(y)\,. \tag{2.21}$$

Given the Dirac pulse coefficients of I_{Sampled}, the signal reconstruction can also be written with help of a sum over all sensor elements $(i,j) \in S$ of the coefficients $C_{\text{S},i,j}$, weighted with the reconstruction function:

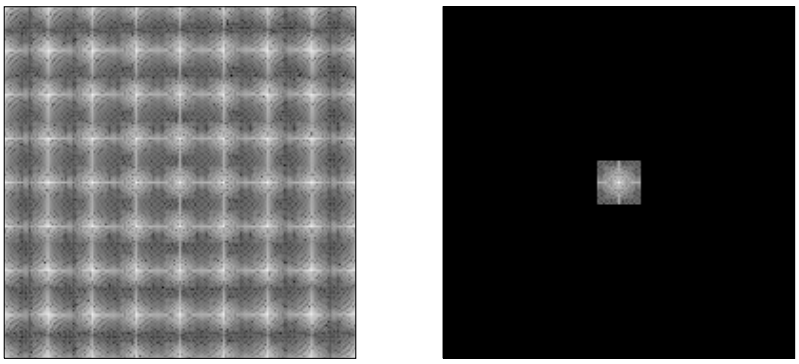
$$I_{\text{Reco}}(x,y) = \sum_{(i,j)\in S} C_{\text{S},i,j} \cdot \Phi(x-i, y-j) \tag{2.22}$$

The Fourier transform of the ideal clipping function $\sqcap(x)$ is the sinc function $\text{sinc}(x) = \frac{\sin(x)}{x}$, which has an infinitely large domain. Thus the reconstruction with the sinc-function leads to a computational effort that is far too large to be

considerable, and useful approximations need to be made. Fig. 2.13b shows the clipped signal in the Fourier space with a rect function.

Aliasing: Given the above equations, an ideal reconstruction is only possible if the repetitions of the signal are not overlapping. As the repetitions occur with a frequency of $\frac{1}{D_y}$ and $\frac{1}{D_x}$, all frequency components higher than $\frac{1}{2}\frac{1}{D_y}$ and $\frac{1}{2}\frac{1}{D_x}$ should be 0. This is known as the Nyquist-Shannon sampling theorem. All non zero frequency components that do not fulfill this condition will be present in the clipped signal as components of lower frequencies. The image degradation caused by these high frequencies are called aliasing effects. The frequencies $\frac{1}{2}\frac{1}{D_x}$ and $\frac{1}{2}\frac{1}{D_y}$ are also named the *Nyquist frequencies*. To avoid aliasing effects, the signal ($\hat{I}_{\text{World}} \cdot \hat{S}_{PSF}$) should not contain any high frequency components above the Nyquist frequencies and as a consequence the signal has to be low pass filtered *before* the sampling.

In practice the PSF of the imaging system (optics and imager) $\hat{S}_{PSF}$ damps the high frequency components to a degree that the remaining aliasing effects do not disturb the image signal. However, the PSF is not ideal and the lower frequencies are therefore damped as well. The trade off between these two effects is considered for each optical system separately with respect to its field of application. Further, the reconstruction function Φ can try to regain the damped low frequencies in the reconstruction process.



(a) Fourier transform of Sampled signal

(b) Reconstruced image in Fourier space

Figure 2.13.: Fourier transform of the sampled signal and reconstruction in the Fourier space by clipping. Image data represented as $\log(|f(\cdot)|)$.

2.6.3. Typical Realizations for the Reconstruction

Besides the reconstruction with the sinc function, a linear or cubic reconstruction function is more common for typical image processing algorithms. Fig. 2.14 shows

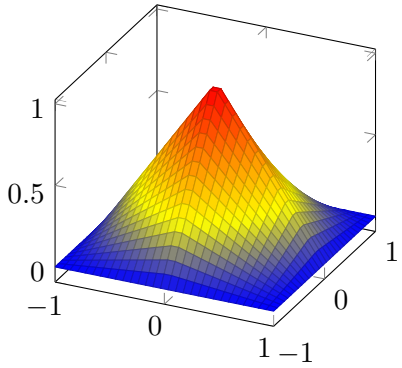
a plot of the corresponding 2D reconstruction functions, first the bilinear reconstruction filter ($\Phi = \mathrm{BL}$, fig. 2.14a) and second the bicubic reconstruction filter (fig. 2.14b). The bicubic reconstruction filter is realized as a Mitchell Netravalli Filer [26] with the parameters $B = 1.0$ and $C = 0.0$ ($\Phi = \mathrm{MN}$). The sinc filter is shown in fig. 2.14c ($\Phi = \mathrm{sinc}$). Also visible in the images is the covered area for the reconstruction filters with 2×2 for the bilinear filter, 4×4 for the bicubic filter. For the sinc filter the area is infinite large and the representation has been clipped. As mentioned in eq. 2.21, the reconstruction functions are separable and the above mentioned definitions for the BL and MN reconstruction are defined as:

$$\mathrm{BL}(x) = \max(1 - |x|, 0) \tag{2.23}$$

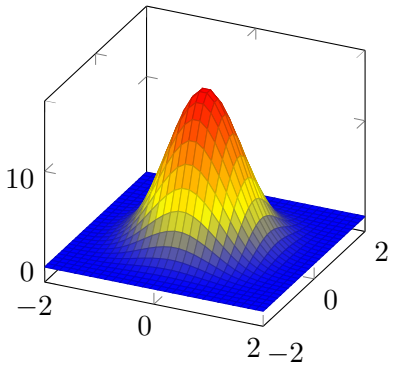
$$\mathrm{MN}(x) = \frac{1}{6} \begin{cases} (12 - 9B - 6C)|x|^3 + (-18 + 12B + 6C)|x|^2 \\ \qquad +(6 - 2B) & \text{, if } |x| < 1 \\ (-B - 6C)|x|^3 + (6B + 30C)|x|^2 \\ +(-12B - 48C)|x| + (8B + 24C) & \text{, if } 1 \leq |x| < 2 \\ 0 & \text{else} \end{cases} \tag{2.24}$$

However, not only the signal itself, but also its first and second spatial derivatives are of great interest as these can be used to represent the edges of the intensity signal. Considering the bilinear reconstruction only the first derivative exists. Using the cubic reconstruction the second derivative can be reconstructed as well, while for the sinc-based reconstruction, all derivatives can be reconstructed in principle.

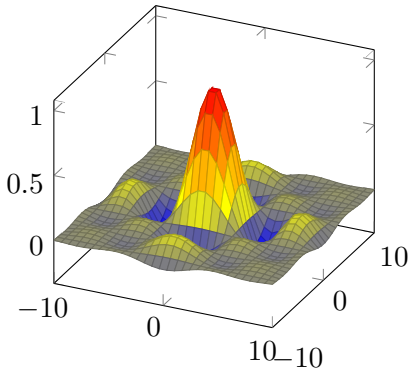
Fig. 2.15 shows the result of the different reconstruction filters and their first and second derivatives. The bilinear reconstruction shows a blurry image as higher frequencies are not reconstructed (fig. 2.15a). Further one can observe some reconstruction artifacts that correspond to the sample points of the light sensors. As expected, the first derivative (fig. 2.15b) has constant areas that correspond to the piecewise constant areas of the reconstruction filter. The second derivative is trivial and equal to zero for the whole signal area (fig. 2.15c). The bicubic (MN) reconstruction shows a much smoother image because the higher frequencies are damped stronger in this reconstruction mode and further, no reconstruction artifacts show up like they do for the bilinear reconstruction. The first and second derivative then look as expected (fig. 2.15e and fig. 2.15f). For the sinc reconstruction we see again reconstruction artifacts for the sample positions, but the image in general tends to be not as blurred as for the other reconstruction modes (fig. 2.15g). This is an effect of the almost ideal frequency cut in the Fourier space which does not lead to unwanted damping in the remaining high frequencies. With concern to the first and second derivatives, the reconstruction artifacts show up and give signals that are less well reconstructed than for the MN reconstruction (see fig. 2.15e and fig. 2.15f). This results from the fact that the sinc functions reconstruction needs to consider sampling points that lie outside the actual signal. The problems with this reconstruction mode are also discussed by Wolf [24].



(a) Bilinear ($\Phi = \mathrm{BL}$) reconstruction filter



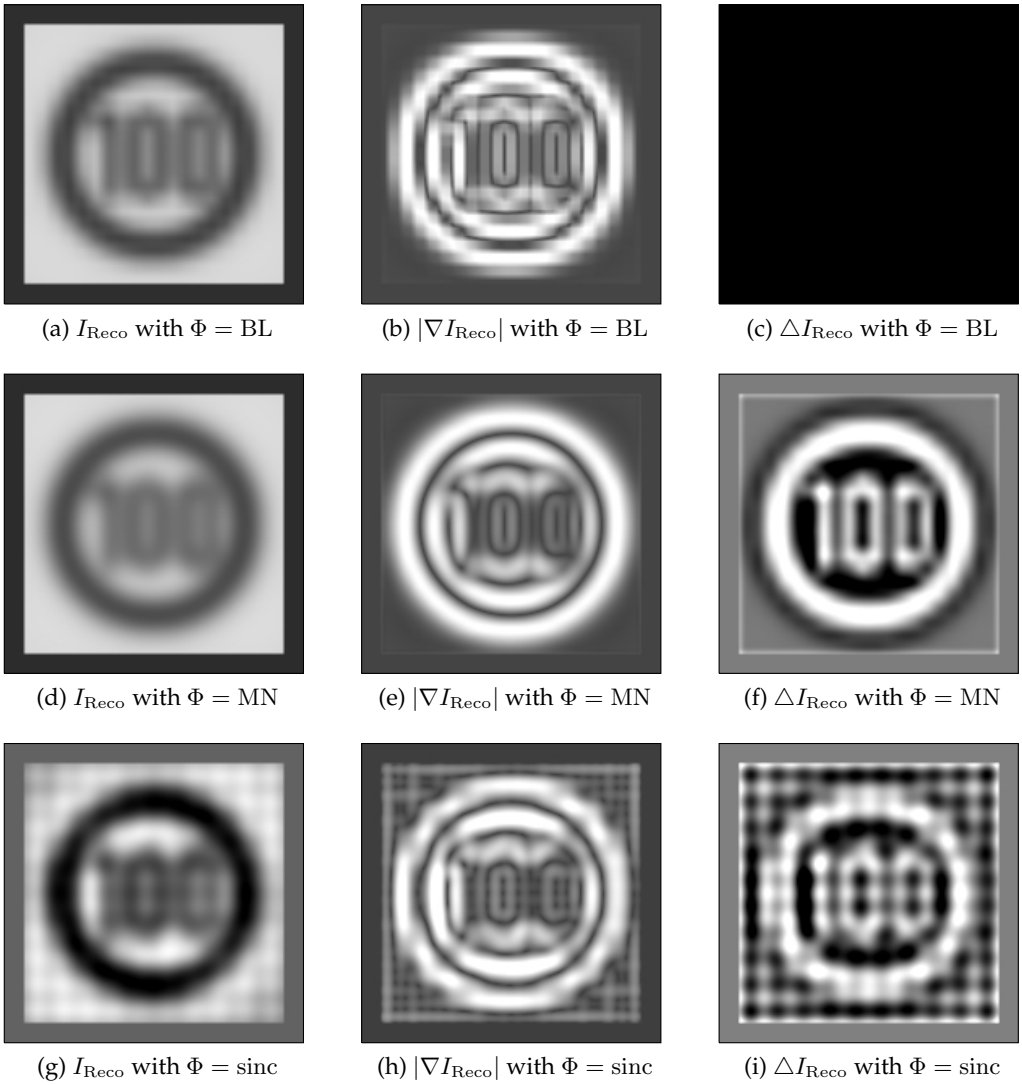
(b) Bicubic ($\Phi = \mathrm{MN}$) reconstruction filter



(c) Sinc ($\Phi = \mathrm{sinc}$ reconstruction filter

Figure 2.14.: Schematics of the sampling of the continuous light signal

Concluding, the MN reconstruction tends to give the best signals with concern to the second derivative. For the first derivative of the BL and MN reconstruction the applications have to decide which property is more important: higher amplitude in the higher frequencies or smoother signal in the first derivatives. Further one has to keep in mind that these signals are simulated results which may not occur in this way in a real world application, because other (and non-ideal) PSFs are more likely.



(a) I_{Reco} with $\Phi = \mathrm{BL}$ (b) $|\nabla I_{\mathrm{Reco}}|$ with $\Phi = \mathrm{BL}$ (c) $\triangle I_{\mathrm{Reco}}$ with $\Phi = \mathrm{BL}$

(d) I_{Reco} with $\Phi = \mathrm{MN}$ (e) $|\nabla I_{\mathrm{Reco}}|$ with $\Phi = \mathrm{MN}$ (f) $\triangle I_{\mathrm{Reco}}$ with $\Phi = \mathrm{MN}$

(g) I_{Reco} with $\Phi = \mathrm{sinc}$ (h) $|\nabla I_{\mathrm{Reco}}|$ with $\Phi = \mathrm{sinc}$ (i) $\triangle I_{\mathrm{Reco}}$ with $\Phi = \mathrm{sinc}$

Figure 2.15.: Representation of different reconstructed signals and their derivatives, using the above reconstruction functions Φ.

2.6.4. Combining Sensor Model and Signal Reconstruction

This section provides a combination of the sensor model from sec. 2.4 and the sampling and reconstruction from sec. 2.6.1. The EMVA1288 linear sensor model of eq. 2.13 gives the expectation value of the sensor output for a light stimulus E on the photosensitive area A:

$$\mu_{\mathrm{y},i,j}[DN] = a_{i,j}E + b_{i,j} \tag{2.25}$$

With the knowledge from sec. 2.6.1, we know that the averaged light signal on the sensor area is the coefficient of the sampled signal $C_{\mathrm{S},i,j}$. Further, the sensors will of course never measure their exact expectation values. If the sensor model is correct in its assumptions and all systematical errors are covered by the gain and offset coefficients a and b, then for a real measurement a mean free stochastic error $\chi_{i,j}$ has to be added. The *measured* coefficients of the sampled signal $C_{\mathrm{M},i,j}$ are then defines as:

$$C_{\mathrm{M},i,j} = a_{i,j}C_{\mathrm{S},i,j} + b_{i,j} + \chi_{i,j}\,, \tag{2.26}$$

and inverting this equation, the signal reconstruction formula develops to:

$$C_{\mathrm{S},i,j} = \frac{1}{a_{i,j}}(C_{\mathrm{M},i,j} - b_{i,j} - \chi_{i,j}) \tag{2.27}$$

$$I_{\mathrm{Reco}}(x,y) = \sum_{(i,j)\in S} \frac{1}{a_{i,j}}(C_{\mathrm{M},i,j} - b_{i,j} - \chi_{i,j}) \cdot \Phi(x-i, y-j) \tag{2.28}$$

This equation will be used as starting point for the development of the nonuniformity estimation methods.

Actually, each of the above measurements happens during the exposure times of the sensors and therefore can be considered a 1D sampling of the signals in the temporal domain [16]. This means that C_{S} and C_{M} are explicitly time dependent coefficients of the temporally sampled signal at the times t_n: C_{S,i,j,t_n} and C_{M,i,j,t_n}.

3. Image Nonuniformities and their Quantification

3.1. Definition of DSNU and PRNU

Given the physical background on light sensors and image sensors from chapter 2, this chapter introduces the image nonuniformities and their quantification. As explained in sec. 2.5, image sensors are composed of millions of light sensors and a linear model of the light sensors, as introduced by the EMVA1288 standard, is sufficient for typical CMOS and CCD sensors. To reapeat, the linear sensor model of eq. 2.13 develops into the measurement model in the form of eq. 2.26:

$$C_{\mathrm{M},i,j} = a_{i,j} C_{\mathrm{S},i,j} + b_{i,j} + \chi_{i,j} \,.$$

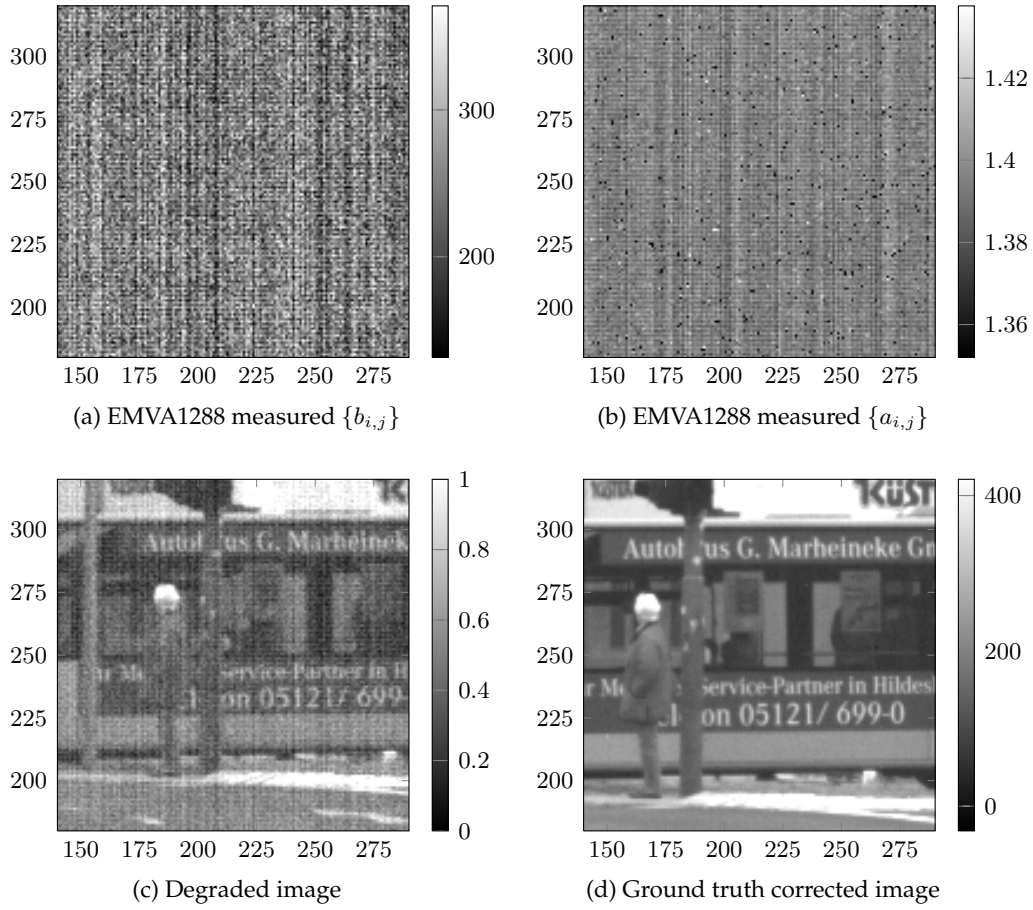
The coefficients $a_{i,j}$ are named *gain*s and the coefficients $b_{i,j}$ are named *offset*s, while the $\chi_{i,j}$ represent the stochastic errors that are considered independent to each other. Each sensor (i, j) will have a slightly different gain and offset parameters due to imperfections in the manufacturing of the circuity and/or the silicon doping processes. All these errors influence the measured intensity and the gain and offset imperfections will present as systematic errors in the reconstructed light signals. While in a single image the systematic errors are not distinguishable from random noise χ, the evaluation of many frames allows to separate the systematic from the stochastic errors.

In this thesis, sets of indexed variables that consider all light sensors of an image sensor are defined in the short notation:

$$\{x_{i,j}\} = \{x_{i,j} | i, j \in S\} \quad \text{with: } S \text{ all sample positions of the image sensor} \tag{3.1}$$

If many images are recorded at complete darkness, the averaged image will show the nonuniformity of the set of offsets $\{b_{i,j}\}$. This appearance is called Dark Signal Nonuniformity, DSNU. For recordings at a uniform illumination, the averaged image will show a superposition of the nonuniformities caused by the offsets and the gains. As the sensor model is linear, the averaged image can be corrected by the measured DSNU coefficients $\{b_{i,j}\}$ and the remaining non-uniformity results from the different gain coefficients $\{a_{i,j}\}$. This *gain based* nonuniformity is called Photo Response Nonuniformity, PRNU. Figure 3.1 shows the measured offset set $\{b_{i,j}\}$ and gain set $\{a_{i,j}\}$ of a Photonfocus MV1-D1312-160-CL camera. The measurements are according to the EMVA1288 standard and therefore with switched off

internal nonuniformity corrections. A recorded image that is degraded by these nonuniformities and the nonuniformity corrected image are shown as well to give an impression for the effects of the degradations. Especially the horizontally and vertically correlated degradations are disturbing the visual impression.



(a) EMVA1288 measured $\{b_{i,j}\}$

(b) EMVA1288 measured $\{a_{i,j}\}$

(c) Degraded image

(d) Ground truth corrected image

Figure 3.1.: Measured PRNU and DSNU parameter sets and a recorded, degraded image in comparison to the nonuniformity corrected image

3.2. EMVA1288 Nonuniformity Quantization

The PRNU and DSNU are defined by their sets of gain and offset coefficients $\{a_{i,j}\}$ and $\{b_{i,j}\}$, but these sets are too large to give an easy way to quantify the image signal degradation. EMVA1288 [1] therefore defines a measure for the DSNU and PRNU that results in one real number for the DSNU and one real number for the PRNU. These numbers are similar to the definitions found in the various literature [27, 28]. Additionally, to the definitions EMVA1288 also standardizes the measurement conditions and camera setups to record the images for the quantization [1]. Two measurement setups are required to calculate the DSNU and PRNU

quantifications: One measurement that gives the expectation value of the signal in the case of no illumination (EMVA value: y_{dark}). And one further measurement at 50% saturation with a constant uniform illumination (EMVA value y_{50}).

Spatial and Temporal Statistical Moments: To be able to define the needed signals, EMVA1288 separates the spatial and temporal statistical moments. For a given discrete temporal spatial signal $X(i, j, t_n)$, or the corresponding set $\{X_{i,j,t_n}\}$, the times t_n are the temporal sample points and the spatial positions (i, j) are the spatial sample points. The limits of the sums in the below definitions have to be adjusted according to the number of light sensors per spatial dimension (i, j) and the number of temporal samples t_n. However, assuming that there are T_N temporal sample points and (M, N) spatial sample points, then the statistical definitions can be simplified.

The temporal expectation is defined as:

$$\mu(X) = \mu(\{X_{i,j,t_n}\}) = \frac{1}{\sum_n 1} \sum_n X_{i,j,t_n} = \frac{1}{T_\mathrm{N}} \sum_n X_{i,j,t_n} \,. \tag{3.2}$$

The spatial expectation value is defined as:

$$e(X) = e(\{X_{i,j,t_n}\}) = \frac{1}{\sum\limits_{i,j} 1} \sum_{i,j} X_{i,j,t_n} = \frac{1}{MN} \sum_{i,j} X_{i,j,t_n} \,. \tag{3.3}$$

The temporal variance is defined as:

$$\sigma^2(X) = \sigma^2(\{X_{i,j,t_n}\}) = \frac{1}{-1 + \sum_n 1} \sum_n (X_{i,j,t_n} - \mu(X))^2 \tag{3.4}$$

$$= \frac{1}{T_\mathrm{N} - 1} \sum_n (X_{i,j,t_n} - \mu(X))^2 \,. \tag{3.5}$$

And last, the spatial variance is defined as:

$$s^2(X) = s^2(\{X_{i,j,t_n}\}) = \frac{1}{-1 + \sum\limits_{i,j} 1} \sum_{i,j} (X_{i,j,t_n} - e(X))^2 \tag{3.6}$$

$$= \frac{1}{MN - 1} \sum_{i,j} (X_{i,j,t_n} - e(X))^2 \,. \tag{3.7}$$

The measures for DSNU and PRNU as described in EMVA1288 can now be expressed with these definitions. Each measurement of the image sensor results in a set of coefficients $\{C_{\mathrm{M},i,j}\}$, assigned to a discrete point in time t_n. The defined sensor model for the measurements (eq. 2.26) gives the values $C_{\mathrm{M},i,j}$ as an estimate of the expectation values of the signal for the chosen exposure time (see eq. 2.12).

Thus, averaging many measurements leads to a better estimate of the real *temporal* expectation value of the constant signal.

Definition of DSNU: EMVA1288 defines the value of the DSNU as the spatial deviation of the temporal expectation value of the signal measured at darkness: $\{C_{\mathrm{M},i,j,\mathrm{dark}}\}$. The temporal expectation value limits also the effects of the mean free stochastic errors of χ and the spatial deviation gives a good number on how much noise is caused by the offsets in the image, as these parameters usually obey a Gaussian distribution.

With help of eq. 3.2-3.7 the definition is expressed as:

$$DSNU_{\mathrm{EMVA1288}} = \frac{s(\mu(\{C_{\mathrm{M},i,j,\mathrm{dark}}\}))}{K}\,. \tag{3.8}$$

Of course the actual temporal expectation value cannot be measured and a theoretical remain of the mean free stochastic errors may still be present. However, the standard gives a reasonable number of frames that have to be averaged to achieve good results (see [1] for more details).

Transforming all the demands to the EMVA1288 measurement process into assumptions, results in the following mathematical formulae:

$$C_{\mathrm{S},i,j,\mathrm{dark}} = 0 \quad \text{measurements at darkness}\,, \tag{3.9}$$
$$\mu(\{\chi_{i,j}\}) = 0 \quad \forall(i,j) \quad \text{as mean free random variable}\ , \tag{3.10}$$
$$\mu(\{b_{i,j}\}) = b_{i,j} \quad \forall(i,j) \quad \text{as temporal constants}\,. \tag{3.11}$$

In detail the above equation state that the measurements were conducted at darkness, that the remaining stochastic error is mean free and that the offsets are constants in time. Combined with the definitions of $C_{\mathrm{M},i,j,\mathrm{dark}}$, the following transformations express the DSNU by its offset coefficients (see [29] for the mathematical foundations):

$$DSNU_{EMVA1288} = \frac{s(\{\mu(a_{i,j}C_{\mathrm{S},i,j,\mathrm{dark}} + b_{i,j} + \chi_{i,j}\}))}{K} \tag{3.12}$$
$$= \frac{s(\mu(\{a_{i,j}C_{\mathrm{S},i,j,\mathrm{dark}}\}) + \mu(\{b_{i,j}\}) + \mu(\{\chi_{i,j}\})))}{K} \tag{3.13}$$
$$= \frac{s(\{b_{i,j}\})}{K} \tag{3.14}$$

Definition of PRNU: For the definition of PRNU the signal at 50% saturation ($C_{\mathrm{M},i,j,50\%}$), recording a uniform illumination source has to be measured. Given

a dynamic range of d quantization steps, $C_{\mathrm{M},i,j,50\%}$ is defined as the signal that obeys[1]:

$$C_{\mathrm{M},i,j,50\%}: \quad \frac{d}{2} = e(\mu(\{C_{\mathrm{M},i,j,50\%}\})) \tag{3.15}$$

EMVA1288 then defines the PRNU as:

$$PRNU_{\mathrm{EMVA1288}} = \frac{\sqrt{s^2(\mu(\{C_{\mathrm{M},i,j,50\%}\})) - s^2(\mu(\{C_{\mathrm{M},i,j,\mathrm{dark}}\}))}}{e(\mu(\{C_{\mathrm{M},i,j,50\%}\})) - e(\mu(\{C_{\mathrm{M},i,j,\mathrm{dark}}\}))} \cdot 100\% \,. \tag{3.16}$$

Appendix A.1 shows how this definition can be transformed into the spatial *coefficient of variation* of the gain parameters $\{a_{i,j}\}$:

$$PRNU_{\mathrm{EMVA1288}} = \frac{s(\{a_{i,j}\})}{e(\{a_{i,j}\})} \cdot 100\% \tag{3.17}$$

3.2.1. Adaptations for Scene Based Estimation Methods

While the DSNU and PRNU estimation (eq. 3.8 and eq. 3.16) are based on the actual measurements, the mathematical transformations of eq. 3.17 and eq. 3.14 utilize the sensor model's individual gain and offset factors $\{a_{i,j}\}$ and $\{b_{i,j}\}$.

The transformed definition of PRNU is calculated directly from these values, but the DSNU depends further on the overall system gain K from the EMVA1288 standard. As mentioned in sec. 2.4, K gives the transformation of the measured digital numbers, that represent the image, into the physical number of electrons inside the capacitance of the light sensor. However, a scene based nonuniformity estimation can only depend on the signal in digital numbers, which makes the gain factor K obsolete. Therefore a DSNU with the unit of digital numbers [DN] is defined by neglecting the factor K from the previous definition:

$$DSNU_{\mathrm{EMVA1288,DN}} = s(\{b_{i,j}\}) \tag{3.18}$$

3.2.2. Measurement of the Parameter Sets

The parameter sets $\{a_{i,j}\}$ and $\{b_{i,j}\}$ have to be obtained from the measurements $\{C_{\mathrm{M},i,j,\mathrm{dark}}\}$ and $\{C_{\mathrm{M},i,j,50\%}\}$. With the assumptions from eq. 3.10-3.11, the $\{b_{i,j}\}$ are the direct result of the temporal averaging of the $C_{\mathrm{M},i,j,\mathrm{dark}}$ measurements:

$$\begin{aligned} \mu(C_{\mathrm{M},i,j,\mathrm{dark}}) &= \mu(a_{i,j}C_{\mathrm{S},i,j,\mathrm{dark}} + b_{i,j} + \chi_{i,j}) && (3.19)\\ &= \mu(a_{i,j}C_{\mathrm{S},i,j,\mathrm{dark}}) + \mu(b_{i,j}) + \mu(\chi_{i,j}) && (3.20)\\ &= b_{i,j} && (3.21) \end{aligned}$$

[1] Assuming the does not saturate with the maximum number of quantization steps

The definition of the $\{a_{i,j}\}$ can be given with the knowledge of the $\{b_{i,j}\}$ by using the definitions of sec. 3.2. It follows for the individual $a_{i,j}$

$$\mu(C_{\mathrm{M},i,j,50\%}) = \mu(a_{i,j}C_{\mathrm{S},i,j,50\%} + b_{i,j} + \chi_{i,j}) \tag{3.22}$$

$$= \mu(a_{i,j})\mu(C_{\mathrm{S},i,j,50\%}) + \mu(b_{i,j}) + \mu(\chi_{i,j}) \tag{3.23}$$

$$\Leftrightarrow \mu(C_{\mathrm{M},i,j,50\%}\}) - \{b_{i,j}\} = a_{i,j}\frac{d}{2} \tag{3.24}$$

$$\Leftrightarrow a_{i,j} = \frac{2}{d} \cdot \left(\mu(C_{\mathrm{M},i,j,50\%}) - b_{i,j}\right) , \tag{3.25}$$

where the last equation allows to calculate the gain coefficients $\{a_{i,j}\}$ directly from the measurements.

3.3. Further Quantization Methods and Limits

The above described EMVA1288 nonuniformity quantification is one among several other methods to measure the nonuniformities in a defined laboratory environment [30, 1]. Some of the other calibration methods use more complex (e.g. nonlinear) models of the light sensors to adapt for special sensors [28]. For the DSNU parameters, the exponential temperature dependency according to eq. 2.9 should also be considered during the calibration. In general, all these methods rely on a defined environment with uniform illuminations to extract the DSNU and PRNU parameters.

The limits for the laboratory calibrations are given by the temporal drift effects from sec. 2.3.2. These make a more or less frequent recalibration of the sensors necessary to guarantee the best image quality. However, for applications like military infrared and thermal imaging, as well as video based driver assistance systems, a recalibration in a defined laboratory environment is not possible or too expensive and too complex. While for thermal imaging a calibration is very complex, for video based driver assistance systems a lifetime of the systems of $10 - 20$ years without a recalibration in a laboratory environment has to be considered.

Therefore scene based nonuniformity estimation methods are necessary in those situations. The nonuniformity variables $\{a_{i,j}\}$ and $\{b_{i,j}\}$ depend on exposure time and temperature as already shown in the EMVA1288 sensor model (eq.2.12). However, most scene based methods assume the parameters $\{a_{i,j}\}$ and $\{b_{i,j}\}$ constant. This is actually a sufficient assumption as in many applications, the exposure time is not adapted very often and the temperature changes usually very slowly compared to the number of captured frames and the actual parameter drifts. Scene based methods use further assumptions to extract the nonuniformity parameters and sec. 5 describes the most relevant scene based methods in further detail. The estimated nonuniformities should be evaluated against a ground truth nonuniformity measurement like the EMVA1288 based estimation to measure their performances.

3.4. Quality Metrics for Scene Based Estimated Nonuniformities

With the upcoming scene based nonuniformity estimations, the need for an evaluation of the new estimate against a ground truth reference is needed. The laboratory estimation parameters are therefore considered as the ground truth reference and literature provides several norms to calculate the quality scene based estimations. The common methods in state of the art publications compare the ground truth corrected images against the images corrected with the current estimate [7, 5, 4, 6]. Given an estimated set of parameters ($\{a_{i,j,\text{est.}}\}$ and $\{b_{i,j,\text{est.}}\}$) and a set of ground truth reference parameters ($\{a_{i,j,\text{ref.}}\}$ and $\{b_{i,j,\text{ref.}}\}$) the estimated and ground truth corrected images are calculated with:

$$C_{i,j,\text{ref.}} = \frac{1}{a_{i,j,\text{ref.}}}(C_{\text{M},i,j} - b_{i,j,\text{ref.}}) \tag{3.26}$$

$$C_{i,j,\text{est.}} = \frac{1}{a_{i,j,\text{est.}}}(C_{\text{M},i,j} - b_{i,j,\text{est.}})\,. \tag{3.27}$$

In the publications the mean absolute error (MAE), the root mean squared error (RMSE) and the signal to noise ration (SNR) are the mainly used measures for the quality of the nonuniformity reduction. Given the definitions of R. Gonzalez and R. Woods [31] these error measures are defined as:

$$\text{SNR} = 10 \cdot \log_{10}\left(\frac{\sum_{i,j} C_{i,j,\text{rem.}}^2}{\sum_{i,j}\left(C_{i,j,\text{rem.}} - C_{i,j,\text{est.}}\right)^2}\right) \tag{3.28}$$

$$\text{MAE} = \frac{1}{\sum_{i,j} 1}\sum_{i,j}|C_{i,j,\text{rem.}} - C_{i,j,\text{est.}}| \tag{3.29}$$

$$\text{RMSE} = \sqrt{\frac{1}{\sum_{i,j} 1}\sum_{i,j}\sum_{i,j}\left(C_{i,j,\text{rem.}} - C_{i,j,\text{est.}}\right)^2} \tag{3.30}$$

The biggest disadvantage of the reported quality metrics is their dependency on the used image material, as only the corrected images are compared and not the estimations of the actual sensor nonuniformities. For example a dark sequence will mainly be degraded by the DSNU component, while bright images are degraded by PRNU. Consequently, the calculated error measures are only valid if the used sequences are *typical* with respect to the intended application. Furthermore, it is not possible to calculate an improvement of the DSNU and PRNU effects separately, without an influence of the current image.

3.5. A Quality Metric Against a Laboratory Ground Truth

With the mentioned drawbacks of the common quality metrics for DSNU and PRNU, a new quality metric that allows to calculate an improvement of the DSNU and PRNU parameters as a percentual value has been developed. Therefore the performance of any given estimation method is evaluated with help of ground truth reference patterns $\{a_{i,j,\mathrm{ref.}}\}$ and $\{b_{i,j,\mathrm{ref.}}\}$.

In a first step the estimated set of parameters of most methods from literature needs to be scaled to achieve a comparable performance. This is necessary because the methods assume the nonuniformities with ideal mean values (e.g. $e(\{b_{i,j}\}) = 0$ and $e(\{a_{i,j}\}) = 1$), which in reality is usually not the case. As a consequence the produced images with the estimated nonuniformities reside on different gray value levels, when compared to the reference corrected image. This also effects the calculated numbers of the PRNU and DSNU values according to eq. 3.17 and 3.18.

The estimated parameters $\{a_{i,j,\mathrm{est.}}\}, \{b_{i,j,\mathrm{est.}}\}$ are therefore transformed in a way that the corrected image approximately fits the ground truth image in its two spatial statistical moments e and s^2:

$$e(C_{i,j,\mathrm{est.}}) = e(C_{i,j,\mathrm{ref.}}) \tag{3.31}$$

$$s^2(C_{i,j,\mathrm{est.}}) = s^2(C_{i,j,\mathrm{ref.}}) \tag{3.32}$$

Due to the nonlinear influence of $a_{i,j}$ in the image correction process, this transformation is done only approximately by exploiting the transformations of random variables as described by Rinne [29]. The details are given in detail in appendix A.2 and as a result, the *scaled* estimated parameters can be calculated from the statistical measures of the sets: Measured image $\{C_{\mathrm{M},i,j}\}$, estimated DSNU $\{b_{i,jest.}\}$, estimated PRNU $\{a_{i,j,est.}\}$, reference DSNU $\{b_{i,j,ref}\}$ and reference PRNU $\{a_{i,j,ref.}\}$:

$$a_{i,j,\mathrm{sc.}} = h_{\mathrm{sc.,PRNU}}(V; a_{i,j,\mathrm{est.}}) \tag{3.33}$$

$$\text{and:} \quad b_{i,j,\mathrm{sc.}} = h_{\mathrm{sc.,DSNU}}(V; b_{i,j,\mathrm{est.}})\,, \tag{3.34}$$

V is used as a shortcut of statistical moments of the known variables in this notations.

As a follow up step of the scaling, a set of remaining correction factors $\{a_{i,j,\mathrm{rem.}}\}$ and $\{b_{i,j,\mathrm{rem.}}\}$ can be calculated as

$$a_{i,j,\mathrm{rem.}} = \frac{a_{i,j,\mathrm{ref.}}}{a_{i,j,\mathrm{sc.}}} \quad \forall(i,j) \tag{3.35}$$

$$b_{i,j,\mathrm{rem.}} = b_{i,j,\mathrm{ref.}} - b_{i,j,\mathrm{sc.}} \quad \forall(i,j)\,. \tag{3.36}$$

For a good estimation the $\{a_{i,j,\mathrm{rem.}}\}$ should be close to 1.0, and the $\{b_{i,j,\mathrm{rem.}}\}$ should be close to 0. A remaining PRNU or DSNU is calculated from the $a_{i,j,\mathrm{rem.}}$ or $b_{i,j,\mathrm{rem.}}$

parameters according to eq. 3.18 or 3.17 and a percentual correction factor $p_{c,X}$ is defined as:

$$p_{c,X} = (X_{\text{ref.}} - X_{\text{rem.}})/X_{\text{ref.}} \cdot 100\% \quad \text{with} \quad X \in \{\text{PRNU}, \text{DSNU}\} \tag{3.37}$$

Given a $p_{c,X} = 0\%$ the remaining parameters did not lead to an improving correction, while $p_{c,X} = 100\%$ results from estimated parameters that equal the reference. Note that values of $p_{c,X} < 0\%$ are possible if the image was deteriorated by the corresponding nonuniformity correction. $p_{c,\text{PRNU}}$ and $p_{c,\text{DSNU}}$ are independent of any image information and purely correspond to the physical properties of the image sensor.

Nonuniformity correction methods that correct both, DSNU and PRNU, need a combined correction measure to evaluate their performance. To define such a combined correction rate the effect of the PRNU in digital numbers is needed, as it allows to compare the effects of DSNU and PRNU directly. Of course the effect of the PRNU is dependent on the typical illumination of the image sensor in the used scenario. Therefore a *typical illumination* $I_{\text{typ.,DN}}$ is defined by the average of all light sensor measures over a representative set of image sequences. For example, if the performance evaluation for a certain sequence shall be measured, then the typical illumination is calculated as the average of all pixels of the given sequence:

$$I_{\text{typ.,DN}} = s(\mu(C_{\text{M},i,j,t_n})) \quad \forall\, C_{\text{M},i,j,t_n} \in \text{Sequence} \tag{3.38}$$

Of course other definitions could be used as well, for example a weighted average giving lower or higher weights for different illuminations of the scene. The PRNU can now be expressed in its expected effect in digital numbers:

$$PRNU_{\text{EMVA1288,DN}} = PRNU_{EMVA1288} \cdot I_{\text{typ.,DN}} \tag{3.39}$$

A *combined correction* factor is defined by weighting the effects of the ground truth DSNU and PRNU according to their effects in DN:

$$p_{\text{c,Comb.}} = \frac{DSNU_{\text{EMVA1288,DN}} \cdot p_{\text{c,DSNU}} + PRNU_{\text{EMVA1288,DN}} \cdot p_{\text{c,PRNU}}}{PRNU_{\text{EMVA1288,DN}} + DSNU_{\text{EMVA1288,DN}}}\,. \tag{3.40}$$

Of course this combined correction relies on the fact that the typical illumination $I_{\text{typ.,DN}}$ is measured or estimated correctly. On the other hand it gives the chance to evaluate the performance of an assumed algorithm with respect to future assumed illuminations. For example the performance $p_{c,Combined}$ of an algorithm could be evaluated for night sequences ($I_{\text{typ.,DN}}$ is small) and daytime image sequences ($I_{\text{typ.,DN}}$ is high) separately.

3.6. Spatial Properties of the Corrections

Besides the above introduced scalar measurements of the DSNU and PRNU corrections, the remaining sets $a_{i,j,\mathrm{rem.}}$ and $b_{i,j,\mathrm{rem.}}$ allow a quick visual impression of the remaining nonuniformities. Fig. 3.2 shows the remaining DSNU parameters of estimations after 5, 50 and 500 frames for the energy minimization method introduced in chapter 6. The achieved percentual correction rate is shown below the images. For only 5 learned frames one clearly observes remaining artifacts which are referred to as ghosting in some literature (e.g. [5]). This name comes from the fact that those artifacts may be false-classified as real objects by image processing algorithms. With concern to Rossi [32], ghosting artifacts occur from objects that provide strong edges or high intensity for a short time. Fig. 3.2a to 3.2c show that the ghosting artifacts reduce with more included frames into the estimation for the presented algorithm.

However in fig. 3.2b after 50 frames, some remaining artifacts are still visible, although the measured percentual correction rate is already up to 71%. In fig. 3.2c after 500 frames, the remaining artifacts vanished almost completely and only a few spatial low frequency components are left. The correction rate has increased to 77%. Although the increase from 50 to 500 frames in the percentual measure is *only* 6%, the decrease in the visually percepted artifacts seems larger.

Another spatial property of the correction patterns is the *coarseness* of the remaining parameter sets. If the remaining patterns look very smooth, it can be assumed that the visually prominent high frequencies have been corrected successfully. For example, the correction pattern in fig.3.2c after 500 frames, gives a very smooth impression of the remaining pattern.

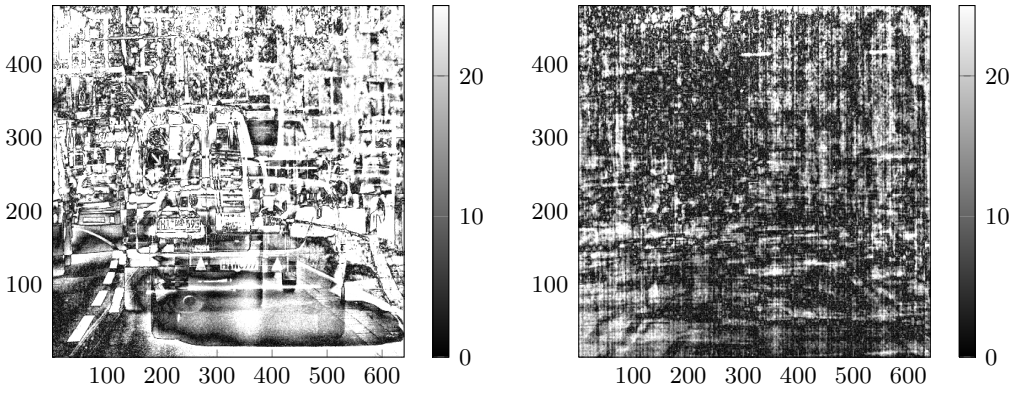
3.7. Spectral Properties of the Corrections

A spectral analysis will give more insight of the spatial frequency dependency of the corrections. A Fourier transformation of the remaining and the ground truth patterns allows to visualize the correction dependency on the spacial frequency components for the different methods:

$$\{a_{i,j}\} \circ\!\!-\!\!- \{\hat{a}_{\mu,\nu}\} \tag{3.41}$$

$$\{b_{i,j}\} \circ\!\!-\!\!- \{\hat{b}_{\mu,\nu}\} \tag{3.42}$$

For visualization it is common to provide the logarithm of the amplitude of the corresponding frequency component (e.g. $\log|\hat{b}_{\mu,\nu}|$). This is shown in fig. 3.3 for a measured reference DSNU pattern. The typical dominant column nonuniformities of this camera are visible in fig. 3.3b as a vertical line in the middle of the spectrogram. As mentioned before, such column and row wise components are typical [2].

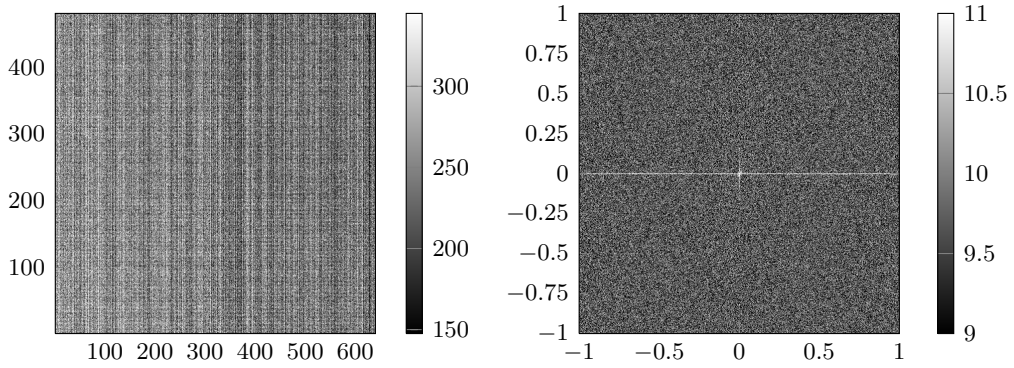


(a) $|\{b_{i,j,\mathrm{rem.}}\}|$ with input: 5 frames, and $p_{c,\mathrm{DSNU}}$=13%

(b) $|\{b_{i,j,\mathrm{rem.}}\}|$ with input: 50 frames, and $p_{c,\mathrm{DSNU}}$=71%

(c) $|\{b_{i,j,\mathrm{rem.}}\}|$ with input: 500 frames, and $p_{c,\mathrm{DSNU}}$=77%

Figure 3.2.: Spacial impression of the remaining DSNU after different frames of a sequence.



(a) $\{b_{i,j}\}$, measured

(b) Fourier transformed components as $\log|\hat{b}_{\mu,\nu}|$

Figure 3.3.: Spatial and Fourier representation of a DSNU nonuniformity pattern with a prominent column wise degradation. The Fourier representation is normalized to the Nyquist frequencies $\in (-1, 1)$

With consideration to the correction performance the reduction of spatial high frequencies is of more interest than that of low spatial frequencies. Especially because common computer vision algorithms usually are based on the analysis of edges, which present as spatial high frequencies.

In addition to the spectral analysis of the estimated and remaining nonuniformity components, a damping factor of the power spectral density can be calculated:

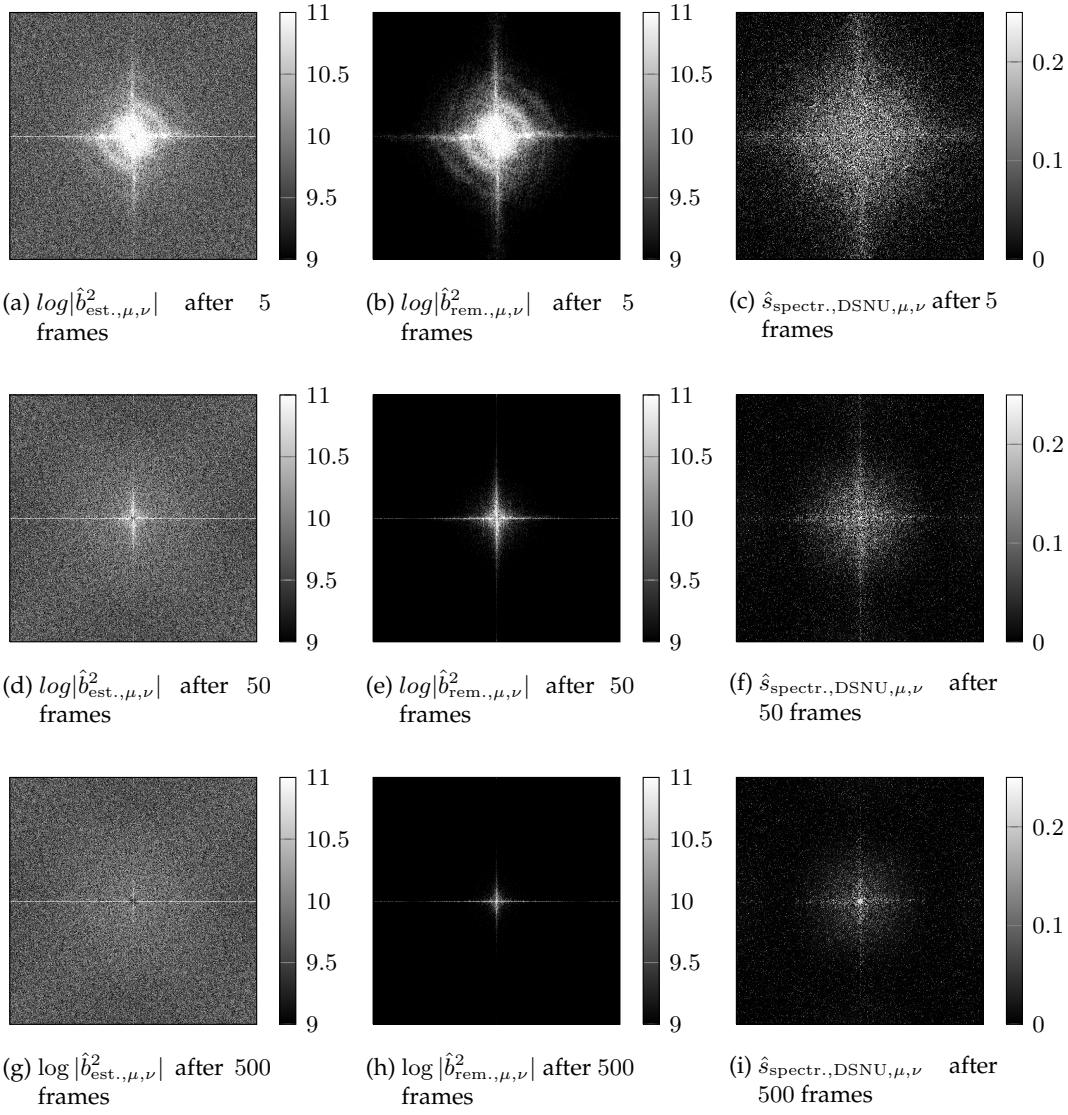
$$\hat{s}_{\text{spectr.,PRNU},\mu,\nu} = \frac{\hat{a}^2_{\text{rem.},\mu,\nu}}{\hat{a}^2_{\text{ref.},\mu,\nu}} \in [0, 1] \tag{3.43}$$

$$\hat{s}_{\text{spectr.,DSNU},\mu,\nu} = \frac{\hat{b}^2_{\text{rem.},\mu,\nu}}{\hat{b}^2_{\text{ref.},\mu,\nu}} \in [0, 1] \tag{3.44}$$

This definition is motivated by the fact that $\hat{x}^2_{\text{rem.},\mu,\nu}$ represents an sort-of equivalent for the physical energy of the electromagnetic field. Which then allows to interpret $\hat{s}$ as the damping coefficient for the corresponding frequency components, with respect to its energy equivalent. The resulting numbers of $\hat{s}$ should reside between $[0, 1]$, where a 0 states a complete correction of the frequency component, while a 1 indicates no correction. Malfunction of the nonuniformity correction algorithm may however result in an amplification of some spectral frequencies which then leads to numbers that are larger than 1.

Fig. 3.4 shows the typical behavior of a scene based nonuniformity correction method for $5, 50$ and 500 given input frames in the spectral analysis of the estimated and remaining nonuniformities as well as the resulting damping factors $\hat{s}$. The high frequency components are typically corrected first which is visible by the darker areas in the border regions of the representations. With more frames, the lower frequencies are corrected as well, leading to the observed remainders

of bright pixels in the center of the representations. In general all Fourier based representations in this thesis are normalized onto the corresponding Nyquist frequencies $\in (-1, 1)$ for each dimension. An interesting point to observe is the power spectral damping factor $\hat{s}$ according to eq. 3.43, which rarely goes above the 0.2 mark. This indicates that a damping of the nonuniformities takes place for all spectral frequencies, even in the case of only 5 frames input.



(a) $log|\hat{b}^2_{\text{est.},\mu,\nu}|$ after 5 frames
(b) $log|\hat{b}^2_{\text{rem.},\mu,\nu}|$ after 5 frames
(c) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ after 5 frames
(d) $log|\hat{b}^2_{\text{est.},\mu,\nu}|$ after 50 frames
(e) $log|\hat{b}^2_{\text{rem.},\mu,\nu}|$ after 50 frames
(f) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ after 50 frames
(g) $\log|\hat{b}^2_{\text{est.},\mu,\nu}|$ after 500 frames
(h) $\log|\hat{b}^2_{\text{rem.},\mu,\nu}|$ after 500 frames
(i) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ after 500 frames

Figure 3.4.: Spectral representations for the estimated and remaining DSNU components, as well as the power spectral damping factor $\hat{s}$ according to eq. 3.43.

4. Image Data for Evaluations

The newly developed methods and the reference methods are evaluated against each other with help of defined image material recorded by 4 different types of cameras. The recorded image material covers a variety of typical use cases for camera applications. It contains high speed consecutive frame sequences, typical frame rate camera panning sequences and random frame sequences. With concern to the nonuniformities, the degradation of the images varies from hardly visible degradations to a strong degradation, depending on the type of camera used. In the original papers the reference methods [33, 4, 5, 6] were evaluated only on a few sequences. Therefore the new evaluations also give a deeper insight into the performance of the reference methods with this new experimental setups.

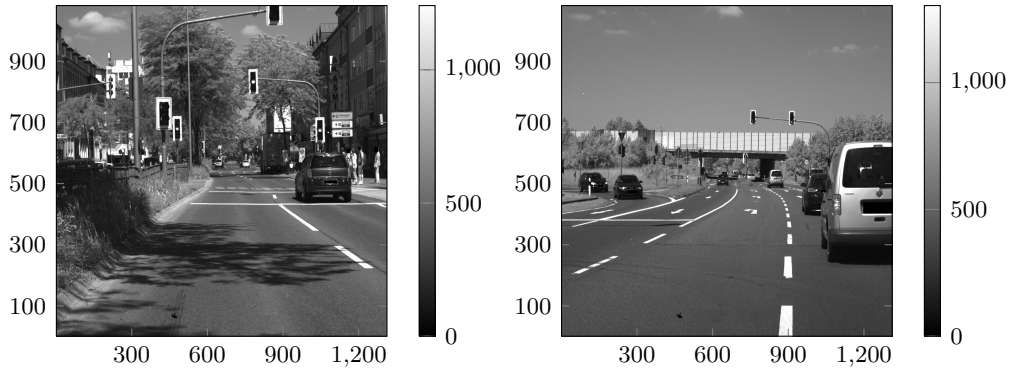
The used cameras are a Basler A602f, a photon-focus MV-D640-66-CL, a photon-focus MV-D640-66-CL-LINLOG, and a photon-focus MV1-D1312-160-CL camera. With the MV1-D1312-160-CL camera many sequences have been recorded while it was mounted inside a vehicle that was driving in public traffic (see sec. 4.1.1). Further scenes of a road scene in front of the lab have been recorded with all mentioned cameras by camera panning. These scenes have been captured at different exposure times and different temperatures and serve as a reference to evaluate the functionality of the algorithms in these setups. For each camera setup, the nonuniformities have been calibrated in a laboratory setup according to the EMVA1288 standard. The results are shown in the following sections of this chapter.

4.1. Description of the Recorded Data

4.1.1. Car Front Camera Scenes

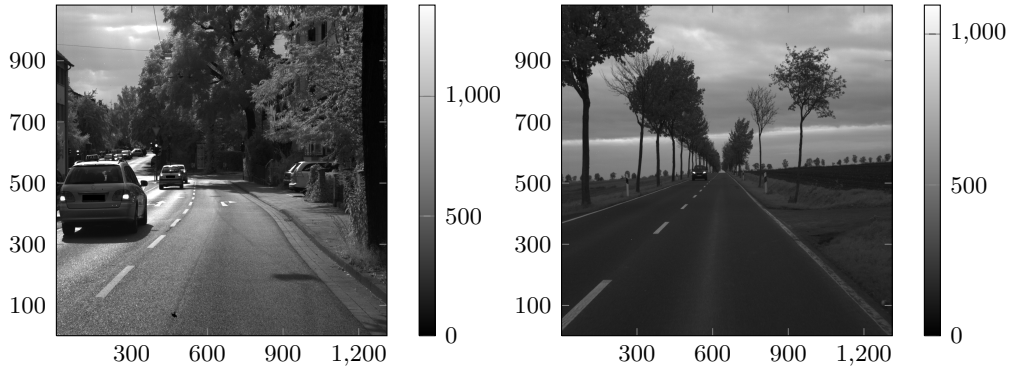
The scenes with the description *Car Front Camera* have been recorded with the right camera of a stereo camera setup developed at the HCI in Heidelberg [34]. Two MV1-D1312-160-CL cameras have been used to record sequences as the system was mounted to look out through the windshield in front of the vehicle. The used image sequences have a length of 500 consecutive frames and were recorded with a frame rate of 100 Hz. In total 108 sequences have been recorded at 27 waypoints on roads around the city of Hildesheim on the 4 days: 15.06.2010, 16.06.2010, 21.07.2010 and 22.10.2010. The image data shows day-to-day driving situations like highway, inner-city, pedestrians or busy crossings under different lighting and weather conditions. Fig. 4.1 shows some images out of these scenes and tab. 4.1

shows the summary of the recorded data. The markings ✓$_{Calib.}$ and ✓ are explained in sec. 4.2 and separate the sequences in a training and evalutaion data set.



(a) Inner city scene with pedestrians

(b) Feeder road near the Autobahn



(c) Residential area traffic

(d) Empty road

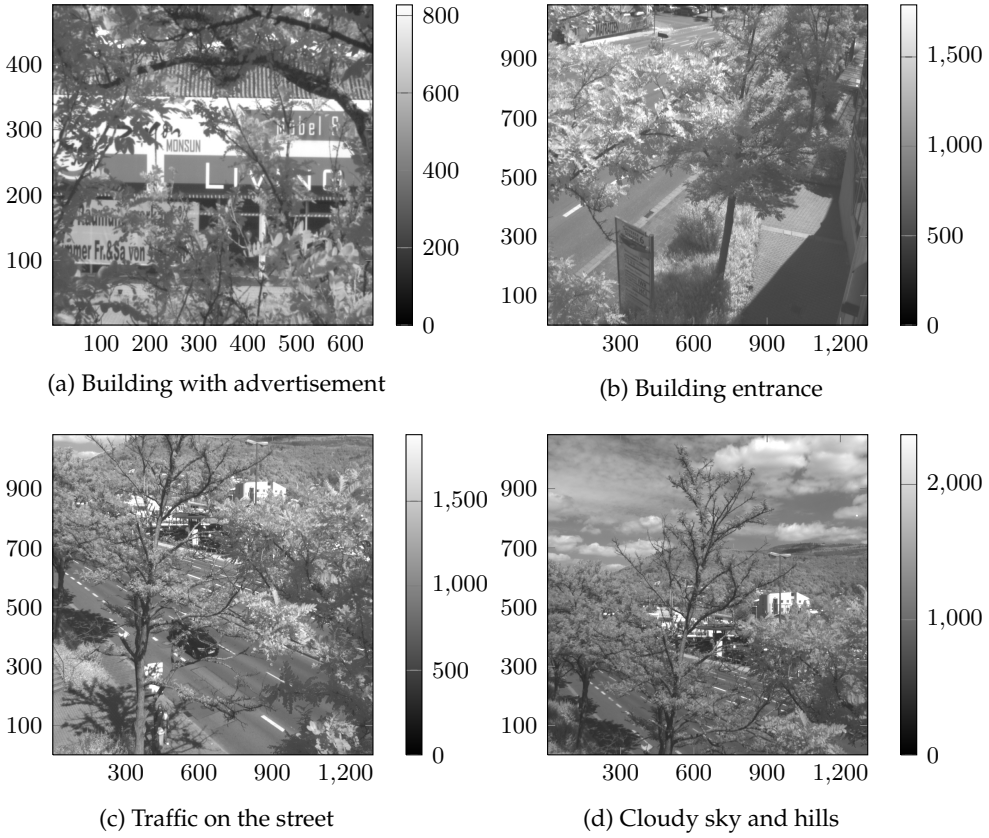
Figure 4.1.: Typical images from the *car front camera* sequences (see tab. 4.1)

Scene Content	15.06.2010	16.06.2010	21.07.2010	22.10.2010
Fuel Station	✓	✓	✓ $_{Calib.}$	✓ $_{Calib.}$
Bridge and highway entry	✓	✓	✓ $_{Calib.}$	✓
Busstop	✓	✓	✓	✓
Street in a village	✓	✓	✓	✓ $_{Calib.}$
Crossroads	✓	✓	✓	✓ $_{Calib.}$
Country Road Crossing	✓	✓ $_{Calib.}$	✓	✓
Trees	✓	✓	✓	✓
Pedestrians at high street	✓	✓	✓ $_{Calib.}$	✓
Small Country Road	✓	✓	✓	✓ $_{Calib.}$
Complex Crossroads	✓ $_{Calib.}$	✓	✓	✓
Highway with 6 lanes	✓ $_{Calib.}$	✓	✓	✓
Highway underpass	✓	✓ $_{Calib.}$	✓	✓
long road turn	✓ $_{Calib.}$	✓	✓	✓
Country road 4 lanes	✓	✓	✓	✓ $_{Calib.}$
Trees and bushes besides the road	✓	✓	✓ $_{Calib.}$	✓
Village passing of Giesen 1	✓	✓	✓	✓ $_{Calib.}$
Village passing of Giesen 2	✓	✓	✓ $_{Calib.}$	✓
Between high buildings	✓	✓	✓	✓ $_{Calib.}$
Car parking at roadside	✓ $_{Calib.}$	✓	✓	✓
Pedestrian crossings	✓	✓ $_{Calib.}$	✓	✓
Inner city	✓	✓ $_{Calib.}$	✓	✓
Right turn before Giesen	✓	✓	✓	✓
Avenue	✓	✓ $_{Calib.}$	✓ $_{Calib.}$	✓
Emmerke train crossing	✓ $_{Calib.}$	✓	✓	✓
Residential area	✓	✓	✓	✓
Underpass central train station	✓	✓ $_{Calib.}$	✓	✓
Country site	✓	✓ $_{Calib.}$	✓	✓

Table 4.1.: Overview of the recorded sequences with the Photonfocus MV-D640-66-CL samera. The markings ✓ and ✓ $_{Calib.}$ indicate the membership of the sequence to the training or evaluation data set.

4.1.2. L600a in HD

The scenes with the description *L600a in HD* contain images that have been captured by a camera panning over the street *L600a* in Heidelberg. In the city of Heidelberg the street is also known by the name Speyerer Strasse and the images have been recorded out of the window in the laboratory of building number 6 (HCI building). The captured scene contains the road outside the lab with its traffic as well as the hills of Heidelberg in the background and a short view into the cloudy sky. Fig. 4.2 shows some typical images out of these scenes. These scenes have been recorded with a frame rate of 20 Hz.



(a) Building with advertisement

(b) Building entrance

(c) Traffic on the street

(d) Cloudy sky and hills

Figure 4.2.: Typical images in the scene *L600a in HD*

4.2. Generation of Evaluation and Calibration Sequences

As most of the image material is recorded with the MV1-D1312-160-CL camera, this material was used for the algorithm development and the statistical analysis of its performance. The camera panning sequences of the other cameras have been used to evaluate the general ability of the algorithms to deal with other types of nonuniformities and other types of cameras and frame rates.

The data was divided into an evaluation and a calibration set, where the calibration data is used to train the free parameters of the algorithms to reach an optimal performance. The evaluation data is then used to calculate the performance measures including their statistical deviations with a given set of fixed parameters. Two types of image sequences are possible to be generated out of the MV1-D1312-160-CL recordings: Sequences with consecutive frames and sequences with randomly chosen frames. Both are scenarios that can be created by car front cameras to calibrate the cameras. All the image material, sequence definitions and source code snippets are available in the archives of the the *HCI Ground Truth Group*, url: `hci.iwr.uni-heidelberg.de`[1].

Consecutive Frames: Consecutive frame sequences are useful to simulate scenarios where a very fast drifting nonuniformity is present in the imaging system. A fast adaptation is then needed, but of course the information in high frame rate consecutive frames may be limited. For example, the 500 frame sequences of the MV1-D1312-160-CL camera have been recorded at $100\,\mathrm{Hz}$ frame rate, which gives only 5 seconds of information. Given a typical driving situation, this information may not be sufficient to correct the nonuniformities and the algorithms will fail in these sequences for several reasons, e.g. if an overexposed area stays for these 5 seconds in the image, or if parts of the scene are not moving enough.

For the consecutive frames, the *Car Front Camera* scenes have been randomly divided into 25 calibration and 83 evaluation sequences. The choice of the sequences is shown in tab. 4.1, where the calibration sequences have been marked with $\checkmark_{Calib.}$. The 83 evaluations sequences are marked ✓ and are used for the statistical evaluations.

At some points in the research the evaluation of longer sequences seems necessary and two randomly chosen sequences have been concatenated to produce 1000 frame sequences. 25 sequences each have been produced, but at the point where the sequences are concatenated, the scene changes suddenly which may not happen that often for actual consecutively recorded data.

Random Frames: Random frame sequences are used to simulate scenarios where either a very slow capture rate is present, or scenarios in which a choice of some random frames out of the last N recorded frames is used for calibration. These sequences are especially useful if the nonuniformity does not change very fast. This is often the case for classical CMOS near infrared cameras that are based on thicker layers of p-doted-silicon. As the typical scene based nonuniformity correction algorithms do not need a constant frame rate input, random sequences are the better choice due to its variety of independent information.

[1] Person of contact is Daniel Kondermann

The random frame sequences have been created by selecting N frames randomly out of all the recordings with the MV1-D1312-160-CL camera on one day, as presented in tab. 4.1. Similar to the consecutive sequences, there have been created 25 calibration sequences and 83 evaluation sequences with a length of 500 frames and 25 sequences with the length of 1000 frames.

4.3. Description of the Cameras and their Nonuniformities

(a) The Photonfocus MV1-D1312-160-CL Camera (b) The Basler A602f Camera (c) The Photonfocus MV-D640-66-CL and MV-D640-66-CL-LinLog Camera

Figure 4.3.: Pictures of the 4 used camera types

Fig. 4.3 shows the outer appearance of the four used cameras, where the MV-D640-66-CL and the MV-D640-66-CL-LINLOG have the same housings. Tab. 4.2 gives an overview of the typical camera parameters like resolution, pixel/light sensor sizes, minimum and maximum operation temperatures and the maximum achievable frame rate $\mathrm{fps}_{\max}$.

Camera	Resolution	Pixel [µm^2]	$\mathrm{T}_{\min}[°\mathrm{C}]$	$\mathrm{T}_{\max}[°\mathrm{C}]$	$\mathrm{fps}_{\max}$
MV1-D1312-160-CL	1312×1082	8.0×8.0	0	50	108
MV-D640-66-CL	640×480	9.9×9.9	0	60	200
MV-D640-66-CL-LINLOG	640×480	9.9×9.9	0	60	200
A602f	656×491	9.9×9.9	0	50	100

Table 4.2.: Overview of some typical parameters of the used cameras

These and further information about the cameras can be found at the homepages of the manufacturing companies Photonfocus [35] and Basler AG [36], and in the manuals of the cameras [37, 38, 39].

Prior to the recordings of the sequences, an EMVA1288 compliant ground truth calibration of the cameras has been conducted. A CoolTronic TC3224 temperature controller [40] was used to adjust the cameras' temperatures with help of peltier

elements. The calibrations then took place with help of an integrating sphere with a set of colored and calibrated LED light sources. Further information about the calibration procedure can be found in the EMVA1288 document [1] and the measurement equipment is also well described in the PhD thesis of Michael Erz [41]. Fig. 4.4 gives an impression of the calibration setup.

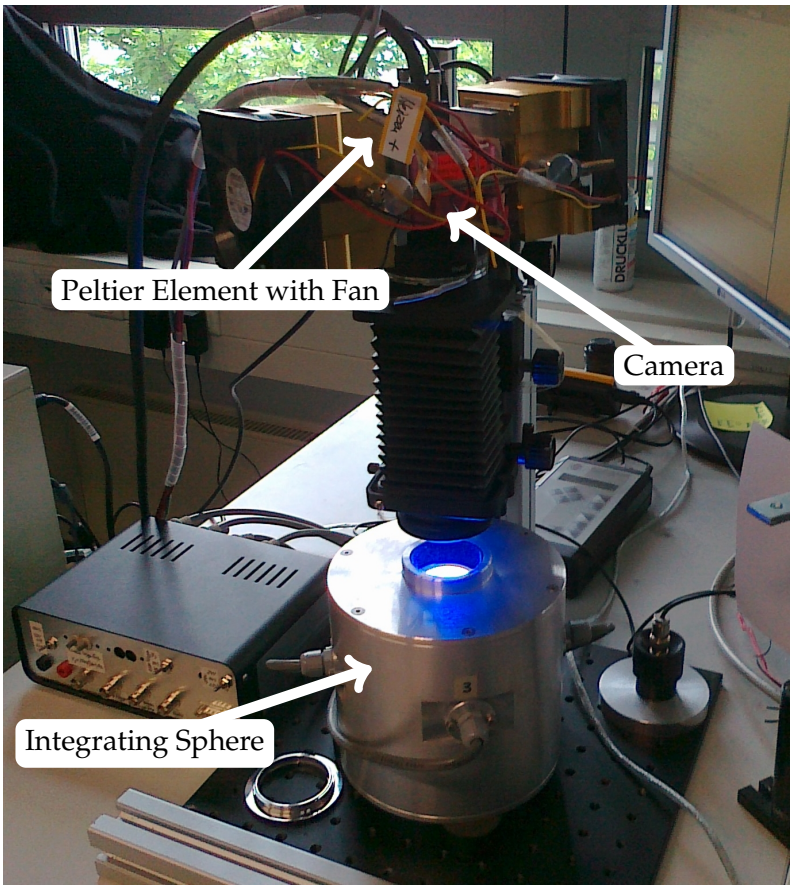


Figure 4.4.: Calibration setup for the EMVA1288 DSNU and PRNU estimation, depicting the peltier elements for temperature regulation and the integrating sphere to provide uniform illumination

4.3.1. Photonfocus MV1-D1312-160-CL

Tab. 4.2 and fig. 4.3a have already provided a first overview of this camera. A further very important feature of this camera is the possibility to switch off all the image enhancing techniques. This allows to observe the sensor responses directly. For the development of the nonuniformity estimation methods this feature will lead to a best possible fit of the sensor models to the actual sensor responses. The EMVA1288 calibration was performed prior to the recording of the scenes at different temperatures and exposure times. The temperature limit was chosen ac-

cording to the specifications of the manual at $50\,°\mathrm{C}$ [39]. The exposure time was limited to not more than $10\,\mathrm{ms}$ due to occurring overexposure even with fully closed aperture of the used optics.

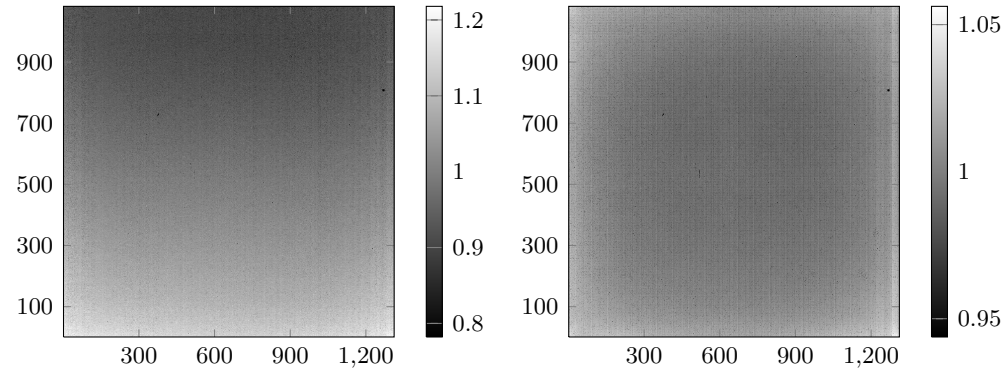
Sequence Name	$t_{\mathrm{exp.}}$ [ms]	T [°C]	Scene	$\mathrm{DSNU}_{\mathrm{DN}}$	PRNU
MV1-D1312-CarFront-X ($X \in [1, 108]$)	1.2	23	Car Front Camera	49.84	1.88
MV1-D1312-1a,b	0.05	25	L600a in HD	50.80	7.26
MV1-D1312-1a,b	5	25	L600a in HD	50.81	1.92
MV1-D1312-1a,b	10	25	L600a in HD	50.83	1.88
MV1-D1312-1a,b	0.05	49	L600a in HD	51.07	7.39
MV1-D1312-1a,b	5	49	L600a in HD	51.19	1.91
MV1-D1312-1a,b	10	49	L600a in HD	51.44	1.87

Table 4.3.: Overview of the recorded sequences with the Photonfocus MV1-D1312-160-CL camera

Tab. 4.6 summarizes the different temperature and exposure time setups for the experimental setups as well as the measured DSNU and PRNU values. As the table states, there have been 108 sequences recorded as the camera was mounted inside a vehicle in public traffic, looking through the windshield in front (Car-Front-Camera). These images represent the majority of the used image data for this thesis. A detailed description of these scenes is provided in sec. 4.1.1. Additionally 12 sequences have been recorded by a *camera panning* of the scene *L600a in HD* at different temperatures and exposure times. This scene is described in sec. 4.1.2. With concern to the measured DSNU and PRNU parameters, the behavior of the DSNU increases with temperature and exposure time as expected. The large number of $\mathrm{DSNU}_{\mathrm{DN}} \approx 50\,\mathrm{DN}$ hints towards a visible image degradation. The PRNU however shows a severe increase for the minimum exposure time of $0.05\,\mathrm{ms}$, which most likely emerges from effects of the analogue amplification circuity and the short reset times. This effect can also be observed for the other cameras, if the exposure time was limited below $1\,\mathrm{ms}$.

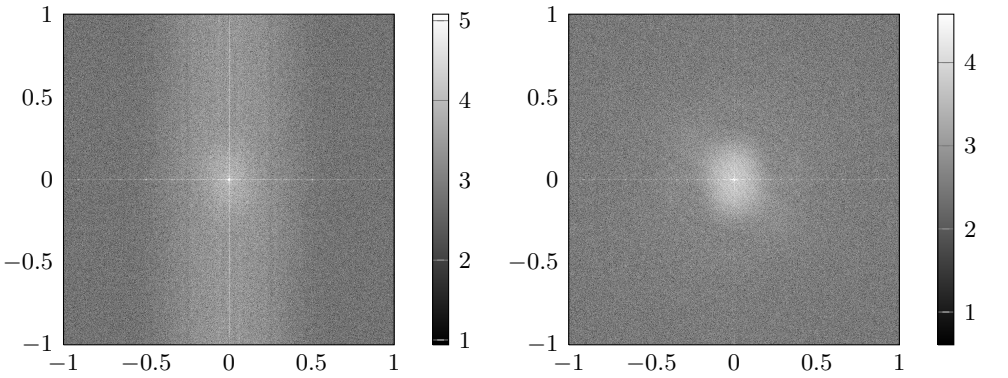
Fig. 4.5 and 4.6 show the analysis of the PRNU and DSNU as described in sec. 3.7 and 3.6. For the nonuniformity parameter sets $\{a_{i,j}\}$ and $\{b_{i,j}\}$ the most differing parameters are shown ($10\,\mathrm{ms}\,@\,50\,°\mathrm{C}$ and $0.05\,\mathrm{ms}\,@\,25\,°\mathrm{C}$) and the representations of the remaining evaluations can be found in appendix B.1.

The visual impression of the PRNU shows a significant spatial increase of the gain factors towards the lower end of the image (see fig. 4.5a), which is not present at longer exposure times (see fig. 4.5b and appendix B.1). The PRNU shows also dominant row and column wise components, with a higher amplitude in the column wise component. These components are visible as well in the Fourier analysis in fig. 4.5c and 4.5d by the central horizontal and vertical lines of 1 pixel width. In addition to the row and column wise components, the PRNU seems to express dominant low spatial frequencies, which usually are the result of the vignetting



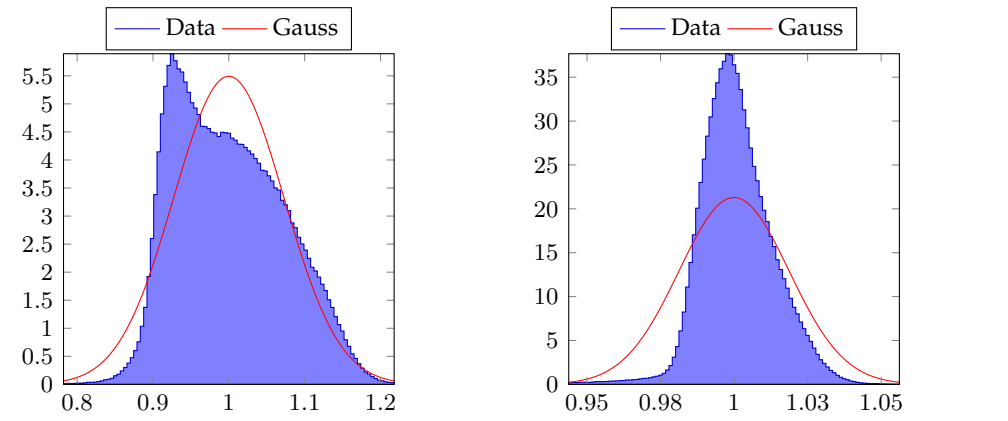
(a) Spatial PRNU for 0.05 ms @ 25 °C

(b) Spatial PRNU for 10 ms @ 50 °C



(c) Spectral: $\log|\hat{a}_{\mu,\nu}|$ for 0.05 ms @ 25 °C

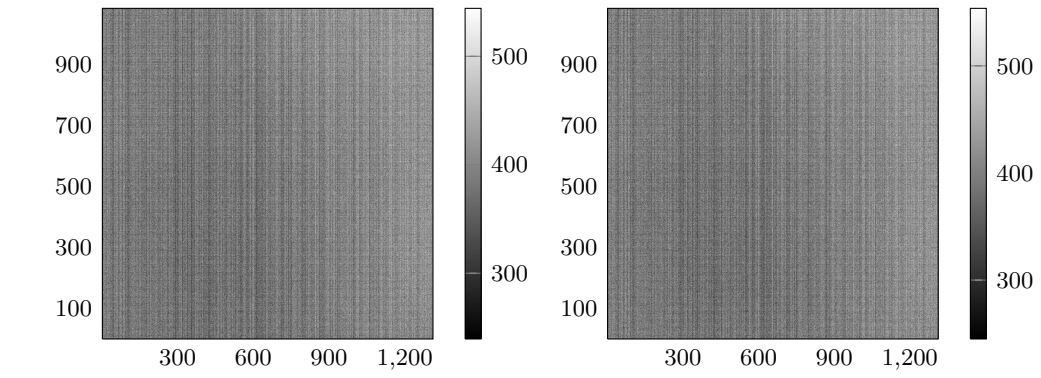
(d) Spectral: $\log|\hat{a}_{\mu,\nu}|$ for 10 ms @ 50 °C



(e) Histogram of $\{a_{i,j}\}$ for 0.05 ms @ 25 °C

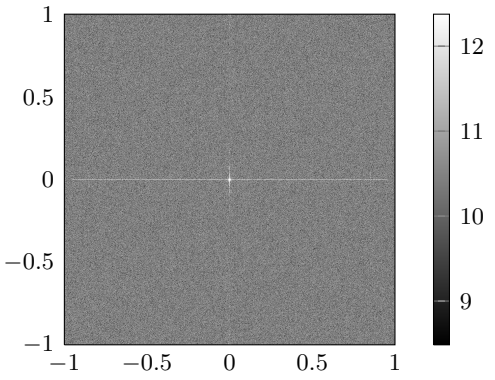
(f) Histogram of $\{a_{i,j}\}$ for 10 ms @ 50 °C

Figure 4.5.: Representations of the $\{a_{i,j}\}$ correction patterns of the EMVA1288 calibration measurements with the Photonfocus MV1-D1312-160-CL camera. Showing the spatial and the spectral analysis as well as histograms at the given temperatures and exposure times.

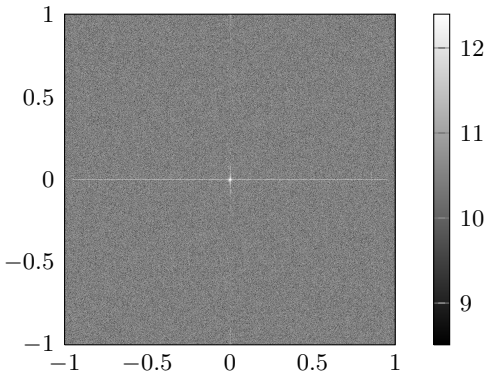


(a) Spatial DSNU for 0.05 ms @ 25 °C

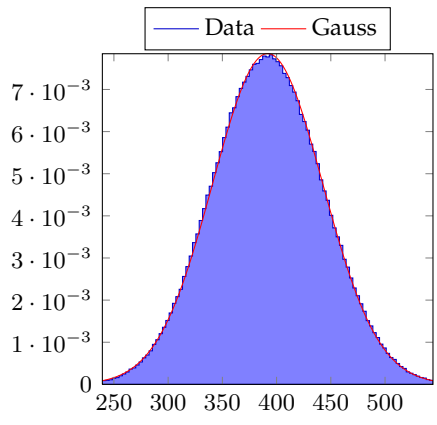
(b) Spatial DSNU for 10 ms @ 50 °C



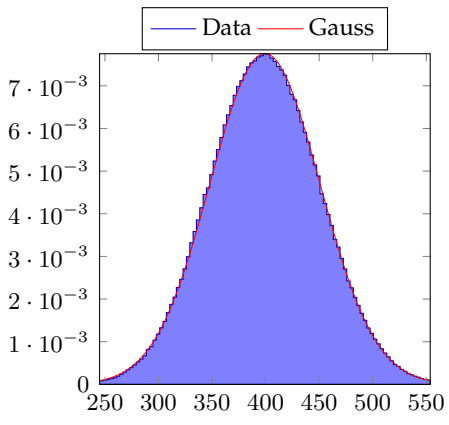
(c) Spectral:log $|\hat{b}_{\mu,\nu}|$ for 0.05 ms @ 25 °C



(d) Spectral:log $|\hat{b}_{\mu,\nu}|$ for 10 ms @ 50 °C



(e) Histogram of $\{b_{i,j}\}$ for 0.05 ms @ 25 °C



(f) Histogram of $\{b_{i,j}\}$ for 10 ms @ 50 °C

Figure 4.6.: Representations of the $\{b_{i,j}\}$ correction patterns of the EMVA1288 calibration measurements with the Photonfocus MV1-D1312-160-CL camera. Showing the spatial and the spectral analysis as well as histograms at the given temperatures and exposure times

from the optic's mount. The optics themselves are removed for EMVA1288 calibration. The histograms of $\{a_{i,j}\}$ do not completely fit Gaussian distributions (see fig. 4.5f and 4.5e), but due to the large variations between the camera types and exposure times, a Gaussian distribution seems to be the best fit for the later developments (see chapter 8).

The visual impression of the DSNU shows dominant row and column wise components as well (see fig. 4.6a and fig. 4.6b), with a higher amplitude in the column wise component. The spectral analysis shows that besides the row and column wise components, the $\{b_{i,j}\}$ distributes like white noise (see fig. 4.6c and fig. 4.6d). The histograms further show that the spatial DSNU distribution obeys almost perfectly a Gaussian distribution (see fig. 4.6e and fig. 4.6f), which supports the assumptions used in chapter 6.

4.3.2. Basler A602f

The Basler A602f is an industrial camera with a fire wire connection (IEEE 1394a). Tab. 4.2 and fig. 4.3b have already provided a first overview of this camera. The Basler company is one of the few companies that publishes the results of the EMVA1288 calibration of their cameras according to EMVA1288 [1]. For the A602f camera, the results for the nonuniformity measures are:

$$\mathrm{DSNU}_{EMVA1288} = (53.1 \pm 2.3)\mathrm{e}^- \tag{4.1}$$

$$\mathrm{PRNU}_{EMVA1288} = (1.0 \pm 0.2)\% \tag{4.2}$$

$$K = \frac{1}{59.6}\left[\frac{\mathrm{DN}}{\mathrm{e}^-}\right] \tag{4.3}$$

Further information can be found online in the EMVA1288 documentation for the A602f [42]. As mentioned already in sec. 3.2.1 for the scene based nonuniformity estimations, the measure for the DSNU has to be given in DN and not in e^- for scene based NUC. Therefore the DSNU measure needs to be multiplied with the overall system gain which results in:

$$\mathrm{DSNU}_{EMVA1288,DN} = (0.89 \pm 0.03)\mathrm{DN} \tag{4.4}$$

This is a DSNU that will obviously be hard to detect in a recorded frame, because if $\{b_{i,j}\}$ obey approximately a Gaussian distribution, then more than $68.27\,\%$ of all $\{b_{i,j}\}$ realizations are valued less than $1\,\mathrm{DN}$.

For the evaluations in this thesis an EMVA1288 calibration was performed prior to the recording of the scenes. Tab. 4.4 summarizes the different temperature and exposure time setups that were used as well as the measured DSNU and PRNU values. Higher temperatures than $50\,^\circ\mathrm{C}$ have not been allowed in the camera specification and longer exposure times than $10\,\mathrm{ms}$ led to overexposed images, even with fully closed aperture of the optics.

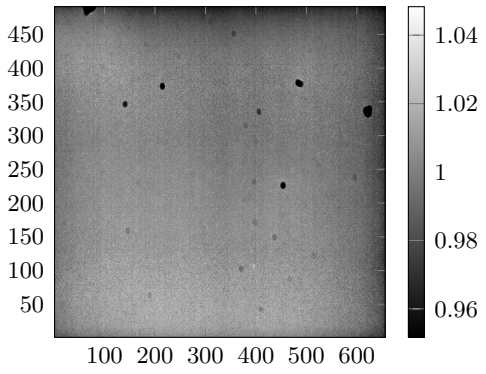
Sequence Name	$t_{\text{exp.}}$ [ms]	T [°C]	Scene	DSNU_{DN}	PRNU
A602f-1a, A602f-1b	1	22	L600a in HD	1.34	1.61
A602f-4a, A602f-4b	10	22	L600a in HD	1.29	1.43
A602f-2a, A602f-2b	1	30	L600a in HD	0.89	1.58
A602f-5a, A602f-5b	10	30	L600a in HD	0.75	1.44
A602f-3a, A602f-3b	1	50	L600a in HD	2.09	1.78
A602f-6a, A602f-6b	10	50	L600a in HD	2.33	1.53

Table 4.4.: Overview of the recorded Sequences with the Basler A602f Camera

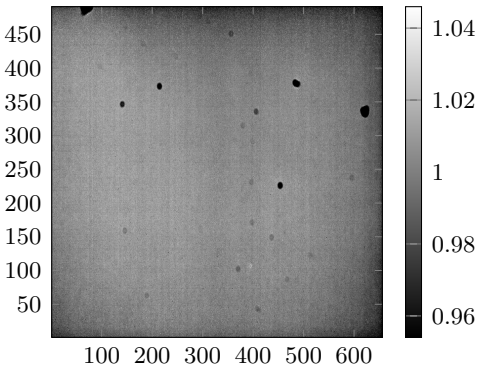
As the tab. 4.4 states, 2 sequences have been recorded per setup by a *camera panning* of the scene *L600a in HD*. A full description of the scene is given in sec. 4.1.2. With concern to the measured DSNU and PRNU parameters, the expected increase with exposure time and temperature could not be observed in all cases. The Basler EMVA1288 document [42] does not provide these information either. It is very likely that an internal temperature and exposure time dependent nonuniformity correction takes place that results in the low numbers we have measured. The measured DSNU numbers fit the expected $0.9\,\text{DN}$ at $30\,^{\circ}\text{C}$ room temperature. The measured PRNU numbers are outside the given $(1.0 \pm 0.2)\%$ range, which is most likely a result of clusters of defective pixels or dust particles as discussed below.

Fig. 4.7 and 4.8 show the nonuniformity parameter sets $\{a_{i,j}\}$ and $\{b_{i,j}\}$ for the two samples at $(1\,\text{ms} \,@\, 22\,^{\circ}\text{C})$ and $(10\,\text{ms} \,@\, 50\,^{\circ}\text{C})$. The visual representations of the other evaluations can be found in appendix B.2. The PRNU shows particle-like degradations as well as column and row wise components, with a lower amplitude in the row wise components. Before the measurements pressure air was used to clean the image sensor, thus we believe the particle-like degradations have to be considered as clusters of degraded pixels. For the row and column wise components, the spectral analysis in fig. 4.7c and 4.7d show a severe increase in all vertical low frequencies. Furthermore, at ± 0.5 of the column Nyquist frequency, sharp vertical row-wise components show up. The histograms do not fit a Gaussian distribution perfectly and look more like a skewed Gaussian distribution.

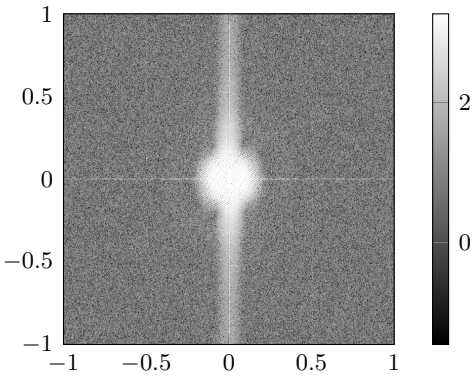
The analysis for the $\{b_{i,j}\}$ shows a different behavior to the rest of the array for the upper rows. Fig. 4.7d shows a small additional increase in the spatial low frequencies' amplitudes. For the spectral analysis, the components for ± 0.5 of the column Nyquist frequency show up again in fig. 4.8c and 4.8d. The histograms for the DSNU show a good similarity to the Gaussian distribution (see fig. 4.8e and 4.8f).



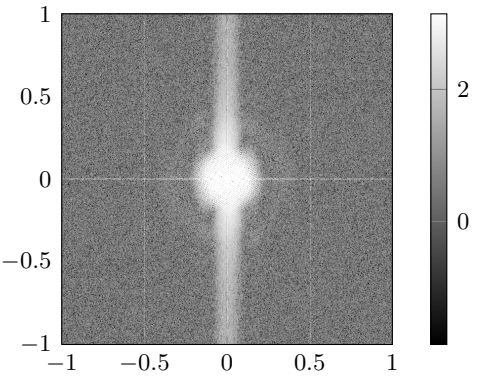
(a) Spatial PRNU: (1 ms @ 22 °C)



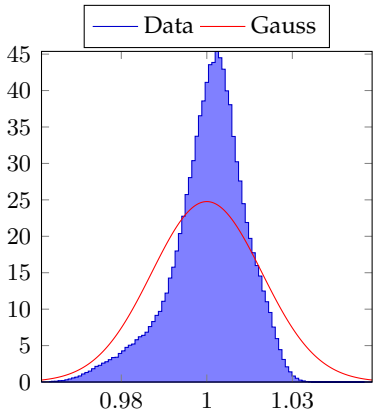
(b) Spatial PRNU: (10 ms @ 50 °C)



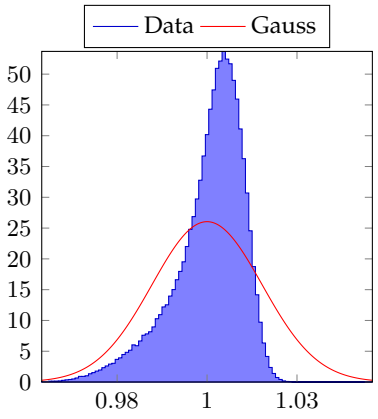
(c) Spectral:log $|\hat{a}_{\mu,\nu}|$: (1 ms @ 22 °C)



(d) Spectral:log $|\hat{a}_{\mu,\nu}|$ with (10 ms @ 50 °C)

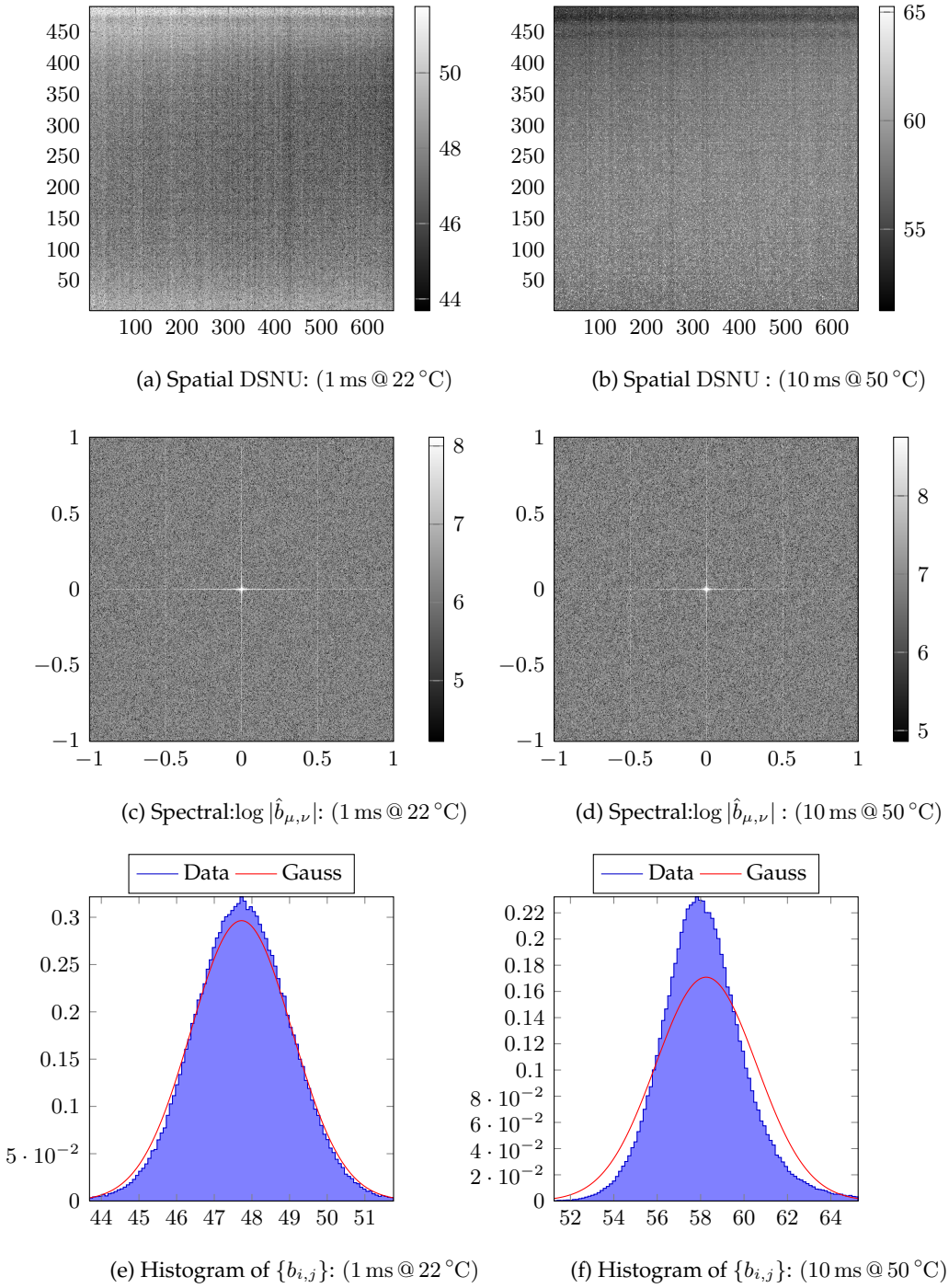


(e) Histogram of $\{a_{i,j}\}$: (1 ms @ 22 °C)



(f) Histogram of $\{a_{i,j}\}$: (10 ms @ 50 °C)

Figure 4.7.: Representations of the $\{b_{i,j}\}$ correction patterns of the EMVA1288 calibration measurements with the Basler A602f camera. Showing the spatial and the spectral analysis as well as a histograms at the given temperatures and exposure times.



(a) Spatial DSNU: (1 ms @ 22 °C)

(b) Spatial DSNU : (10 ms @ 50 °C)

(c) Spectral:log $|\hat{b}_{\mu,\nu}|$: (1 ms @ 22 °C)

(d) Spectral:log $|\hat{b}_{\mu,\nu}|$: (10 ms @ 50 °C)

(e) Histogram of $\{b_{i,j}\}$: (1 ms @ 22 °C)

(f) Histogram of $\{b_{i,j}\}$: (10 ms @ 50 °C)

Figure 4.8.: Representations of the $\{b_{i,j}\}$ correction patterns of the EMVA1288 calibration measurements with the Basler A602f camera. Showing the spatial and the spectral analysis as well as histograms at the given temperatures and exposure times.

4.3.3. Photonfocus MV-D640-66-CL

The Photonfocus MV-D640-66-CL is an industrial camera with a camera link connection. Tab. 4.2 and fig. 4.3c have already provided a first overview of this camera. Further information on the camera can be found at the manufacturer's website and the manual [35, 38].

For the EMVA1288 calibrations the limits of the given specifications were used to determine the upper temperature limit. Exposure times longer than $10\,\mathrm{ms}$ led to overexposed images, even with fully closed aperture of the optics. Tab. 4.5 summarizes the different temperature and exposure time setups that were used for the recordings as well as the measured DSNU and PRNU values.

Sequence Name	$t_{\mathrm{exp.}}$ [ms]	T [°C]	Scene	$\mathrm{DSNU}_{\mathrm{DN}}$	PRNU
MV-D640-66-CL-1	5	25	L600a in HD	0.50	1.57
MV-D640-66-CL-2	10	25	L600a in HD	0.55	1.24
MV-D640-66-CL-3	5	50	L600a in HD	1.27	1.50
MV-D640-66-CL-4	10	50	L600a in HD	1.39	1.24

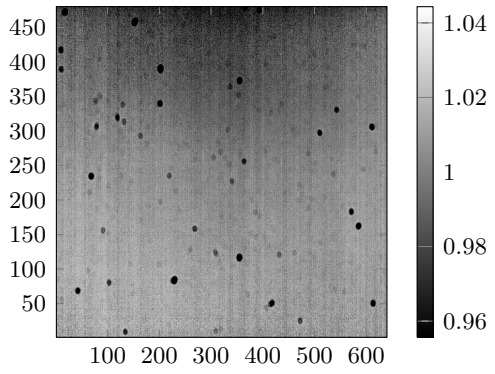
Table 4.5.: Overview of the recorded sequences with the Photonfocus MV-D640-66-CL camera

As tab. 4.5 states, there are 4 recorded sequences by *camera panning* of the scene *L600a in HD*. A full description of the scene is given in sec. 4.1.2. With concern to the measured DSNU and PRNU parameters, the temperature and exposure time depended behavior fits better than for the Basler A602f camera, but here as well, some internal correction seems to influence at least the PRNU measurements.

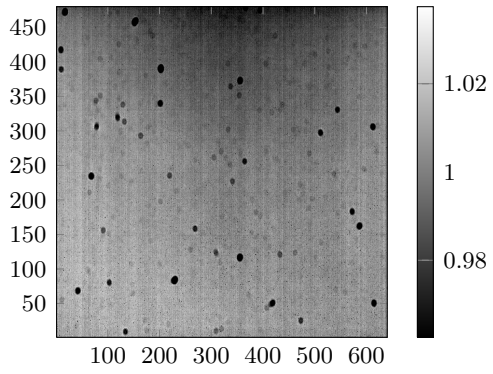
Fig. 4.9 and 4.10 show the nonuniformity parameter sets $\{a_{i,j}\}$ and $\{b_{i,j}\}$ for the most differing setups at ($10\,\mathrm{ms}$ @ $50\,°\mathrm{C}$) and ($5\,\mathrm{ms}$ @ $25\,°\mathrm{C}$). The visual representations of the other evaluations can be found in appendix B.3. The visual impression of the PRNU shows many more particle like degradations than the Basler camera does. Again, the cleaning with pressure air did not remove those dust particle-like degradations. The PRNU in general seems to be very stable with respect to exposure time and temperature (see fig. 4.9a and fig. 4.9b), however the critical measurements at extremely short exposure times were not possible.

Dominant row and column wise components, are visible again and show up in the Fourier analysis in fig. 4.9c and fig. 4.9d as discussed before. In addition to the row and column wise components, the PRNU expresses dominant low spatial frequencies that may be the effects of the optic's mount.

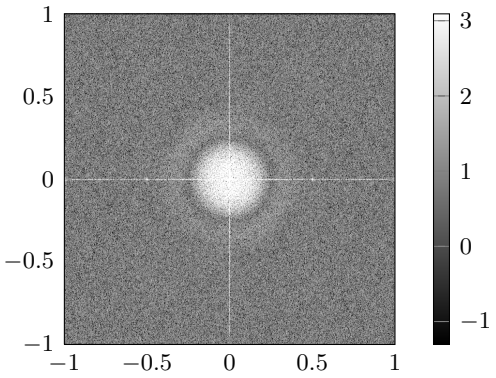
The visual impression of the DSNU contains row and column wise components (see fig. 4.10a and fig. 4.10b). The spectral analysis in fig. 4.10c and 4.10d then represents the row and column wise components as the centered horizontal and vertical lines again. The DSNU histograms further show that there is no similarity to a Gaussian distribution for low temperatures, while higher temperatures provide a better Gaussian approximation (fig. 4.10e and fig. 4.10f).



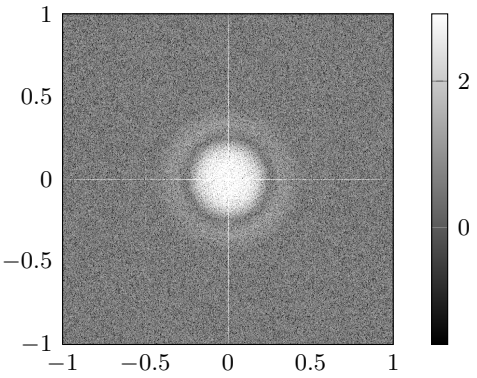
(a) Spatial PRNU for 5 ms @ 25 °C



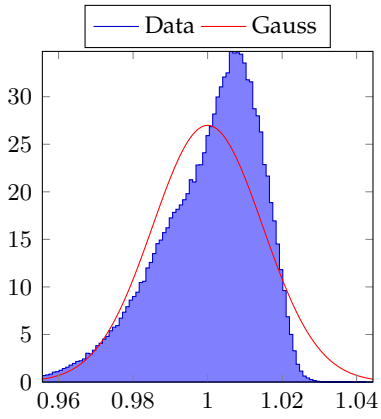
(b) Spatial PRNU for 10 ms @ 50 °C



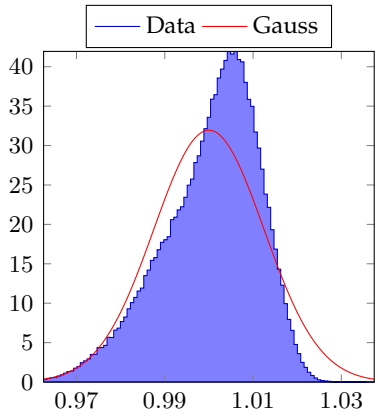
(c) Spectral: $\log|\hat{a}_{\mu,\nu}|$ for 5 ms @ 25 °C



(d) Spectral: $\log|\hat{a}_{\mu,\nu}|$ for 10 ms @ 50 °C

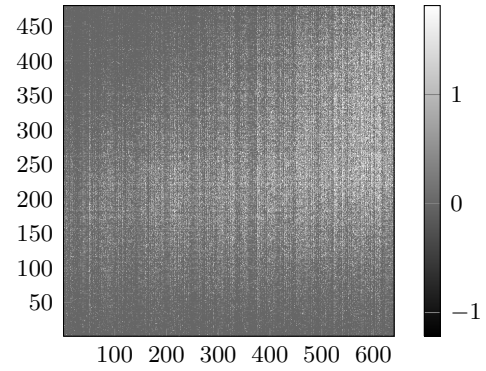


(e) Histogram of $\{a_{i,j}\}$ for 5 ms @ 25 °C

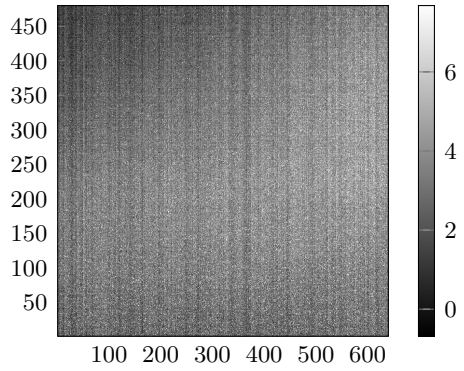


(f) Histogram of $\{a_{i,j}\}$ for 10 ms @ 50 °C

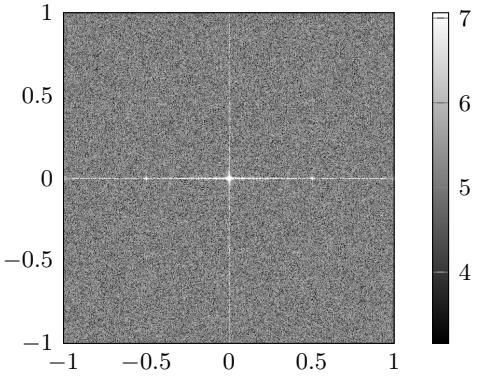
Figure 4.9.: Representations of the $\{a_{i,j}\}$ correction patterns of the EMVA1288 calibration measurements with the Photonfocus MV-D640-66-CL camera. Showing the spatial and the spectral analysis as well as histograms at the given temperatures and exposure times.



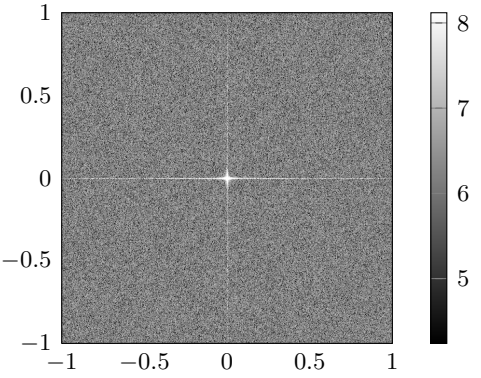
(a) Spatial DSNU for 5 ms @ 25 °C



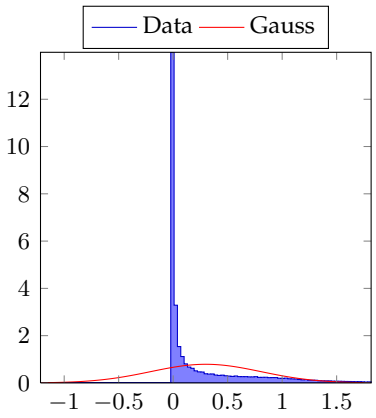
(b) Spatial DSNUfor 10 ms @ 50 °C



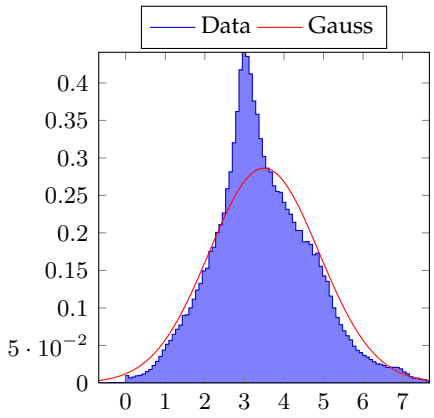
(c) Spectral: $\log|\hat{b}_{\mu,\nu}|$ for 5 ms @ 25 °C



(d) Spectral: $\log|\hat{b}_{\mu,\nu}|$ for 10 ms @ 50 °C



(e) Histogram of $\{b_{i,j}\}$ for 5 ms @ 25 °C



(f) Histogram of $\{b_{i,j}\}$ for 10 ms @ 50 °C

Figure 4.10.: Representations of the $\{b_{i,j}\}$ correction patterns of the EMVA1288 calibration measurements with the Photonfocus MV-D640-66-CL camera. Showing the spatial and the spectral analysis as well as histograms at the given temperatures and exposure times.

4.3.4. Photonfocus MV-D640-66-CL-LinLog

The Photonfocus MV-D640-66-CL-LinLog is an industrial camera with a camera link connection. Tab. 4.2 and fig. 4.3c have already provided a first overview of this camera. Further information on the camera can again be found at the manufacturer's website and the manual [35, 38]. The LinLog camera actually differs from the MV-D640-66-CL only in its sensor model, which does *not* provide a linear dependency between the light's intensity and the measured digital numbers. Therefore this camera contributes only to show the possible limitations with respect to the sensor model of the algorithms.

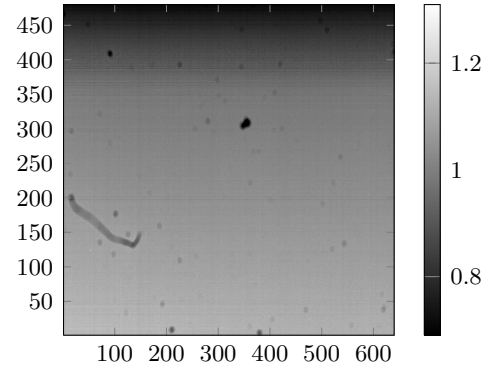
The limits of the specifications were used to determine the upper temperature limit at $50\,^{\circ}\mathrm{C}$. Exposure times longer than $10\,\mathrm{ms}$ were not possible due to overexposure of the images, as mentioned for the other cameras before. Tab. 4.6 summarizes the different temperature and exposure time setups of the camera that were used in the calibrations and recordings as well as the measured DSNU and PRNU values. As the table states, 6 sequences sequences have been recorded by a *camera panning* of the scene *L600a in HD*. A full description of the scene is given in sec. 4.1.2.

Sequence Name	$t_{\mathrm{exp.}}$ [ms]	T [°C]	Scene	$\mathrm{DSNU}_{\mathrm{DN}}$	PRNU
MV-D640-LinLog-1	0.082	25	L600a in HD	1.54	10.32
MV-D640-LinLog-2	5	25	L600a in HD	1.55	1.86
MV-D640-LinLog-3	10	25	L600a in HD	1.55	1.81
MV-D640-LinLog-4	0.082	50	L600a in HD	1.97	10.71
MV-D640-LinLog-5	5	50	L600a in HD	1.95	1.87
MV-D640-LinLog-6	10	50	L600a in HD	2.13	1.86

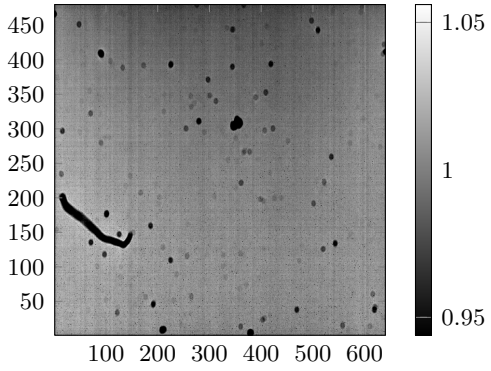
Table 4.6.: Overview of the recorded sequences with the Photonfocus MV-D640-66-CL-LinLog camera

Fig. 4.11 and 4.12 show the nonuniformity parameter sets $\{a_{i,j}\}$ and $\{b_{i,j}\}$ for the most differing setups for $(10\,\mathrm{ms}\,@\,50\,^{\circ}\mathrm{C})$ and $(0.082\,\mathrm{ms}\,@\,25\,^{\circ}\mathrm{C})$. The visual representations of the other evaluations can be found in appendix B.4.

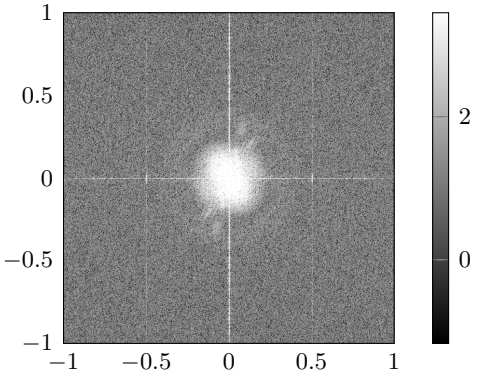
The visual impression of the PRNU shows again non removable dust particles and an instability if the minimum exposure time is used. Dominant row and column wise frequencies are visible as well as some spatial low frequency components. With concern to the histograms, the distribution converges towards a Gaussian distribution with increasing temperatures. For the DSNU the row and column wise components show up clearly as discussed for the other camera types and the histograms fit the Gaussian distribution very well.



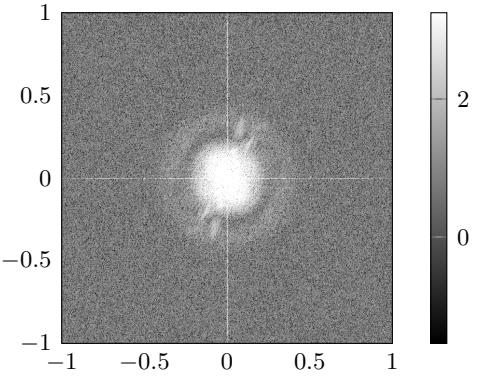
(a) Spatial PRNU for (0.082 ms @ 25 °C)



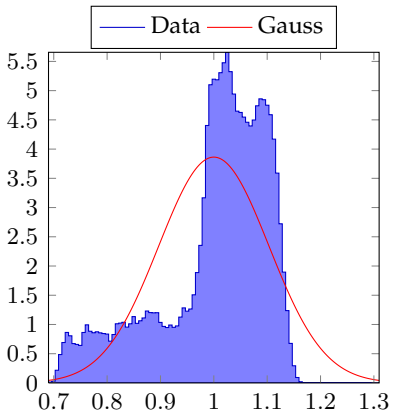
(b) Spatial PRNU for (10 ms @ 50 °C)



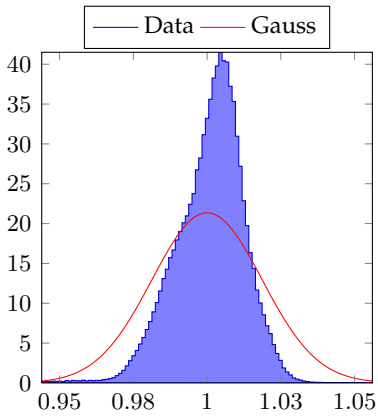
(c) Spectral: $\log|\hat{a}_{\mu,\nu}|$ for (0.082 ms @ 25 °C)



(d) Spectral: $\log|\hat{a}_{\mu,\nu}|$ for (10 ms @ 50 °C)

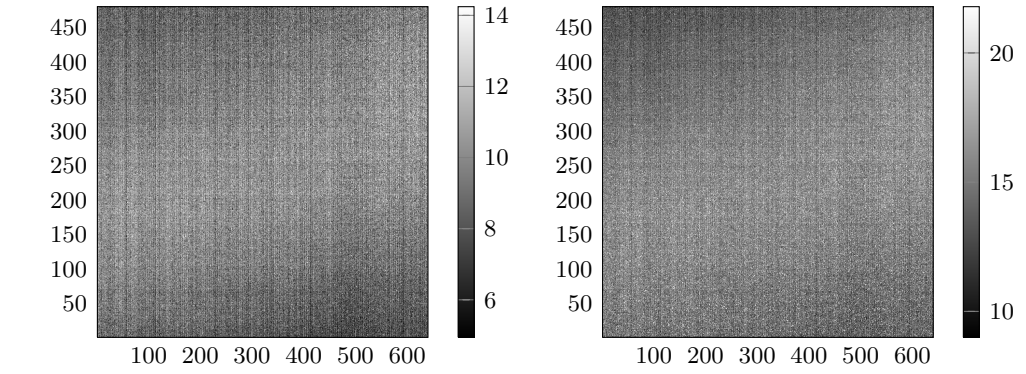


(e) Histogram of $\{a_{i,j}\}$ for (0.082 ms @ 25 °C)

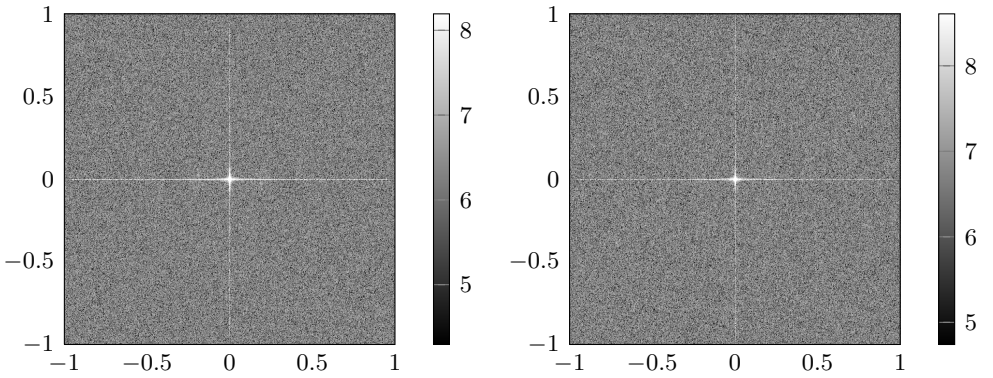


(f) Histogram of $\{a_{i,j}\}$ for (10 ms @ 50 °C)

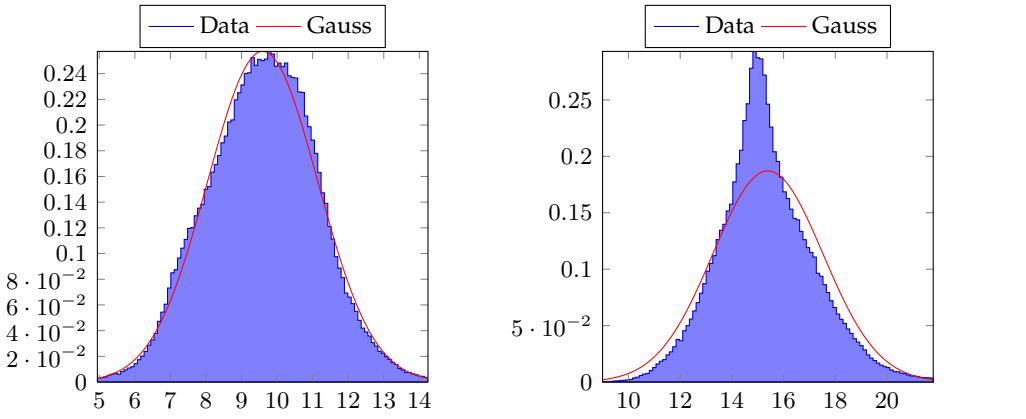
Figure 4.11.: Representations of the $\{a_{i,j}\}$ correction patterns of the EMVA1288 calibration measurements with the Photonfocus MV-D640-66-CL-LinLog Camera. Showing the spatial and the spectral analysis as well as histograms at the given temperatures and exposure times.



(a) Spatial DSNU for (0.082 ms @ 25 °C) (b) Spatial DSNU for (10 ms @ 50 °C)



(c) Spectral: $\log|\hat{b}_{\mu,\nu}|$ for (0.082 ms @ 25 °C) (d) Spectral: $\log|\hat{b}_{\mu,\nu}|$ for (10 ms @ 50 °C)



(e) Histogram of $\{b_{i,j}\}$ for (0.082 ms @ 25 °C) (f) Histogram of $\{b_{i,j}\}$ for (10 ms @ 50 °C)

Figure 4.12.: Representations of the $\{b_{i,j}\}$ correction patterns of the EMVA1288 calibration measurements with the Photonfocus MV-D640-66-CL-LinLog Camera. Showing the spatial and the spectral analysis as well as histograms at the given temperatures and exposure times.

5. Related Work and State of the Art Methods

In the area of nonuniformity corrections a huge variety of different papers is published. However only a small subset of these publications target the scene based nonuniformity correction methods. While some of the literature is specialized on general sensor response models, some publications concentrate on the Fourier analysis of the occurring nonuniformities [22, 28, 2]. The sensor model plays an important role in the nonuniformity correction process but all the publications that do not include the actual captured images cannot be described as *scene based* nonuniformity correction methods. Many of the online correction methods use dark-pixel evaluations or CMOS integrated circuity to correct the light sensors directly in hardware.

The first scene based methods were published in the early 1990s by Scribner [3] and describe very basic techniques for nonuniformity corrections, based on the knowledge of the vertebra retina back at the time. The first technique is a temporal high pass filter, which assumes that the temporal low frequencies in the images are the less important information and thus can be considered as the systematic errors of the system. This method has been developed further into the constant statistics algorithm by Harris [7] and a gating threshold was later added by Hardie [5] to improve the performance even further.

The second method by Scribner utilizes a least mean square algorithm (LMS) to minimize the error image between the corrected image and an assumed original image. The main structure of the LMS algorithm has been nearly unchanged until today, and further developments exploit the retinomorphic hypothesis of IR-cameras by the use of an *adaptive* least mean square algorithm [33, 4]. The most recent improvements to this method are a signal gating technique [5] and a generalization called *total variation approach* [6].

Other research directions include Kalman filters [8, 9] to adapt to the drifts of the nonuniformity parameters. Their basic algorithmic principles for the nonuniformity estimation are however covered by the constant statistics and least mean square algorithms. Thus the Kalman filter approaches mainly possess an advanced model for the temporal and temperature parameter drifts. However, the newly developed methods in this thesis consider the parameter drifts only implicitly by the effects from a changing input over time (see chapter 6). This is a valid consideration for the common slow parameter drifts in most applications. The

Kalman filter methods therefore form a method for specialized applications and are excluded from further comparisons.

All the presented methods correct successfully for nonuniformities but their performances can still be improved. Further, as mentioned in the introduction, many methods miss a motivation of their assumptions (e.g. the way of estimating the original image in [5, 33, 4, 6]) and the available statistical information of the camera's nonuniformities (e.g. mean and standard deviations) has not been considered, which leads to corrected images on different gray value scales.

The improved constant statistics method and the Least Mean Square Algorithm methods are highly developed scene based methods that allow a good a comparison against the newly developed methods. Their very recent developments [6] give the best results found in literature at the time of this text. Further, those methods are the most referenced methods in literature to compare against and form the *gold standard* for scene based NUC methods.

5.1. The Constant Statistics Method

The constant statistics method of Harris [7] is the latest development of this type of method, that was originally introduced by Scribner [3]. While in Scribner's publications the temporal low pass signal was considered as the nonuniformity signal, the idea was theoretically expanded into a constant statistics assumption. This constant statistics assumption states that each pixel's mean and standard deviation is constant over time and space. Mathematically expressed, the basic constant statistics algorithm (CS) states, that if enough information is gathered, the difference in the mean of each pixel corresponds to the DSNU and the difference in the standard deviations corresponds to the PRNU signal. The statistical parameters of the PRNU ($\{a_{i,j}\}$) and DSNU ($\{b_{i,j}\}$) are already assumed in this method by $e(\{a_{i,j}\}) = 1$ and $e(\{b_{i,j}\}) = 0$.

An estimate in the paper demands the order of $10'000$ frames for the algorithm to converge. For video based driver assistance systems and many other applications less input frames (≈ 500) would be a more desirable target as explained below. Harris also states in his paper that results from about 720 frames have been sufficient to visually correct the nonuniformities.

The Basic Constant Statistic Algorithm: The very basic constant statistics algorithm is then expressed mathematically as

$$\begin{aligned} b_{\mathrm{CS},i,j} &= \mu(C_{\mathrm{M},i,j}) \\ a_{\mathrm{CS},i,j} &= \sigma(C_{\mathrm{M},i,j})\,. \end{aligned} \tag{5.1}$$

Harris et al. state that the mean deviation performs better than the standard deviation and thus the estimation leads immediately to the following improvement:

$$a_{\mathrm{CS}^*,i,j} = \frac{1}{-1+\sum_n 1} \sum_n |C_{\mathrm{M},i,j,t_n} - \mu(C_{\mathrm{M},i,j,t_n})| \tag{5.2}$$

To calculate the temporal mean and temporal mean deviation, the following recursive formulas are given in the paper (where n denotes the number of the input image):

$$b_{\mathrm{CS},i,j}(n) = \frac{1}{n}\left(C_{\mathrm{M},i,j}(n) + (n-1)b_{\mathrm{CS},i,j}(n-1)\right) \tag{5.3}$$

$$a_{\mathrm{CS},i,j}(n) = \frac{1}{n}\left(|C_{\mathrm{M},i,j}(n) - b_{\mathrm{CS},i,j}(n)| + (n-1)a_{\mathrm{CS},i,j}(n-1)\right) \tag{5.4}$$

The $b_{\mathrm{CS},i,j}(0)$ and $a_{\mathrm{CS},i,j}(0)$ provide initial values that could be the values $b_{\mathrm{CS},i,j}(0) = 0$ and $a_{\mathrm{CS},i,j}(0) = 1 \quad \forall i,j$ or any other useful estimate. The calculations of eq. 5.3 and 5.4 result in the addition of values that differ by the factor $(n-1)$, which may reach several magnitudes. Fewer input frames are therefore preferred to safe computational performance and to avoid problems in the accuracy.

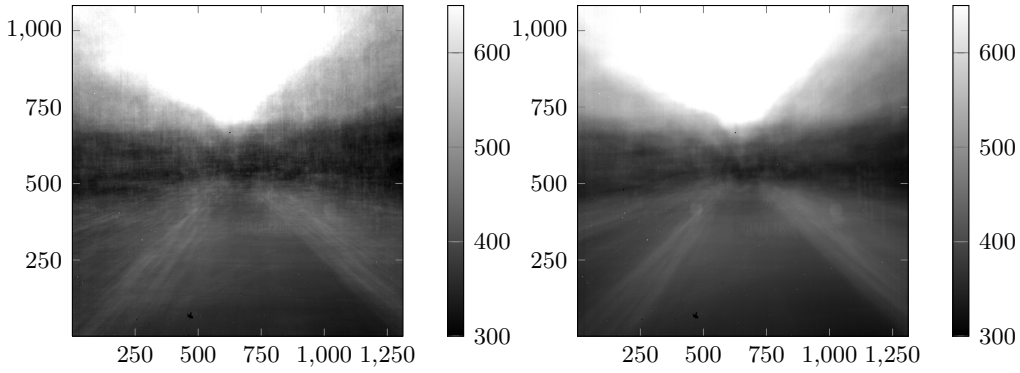
Harris et al. mention further that to adapt for slow temporal or temperature drifts in the parameters a continuous operation with $100'000$s of averaged frames should be considered. The mean and standard deviation should then be calculated with help of an exponential window to give lower weights to old frames, which implicitly allows an adaptation of the estimate to the parameter drifts.

The Gated Constant Statistic Algorithm: Harris introduced in [7] an improved CS algorithm that uses a gating threshold. His idea is also used by Hardie [5] and named the gated CS method (CS_G). The idea is triggered by the fact that ghosting artifacts occur in parts of the scene without motion. However, if motion is present, it is usually coupled to a change of the measured signal in the given area. With this assumption, a gate threshold is introduced that blocks updates from regions with no scene motion by allowing an update only if the gray value of the incoming frame provides a high enough difference compared to the previous frame. Given the threshold value T, the update rules for the gated constant statistics method are provided as

$$b_{\mathrm{CS}_{g.},i,j}(n) = \begin{cases} b_{\mathrm{CS},i,j}(n) \quad \text{(see eq.5.3)} & \text{if} \quad |C_{\mathrm{M},i,j}(n) - C_{\mathrm{M},i,j}(n-1)| < T \\ b_{\mathrm{CS}_{g.},i,j}(n-1) & \text{otherwise} \end{cases} \tag{5.5}$$

$$a_{\mathrm{CS}_{g.},i,j}(n) = \begin{cases} a_{\mathrm{CS},i,j}(n) \quad \text{(see eq.5.4)} & \text{if} \quad |C_{\mathrm{M},i,j}(n) - C_{\mathrm{M},i,j}(n-1)| < T \\ a_{\mathrm{CS}_{g.},i,j}(n-1) & \text{otherwise} \end{cases} . \tag{5.6}$$

Harris et al. suggest to use a threshold between $10\,\%$ to $50\,\%$ of the full dynamic range of the sensor, given in DN.



(a) Average of 500 random frames (b) Average of 10′000 random frames

Figure 5.1.: The constant statistics assumptions is not fulfilled for the video based driver assistance systems. If it would be, the differences in the averaged frames should converge to 0.

5.1.1. Discussion of the Constant Statistics Algorithm

There are two main disadvantages of the constant statistics algorithm. First, the constant statistic assumption on which the algorithm relies is not fulfilled in many scenarios. For example in video based driver assistance systems, the area that is covered by the road will have at least a different mean value with respect to the brighter sky. Fig. 5.1 shows an example of 10′000 averaged frames, with ground truth corrected nonuniformities to illustrate this problem.

Second, the algorithm does not take into account the neighboring pixels and corrects only on a per-light-sensor basis. However, in the measurement process the convolutions of each sample point with the light sensors' PSF and the optic's PSF lead to a dependency between neighboring pixels that should be exploited (see eq. 2.16).

5.2. Least Mean Square Algorithms

While the *CS* algorithms provide a correction using only the simple constant statistics assumption, the second method for nonuniformity correction of Scribner's first paper (the LMS method, [3]) exploits more assumptions. The algorithm uses an estimate, or *best guess* for the ideal sampled signal C_S: $C_{\mathrm{S},i,j,\mathrm{guess}}$ and minimizes the difference between this estimate and the current correction in an iterative frame-to-frame algorithm. In Scribner's paper the estimate is calculated

as the spatial averaged output of the light sensors in the neighborhood $N(i,j)$ around the current sensor i,j:

$$C_{\mathrm{S},i,j,\mathrm{guess}} = \frac{1}{\sum_{k,l} 1} \sum_{k,l \in N(i,j)} C_{\mathrm{M},k,l} \tag{5.7}$$

This estimate is motivated by an "always moving scene with respect to the light detectors" [3]. In Vera's later development [4], this motivation is renewed by the retinomorphic hypothesis, which states that "[...] there is a higher probability for a given detector and its surrounding neighbors, of being illuminated by the same infrared irradiance level". Next, Hardie [5] states that just a spatial low pass filter is necessary to produce a suitable "desired" image, without any further explanation.

Eq. 5.7 can in fact be written as a low pass filter, as it represents the discrete convolution of the signal $C_{\mathrm{M},k,l}$ with a BOX filter. The authors use different BOX filter sizes up to 21×21, but for this thesis only the computationally affordable sizes 3×3 and 7×7 are used. The definition of the a 3×3 neighborhood BOX filter writes as:

$$BOX(3) = \frac{1}{9}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} .$$

For a simplified notation, we consider the expressions without the indicies i,j as the matrices that contain all the elements indiced with i,j:

$$A = \begin{pmatrix} A_{1,1} & A_{1,2} & \dots & A_{1,N} \\ \vdots & \vdots & \ddots & \vdots \\ A_{M,1} & A_{M,2} & \dots & A_{M,N} \end{pmatrix} \tag{5.8}$$

The estimate of the original signal can then be expresses with a discrete convolution ($*$) as:

$$C_{\mathrm{S,guess}} = BOX(n) * C_M \quad n \in 3,7 \tag{5.9}$$

This new matrix like notation allows a quicker idea on how to parallelize and compare the later algorithms.

Considering the methods assumption, $C_{\mathrm{S,guess}}$ can only be close to the real C_S if the recorded image is blurred to the extend of the size of the BOX filter, in addition to a mean free nonuniformity degradation for the DSNU, and a degradation by a PRNU with an expectation value of 1. Although these demands on the degradations are usually not met, the LMS methods will provide a very good correction rate under certain circumstances. For the further development of the method to estimate the $\{a_{i,j}\}$ and $\{b_{i,j}\}$ parameters, $C_{\mathrm{S,guess}}$ is considered identical with the estimated corrected image $C_{\mathrm{S,est}}$. To avoid the nonlinear dependency of $\frac{1}{a_{i,j}}$ in the

correction process, the authors use a linearisation into a weight and offset parameter. This linearisation is common to all LMS methods:

$$C_{\mathrm{S},i,j,\mathrm{est.}} = \frac{1}{a_{i,j,\mathrm{est.}}}(C_{\mathrm{M},i,j} - b_{i,j,\mathrm{est.}}) \tag{5.10}$$

$$= w_{i,j}C_{\mathrm{M},i,j} + o_{i,j} \tag{5.11}$$

$$\text{with:} \quad w_{i,j} = \frac{1}{a_{i,j,\mathrm{est.}}} \tag{5.12}$$

$$o_{i,j} = -\frac{b_{i,j,\mathrm{est.}}}{a_{i,j,\mathrm{est.}}} \tag{5.13}$$

As next step, the methods define an error measure E that compares the guessed original image to the estimated image. To be able to express the method smoothly in matrix notation, the *Hadamard* (element-wise) multiplication ($\otimes$) and division ($\oslash$) of matrices is used:

$$\text{assume:} \quad C_{\mathrm{S},i,j,\mathrm{est.}} \approx C_{\mathrm{S,guess}} \tag{5.14}$$

$$E = (C_{\mathrm{est}} - C_{\mathrm{S,guess}}) \otimes (C_{\mathrm{est}} - C_{\mathrm{S,guess}})$$

$$= (C_{\mathrm{est}} - BOX * C_M) \otimes (C_{\mathrm{est}} - BOX * C_M) \tag{5.15}$$

$$E_{i,j} = ((C_{\mathrm{est}} - BOX * C_M)_{i,j})^2 \tag{5.16}$$

This error measure E is given as a quadratic form and is minimized with respect to the variables w and o by using the least mean square algorithm [43]. To accomplish this, the gradient in the variables w and o is needed which is defined by the ordered vector of the partial derivatives:

$$\nabla_x F(x_{1,1}, \ldots, x_{M,N}) = \left(\partial_{x_{1,1}} F, \ldots, \partial_{x_{M,N}} F\right)^T \tag{5.17}$$

This vector notation is difficult to handle if the algorithm is to be expressed in the desired matrix notation with dominant element wise operations. Therefore the $\hat{\nabla}$ operator is introduced, which results in a matrix with the (i,j)th partial derivative as the (i,j)th entry of the matrix:

$$\hat{\nabla}_x F = \begin{pmatrix} \partial_{1,1}F & \partial_{1,2}F & \ldots & \partial_{1,N}F \\ \vdots & \vdots & \ddots & \vdots \\ \partial_{M,1}F & \partial_{M,2}F & \ldots & \partial_{M,N}F \end{pmatrix} \tag{5.18}$$

The gradients of the error function E can then be rewritten as:

$$\hat{\nabla}_w E = 2C_M \otimes (C_{\mathrm{est}} - BOX * C_M) \tag{5.19}$$

$$\hat{\nabla}_o E = 2(C_{\mathrm{est}} - BOX * C_M) \tag{5.20}$$

It is to note that the derivation is only partially executed with concern on the real variables a and b as the dependencies $w(a)$ and $o(a,b)$ are neglected. The gradient's direction points consequently not into the direction of the physical gain and offset

parameters a and b, but approximates this direction if the values of $a_{i,j}$ are close to the ideal value $a_{i,j} = 1$.

Next the *Least Mean Square Algorithm* [43] is used to minimize the error measure E in form of a frame-to-frame update of the parameters $o(n)$ and $w(n)$ The parameter α is introduced as a global and fixed step size and n denotes the frame number:

$$\begin{aligned} o(n+1) &= o(n) - \alpha \cdot \hat{\nabla}_w E \\ &= o(n) - 2(C_{\text{est}} - BOX * C_M) \\ w(n+1) &= w(n) - \alpha \cdot \hat{\nabla}_o E \\ &= w(n) - 2C_M \otimes (C_{\text{est}} - BOX * C_M) \end{aligned} \tag{5.21}$$

The correction factors a and b can be finally calculated using the inverse of eq. 5.12 and 5.13:

$$a_{\text{LMS},i,j} = \frac{1}{w_{i,j}} \tag{5.22}$$

$$b_{\text{LMS},i,j} = -\frac{o_{i,j}}{a_{\text{LMS},i,j}}\,. \tag{5.23}$$

The final parameters of this method are the size of the neighborhood and the step size α. With the already fixed neighborhood sizes of 3×3 and 7×7 two variants of the algorithm are set, which have only the step size parameter α left to tune: The $LMS_{3\times3}$ and $LMS_{7\times7}$ methods. The LMS algorithm gives a good correction performance, especially for DSNU corrections and thus it has been subject to many improvements over the last two decades. The most successful variants are now discussed.

5.2.1. High Pass Variant of the LMS Algorithm

A variant of the algorithm appears if the low pass filter is applied onto the *current estimate* $C_{\text{est}}(n)$ instead of the current measurement:

$$C_{\text{S,guess}} = BOX * C_{\text{est}} \tag{5.24}$$

$$E = (C_{\text{est}} - BOX * C_{\text{est}}) \otimes (C_{\text{est}} - BOX * C_{\text{est}}) \tag{5.25}$$

$$= (HP * C_M) \otimes (HP * C_M) \tag{5.26}$$

E is then calculated with help of the current estimate, subtracted by its own low pass version. This rewrites as the convolution with a High Pass Filter (HP) which

is defined in this case as $HP = \delta - BOX$, with δ the discrete version of the Dirac pulse. For a 3×3 neighborhood this results in:

$$HP(3) = \frac{1}{9}\begin{pmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{pmatrix} \tag{5.27}$$

When this definition of the error image is compared to the above presented in eq. 5.15, then the numerical differences vanish, especially in cases of moderate to low magnitudes of the nonuniformity parameters. The immediate consequence of this approach is the application of a single high pass filter HP instead of the subtraction of the BOX filtered version of the image.

The update rules for o and w of this variant present as

$$\begin{aligned} o(n+1) &= o(n) - \alpha \cdot \hat{\nabla}_w E && (5.28)\\ &= o(n) - 2\alpha(C_{\text{est}} * HP) && (5.29)\\ w(n+1) &= w(n) - \alpha \cdot \hat{\nabla}_o E && (5.30)\\ &= w(n) - 2\alpha C_{\text{est}} \otimes (C_{\text{est}} * HP)\,, && (5.31) \end{aligned}$$

and are computationally less expensive to implement. However this method is not mentioned in the literature and serves therefore as a theoretical intermediate step, which is included in the below improvements of the method.

5.2.2. The Fast Adaptive LMS Algorithm

An important improvement to Scribners basic LMS algorithm is the *fast adaptive* version of the LMS algorithm, introduced by Vera [4]. Vera improved two parts of the algorithm: The estimation of the guessed original image $C_{\text{S,guess}}$ and a spatial adaption of the step size α.

Given the fact, that after a few iterations, the parameters w and o have a better estimate than before, the guess of the original image should perform better if the low pass filter is applied onto the *current estimate* $C_{\text{est}}(n)$ instead of the current measurement C_M, which was already introduced in the high pass variant in eq. 5.28 .

The second improvement concerns the step size parameter α where Vera adapts the step size dynamically for each light sensor. He uses the observation that edges of high intensity cause ghosting artifacts in combination with the retinomorphic hypothesis which states that "[...] there is a higher probability for a given detector and its surrounding neighbors, of being illuminated by the same infrared irradiance level" [4]. As a consequence, the retinomorphic hypothesis is used to detect edges and if the retinomorphic hypothesis is fulfilled, larger update steps are used for the light sensor, while the update step size is reduced for areas where the hypothesis is not fulfilled. As measure for the retinomorphic hypothesis, Vera first

proposed in 2003 the variance of the measurement C_M within a small neighborhood [33]. Then in 2005 he presented the standard deviation as a better measure [4].

The following definition of the step size matrix is an equivalent formulation of the one used by Vera. It provides a larger step sizes for small standard deviations and vice versa:

$$\alpha_{i,j} = \frac{\alpha_{\max}}{1 + s(C_{\mathrm{M},N(i,j)})} \tag{5.32}$$

$$\mathcal{S} = \left(s(C_{\mathrm{M},N(i,j)})\right)_{i,j} \tag{5.33}$$

$$\alpha = \alpha_{\max} \oslash (1 + \mathcal{S}) \tag{5.34}$$

In the equation $s(C_{\mathrm{M},N(i,j)})$ represents the spatial standard deviation of C_M in the given neighborhood N around the pixel position (i, j). $\mathcal{S}$ is the corresponding matrix.

The resulting parameters for this method are the maximum allowed step size $\alpha_{\max}$ and the size of the neighborhood, leading to the two adaptive algorithms: $LMS_{A,3\times3}$ and $LMS_{A,7\times7}$. The update rules can be adapted from the original LMS algorithm by exchanging the step size. They result in the expressions:

$$o(n+1) = o(n) + 2\alpha_{\max} \oslash (1 + \mathcal{S}(n)) \otimes (C_{\mathrm{est}}(n) * HP) \tag{5.35}$$

$$w(n+1) = w(n) + 2\alpha_{\max} C_{\mathrm{est}}(n) \oslash (1 + \mathcal{S}(n)) \otimes (C_{\mathrm{est}}(n) * HP)\,. \tag{5.36}$$

5.2.3. The Gated Adaptive LMS Algorithm

Hardie [5] provided one further update for the adaptive LMS algorithm to avoid ghosting artifacts in special situations. While Vera's adaptive LMS algorithm is able to prevent edges of high intensity to burn in, a lack of motion in the input frames still results in ghosting artifacts. Therefore Hardie proposes to stop the learning in regions without any detected motion. This is achieved by an adaptive step size that is set to 0 if a so-called *change threshold* is not reached. In contrast to the constant statistic method's change threshold, Hardie uses the measured differences of the guessed original image to stop the learning. Mathematically the update step expresses as:

$$\alpha_{\max,i,j} = \begin{cases} \frac{\alpha_{\max}}{1+s(C_{\mathrm{M},N(i,j)})} & |C_{\mathrm{S,guess}} - Z|_{i,j} > T \\ 0 & \text{else} \end{cases} \tag{5.37}$$

$$Z_{i,j}(n+1) = \begin{cases} C_{\mathrm{S},i,j,\mathrm{guess}}(n) & |C_{\mathrm{S,guess}} - Z|_{i,j} > T \\ Z_{i,j}(n) & \text{else} \end{cases} \tag{5.38}$$

Hardie showed already in his paper [5], that his method provides an advantage only in regions without motion. Otherwise the performance was not significantly

improved compared to the $LMS_{A,3\times3}$ and $LMS_{A,7\times7}$ methods. Besides the change in the step size parameters, the update rules are the same as for the adaptive least mean square algorithms. The two new *gated adaptive* LMS methods are named: $LMS_{GA,3\times3}$ and $LMS_{GA,7\times7}$.

5.2.4. Total variation LMS algorithm

A very recent improvement of the LMS algorithm is introduced by Vera [6]. He changes the error estimate E which is minimized by the Least Mean Square Algorithm as an approximation of the absolute value of the gradient of the estimated image:

$$E \approx |\nabla C_{\text{est}}| \tag{5.39}$$

To approximate ∇C_{est} two alternative definitions called ∇ for forward differences and ∇^* for backward differences are given by Vera as:

$$\nabla C_{i,j,\text{est}} = \sqrt{(C_{i,j,\text{est}} - C_{i-1,j,\text{est}})^2 + (C_{i,j,\text{est}} - C_{i,j-1,\text{est}})^2} \tag{5.40}$$

$$\nabla^* C_{i,j,\text{est}} = \sqrt{(C_{i,j,\text{est}} - C_{i+1,j,\text{est}})^2 + (C_{i,j,\text{est}} - C_{i,j+1,\text{est}})^2} \tag{5.41}$$

His approach can also be transformed into the matrix notations introduced above and more information can be found in his publication [6]. The equivalent formulation leads to two filter masks according to the used definitions of either $\nabla C_{i,j,\text{est}}$ or $\nabla^* C_{i,j,\text{est}}$. The two sets of update rules can be expressed mathematically identical to the original paper with help of the new matrix notation:

$$HP_{\text{TV}} = \frac{1}{3}\begin{pmatrix} 0 & -1 & 0 \\ -1 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$HP_{\text{TV}^*} = \frac{1}{3}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & -1 \\ 0 & -1 & 0 \end{pmatrix} \tag{5.42}$$

$$o_{i,j}(n+1) = o_{i,j}(n) - \frac{\alpha_{\text{TV}}}{\nabla C_{\text{S},i,j,\text{est.}}} \cdot (HP_{\text{TV}} * C_{\text{est}}) \tag{5.43}$$

$$w_{i,j}(n+1) = w_{i,j}(n) - \frac{\alpha_{\text{TV}}}{\nabla C_{\text{S},i,j,\text{est.}}} \cdot C_{\text{M},i,j} \cdot (HP_{\text{TV}} * C_{\text{est}}) \tag{5.44}$$

$$o_{i,j}(n+1) = o_{i,j}(n) - \frac{\alpha_{\text{TV}}}{\nabla^* C_{\text{S},i,j,\text{est.}}} \cdot (HP_{\text{TV}^*} * C_{\text{est}}) \tag{5.45}$$

$$w_{i,j}(n+1) = w_{i,j}(n) - \frac{\alpha_{\text{TV}}}{\nabla^* C_{\text{S},i,j,\text{est.}}} \cdot C_{\text{M},i,j} \cdot (HP_{\text{TV}^*} * C_{\text{est}}) \tag{5.46}$$

This method provides the same structure like the high pass versions of the LMS algorithms: A High-Pass-Filtered version of C_{est} is weighted with an (adaptive) step size and used to update the current estimations. This time however, the step size is not determined by the variance of the input signal, but by the occurring

gradients. This is similar but not identical to the measure of the standard deviation of the adaptive LMS methods. In the limiting case of standard deviation of 0, the gradient measure will however be 0 as well.

If implemented as introduced in the paper, this method has a potential for an undefined zero devision by zero if the gradient in a certain region is zero (for example due to occurring local overexposure). To avoid this, the method is adapted by the same safeguards that are used in the adaptive LMS method's step size:

$$o_{i,j}(n+1) = o_{i,j}(n) - \frac{\alpha_{\mathrm{TV}}}{1 + \nabla C_{\mathrm{S},i,j,\mathrm{est.}}} \cdot (HP_{\mathrm{TV}} * C_{\mathrm{est}}) \tag{5.47}$$

$$w_{i,j}(n+1) = w_{i,j}(n) - \frac{\alpha_{\mathrm{TV}}}{1 + \nabla C_{\mathrm{S},i,j,\mathrm{est.}}} \cdot C_{\mathrm{M},i,j} \cdot (HP_{\mathrm{TV}} * C_{\mathrm{est}}) \tag{5.48}$$

$$o_{i,j}(n+1) = o_{i,j}(n) - \frac{\alpha_{\mathrm{TV}}}{1 + \nabla^* C_{\mathrm{S},i,j,\mathrm{est.}}} \cdot (HP_{\mathrm{TV}^*} * C_{\mathrm{est}}) \tag{5.49}$$

$$w_{i,j}(n+1) = w_{i,j}(n) - \frac{\alpha_{\mathrm{TV}}}{1 + \nabla^* C_{\mathrm{S},i,j,\mathrm{est.}}} \cdot C_{\mathrm{M},i,j} \cdot (HP_{\mathrm{TV}^*} * C_{\mathrm{est}}) \tag{5.50}$$

5.2.5. Discussion of the LMS Algorithms

Concerning the LMS algorithms three points should be addressed. First, the algorithms achieve a good correction rate by minimizing an error image in a given norm (sec. 5.3). Unfortunately this norm is not motivated and is maybe not chosen optimally. Extending this point, the guessed original signal is always set as a low-pass filtered version of the measured image, which works well, but again a motivation of which low-pass version is the best choice is missing.

The newly introduced comparable way of writing the methods with help of high-pass-filters shows how all these methods base on the same principles. Further research will lead to a deeper understanding why the low-pass or blurred version of the measured image is a good choice to estimate the original light signal with respect to signal theory. One way to achieve this is to use physical properties of the actual light signal that are affected by the nonuniformities. This will result in defined high-pass filters which are not arbitrary any more, as introduced in chapter 6.

5.3. Experiments and Results

The algorithms introduced above represent the state-of-the art in literature. As the most recent improvements are included, these methods form a good standard to compare new developments against them. The methods have been implemented in C++ using NVidia's CUDA technology [44] to accelerate the computations. At this point the matrix notation proved very helpful and allowed to implement the methods straight forward on a parallel architecture.

To evaluate the methods' performance, a parameter search is executed. This is necessary as the given parameters from the publications lead to significantly worse results when compared to an optimized parameter set. In a second step, the algorithms' performances were evaluated and discussed with respect to the introduced quality norms of sec. 3.4. The spatial and spectral quantities of the remaining nonuniformities are also investigated as introduced in chapter 3.

5.3.1. Parameter Search

Besides the basic constant statistic method (CS), all algorithms have at least one parameter to be tuned. For the improved version of the CS method, the gated constant statistics method (CS_G) the gate threshold needs to be adapted. Considering the LMS algorithms, the parameter is the step size α and for the fast adaptive LMS algorithm the parameter $\alpha_{\max}$ needs to be optimized. The gated adaptive LMS methods depend on a threshold T which needs to be known in addition to the step size parameter $\alpha_{\max}$. As the gated adaptive LMS methods complement the adaptive LMS methods, the threshold is regarded the only parameter to tune and the $\alpha_{\max}$ parameter of the corresponding adaptive LMS methods are used. For the total variation version only the step size parameter α_{TV} needs to be adjusted.

Concerning the change threshold T, the search is performed as a *binary search* over the whole dynamic range of the corresponding camera. For the step size values the interval $[0, \infty]$ is valid and a modified binary search with exponentially divided search ranges is used. This exponentially stepping enables the search to find the correct order of the parameter very fast and the starting interval is limited to the reasonable parameters $[10^{-8}, 10^{3}]$. This type of parameter search assumes of course that only one minimum exists inside the search area, which is an assumption that cannot be proven mathematically. Experimental results however showed that the found parameter sets seem to be optimal.

The parameter search is conducted with all the 25 consecutive frame and random frame *calibration* sequences introduced in sec. 4.2. Their performance was evaluated every 50 frames and the results of the last $25\,\%$ of a given sequence are averaged and considered the performance of the method. The averaging is necessary to balance eventual performance fluctuations. It is obvious that each sequence has an individual best parameter and an individual best correction rate. To illustrate this, fig. 5.2 shows the different outcomes of the parameter search for the $LMS_{A,7\times7}$-method with respect to the newly introduced percentual correction measure $p_{\mathrm{c,Comb.}}$.

The averages of the found best parameters are considered as the optimal parameters while the standard deviations of the gathered values give an estimate of the respective errors. Tab. 5.1 summarizes the used abbreviations for the methods and tab. 5.2 and 5.3 show the resulting parameter sets for the random frame and the consecutive frame sequences. Tab. 5.4 shows for comparison the parameters that have been proposed in the publications by the methods' authors.

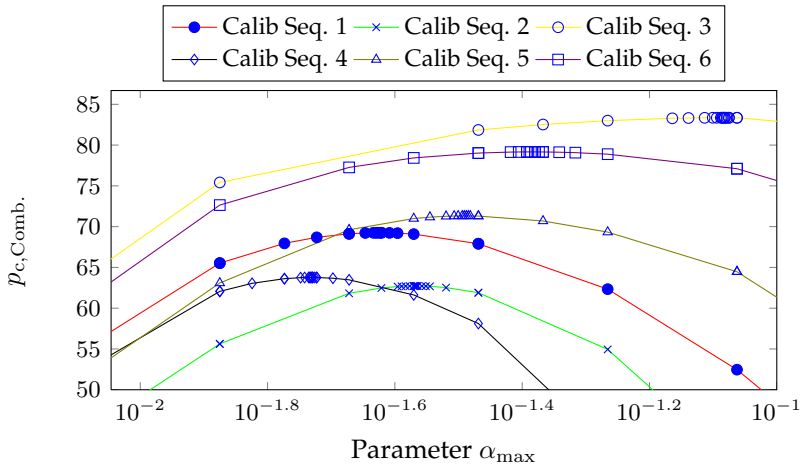


Figure 5.2.: Parameter search outcomes for the $LMS_{A,7\times7}$ box filter.

Method	Abbreviation
Constant Statistics (see sec. 5.1)	CS
Gated Constant Statistics (see sec. 5.1)	CS_G
Least Mean Square Method ... with 3×3 mask (see sec. 5.2)	$LMS_{3\times3}$
... with 7×7 mask (see sec. 5.2)	$LMS_{7\times7}$
Adaptive Least Mean Square Method ... with 3×3 mask (see sec. 5.2)	$LMS_{A,3\times3}$
... with 7×7 mask (see sec. 5.2.2)	$LMS_{A,7\times7}$
Gated Adaptive Least Mean Square Method ... with 3×3 mask (see sec. 5.2)	$LMS_{GA,3\times3}$
... with 7×7 mask (see sec. 5.2.3)	$LMS_{GA,7\times7}$
Total variation LMS Algorithm (see sec. 5.2.4)	LMS_{TV}

Table 5.1.: Implemented methods and their abbreviations

Method	Parameter
CS_G	$613.1 \pm (91.02, 14.85\%)$
$LMS_{3\times3}$	$0.01996 \pm (0.01144, 57.31\%)$
$LMS_{7\times7}$	$0.007149 \pm (0.0002339, 3.272\%)$
$LMS_{A,3\times3}$	$1.758 \pm (0.08248, 4.691\%)$
$LMS_{A,7\times7}$	$1.526 \pm (0.2534, 16.61\%)$
$LMS_{GA,3\times3}$	$0.0194 \pm (0.03609, 186\%)$
$LMS_{GA,7\times7}$	$0.04898 \pm (0.0823, 168\%)$
LMS_{TV}	$2.841 \pm (0.4092, 14.4\%)$

Table 5.2.: Optimized parameter set for random frame sequences.

Method	Parameter
CS_G	$23.95 \pm (16.2, 67.63\%)$
$LMS_{3\times3}$	$0.008654 \pm (0.007004, 80.93\%)$
$LMS_{7\times7}$	$0.005877 \pm (0.005487, 93.37\%)$
$LMS_{A,3\times3}$	$1.069 \pm (0.4134, 38.67\%)$
$LMS_{A,7\times7}$	$0.6192 \pm (0.1334, 21.55\%)$
$LMS_{GA,3\times3}$	$0.8241 \pm (3.364, 408.2\%)$
$LMS_{GA,7\times7}$	$0.1544 \pm (0.2963, 191.9\%)$
LMS_{TV}	$2.291 \pm (0.735, 32.09\%)$

Table 5.3.: Optimized parameter set for consecutive frame sequences

Method	Parameter
CS_G	$800 \pm (400, 50\%)$
$LMS_{3\times3}$	$0.055 \pm (0.54, 98\%)$
$LMS_{7\times7}$	$0.055 \pm (0.54, 98\%)$
$LMS_{A,3\times3}$	$0.1 \pm (0.25, 25\%)$
$LMS_{A,7\times7}$	$0.1 \pm (0.25, 25\%)$
$LMS_{GA,3\times3}$	$800 \pm (400, 50\%)$
$LMS_{GA,7\times7}$	$800 \pm (400, 50\%)$
LMS_{TV}	Optimized

Table 5.4.: Parameter set extracted from the publications in literature

A comparison between the parameter sets for random and consecutive frame sequences in case of the CS_G method shows that the optimal thresholds differ by a factor of ≈ 26. This effect can be explained as the consecutive frame sequences are recorded with a high frame rate of $100\,\mathrm{Hz}$ and thus the images are very similar. Consequently a large change in gray values does not occur often and thus a large threshold does not allow to learn enough images in those regions to achieve a good performance. Therefore a smaller threshold gives a better result for consecutive frame sequences, which however also limits the desired improvements. For the random frame sequences a larger threshold is possible in comparison to the consecutive frame sequences as the images naturally change a lot.

For the $LMS_{3\times3}$, $LMS_{7\times7}$ and the LMS_{TV} methods, the parameter values are identical for random and consecutive frame sequences if one considers the estimated errors. However, for the $LMS_{A,3\times3}$ and $LMS_{A,7\times7}$ methods, the optimal values are significantly lower for consecutive frame sequences. This makes sense as a lower value avoids burn-in artifacts [4], however it also results in a lower convergence speed of these methods.

The $LMS_{GA,3\times3}$ and $LMS_{GA,7\times7}$ provide an optimized threshold of $0.01\,\mathrm{DN}$ to $0.8\,\mathrm{DN}$, which is effectively a value that lies inside the nonuniformity noise of $50\,\mathrm{DN}$. As the threshold of 0 has also been considered in this search, a marginal improvement of $\pm 0.01\,\%$ can be confirmed for the gating technique. As mentioned above, this

gating technique mainly has advantages for sequences with parts of no motion, which result in constant illumination values. However, areas of constant illumination exist only in the consecutive frame sequences and even there small changes in illumination are still typical due to noise. Consequently, only larger thresholds would result in advantages but tend to stop the learning too often in the used 500 frame sequences to let the method achieve good performances. Consequently, these effects lead to choices of the threshold value that unfortunately allow only a neglectable performance increase compared to the $LMS_{A,3\times3}$ and $LMS_{A,7\times7}$ methods as shown below.

A comparison of the optimized values towards the values published in literature (tab. 5.4) shows that the values match only for the CS_G method in the case of random frame sequences. For all other methods, even if the error boundaries are considered, the parameters do not match and it is to expect that the literature parameters will perform worse than the optimized parameters. It should be further mentioned, that the literature parameter are optimized for different camera types and supposing, they were further optimized onto the very sequences in the mentioned publications.

5.3.2. Analysis of the Best-Case Performances

The results of the parameter search of sec. 5.3.1 allow to evaluate a best-case performance of the methods. As mentioned above, the average of the best parameters per sequence is considered the optimal set of parameters. The according measures of the nonuniformity quality metrics are logged along with the parameters as well. These metrics' values represent the best possible result for a specific sequence and therefore we consider their average over all the calibration sequences as a measure for the optimal performance that can be achieved. The standard deviation of this *upper border* values is again considered as an estimate of the measurement error.

As for the CS method there is no parameter to optimize, each sequences' optimal performance is given by a single evaluation of the method. The same procedure is done for the evaluations with the parameter set from literature, where all methods are evaluated once with the given parameters. The final best-case performances are obtained by averaging over the individual evaluations of the sequences.

There are four different types of evaluation possible for the best-case performance analysis at this point: The optimized parameter sets applied to the random and consecutive frame sequences and the literature parameter set applied onto the random and the consecutive frame sequences. Tab. 5.5 gives an overview of these combinations and introduces their abbreviations in the plots.

Fig. 5.3 shows the best-case performance analysis of the different methods with respect to the $p_{\mathrm{c,Comb.}}$ quality metric. The CS and CS_G methods achieve only large negative correction percentages of less than $-100\,\%$ and their results are cut in the plots to provide a better comparability of the LMS methods. The results of the basic $LMS_{3\times3}$ method are very similar for the consecutive and random frame

Abbreviation	Evaluation
Opt. Random (tab. 5.2)	Random frame sequences with the optimized parameters for random sequences
Opt. Consec (tab. 5.3)	Consecutive frame sequences with the optimized parameters for consecutive sequences
Lit. Random (tab. 5.4)	Random frame sequences with the literature given parameter set
Lit. Consec (tab. 5.4)	Consecutive frame sequences with the literature given parameter set

Table 5.5.: The four combinations of parameter sets and sequence types that are evaluated for the best-case performance analysis

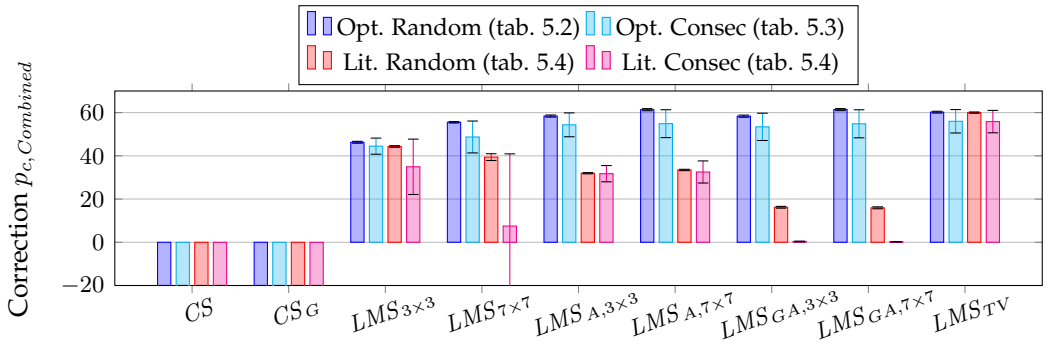


Figure 5.3.: Best-case performance with respect to the $p_{\mathrm{c,Comb.}}$ quality metric

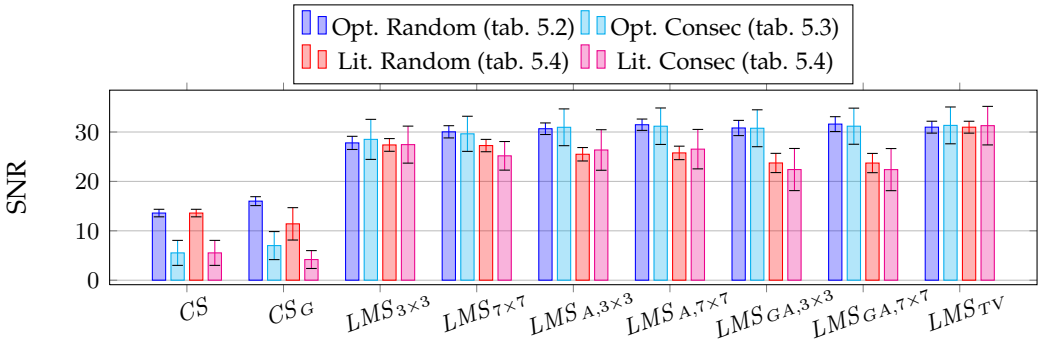


Figure 5.4.: Best-case performance with respect to the SNR quality metric

sequences and reach correction rates that are slightly above $40\,\%$. For these methods the results that are obtained from the parameters given in literature compete with the ones obtained with the optimized parameter set. The $LMS_{7\times7}$ method shows the results for the increased box filter mask, which gives an increased performance to almost $60\,\%$ with the optimized parameter set for the random frame sequences. This performance increase for the larger mask is in contrast to the results stated in [4], where larger masks were reported to produce images that are "noisier than the original $26dB$ level of the test sequence". The result, that larger masks produce the better results seems however more convincing, of course for giant masks of (e.g. 100×100) the improving effects will not show up any more. The performance decrease for the consecutive frame sequences for this method, and for the parameters published in literature an even further decrease in the correction performance is observed. This effect can also be observed for the $LMS_{A,3\times3}$, $LMS_{A,7\times7}$, $LMS_{GA,3\times3}$ and $LMS_{GA,7\times7}$ methods. In general the literature parameter set gives a severely decreased performance and reaches in the average only $\approx 40\,\%$. Therefore this parameter set is not considered in the further evaluations and can merely be considered as a *proof-of-concept* parameter set. The LMS_{TV} method already states that the parameters have to be optimized, which leads to the doubled results as shown in the figure. For all the discussed methods, the consecutive frame sequences' evaluations result in a wider error range, which is the expected behavior as the consecutive frame sequences present a variety of different scenes. This variation in the presented information is diminished by the mix of information of the random frame sequences.

An increase in the correction performance can only be observed until the development of the adaptive LMS method. After that stagnation occurs, which is still unbroken even for the latest development, the LMS_{TV} method. The highest observed correction rate is about $60\,\%$ and in general the performance for random frame sequences is higher than for the consecutive frame sequences. With respect to the error measure, the random frame sequences produce also more stable results which can be seen by the wide error bars for the consecutive frame evaluations. The described results qualitatively occur throughout all the available quality metrics: MAE, RMSE and SNR. The SNR quality metric is shown in fig. 5.4 for example and the MAE and MRSE metrics can be found in appendix C.1. A possible advantage of the SNR metric is its logarithmic scale, however its disadvantage still is its dependency on the image data. Further, the new percentual metrics allow to separate the results with concern to the DSNU and PRNU as it was described in sec. 3.5.

From the publications it becomes obvious that the methods have been developed to correct infrared images, which possess a dominant DSNU. This raises the question, if the correction performance of the algorithms concentrates on the DSNU only and maybe neglects the PRNU, which can not be answered by evaluations in the literature proposed quality metrics. The newly introduced $p_{\mathrm{c,DSNU}}$ and $p_{\mathrm{c,PRNU}}$ metrics however allow this quantification and fig. 5.5 shows the $p_{\mathrm{c,DSNU}}$ correction performance while fig. 5.6 gives the $p_{\mathrm{c,PRNU}}$ correction performance. Surprisingly, the CS and CS_G methods achieve a positive correction rate for the

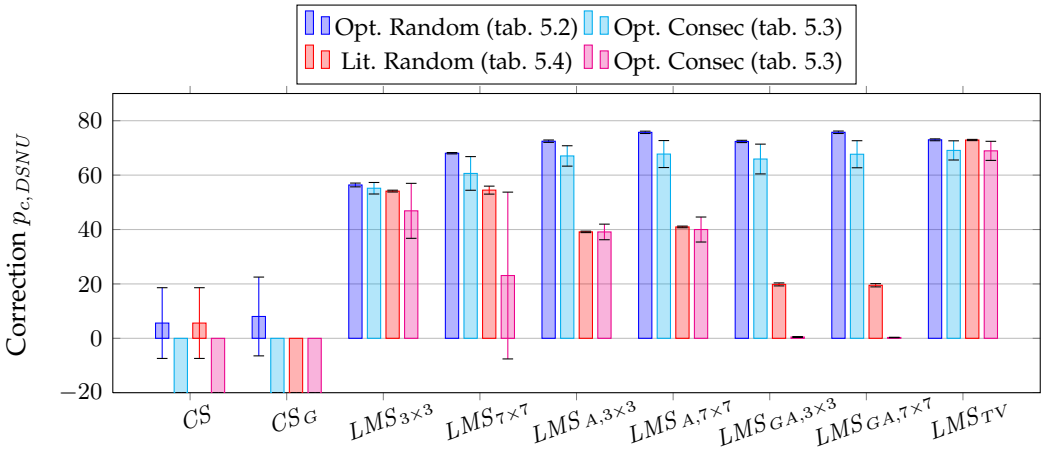


Figure 5.5.: Best-case performance with respect to the $p_{\mathrm{c,DSNU}}$ quality metric

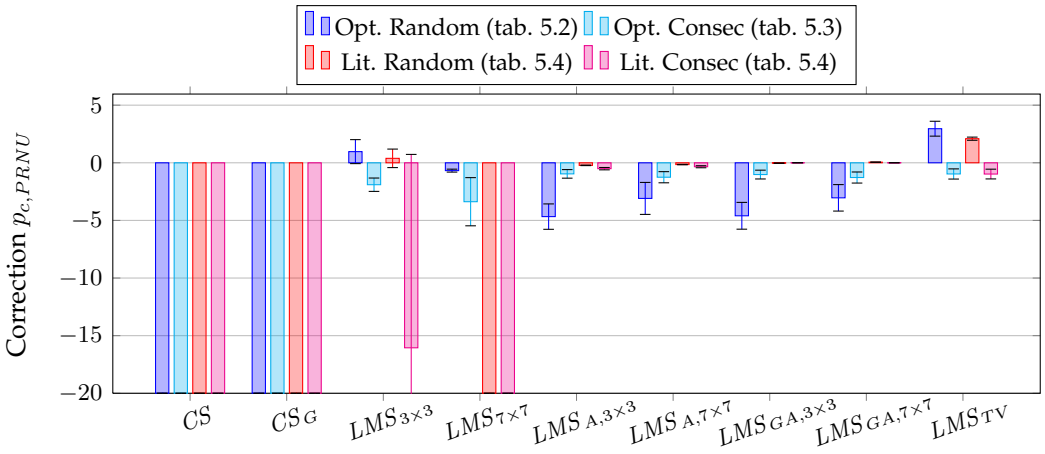


Figure 5.6.: Best-case performance with respect to the $p_{\mathrm{c,PRNU}}$ quality metric

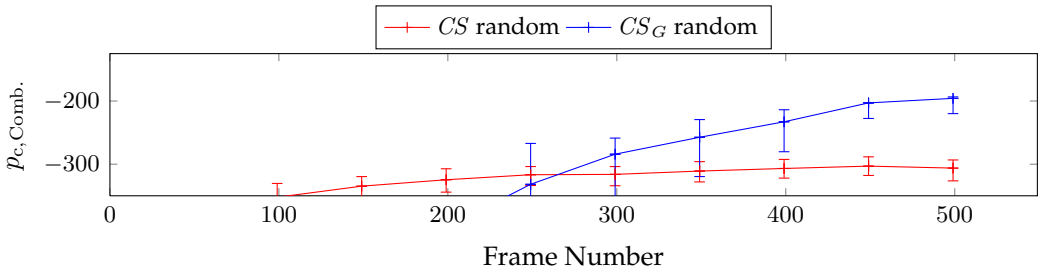
analyzed random frames sequences in the DSNU-only evaluation. The PRNU on the other hand gives a large negative correction rate, which results in the combined negative performance shown in $p_{\mathrm{c,Comb.}}$. The LMS methods show as well a positive DSNU correction rate and for the adaptive LMS methods this correction rate reaches more than $70\,\%$. The PRNU correction rate achieves however only $1\,\%$ to $2\,\%$ and results party in negative PRNU correction values and here the literature proposed parameter set results in less negative $p_{\mathrm{c,PRNU}}$ values. Concluding, the parameter optimization reduces the PRNU correction but increases the DSNU correction when compared to the literature parameter set. The negative correction rates yield in $-1\,\%$ to $-5\,\%$ for the adaptive LMS methods. Only LMS_{TV} is able to achieve a positive PRNU correction rate in combination with a high DSNU correction rate.

Given the performance analysis, LMS_{TV} and $LMS_{A,7\times7}$ are considered the best methods in this comparison. To summarize further, the best correction performance is achieved with random frame sequences as input and consecutive frame input decreases and destabilizes the performance of the methods. Therefore, in the following evaluations the random frame sequences are evaluated first, while an analysis for consecutive frame sequences and the image material with low valued nonuniformities is presented in sec. 10.1 and 10.2.

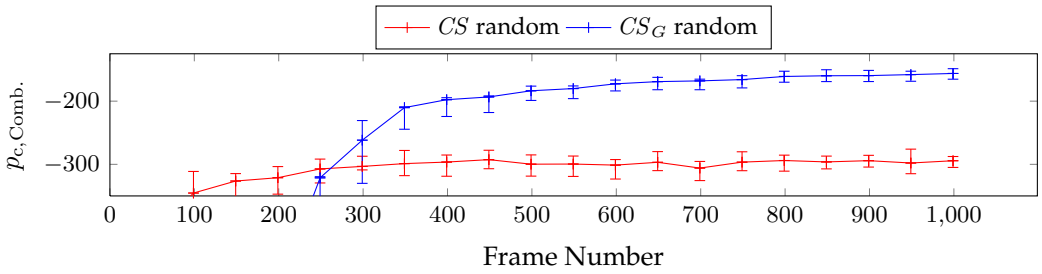
5.3.3. Temporal Performance Analysis for the CS methods

Besides this general overview of the methods in the best-case performance analysis, the temporal performance of the methods gives further insight for the intended use in real applications. For the evaluation of the temporal performance, the 80 random frame evaluation sequences have been used and for specialized evaluations, the 25 sequences with a length of 1000 random frames were used as well. The fixed optimal parameters are used, which resembles the method's usage in real applications. The algorithms' performances are measured in the quality norms $p_{c,Combined}$, $p_{\mathrm{c,DSNU}}$, $p_{\mathrm{c,PRNU}}$, SNR, MAE and $RMSE$. The sampling took place in steps of 50 frames and the $30\,\%$ and $70\,\%$ percentiles are shown as error bars in the following plots.

For the CS and CS_G methods, fig. 5.7a shows this temporal performance analysis for the 500 frame sequences. The methods result in negative percentual correction rates and throughout all quality metrics a huge gap between their performances towards and other methods' performances is present. However, an increase in the method's performance can be seen if the sequence length is increased from 500 to 1000 frame sequences (see fig. 5.7b) but as the figure shows, the method already starts saturating. In general the performance of the constant statics based method seems not to be satisfying given the evaluations. Surprisingly, in the spatial evaluations below the visual correction performance will be quite good.



(a) CS and CS_G evaluations in the $p_{c,\mathrm{Comb.}}$ metric for 500 frame random sequences



(b) CS and CS_G evaluations in the $p_{c,\mathrm{Comb.}}$ metric for 1000 frame random sequences

Figure 5.7.: Optimal performance with respect to the $p_{c,\mathrm{Comb.}}$ quality metric for random frame sequences of 500 and 1000 frame length.

5.3.4. Temporal Performance Analysis for the LMS methods

The temporal evaluations of the remaining LMS methods are presented in fig. 5.8 where all methods show an asymptotic behavior that starts to saturate after 400 frames. The methods' ranking is without any severe surprises: The lowest performances in this group are achieved by the pure LMS methods $LMS_{3\times3}$ and $LMS_{7\times7}$, which are outperformed by their adaptively improved methods $LMS_{A,3\times3}$ and $LMS_{A,7\times7}$. With concern to the mask size, the $LMS_{7\times7}$ methods takes a bit longer to reach its full potential but outperforms the $LMS_{3\times3}$ method after 150 frames. The LMS_{TV} method separates from the other methods as it performs almost as good as the $LMS_{A,7\times7}$ method but is based upon a 3×3 masks only. As the error bars almost vanish and do not overlap, $LMS_{A,7\times7}$ results as the statistically significantly best method, followed by the LMS_{TV} method. The gated adaptive LMS methods are excluded from further investigations, as their performance is almost indistinguishable to the adaptive LMS methods as shown in fig. 5.9. This result is no surprise as for the random frame sequences there is reason or possibility to block out unwanted image information with this technique.

For completeness, fig. 5.10 shows the SNR quality metric while the MAE and RMSE metrics are shown in appendix C.2. The SNR measure seems not to stabilize on a constant value, which is expected due to the image data dependency and the random frame inputs. Of course these fluctuations are damped by averaging the 80 sequences, however they are still visible in fig. 5.10. Anyway, given these plots, it is obvious that the discussed results are valid in all the quality measures, but

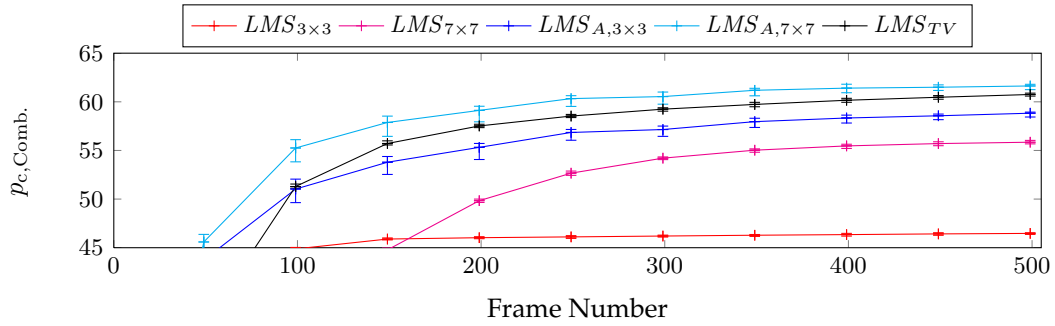


Figure 5.8.: The LMS methods' evaluations for the $p_{\mathrm{c,Comb.}}$ metric with random frame sequences as input.

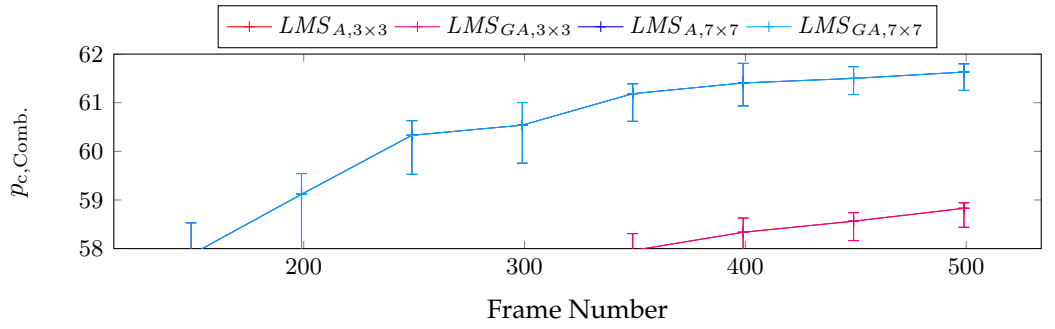


Figure 5.9.: Comparison between the gated adaptive and the adaptive LMS methods in the $p_{c,Combined}$ quality metric for random frame sequences. The differences between the two methods is not notable, as only 2 plot lines are separable.

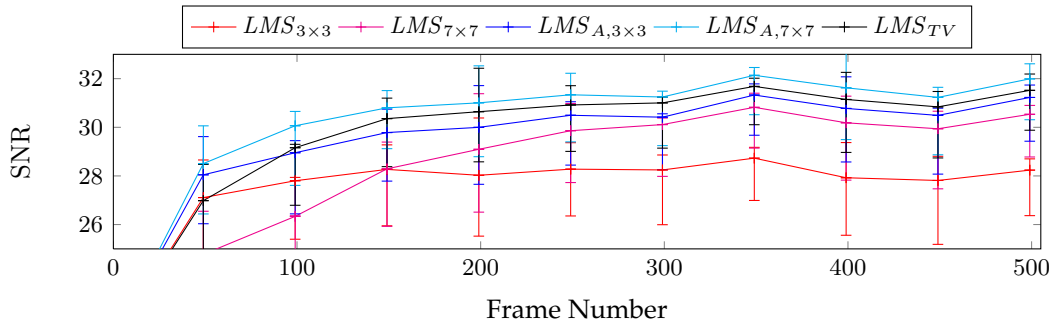


Figure 5.10.: The LMS methods' evaluations for the SNR metric with random frame sequences as input.

only the $p_{\mathrm{c,Comb.}}$ metric is image data independent and with the available $p_{\mathrm{c,DSNU}}$ and $p_{\mathrm{c,PRNU}}$ the separation of the performance onto the effects of DSNU and PRNU is possible (sec. 3.4).

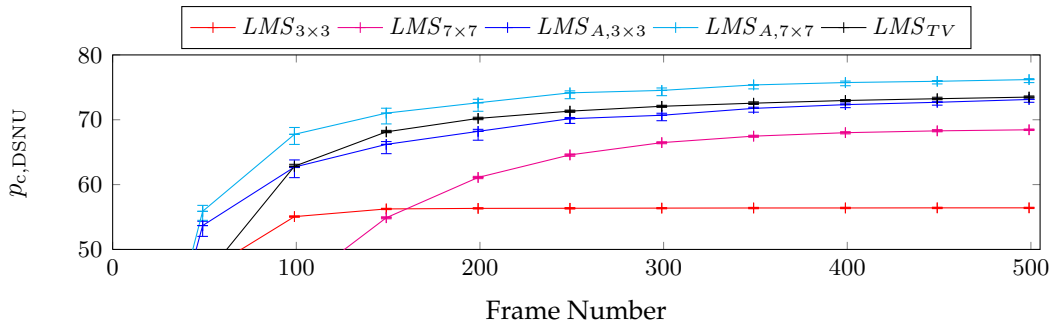


Figure 5.11.: The LMS methods' evaluations for the $p_{\mathrm{c,DSNU}}$ metric with random frame sequences as input.

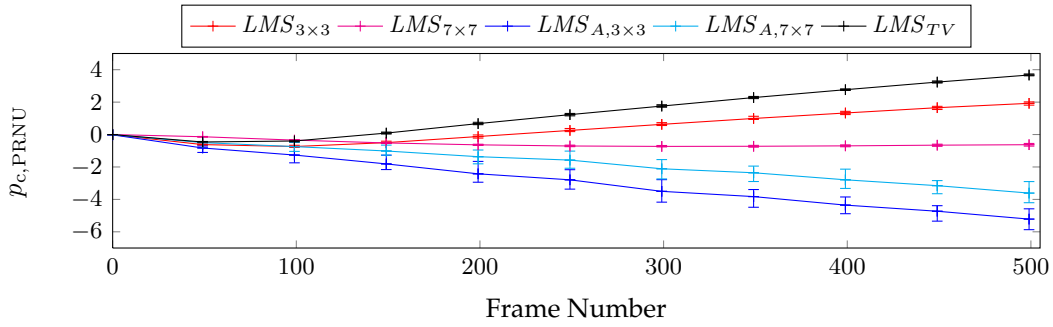


Figure 5.12.: The LMS methods' evaluations for the $p_{\mathrm{c,PRNU}}$ metric with random frame sequences as input.

The separated view onto the DSNU and PRNU corrections is shown in fig. 5.11 and 5.12. As expected, the DSNU correction is similar to the combined correction rate while the PRNU correction results in negative rates for $LMS_{A,3\times3}$ and $LMS_{A,7\times7}$. These results confirm the observations from the best-case performance evaluations shown in sec. 5.3.2. With respect to the PRNU correction the $LMS_{3\times3}$ and LMS_{TV} methods result in the best performance but still only correct about $4\,\%$ of the PRNU. Further, the $p_{\mathrm{c,PRNU}}$ seems not to convergence for the 500 frame sequences and the 1000 frame sequences have also been evaluated and their $p_{\mathrm{c,Comb.}}$ and $p_{\mathrm{c,PRNU}}$ performances are shown in fig. 5.13 and 5.14. The results do not differ much from the above described 500 frame sequences. Especially the non-convergence of the $p_{\mathrm{c,PRNU}}$ is still present after 1000 frames. For the other quality norm results see appendix C.2.

With concern to the literature parameter set, fig. 5.3.4 gives the expected results of a decreased performance when compared to the optimized parameters shown in fig. 5.8. Especially the adaptive LMS methods $LMS_{A,7\times7}$ and $LMS_{A,3\times3}$ are not fully converging within the 500 frames using this parameter set.

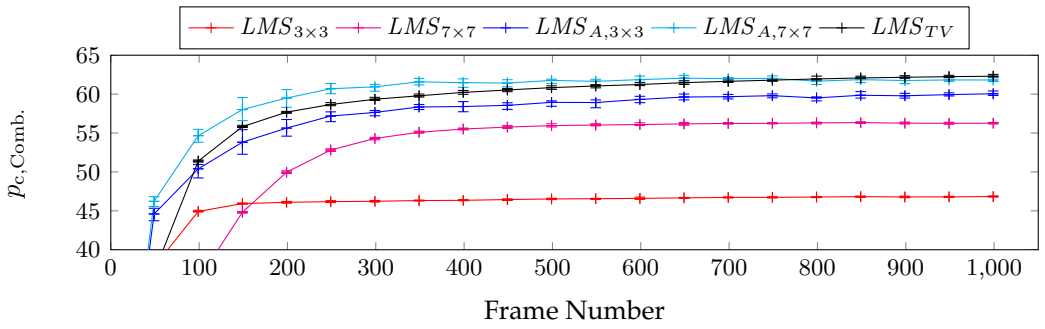


Figure 5.13.: The LMS methods' evaluations for the $p_{\mathrm{c,Comb.}}$ metric with random frame sequences of 1000 frames length.

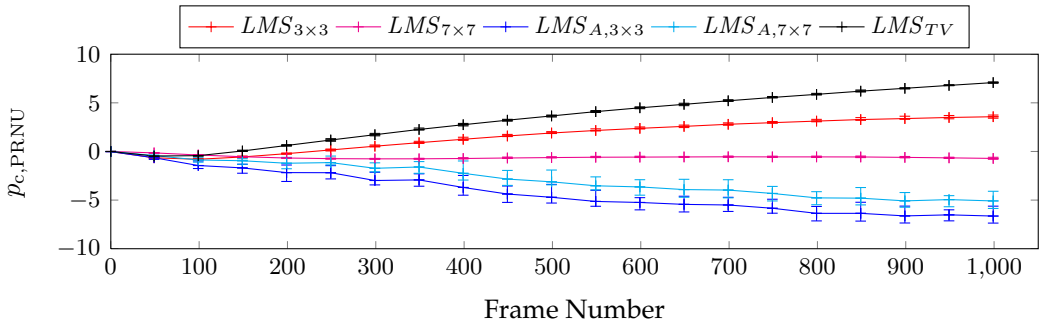


Figure 5.14.: The LMS methods' evaluations for the $p_{\mathrm{c,PRNU}}$ metric with random frame sequences of 1000 frames length.

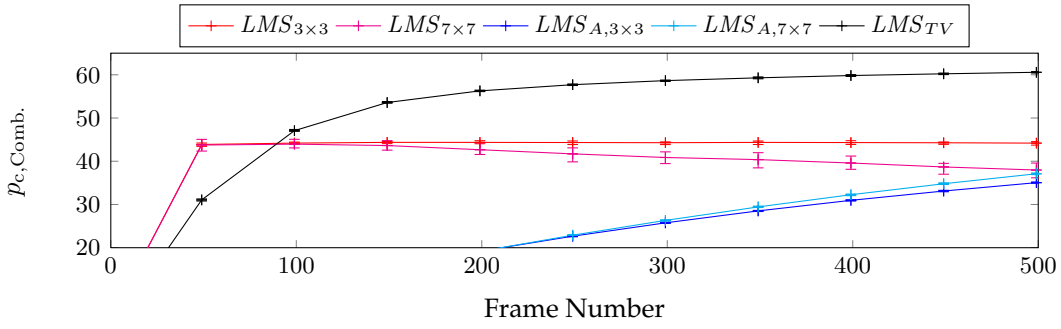
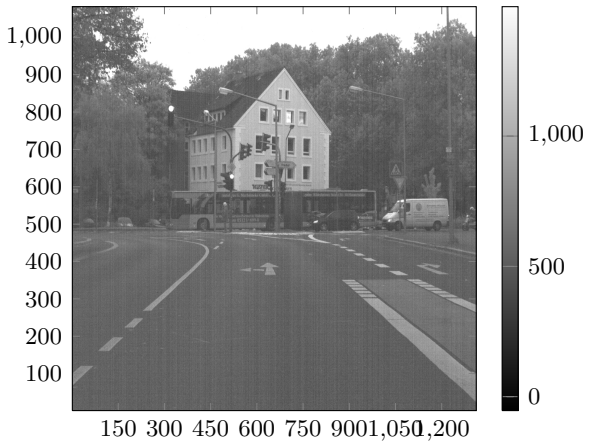


Figure 5.15.: The LMS methods evaluations for the $p_{\mathrm{c,Comb.}}$ metric with the literature parameter set from tab. 5.4 and random frame sequences as input.

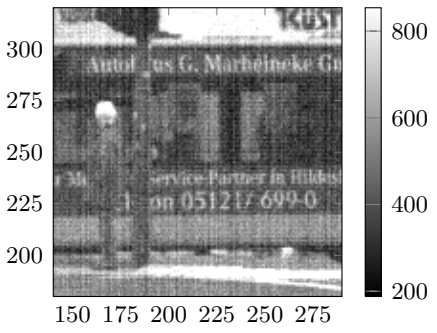
5.3.5. Spatial Performance Analysis

An analysis of the methods with respect to their spatial and spectral performance was already introduced in sec. 3.6. This analysis is now conducted on the last frame of the first random frame evaluation sequence. The last frame forms a good measure, as all methods provide an almost monotone development of their correction performance. The analysis for consecutive frame sequences and the image material with low valued nonuniformities is presented in sec. 10.1 and 10.2.

Fig. 5.16 shows the used image and the spatial high frequency degradations are clearly visible in the uncorrected versions. To provide a detailed view onto the effects of the degradations, a centered region of interest of the image has been chosen for the presentation. Especially the woman in front of the bus is barely visible and the text on the bus is readable only partly in the uncorrected image. However, both tasks can be completed easily in the ground truth corrected image.



(a) Whole recorded scene



(b) Region of interest at the bus

(c) Region of interest from the ground truth corrected image

Figure 5.16.: The used image for the evaluations and the chosen region of interest, as recorded and ground truth corrected version.

(a) Corrected with CS Method

(b) Corrected with CS_G Method

(c) Corrected with $LMS_{3\times3}$ Method

(d) Corrected with $LMS_{7\times7}$ Method

(e) Corrected with $LMS_{A,3\times3}$ Method

(f) Corrected with $LMS_{A,7\times7}$ Method

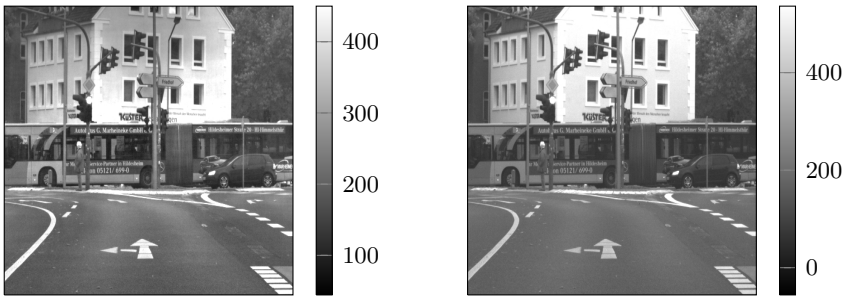
(g) Corrected with $LMS_{GA,7\times7}$ Method

(h) Corrected with LMS_{TV} Method

Figure 5.17.: Spatial correction performance of the reference methods given random frames as input. Parameter set according to tab. 5.2.

Fig. 5.17 shows the image after the corrections with the individual methods. All corrected images give the same mean and standard deviation like the ground truth corrected image as a result of the newly introduced scaling of the correction patterns described in sec. 3.5. For the CS method, the correction looks very good, while the CS_G method provides a deteriorated corrected image. The $LMS_{3\times3}$ and $LMS_{7\times7}$ methods improve the image quality and only few image degradations are left (mainly visible in the constant regions on the bus). Next, the $LMS_{A,3\times3}$ and $LMS_{A,7\times7}$ methods improve the quality even further, and visual degradations are almost not detectable in the given representations. The $LMS_{GA,7\times7}$ method is depicted as well but shows no visual improvement of the quality, as expected. The LMS_{TV} method gives a very good correction as well and a visual difference between the $LMS_{A,3\times3}$, $LMS_{A,7\times7}$ and LMS_{TV} methods is not directly noteable.

With concern to the constant statistic methods, the visual correction of the ROI is good in comparison to their achieved scalar performances. However, degradation for this method becomes visible once a larger ROI is considered as shown in fig. 5.18. The degradation results mainly from spatial low frequency components.



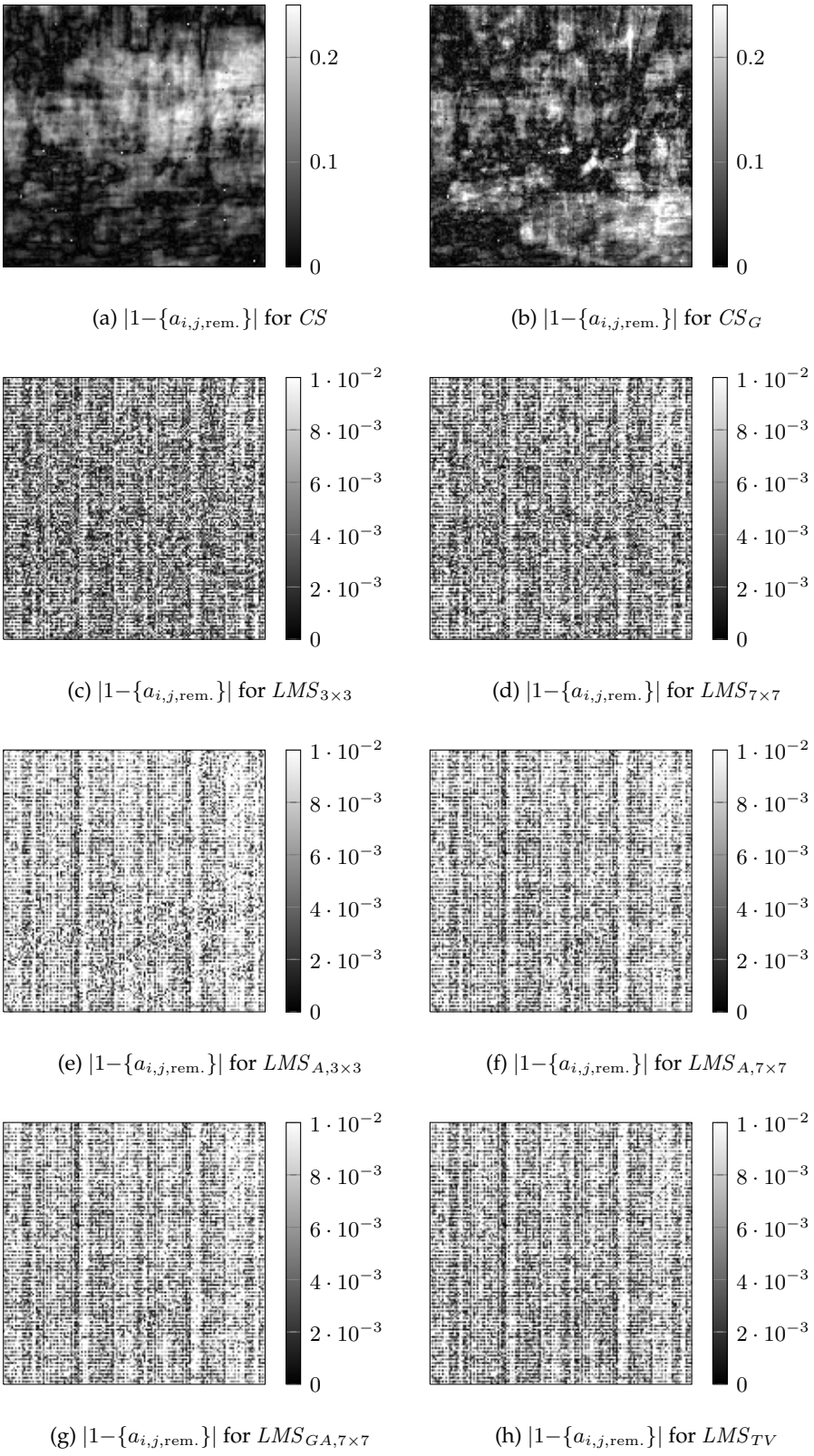
(a) Corrected Image with CS Method

(b) Ground Truth Corrected image

Figure 5.18.: Larger ROI considering the CS method's correction versus the ground truth corrected image. The degradations by spatial low frequencies becomes visible.

More insight in the methods' performances is gained by analyzing the remaining correction patterns. Fig. 5.19 shows the remaining PRNU correction pattern displayed as $|1-a_{i,j,\mathrm{rem.}}|$. The CS and CS_G method show huge remaining patterns that reach the order 0.2 in the given representation, which explains the large negative correction rates. For the LMS methods, the evaluation shows much smaller numbers that range about 10^{-2}. These numbers magnitude are the range of the actual values of $\{a_{i,j}\}$-values measured in sec. 4. Thus the presented result is consistent with the reported low correction rates at the PRNU.

Fig. 5.20 shows the remaining DSNU correction patterns $|\{b_{i,j}\}|$. A view onto the displayed scale reveals that the CS and CS_G methods' remains are a magnitude higher than the actual value of the DSNU, which is about $50\,\mathrm{DN}$. This reveals again reveals the reasons of these methods' bad performances. As mentioned above, the visual correction of the constant statistic based methods is quite good because the



(a) $|1-\{a_{i,j,\mathrm{rem.}}\}|$ for CS

(b) $|1-\{a_{i,j,\mathrm{rem.}}\}|$ for CS_G

(c) $|1-\{a_{i,j,\mathrm{rem.}}\}|$ for $LMS_{3\times3}$

(d) $|1-\{a_{i,j,\mathrm{rem.}}\}|$ for $LMS_{7\times7}$

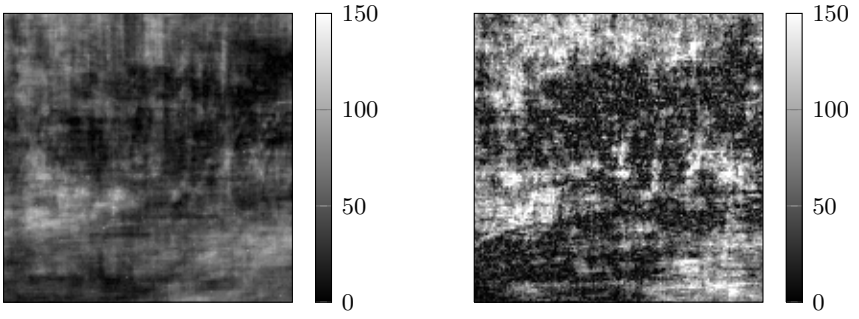
(e) $|1-\{a_{i,j,\mathrm{rem.}}\}|$ for $LMS_{A,3\times3}$

(f) $|1-\{a_{i,j,\mathrm{rem.}}\}|$ for $LMS_{A,7\times7}$

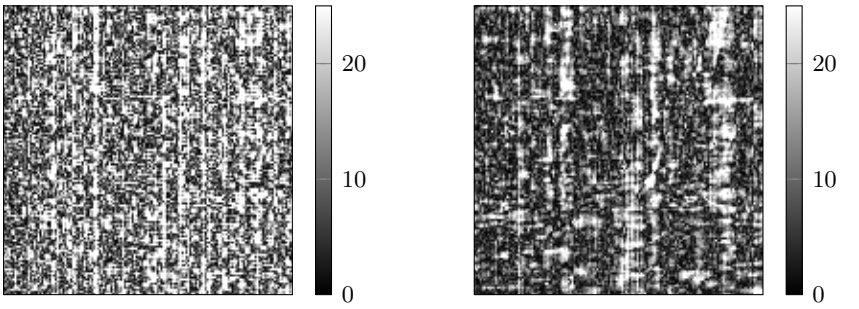
(g) $|1-\{a_{i,j,\mathrm{rem.}}\}|$ for $LMS_{GA,7\times7}$

(h) $|1-\{a_{i,j,\mathrm{rem.}}\}|$ for LMS_{TV}

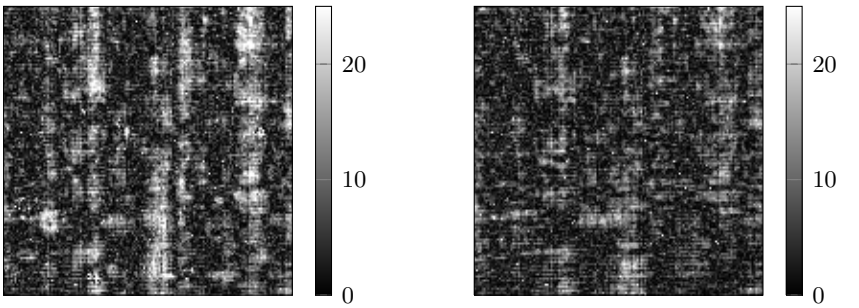
Figure 5.19.: Remaining PRNU patterns for the reference methods given random frames as input. Parameter set according to tab. 5.2.



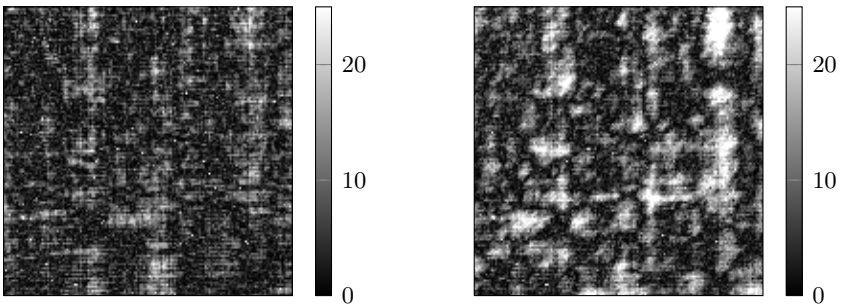
(a) $|\{b_{i,j,\text{rem.}}\}|$ for CS (b) $|\{b_{i,j,\text{rem.}}\}|$ for CS_G



(c) $|\{b_{i,j,\text{rem.}}\}|$ for $LMS_{3\times3}$ (d) $|\{b_{i,j,\text{rem.}}\}|$ for $LMS_{7\times7}$



(e) $|\{b_{i,j,\text{rem.}}\}|$ for $LMS_{A,3\times3}$ (f) $|\{b_{i,j,\text{rem.}}\}|$ for $LMS_{A,7\times7}$



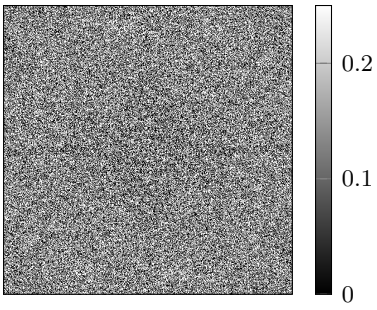
(g) $|\{b_{i,j,\text{rem.}}\}|$ for $LMS_{GA,7\times7}$ (h) $|\{b_{i,j,\text{rem.}}\}|$ for LMS_{TV}

Figure 5.20.: Remaining DSNU patterns for the reference methods given random frames as input. Parameter set according to tab. 5.2.

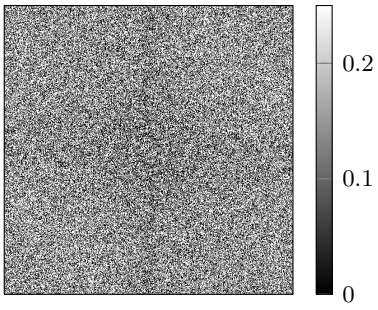
remainders consist mainly of spatial low frequency components. Thus a visually pleasing correction is still achieved and only larger areas of constant gray values are deteriorated. The reasons for the worse impression of the CS_G method are discussed below. With concern to the LMS methods show the same qualities as before: $LMS_{3\times3}$ leaves high amplitude and spatial high frequency components behind, which leads to the visual deterioration in the image. The $LMS_{7\times7}$ method shows these two properties as well, and the $LMS_{A,3\times3}$ and $LMS_{A,7\times7}$ methods are the first two methods with significantly better results. Especially in terms of the spatial higher frequencies, the LMS_{TV} method gives a good correction of these frequencies but leaves remains of a higher amplitude. Although some of the methods' performances are very similar with respect to their scalar quality measures, their remaining DSNU patterns differ significantly from method to method.

The spectral damping analysis of the PRNU correction shows no frequency dependent damping effects (compare in fig. 5.21). Especially the higher frequencies are not damped more than the lower frequencies, which is expected for good corrections. A visual impression like this fits to the results of the almost not present correction of the PRNU as seen in the temporal performance analysis.

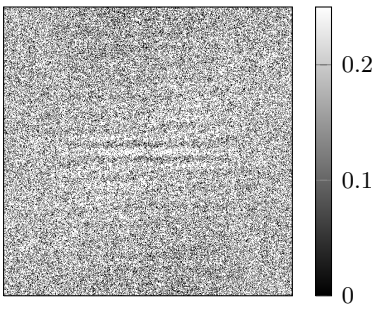
The spectral damping properties for the DSNU are presented in fig. 5.22 and give a better insight in the methods differences. The black pixels of these images represent a very well damped frequency component, while the white parts stand for a damping factor of just 0.25 (or even less damping), compared to the ground truth. The remaining frequency components show up in the center of the spectral analysis image where the spatial lower frequencies are represented. The CS method leaves spectral horizontal and vertical frequencies behind, which are represented by the white cross-like appearance. The CS_G method leaves even more of the spatial higher frequency components behind. While its relative damping properties are comparable to the CS method (e.g. a good correction performance for spatial high frequency components), the average damping factor only ≈ 0.1 and therefore much higher than the one of the CS method. This in general worse performance in the spatial higher frequencies is also the reason for the degraded image quality in the visual inspection in fig. 5.17. For the $LMS_{3\times3}$ method, the damping is not very centered towards the lower frequencies, but on the other hand a good damping of the vertical and horizontal frequencies can be observed. However the $LMS_{3\times3}$ method gives a wide spectrum of uncorrected lower frequencies which results in a worse impression than the CS and CS_G method in this representation, but the lower frequencies are not that critical for the visual impression. On behalf of the $LMS_{A,3\times3}$ and $LMS_{A,7\times7}$ methods even more of the lower frequencies are damped which uncovers a remainder of vertical and horizontal frequency components. The LMS_{TV} method provides less damping in lower frequencies but therefore the vertical and horizontal frequencies are removed much better.



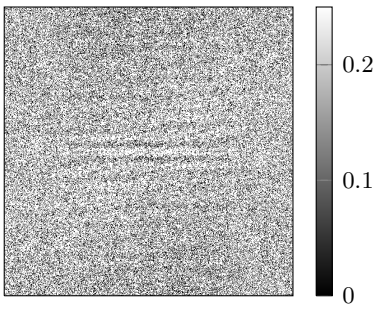
(a) $\hat{s}_{\text{spectr.,PRNU},\mu,\nu}$ for CS



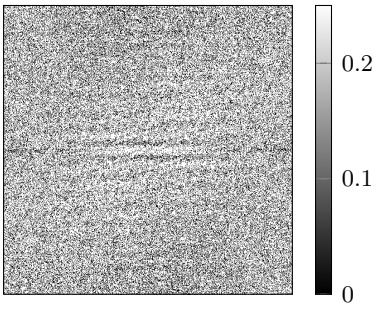
(b) $\hat{s}_{\text{spectr.,PRNU},\mu,\nu}$ for CS_G



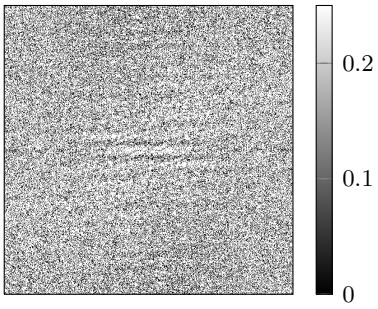
(c) $\hat{s}_{\text{spectr.,PRNU},\mu,\nu}$ for $LMS_{3\times3}$



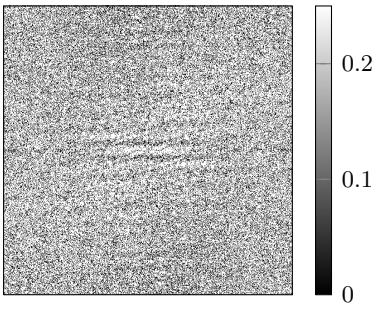
(d) $\hat{s}_{\text{spectr.,PRNU},\mu,\nu}$ for $LMS_{7\times7}$



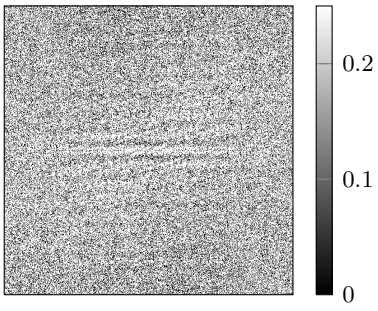
(e) $\hat{s}_{\text{spectr.,PRNU},\mu,\nu}$ for $LMS_{A,3\times3}$



(f) $\hat{s}_{\text{spectr.,PRNU},\mu,\nu}$ for $LMS_{A,7\times7}$

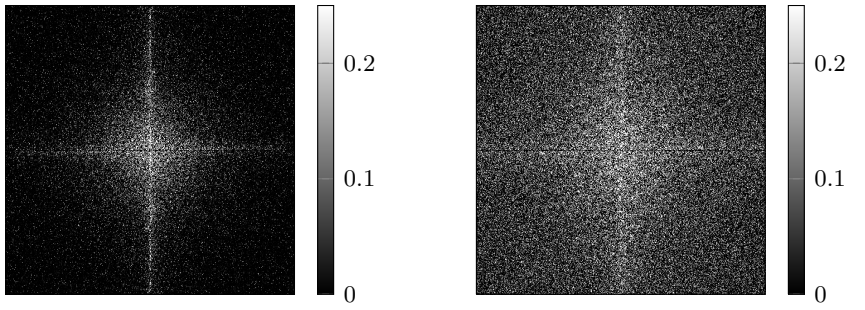


(g) $\hat{s}_{\text{spectr.,PRNU},\mu,\nu}$ for $LMS_{GA,7\times7}$



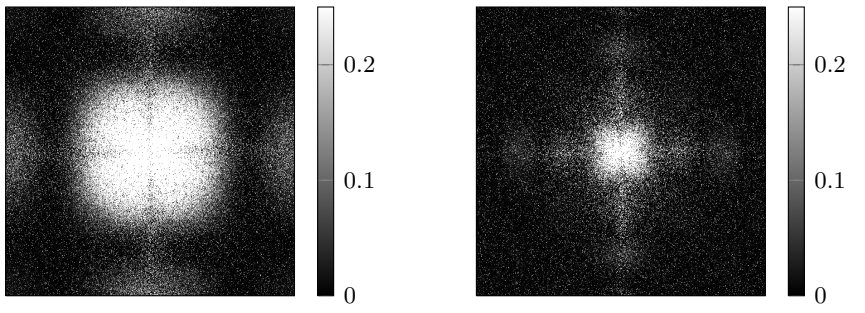
(h) $\hat{s}_{\text{spectr.,PRNU},\mu,\nu}$ for LMS_{TV}

Figure 5.21.: The spectral damping analysis for the PRNU given random frames as input. Parameter set according to tab. 5.2.



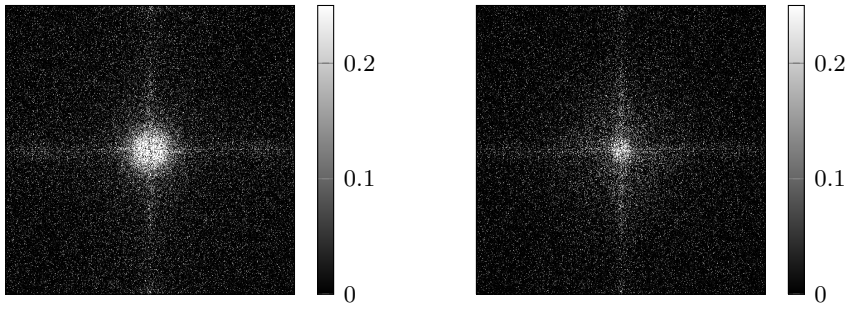
(a) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for CS

(b) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for CS_G



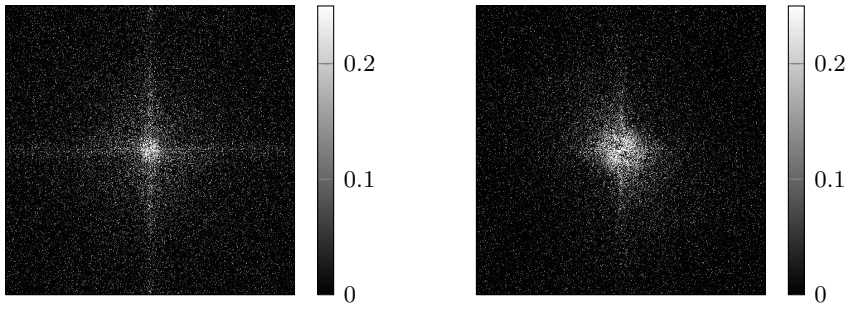
(c) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $LMS_{3\times3}$

(d) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $LMS_{7\times7}$



(e) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $LMS_{A,3\times3}$

(f) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $LMS_{A,7\times7}$



(g) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $LMS_{GA,7\times7}$

(h) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for LMS_{TV}

Figure 5.22.: The spectral damping analysis for the DSNU given random frames as input. Parameter set according to tab. 5.2.

5.4. First conclusions

The observed results confirm the ability of the introduced methods to correct the spatial high frequency components for random frame sequences. The above research however also showed that the choice of the parameter sets for these methods is very critical and that the CS and CS_G methods may not lead to the desired or expected correction performances. The statement that larger filter masks are not leading to an improved performance, which was published in [4], could be shown as inaccurate and was caused by non-optimized parameters.

The removal of visible nonuniformities is possible to some extend by all methods, given the random frame sequence as input. However, further investigations of the remaining DSNU and PRNU patterns and their according spectral damping factors reveale the differences of the methods. It was discussed that the methods in general corrected successfully for DSNU only, which was possible with help of the newly introduced separation into DSNU and PRNU corrections. As the DSNU corrections stagnated, with the development of the adaptive LMS algorithms, a limiting point of the frame-by-frame based LMS algorithms may have been reached.

Further results on the methods behavior on consecutive frame sequences and on the image material with low valued nonuniformities is presented and discussed against the newly introduced methods in sec. 10.1 and 10.2.

6. A New Maximum Likelihood Estimation for DSNU

Given the investigation of the related methods that were described in chapter 5, the possible improvements for a new method are clear. To summarize, a sound motivation for the method's assumptions is needed as well as a defined way to estimate the free parameters. Additionally, most of the yet unconsidered but available statistical information should be used in the new method and an improved PRNU estimation should be possible.

These demands are met by deriving a new method in this chapter, based on a maximum likelihood approach. The approach is formulated for a DSNU-only sensor model and the PRNU estimation will be introduced in chapter 8. As mentioned in chapter 5, the reference methods provide only very limited correction for the PRNU part of the nonuniformities, which will allow first comparisons of the new method in the DSNU-only domain. Further, the mathematical derivation is easier to follow this way. Some applications actually demand for a DSNU-only correction because the part of image degradation by PRNU is in general small in contrast to the typical DSNU degradations (see chapter 4). The method can then be tested and improved first (see chapter 7) before it is completed by a combined DSNU and PRNU estimation in chapter 8. The contents of this chapter have been partly published [12] and the method has a patent pending [14].

6.1. Mathematical Ansatz

With respect to the reference methods, a new method should fulfill the above mentioned demands to overcome the currently known limitations. A maximum likelihood approach is the method of choice, given the demand for a sound motivation of the method's assumptions and especially for the utilization of the available statistical information.

A consideration of the parameters and measurements as realizations of random variables will allow to build the needed mathematics to derive a maximum likelihood approach. To start, a maximum likelihood method targets to find the nonuniformity parameters $\{b_{i,j}\}$ that are the most *likely* ones and for the given task the probability density is considered a good measure for the likelihood. Consequently, the approach is set up to find those parameters that maximize the probability density of a random variable that represents the DSNU. This random variable is

named **DSNU**. In general the following notation on random variables and their realizations is used:

$$\begin{aligned} \mathbf{x} &\quad \text{Random variable} && (6.1)\\ x &\quad \text{realization of } \mathbf{x} && (6.2)\\ \mathbf{x} = x &\quad \mathbf{x} \text{ has the given realization } x. && (6.3)\\ f_{\mathbf{x}}(x) &\quad \text{probability density distribution of } \mathbf{x} && (6.4) \end{aligned}$$

To clarify the notation, **DSNU** is the random variable describing the DSNU in terms of the statistical mathematics. A realization of **DSNU** is given by a defined set $\{b_{i,j}\}$ as $\mathbf{DSNU} = \{b_{i,j}\}$. The measurements of the effects of DSNU are given by the $\mathrm{DSNU}_{\mathrm{EMVA1288,DN}}$ metric, which lies outside the scope of the current statistic mathematical descriptions.

As DSNU is defined by a set of parameters $\{b_{i,j}\}$, **DSNU** has to be dependent to a given realization of these $\{b_{i,j}\}$: $\mathbf{DSNU}(\{b_{i,j}\})$. The probability density is then given as $f_{\mathbf{DSNU}}(\{b_{i,j}\})$ and further details on its definition will follow. A plain and not very smart maximum likelihood approach could now try to solve

$$\arg\max_{\{b_{i,j}\}} f_{\mathbf{DSNU}}(\{b_{i,j}\})\,. \tag{6.5}$$

Of course this would lead to the rather trivial result of the expectation values of the $\{b_{i,j}\}$, because the measurements of the actual light signals have not yet been considered. The above equation needs therefore to be modified by including the measurement of the actual light signal I_{World} that is presented to the sensor. To transfer the light signal into the statistical framework, we consider it as the realization of its random variable $\mathbf{I}_{\mathrm{World}}$.

The maximum likelihood estimation can then be improved to maximize the conditional probability of **DSNU**, given the occurrence of the spatio-temporal signal $\mathbf{I}_{\mathrm{World}} = I_{\mathrm{World}}$:

$$\begin{aligned} &\mathbf{DSNU}|(\mathbf{I}_{\mathrm{World}}{=}I_{\mathrm{World}}) && (6.6)\\ \text{or shorter: } &\mathbf{DSNU}|\mathbf{I}_{\mathrm{World}}{=}I_{\mathrm{World}}\,, && (6.7) \end{aligned}$$

which results in the mathematical task of solving

$$\arg\max_{\{b_{i,j}\}} f_{\mathbf{DSNU}|\mathbf{I}_{\mathrm{World}}=I_{\mathrm{World}}}(\{b_{i,j}\})\,. \tag{6.8}$$

Until now these tasks form only an abstract mathematical ansatz. The probability density of the above conditional probability can indeed not be derived directly, but

the Bayes' theorem [29] gives the probability density distribution of a conditional event as

$$f_{X|Y=y}(x) = \frac{f_X(x) f_{Y|X=x}(y)}{\underbrace{\int_{-\infty}^{\infty} f_X(x) f_{Y|X=x}(y)\, dx}_{N}} . \tag{6.9}$$

This allows to transfer the maximization task into

$$\arg\max_{\{b_{i,j}\}} \frac{1}{N} \cdot f_{\mathbf{DSNU}}(\{b_{i,j}\}) \cdot f_{\mathbf{I}_{\mathrm{World}}|\mathbf{DSNU}=\{b_{i,j}\}}(I_{\mathrm{World}}) . \tag{6.10}$$

The still unknown probability densities of the random variables are now to be defined in the next steps.

6.2. Density Estimations of the Random Variables

6.2.1. The Density of DSNU

The dependency of **DSNU** on the parameters $\{b_{i,j}\}$ is exploited to define the random variable **DSNU**. A random variable is completely defined if its state space as well as its probability density distribution function are given. The state space of **DSNU** is defined as Ω_{DSNU} and covers all the possible realizations given by the parameters $\{b_{i,j}\}$. As the individual $\{b_{i,j}\}$ are real values, their linear span defines the space Ω_{DSNU} and a single realization of $\{b_{i,j}\}$ represents a point ω in the space Ω_{DSNU}:

$$\Omega_{\mathrm{DSNU}} = \mathrm{span}(\{b_{i,j}\}) \quad \omega = \{b_{i,j}\} \in \Omega_{\mathrm{DSNU}}$$

The further definition is straight forward with help of a multivariate random variable, where the values of the parameters $b_{i,j}$ are considered the realizations of their corresponding random variables $\mathbf{b}_{i,j}$ [29]:

$$\mathbf{DSNU} : \Omega \to \mathbb{R}^{M \times N} \tag{6.11}$$

$$(\mathbf{b}_{1,1}, \ldots, \mathbf{b}_{M,N}) \to (b_{1,1}, \ldots, b_{M,N})^T \tag{6.12}$$

Now the random variable **DSNU** is defined almost completely and only its probability density is missing. To be able to define it in a simple way, we assume the random variables $\{\mathbf{b}_{i,j}\}$ to be independent and identically distributed (iid). In this case it follows from the statistical mathematics that a product over the single inde-

pendent densities of the variables $\{\mathbf{b}_{i,j}\}$ forms the combined probability density $f_{\mathbf{DSNU}}$:

$$f_{\mathbf{DSNU}}(\{b_{i,j}\}) = \prod_{i,j} f_{\mathbf{b}_{i,j}}(b_{i,j})\,. \tag{6.13}$$

The assumption of independent identically distributed random variables is common in the literature of scene based nonuniformity estimations and also mentioned implicitly by Hardie [5].

The manufacturing process of the light sensors motivates the iid assumption a bit more, as this manufacturing process can be represented by a random process that generates the individual parameters. Of course this processes should be identical for each produced image sensor and independent from its realization at the other sensor positions. Summarized, the manufacturing process should fulfill the iid assumption for the $\{\mathbf{b}_{i,j}\}$. And even if this does not hold strictly, the manufacturer will aim to make this process as identical and independent for each image sensor as technically possible. However, for the typical row and column wise nonuniformities the iid assumption does not hold and this problem will be discussed in chapter 7.

Assuming *iid* random variables makes it easy to estimate the probability density distribution of the random process. A random sample of the realizations of ground truth measured $b_{i,j}$ of one or more image sensors, allows to fit an empirical model to the histogram of the obtained realizations. As already discussed in sec. 4, and especially for the Photonfocus MV1-D1312-160-CL camera (sec. 4.3.1), the $\{b_{i,j}\}$ obey approximately a Gaussian distribution. Thus we model the random variables for each given (i,j) to obey a Gaussian distribution, with the same parameters μ_b and σ_b for all $\mathbf{b}_{i,j}$. This expresses like

$$\mathbf{b}_{i,j} \sim \mathcal{N}\left(\mu_b, \sigma_b^2\right) \quad \forall (i,j) \tag{6.14}$$

$$f_{\mathbf{b}_{i,j}}(b_{i,j}) = \frac{1}{\sigma_b\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{b_{i,j}-\mu_b}{\sigma_b}\right)^2}\,, \tag{6.15}$$

and the random variable $\mathbf{DSNU}$ is now completely defined.

6.2.2. The Density of $\mathbf{I}_{\mathrm{World}}|\mathbf{DSNU}=\{b_{i,j}\}$

The conditional random variable $\mathbf{I}_{\mathrm{World}}|\mathbf{DSNU}$ gives the probability of the world light signal given an assumed realization of the DSNU pattern. Here the light signal I_{World} is a temporally and spatially continuous signal, which is sampled in the temporal and spatial domain by the light sensors as described in sec. 2.6.

Given these facts, I_{World} is obviously not influenced by the realizations of DSNU and the random variable $\mathbf{I}_{\mathrm{World}}$ and its conditional pendant $\mathbf{I}_{\mathrm{World}}|\mathbf{DSNU}=\{b_{i,j}\}$

are identical. The world signal I_{world} is however not known and cannot be measured due to the nonuniformity degradations that occur in the measurement process. Thus, only the measured and reconstructed version of I_{world} is available, which is the result of the measurement process and therefore dependent on the assumed or given nonuniformities. The best possible reconstruction of I_{world} is therefore used as an approximation to the realizations of the random variable $\mathbf{I}_{\text{World}}$. The reconstruction was already derived in sec. 2 in eq. 2.28 and considers the nonuniformity degradation. Thus the reconstructed signal approximates the real world signal

$$I_{Reco}(\{b_{i,j}\}) \approx I_{\text{World}}\,, \tag{6.16}$$

which allows the signal $I_{Reco}(\{b_{i,j}\})$ to be considered as the realization of a random variable as well. This random variable has of course a dependency on the realizations of the values of $\{b_{i,j}\}$: $I_{Reco}(\{b_{i,j}\})$. As the $\{b_{i,j}\}$ have already been connected to the random variable $\mathbf{DSNU}$, I_{Reco} can be considered as the realization of the conditional random variable:

$$\mathbf{I}_{Reco}|\mathbf{DSNU}{=}\{b_{i,j}\}\,. \tag{6.17}$$

With the similarity of $I_{Reco}(\{b_{i,j}\}) \approx I_{\text{World}}$ it can be concluded that

$$\mathbf{I}_{\text{World}} = \mathbf{I}_{\text{World}}|\mathbf{DSNU}{=}\{b_{i,j}\} \quad \text{and} \quad \mathbf{I}_{\text{World}} \approx \mathbf{I}_{\text{Reco}}|\mathbf{DSNU}{=}\{b_{i,j}\} \tag{6.18}$$

$$\Rightarrow \mathbf{I}_{\text{World}}|\mathbf{DSNU}{=}\{b_{i,j}\} \approx \mathbf{I}_{\text{Reco}}|\mathbf{DSNU}{=}\{b_{i,j}\}\,. \tag{6.19}$$

Extracting Relevant Signal Features

The missing part to define the maximization task is the definition of the probability density of the variable $\mathbf{I}_{\text{Reco}}|\mathbf{DSNU}{=}\{b_{i,j}\}$. As the density of this random variable is not known, further approximations need to be made that exploit a meaningful measurement of the the signal $I_{Reco}(\{b_{i,j}\})$. Such a measurement will allow later to extract the probability density distribution from the histogram of it's realizations. In general many measures for a continuous signal exist, for example the mean or the standard deviation. Even the value of the individual pixels could be used to define such a probability density in a multivariate way. But for the current task those choices seem rather useless, as these properties are not related to the *effects* of the nonuniformities. Therefore features of the realizations of $\mathbf{I}_{\text{Reco}}|\mathbf{DSNU}{=}\{b_{i,j}\}$ that are relevant for the nonuniformity correction task should be used.

The trivial observation that nonuniformities represent as what is commonly referred to as *noise in the image* gives a first idea to define a meaningful measure. Adding noise to an image will *not* smooth the image, or in other words, the probability for an event at which noise smooths the image is obviously neglectable. Therefore adding noise to an image will *increase* spurious edge and gradient like features. Consequently, signal features that consider edges and gradients of the signal should be used for a meaningful measure of the realizations of

$\mathbf{I}_{\mathrm{Reco}}|\mathbf{DSNU}=\{b_{i,j}\}$. To measure the edge and gradient features the fact is used that they are the spatial high frequency components of the image. A signal that expresses edge and gradient features $I_{\hat{L},t_n}$ can therefore be produced by applying an operator $\hat{L}$ with high pass characteristics to the signal:

$$\vec{I}_{\hat{L},t_n} = \hat{L} I_{\mathrm{World}} \tag{6.20}$$

As edges and gradients are the desired measure, the two typical choices of $\hat{L}$ are the Nabla (∇) and the Laplacian ($\triangle$) operator. Both are linear operators but the Nabla operator produces a vector field with two components (e.g. direction and value) for each point of I_{Reco}. Therefore the point wise square of the signal $\vec{I}_{\hat{L},t_n}$ is considered and allows a generalized mathematical formulation for this measure.

$$(\vec{I}_{\hat{L},t_n})^2 = (\hat{L} I_{\mathrm{World}})^2 \tag{6.21}$$

As the nonuniformity parameters are defined for the individual light sensors, the image features should as well allow a sensor-wise interpretation. Such a sensor wise feature can be derived by analyzing the temporal and spatial average of the gradient or edge measure of I_{World}:

$$M_{\hat{L},t_n}(\vec{I}_{\hat{L},t_n}) = \frac{1}{\int_{(x,y)\in A}\int_t 1\, dt\, dx\, dy} \int_{x,y\in A}\int_{t\in(0,T)} \left|\vec{I}_{\hat{L},t_n}(x,y)\right|^2 dt\, dx\, dy \tag{6.22}$$

$$= \frac{1}{N_A N_T} \int_{(x,y)\in A}\int_{t\in(0,T)} \left|\hat{L} I_{\mathrm{World},t_n}(x,y)\right|^2 dt\, dx\, dy\,. \tag{6.23}$$

In the equations, A defines the area of the image sensor on which the signal is sampled. The temporal integral goes over the time from $t = 0$ to $t = T$, which is the recording time of a given sequence. N_A and N_T are the respective normalizing constant for these two integrals and for a better readability of the mathematical equations the integral borders are not always repeated.

For the further transformations, the spatial reconstruction as introduced in eq. 2.22 is used:

$$I_{Reco}(x,y) = \sum_{(i,j)\in S} C_{\mathrm{S},i,j} \cdot \Phi(x-i, y-j)\,, \tag{6.24}$$

where the $C_{\mathrm{S},i,j}$ are the sample coefficient of the world signal, or in other words: The corrected (ideal) sensor responses. Further transformations towards an light sensor wise interpretation lead to:

$$M_{\hat{L}} = \frac{1}{N_A N_T} \int_{(x,y)\in A}\int_t \left(\hat{L} I_{\mathrm{Reco},t_n}(x,y)\right)^2 dt\, dx\, dy \tag{6.25}$$

$$= \frac{1}{N_A N_T} \int_{(x,y)} \int_t \left(\hat{L} \sum_{(i,j)\in S} C_{\mathrm{S},i,j} \cdot \Phi(x-i, y-j) \right)^2 dt\, dx\, dy \tag{6.26}$$

$$= \frac{1}{N_A N_T} \int_t \sum_{(i,j)\in S} \sum_{(k,l)\in S} C_{\mathrm{S},i,j} C_{\mathrm{S},k,l}\, dt \cdot \int_{(x,y)} \hat{L}\Phi(x-i, y-j) \hat{L}\Phi(x-k, y-l)\, dx\, dy \tag{6.27}$$

As the reconstruction functions $\Phi(x, y)$ are symmetrical along the x and y axis, the integral can be simplified to:

$$\int_{x,y\in A} \hat{L}\Phi(x-i, y-j) \hat{L}\Phi(x-k, y-l)\, dx\, dy \tag{6.28}$$

$$= \int_{x,y\in A} \hat{L}\Phi(x, y) \hat{L}\Phi(x-|k-i|, y-|l-j|)\, dx\, dy \tag{6.29}$$

$$= \Psi_{\Phi,\hat{L}}(|k-i|, |l-j|) \in \mathbb{R}\,, \tag{6.30}$$

and $M_{\hat{L}}$ can now be written as:

$$M_{\hat{L}} = \int_t \frac{1}{N} \sum_{(i,j)\in S} \sum_{(k,l)\in S} C_{\mathrm{S},i,j} C_{\mathrm{S},k,l} \Psi_{\Phi,\hat{L}}(|k-i|, |l-j|)\, dt \tag{6.31}$$

$$= \int_t \frac{1}{N} \sum_{(i,j)\in S} C_{\mathrm{S},i,j} \sum_{(k,l)\in S} C_{\mathrm{S},k,l} \Psi_{\Phi,\hat{L}}(|k-i|, |l-j|)\, dt \tag{6.32}$$

In this last form $M_{\hat{L}}$ represents a discrete convolution of C_{S} with a filter mask HP that is defined by $\Psi_{\Phi,\hat{L}}(|k-i|, |l-j|)$ (see eq. 6.30).

$$HP(\hat{L}, \Phi) = \begin{pmatrix} \ddots & \vdots & \vdots & \vdots & \iddots \\ \dots & \Psi_{\Phi,\hat{L}}(1,1) & \Psi_{\Phi,\hat{L}}(0,1) & \Psi_{\Phi,\hat{L}}(1,1) & \dots \\ \dots & \Psi_{\Phi,\hat{L}}(1,0) & \Psi_{\Phi,\hat{L}}(0,0) & \Psi_{\Phi,\hat{L}}(1,0) & \dots \\ \dots & \Psi_{\Phi,\hat{L}}(1,1) & \Psi_{\Phi,\hat{L}}(0,1) & \Psi_{\Phi,\hat{L}}(1,1) & \dots \\ \iddots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{6.33}$$

HP is a symmetric matrix with high pass characteristics that is dependent on the choice of the linear operator $\hat{L}$ and the choice of the reconstruction function Φ. With the Hadamard (element-wise) multiplication $\otimes$ and the definition of the spatial expectation value $e(\cdot)$ eq. 6.32 $M_{\hat{L}}$ can be expressed as:

$$M_{\hat{L}} = \int_t \frac{1}{N_T} e(C_{\mathrm{S}} \otimes (C_{\mathrm{S}} * HP(\hat{L}, \Phi)))\, dt\,. \tag{6.34}$$

At this stage the definition of $M_{\hat{L}}$ is a sum over the sensor dependent, discrete feature points $F_{i,j}(t)$:

$$F_{i,j}(T) = \int_0^T [C_{\mathrm{S}} \otimes (C_{\mathrm{S}} * HP(\hat{L}, \Phi))]_{i,j}\, dt \tag{6.35}$$

$$M_{\hat{L}} = \frac{1}{N_T} e(F_{i,j}(T))\,. \tag{6.36}$$

Each feature point $F_{i,j}(T)$ represents the desired signal property with respect to the individual sensor at position (i, j). The probability of $\mathbf{I}_{\mathrm{Reco}}|\mathbf{DSNU}{=}\{b_{i,j}\}$ is now linked to the realizations of the $F_{i,j}(T)$ features. This means, that the random variable $\mathbf{I}_{\mathrm{Reco}}|\mathbf{DSNU}{=}\{b_{i,j}\}$ can be defined as a multivariate random variable (similar to $\mathbf{DSNU}$).

First, the real valued state space $\Omega_{\mathbf{I}_{\mathrm{Reco}}|\mathbf{DSNU}=\{b_{i,j}\}}$ is given which covers all the possible realizations of the $F_{i,j}(T)$, then $\mathbf{I}_{\mathrm{Reco}}|\mathbf{DSNU}{=}\{b_{i,j}\}$ is considered as a multivariate random variable depending on the realizations of $F_{i,j}$. The $F_{i,j}(T)$ are therefore considered the realizations of their own corresponding random variables $\mathbf{F}_{i,j}(T)$:

$$\mathbf{I}_{\mathrm{Reco}}|\mathbf{DSNU}{=}\{b_{i,j}\} : \Omega_{\mathbf{I}_{\mathrm{Reco}}|\mathbf{DSNU}=\{b_{i,j}\}} \to \mathbb{R}^{M\times N} \tag{6.37}$$

$$(\mathbf{F}_{1,1}, \ldots, \mathbf{F}_{M,N}) \to (F_{1,1}, \ldots, F_{M,N})^T \tag{6.38}$$

Assuming independent and identically distributed features: To define the density of $\mathbf{I}_{\mathrm{Reco}}|\mathbf{DSNU}{=}\{b_{i,j}\}$, we assume again that the random variables $\{\mathbf{F}_{i,j}(T)\}$ are independent and identically distributed (iid). This is a weak assumption which can only be partly justified by theory, but the results of the method will justify its application. Especially neighboring features will express a dependency on each other. However, feature elements with a larger distance can be considered merely independent, as they usually contain information from another object in the image. As most of the possible feature pairs have a large distance, the overall dependencies are rather *weak*, which then partly justifies the assumption of independent features.

The second part of the iid assumption states that the $\mathbf{F}_{i,j}(T)$ obey identical distributions. This is a very common assumption in image processing and has already been exploited in the constant statistics method *CS* described in sec. 5. As the $F_{i,j}(T)$ features measure the gradient or edge strength, they depend on the high frequencies of the image signal. Given the success of the *CS* method, which corrects successfully the high frequency components of the nonuniformities, identical or at least similar probability distributions seem plausible for the $\{\mathbf{F}_{i,j}(T)\}$. Given

the above assumptions, a product over the single independent densities of the variables $\mathbf{F}_{i,j}(T)$ forms the combined probability density $f_{\mathbf{I}_{\mathrm{Reco}}|\mathbf{DSNU}=\{b_{i,j}\}}$:

$$f_{\mathbf{I}_{\mathrm{Reco}}|\mathbf{DSNU}=\{b_{i,j}\}} \approx \prod_{i,j} f_{\mathbf{F}_{i,j}(T)}(F_{i,j}(T))\,. \tag{6.39}$$

This approximation is, as mentioned above, founded on plausible but weak and unverifiable assumptions. However, the good performance of the method justifies these assumptions in the aftermath.

The density of the Feature Variable

The last part missing for a complete definition is the probability density of the feature variable $\mathbf{F}_{i,j}(T)$. The features $F_{i,j}(T)$ depend per definition on the continuous signal $(C_{\mathrm{S}} \otimes (C_{\mathrm{S}} * HP(\hat{L}, \Phi)))$ (def. on p. 43). This signal is however only available as its discrete sample points at the times t_n due to the light sensors' exposure time integration. Using the signal reconstruction formula (eq. 2.22) and an arbitrary reconstruction function $\Phi_{t_n}(t)$ in the temporal domain, then the inclusion of eq. 6.35 allows the following transformations:

$$\text{define}: \quad F_{i,j,t_n} = [C_{\mathrm{S}}(t_n) \otimes (C_{\mathrm{S}}(t_n) * HP(\hat{L}, \Phi)]_{i,j} \tag{6.40}$$

$$F_{i,j}(T) = \int_0^T \sum_{t_n} F_{i,j,t_n} \Phi_{t_n}(t^*)\, dt^* \tag{6.41}$$

$$= \sum_{t_n} F_{i,j,t_n} \underbrace{\int_0^T \Phi_{t_n}(t)\, dt}_{\approx 1} = \sum_{t_n} F_{i,j,t_n}\,. \tag{6.42}$$

It is used that the integral over the reconstruction formula Φ has to be 1 by definition, which is true for most of the functions Φ as they lie completely inside the interval $[0, T]$. In the other cases the assumption will be met at least partly.

The random variable $\mathbf{F}_{i,j}(T)$ depends now on the realizations of the time discrete features F_{i,j,t_n}, which can again be considered as the realizations of the random variables $\mathbf{F}_{i,j,t_n}$. There are typical scenarios where the $\mathbf{F}_{i,j,t_n}$ can be considered independent in their temporal dimension t_n, like randomly chosen frames or frames that are recorded with a very slow frame rate, given a moving scene. In general, the correlation between two arbitrary variables at times t_n and t_{n+m} is only given if m is quite small (e.g. for two consecutive frames, or up to a few hundred frames, depending on capture rate and captured scene). For a whole sequences with N frames, the correlation between two randomly chosen variables $\mathbf{F}_{i,j,t_n}$ may be considered neglectable if the sequence contains enough information. The analysis of the high speed consecutive frames from sec. 4.2 will show the consequences if this assumption is not met (see chapter 10 for these evaluations).

Further to their independence, the random variables for the different times t_n are assumed to obey identical distributions, which can be justified by assuming a temporally stationary random process to create the realizations. Combining the above argumentations, the random variables $\mathbf{F}_{i,j,t_n}$ are assumed iid distributed and the probability density distribution of the $\mathbf{F}_{i,j}(T)$ random variable can then be expressed as the product over the corresponding random variables $\mathbf{F}_{i,j,t_n}$:

$$f_{\mathbf{F}_{i,j}(T)} = \prod_{t_n} f_{\mathbf{F}_{i,j,t_n}}(F_{i,j,t_n})\,. \tag{6.43}$$

The probability density distribution of $\mathbf{I}_{\mathrm{Reco}}|\mathbf{DSNU}=\{b_{i,j}\}$ can consequently be approximated by the random variables $\mathbf{F}_{i,j,t_n}$ and the final approximation of the random variable can be written as

$$f_{\mathbf{I}_{\mathrm{Reco}}|\mathbf{DSNU}=\{b_{i,j}\}} = \prod_{i,j}\prod_{t_n} f_{\mathbf{F}_{i,j,t_n}}(F_{i,j,t_n}) \tag{6.44}$$

$$\text{with} \quad F_{i,j,t_n}(b_{i,j}) = [C_{\mathrm{S}}(b_{i,j},t_n) \otimes (C_{\mathrm{S}}(b_{i,j},t_n) * HP(\hat{L},\Phi)]_{i,j}\,, \tag{6.45}$$

The densities $f_{\mathbf{F}_{i,j,t_n}}(F_{i,j,t_n})$ are yet not defined, but their realizations can be measured on ground truth corrected data. The histograms of those measurements will then allow to extract the probability density distribution of the random variables.

As a consequence of the assumed temporally independent features, the used recorded images $C_S(t_n)$ should be independent. It was mentioned before that this can be achieved approximately by using either a slow capture rate while recording a moving scene, or by selecting the input frames randomly from a given set of consecutive input frames. The final part is now to extract the probability distributions out of the histograms of the realizations F_{i,j,t_n}.

6.2.3. Histogram Evaluations for $\mathbf{F}_{i,j,t_n}$

As a general quality, large values of F_{i,j,t_n} are improbable as they would only occur in very noisy images, which are rare in nature. Smaller values on the other hand seem more likely to some extend, as the image information can be described in many cases by just some edges and gradients. Of course values below 0 do not exist, due to the nature of F_{i,j,t_n} as a squared real number. The choices of $\hat{L}$ in the definition of F_{i,j,t_n} (eq. 6.40) therefore interprete as:

- For $\hat{L} = \nabla$ (with $\nabla = (\partial_x, \partial_y)^T$): A signal I_{World} is considered probable if the spacial gradient in the image is small in most points and a smooth image with no sharp edges (e.g. a constant image) is therefore considered probable.
- For $\hat{L} = \triangle$ (with $\triangle = \partial_x^2 + \partial_y^2$): A signal I_{World} is considered probable if it does not contain sharp edges in most points: This is an extension to $\hat{L} = \nabla$, as we consider constant spacial gradients in I_{World} probable as well (e.g. intensity ramps in the image).

Feature Choice	Resulting Parameter
F_{i,j,t_n} with $HP(\nabla, BL)$	$\lambda = 2.361 \cdot 10^{-5}$
F_{i,j,t_n} with $HP(\nabla, MN)$	$\lambda = 4.652 \cdot 10^{-5}$
F_{i,j,t_n} with $HP(\triangle, MN)$	$\lambda = 1.938 \cdot 10^{-5}$

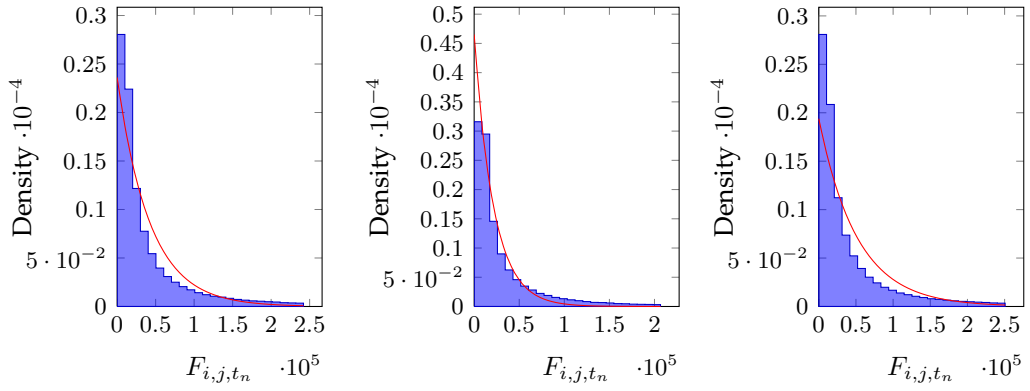
Table 6.1.: Maximum likelihood estimated parameters for the assumed exponential distribution of the feature value distribution of F_{i,j,t_n}.

To define the probability densities of $f_{\mathbf{F}_{i,j,t_n}}(F_{i,j,t_n})$, a histogram over 10^6 samples out of the 25 random frame calibration sequences is created (All this image material was recorded with the MV1-D1312-160-CL camera). To obtain the values of C_S, each frame is corrected with the laboratory ground truth nonuniformity patterns. Fig. 6.1 shows the histograms that result from the different choices of Φ and $\hat{L}$. From a visual inspection the approximation of the probability density distribution as an exponential distribution seems plausible. Although an exponential distribution may not be the best fit, the exponential distribution will lead to a computational easy mathematical expression in the later developments. Fig. 6.1 shows as well the maximum likelihood fit of the assumed exponential distribution and we conclude for the further development of the method that:

$$f_{\mathbf{F}_{i,j,t_n}}(F_{i,j}) = \lambda e^{-\lambda F_{i,j,t_n}} \,. \tag{6.46}$$

The estimated λ parameters according to a maximum likelihood fit are shown in tab. 6.1 and calculate as the expectation value over the samples [29]:

$$\lambda = \frac{1}{\sum_{i,j,t_n} 1} \sum_{i,j,t_n} F_{i,j,t_n} \quad \text{with: } i, j, t_n \in \text{Samples} \,. \tag{6.47}$$



(a) with $HP(\nabla, BL)$ (b) with $HP(\nabla, MN)$ (c) with $HP(\triangle, MN)$

Figure 6.1.: Histograms of the different $\mathbf{F}_{i,j,t_n}$ realizations

6.2.4. Properties of the HP-Filter Masks

For the choices of $\hat{L} \in \{\nabla, \triangle\}$ a high pass filter masks HP is generated as defined in eq. 6.33. To obtain the mask, the reconstruction functions $\Phi \in \{MN, BL\}$ according to sec. 2.6.3 are used. The option with the sinc reconstruction filter is not considered for the reasons mentioned before (e.g. infinite large mask). The numerical representation of the $HP(\nabla, BL)$ filter mask then results as

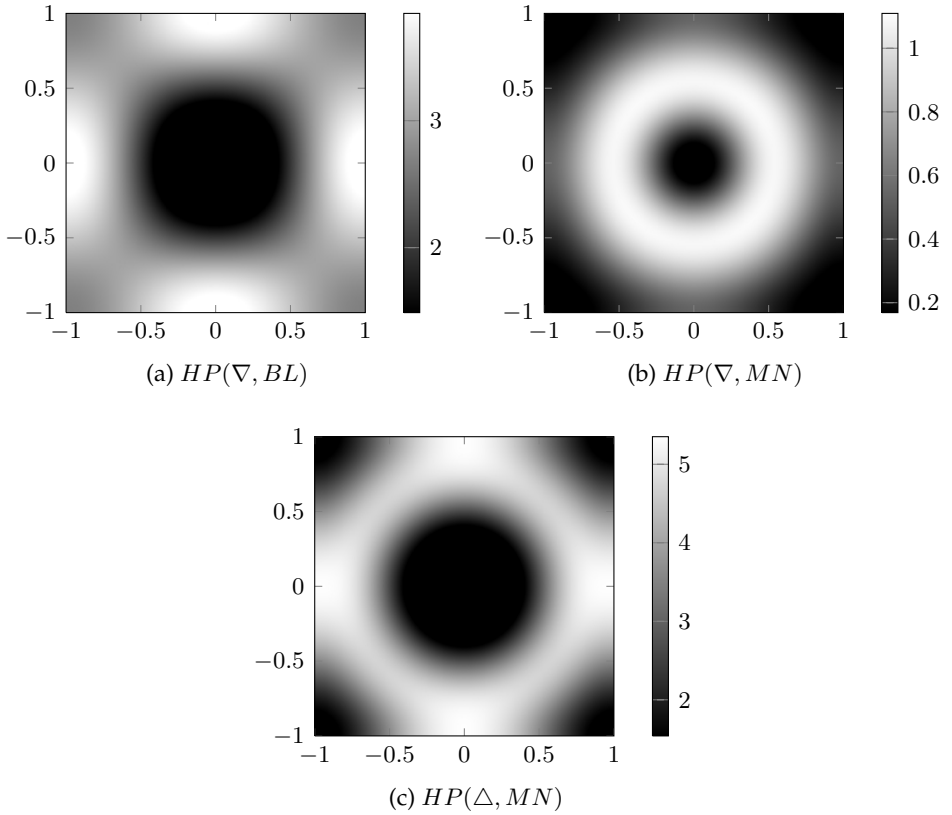
$$HP(\nabla, BL) = \frac{1}{3}\begin{pmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{pmatrix}. \tag{6.48}$$

$HP(\nabla, BL)$ is equal to subtracting a BOX-low-pass filter from a discrete Dirac filter. The LMS reference methods use exactly this mask as discussed in sec. 5.2.1 and fig. 6.2a shows this maks's frequency response by plotting the real part of the Fourier coefficients. The imaginary parts are 0 due to the filter symmetries. The displayed frequencies are normalized to the corresponding Nyquist frequencies $\in (-1, 1)$.

The other masks' numerical representation are as shown below and fig. 6.2b and 6.2c show their frequency responses. All filter masks are high pass filters which can be seen by the damping regions in the center of the image. The influences or the different filter mask properties are discussed in the later sections. To compare the filter masks please consider the different scales of the parameters.

$$HP(\nabla, MN) = \frac{1}{302400}\cdot \begin{pmatrix} -1 & -72 & -603 & -1168 & -603 & -72 & -1 \\ -72 & -2880 & -15192 & -24192 & -15192 & -2880 & -72 \\ -603 & -15192 & -17865 & 29520 & -17865 & -15192 & -603 \\ -1168 & -24192 & 29520 & 193280 & 29520 & -24192 & -1168 \\ -603 & -15192 & -17865 & 29520 & -17865 & -15192 & -603 \\ -72 & -2880 & -15192 & -24192 & -15192 & -2880 & -72 \\ -1 & -72 & -603 & -1168 & -603 & -72 & -1 \end{pmatrix} \tag{6.49}$$

$$HP(\triangle, MN) = \frac{1}{151200}\cdot \begin{pmatrix} 31 & 1104 & 6225 & 10480 & 6225 & 1104 & 31 \\ 1104 & 12096 & 2160 & -30720 & 2160 & 12096 & 1104 \\ 6225 & 2160 & -102465 & -38640 & -102465 & 2160 & 6225 \\ 10480 & -30720 & -38640 & 520960 & -38640 & -30720 & 10480 \\ 6225 & 2160 & -102465 & -38640 & -102465 & 2160 & 6225 \\ 1104 & 12096 & 2160 & -30720 & 2160 & 12096 & 1104 \\ 31 & 1104 & 6225 & 10480 & 6225 & 1104 & 31 \end{pmatrix} \tag{6.50}$$



(a) $HP(\nabla, BL)$ (b) $HP(\nabla, MN)$

(c) $HP(\triangle, MN)$

Figure 6.2.: Frequency responses of the HP filter masks given the choices of $\hat{L} \in \{\nabla, \triangle\}$ and $\Phi \in \{BL, ML\}$. The representation is normalized to the Nyquist frequencies.

6.3. Solving as Log-Likelihood Energy Minimization

With all the above assumptions, the maximum likelihood estimation is fully defined as:

$$\arg\max_{\{b_{i,j}\}} \frac{1}{N} \cdot f_{\mathbf{DSNU}}(\{b_{i,j}\}) \cdot f_{\mathbf{I}_{\text{World}}|\mathbf{DSNU}=\{b_{i,j}\}}(I_{\text{World}}) \tag{6.51}$$

$$\approx \arg\max_{\{b_{i,j}\}} \frac{1}{N} \prod_{i,j} f_{\mathbf{b}_{i,j}}(b_{i,j}) \cdot \prod_{i,j} \prod_{t_n} f_{\mathbf{F}_{i,j,t_n}} \tag{6.52}$$

$$\approx \arg\max_{\{b_{i,j}\}} \frac{1}{N} \prod_{i,j} \frac{1}{\sigma_b \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{b_{i,j}-\mu_b}{\sigma_b}\right)^2} \cdot \prod_{i,j} \prod_{t_n} \lambda e^{-\lambda F_{i,j,t_n}} \,. \tag{6.53}$$

However, the maximization is rather complicated due to the many products and the application of the negative logarithm is a common tool to simplify such tasks.

In the exact solution, the results are identical as the following holds:

$$\arg\max\{g(x)\} \Leftrightarrow \partial_x g(x) = 0 \tag{6.54}$$

$$\arg\min\{-\log g(x)\} \Leftrightarrow \underbrace{\frac{1}{g(x)}}_{\neq 0}\partial_x g(x) = 0 \Leftrightarrow \partial_x g(x) = 0 \tag{6.55}$$

The only drawback of this attempt is that the solutions of iterative solvers will be weighted by $\frac{1}{g(x)}$ when approaching the maximum of the real function, which may result in numerical inaccuracies.

For eq. 6.53 the obtained log-likelihood minimization separates into:

$$\arg\min_{\{b_{i,j}\}}\{\underbrace{-\log\frac{1}{N}}_{\text{const.}}\underbrace{-\log f_{\mathbf{DSNU}}(\{b_{i,j}\})}_{E_\mathrm{P}}\underbrace{-\log f_{\mathbf{I}_\mathrm{World}|\mathbf{DSNU}=\{b_{i,j}\}}(I_\mathrm{World})}_{E_\mathrm{D}}\} \tag{6.56}$$

$$\Leftrightarrow \arg\min_{\{b_{i,j}\}}\{E_\mathrm{P} + E_\mathrm{D} + \text{const.}\} \tag{6.57}$$

Because of the additive behavior of the resulting terms, this problem is similar to the typical kind of energy minimization problems known from physics. The constants are irrelevant for the minimization and the energy term E_P is called *prior-energy* as it contains the prior knowledge of the statistics of the nonuniformities. The energy term is E_D is called *data-energy* as it contains the measurements of the data for the given problem.

With the definitions of eq. 6.53 and eq. 6.56 the energies E_P and E_D can be transformed into:

$$E_\mathrm{P} = -\log f_{\mathbf{DSNU}}(\{b_{i,j}\}) \tag{6.58}$$

$$= \sum_{i,j}\underbrace{\frac{1}{2\sigma_b^2}}_{\beta}(b_{i,j}-\mu_b)^2 + \text{const.} \tag{6.59}$$

$$= \sum_{i,j}\beta\,(b_{i,j}-\mu_b)^2 + \text{const.} \tag{6.60}$$

$$\text{and:}\quad E_\mathrm{D} = -\log f_{\mathbf{I}_\mathrm{World}|\mathbf{DSNU}=\{b_{i,j}\}}(I_\mathrm{World}) \tag{6.61}$$

$$= \sum_{i,j}\sum_{t_n}\lambda F_{i,j,t_n} + \text{const.} \tag{6.62}$$

The intermediate steps for these transformations are explained in appendix D.1.

The definition of F_{i,j,t_n} and the substitution of $C_{\mathrm{S},i,j}$ by the sensor model (eq.2.27, p. 43) complete the definition. The gain factors are thereby assumed as $\{a_{i,j}\} =$

1 because the PRNU is neglected in this DSNU-only approach. The following expressions then represent E_{D}:

$$E_{\mathrm{D}} = \lambda \sum_{i,j} \sum_{t_n} [C_{\mathrm{S}} \otimes (C_{\mathrm{S}} * HP(\hat{L}, \Phi)]_{i,j} + \text{const.} \tag{6.63}$$

$$= \lambda \sum_{t_n} \sum_{i,j} ((C_{\mathrm{M}} - b - \chi) \otimes ((C_{\mathrm{M}} - b - \chi) * HP(\hat{L}, \Phi))) + \text{const.} \tag{6.64}$$

$$= \lambda \sum_{t_n} \sum_{(i,j) \in S} (C_{\mathrm{M},i,j} - b_{i,j} - \chi_{i,j}) \sum_{(k,l) \in S} (C_{\mathrm{M},k,l} - b_{k,l} - \chi_{k,l}) \Psi(\cdot) + \text{const.}\,. \tag{6.65}$$

The energies E_{P} and E_{D} are functions of the set of variables $\{b_{i,j}\}$ which contain as many variables as light sensors on the image sensor. Typical modern sensors have several million light sensors and thus the minimization of this high dimensional problem is a non trivial task. However, there exist many methods in the mathematical literature and the numerical computer sciences that allow to find the minimum of such problems in an iterative way [45]. The simplest of these methods is the iterative steepest descent method. In this method each iteration calculates the gradient at the current point and then updates the point into the direction of the negated gradient. Further developments of the steepest descent method lead to Krylov subset search methods like the conjugate gradient method. Most of the fast iterative methods need to evaluate the gradient of the function to reach the minimum. Further demands of the methods are that the starting point for the iterations needs to be reasonable close to the actual minimum or the function in general needs to be unimodal.

Deriving the Gradient

We derive the evaluation formula for the gradient by splitting the derivative into the individual energy terms:

$$\nabla_{\{b_{i,j}\}}(\text{const.} + E_{\mathrm{P}} + E_{\mathrm{D}}) = \nabla_{\{b_{i,j}\}} E_{\mathrm{P}} + \nabla_{\{b_{i,j}\}} E_{\mathrm{D}} \tag{6.66}$$

$$\text{with:} \quad \nabla_{\{b_{i,j}\}} = (\partial_{b_{1,1}} \dots \partial_{b_{1,N}}, \partial_{b_{2,1}} \dots \partial_{b_{M,N}})^T\,. \tag{6.67}$$

For the prior energy the sum collapses when deriving and one obtains:

$$\nabla_{\{b_{i,j}\}} E_{\mathrm{P}} = \nabla_{\{b_{i,j}\}} (\sum_{i,j} \beta\,(b_{i,j} - \mu_b)^2) + \text{const} \tag{6.68}$$

$$\text{simplified:} \quad \partial_{b_{m,n}} E_{\mathrm{P}} = 2\beta(b_{m,n} - \mu_b)\,. \tag{6.69}$$

To refit the above equation into matrix notation, the $\hat{\nabla}$-operator is used again (as defined in eq. 5.18 on p. 84), which places the (i,j)-th derivative into a matrix at the index (i,j). The gradient of E_P then writes as:

$$\hat{\nabla}_{\{b_{i,j}\}} E_\mathrm{P} = 2\beta(b - \mu_b) \tag{6.70}$$

The gradient of the data energy can also be calculated very easily if one exploits the linearity of the derivative operator and the product rule. (The intermediate steps are shown in appendix D.2):

$$\begin{aligned}
\nabla_{\{b_{i,j}\}} E_\mathrm{D} &= \nabla_{\{b_{i,j}\}} (\sum_{t_n} \sum_{i,j} \lambda (C_\mathrm{M} - b - \chi) \otimes ((C_\mathrm{M} - b - \chi) * HP(\hat{L}, \Phi)) + \text{const.}) \\
\partial_{b_{m,n}} E_\mathrm{D} &= (\sum_{t_n} \sum_{i,j} \lambda (C_\mathrm{M} - b - \chi) \otimes ((C_\mathrm{M} - b - \chi) * HP(\hat{L}, \Phi)) + \text{const.}) \qquad (6.71) \\
&= -2\lambda \sum_{t_n} \sum_{(i,j)\in S} (C_{\mathrm{M},i,j} - b_{i,j} - \chi_{i,j}) \Psi(|i-m|, |j-n|)) \qquad (6.72)
\end{aligned}$$

The sum over t_n suggests to introduce a temporal normalizing factor T_N, which allows to simplify the partial derivative by expressing the variables with their temporal averages. For the temporal random variables (e.g. χ), the temporal average serves as an estimator of the expectation value.

$$T_\mathrm{N} = \sum_{t_n} 1 \tag{6.73}$$

$$\partial_{b_{m,n}} E_\mathrm{D} = -2\lambda T_\mathrm{N} \frac{1}{T_\mathrm{N}} \sum_{t_n} \sum_{(i,j)\in S} (C_{\mathrm{M},i,j} - b_{i,j} - \chi_{i,j}) \Psi(|i-m|, |j-n|)) \tag{6.74}$$

$$= -2\lambda T_\mathrm{N} \sum_{(i,j)\in S} (\underbrace{\frac{1}{T_\mathrm{N}} \sum_{t_n} C_{\mathrm{M},i,j}}_{A_{\mathrm{M},i,j}} - \underbrace{\frac{1}{T_\mathrm{N}} \sum_{t_n} b_{i,j}}_{b_{i,j}} - \underbrace{\frac{1}{T_\mathrm{N}} \sum_{T_\mathrm{N}} \chi_{i,j}}_{\to 0}) \Psi(\cdot)) \tag{6.75}$$

$$= -2\lambda T_\mathrm{N} \sum_{(i,j)\in S} (A_{\mathrm{M},i,j} - b_{i,j}) \Psi(|i-m|, |j-n|)) \tag{6.76}$$

For χ, its property of a *per definition* mean free random variable is used, which lets the temporal average converge towards 0 for many considered frames. b does not have any temporal dependency and the newly introduced variable $A_{\mathrm{M},i,j}$ just represents the temporal average of the light sensor measurements. The sum over the indices i,j is limited by the function $\Psi(|i-m|, |j-n|)$, which becomes 0 outside the direct neighborhood around the index of interest (m,n). This makes the total derivative again the convolution with the high pass filter defined by Ψ and $\hat{L}$ (see sec. 6.2.4). In combination with the $\hat{\nabla}_{\{b_{i,j}\}}$ notation, the gradient can be expressed in matrix notation as

$$\hat{\nabla}_{\{b_{i,j}\}} E_\mathrm{D} = -2\lambda T_\mathrm{N} (A_\mathrm{M} - b) * HP(\hat{L}, \Phi)\,. \tag{6.77}$$

The gradient of the whole energy functional can consequently be calculated by:

$$\hat{\nabla}_{\{b_{i,j}\}} E_{\mathrm{D}} + E_{\mathrm{P}} = -2\lambda T_{\mathrm{N}}(A_{\mathrm{M}} - b) * HP(\hat{L}, \Phi) + 2\beta(b - \mu_b) \tag{6.78}$$

and the equation which has to be solved to find the DSNU pattern with the maximized likelihood is given as the point where the gradient becomes 0:

$$-2\lambda T_{\mathrm{N}}(A_{\mathrm{M}} - b) * HP(\hat{L}, \Phi) + 2\beta(b - \mu_b) = 0\,. \tag{6.79}$$

The linear equation can be solved by a conjugate gradient method or any other iterative solver that exploits the gradient. Meaningful start values like $b = A_{\mathrm{M}}$ or $b = \mu_b$ are used in the experimental section to achieve a faster convergence. Typically a steepest descent solver converges after 150 to 200 iterations for the problems at hand. For all evaluations the maximum limit was set to 250 iterations and results that did not converge to a stable solution were marked.

The derived discrete convolution is mathematically not valid for light sensors of finite extend. However, a mirrored extension of the image is used to obtain results for these light sensors as well. Summarizing, the two demands on the input data are given by a Gaussian distribution for the $\{b_{i,j}\}$ and by A_{M} as the average of *independent* frames, as discussed above.

6.4. First Experimental Results

The different variants in the choices of $\hat{L}$ and Φ to solve eq. 6.79, are investigated using the methods and experiments that have already been introduced in sec. 5.3. Tab. 6.2 summarizes the abbreviations for the variants of the newly introduced methods.

6.4.1. Parameter Search and Best-Case Performance

In the above introduced method, all parameters are obtained by evaluations of ground truth statistics, or ground truth corrected sample data. This gives a fully defined method, but as already mentioned in sec. 6.2.3 the exponential distribution and its parameter λ are founded on weak assumptions, that may not be fully

Method	Description
$\min(\nabla, BL)$	Solution of eq. 6.79 with HP as $HP(\nabla, \mathrm{BL})$
$\min(\nabla, MN)$	Solution of eq. 6.79 with HP as $HP(\nabla, \mathrm{MN})$
$\min(\triangle, MN)$	Solution of eq. 6.79 with HP as $HP(\triangle, \mathrm{MN})$

Table 6.2.: Introduction of used abbreviations for the newly developed methods. The used parameter sets are specified in the given evaluations.

Method	Parameter
$\min(\nabla, BL)$	$4.89 \cdot 10^{-05} \pm (4.61 \cdot 10^{-06}, 9.43\%)$
$\min(\nabla, MN)$	$7.48 \cdot 10^{-05} \pm (5.89 \cdot 10^{-06}, 7.88\%)$
$\min(\triangle, MN)$	$0.0131 \pm (0.0060, 46.0\%)$

Table 6.3.: Optimized set of the parameter λ for random frame sequences.

Method	Parameter
$\min(\nabla, BL)$	$9.67 \cdot 10^{-06} \pm (8.951 \cdot 10^{-06}, 92.5\%)$
$\min(\nabla, MN)$	$2.04 \cdot 10^{-05} \pm (1.319 \cdot 10^{-05}, 64.6\%)$
$\min(\triangle, MN)$	$0.000107 \pm (0.000347, 324\%)$

Table 6.4.: Optimized set of the parameter λ for consecutive frame sequences.

met under certain circumstances. λ can therefore be considered as the free parameter of the new method and will be tuned like the parameters of the reference methods.

In sec. 5.3.1 (p. 90), the parameter search for the reference methods was introduced. This search is now conducted for the λ parameter of the new methods as well. The other parameters of the method: β, μ and T_{N} are taken directly from the statistic ground truth measurements ín sec. 4.1.1.

Tab. 6.3 and 6.4 show the results of the parameter search for the 25 consecutive and the 25 random frame calibration sequences that are defined in sec. 4.2. Fig. 6.3 gives a visual comparison between the experimentally obtained optimal parameters to the parameters that were extracted from the statistics. Except for the $\triangle$-based method, the statistic parameters' values lie within the same order of magnitude, for both the random and consecutive frame sequences. For the $\min(\triangle, MN)$ method, this is only the case for the consecutive frame sequences, while for random frame sequences the optimized parameter is more than 2 magnitudes away from the statistically derived parameter.

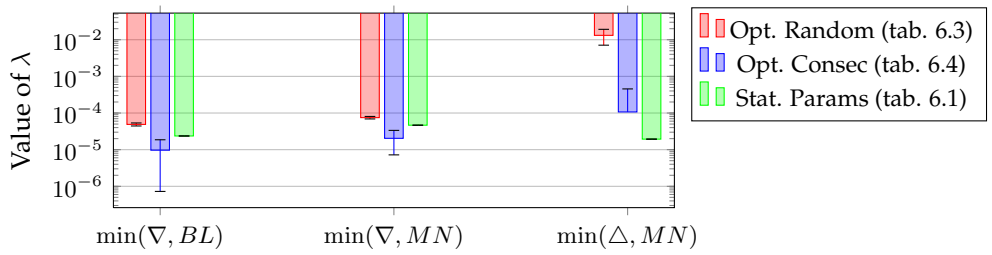


Figure 6.3.: Comparison of the values of the optimized and the statistically derived parameters.

To see the influence of the different parameter choices, the best-case performance analysis is conducted as explained in sec. 5.3.2 (p. 93). Fig. 6.4 shows the results of this evaluation in comparison to the two best reference methods $LMS_{A,7\times7}$ and

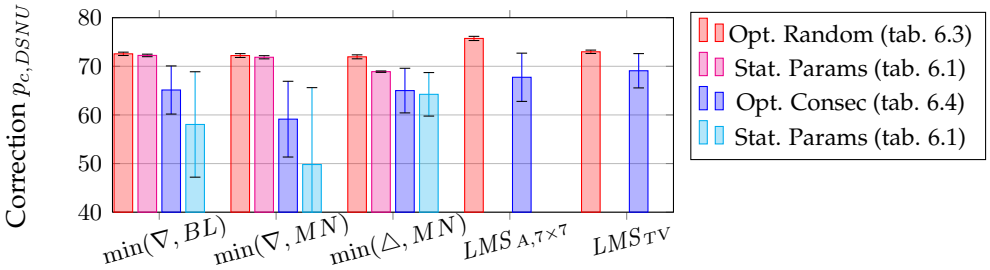


Figure 6.4.: Optimal Performance with respect to the $p_{\mathrm{c,DSNU}}$ quality metric

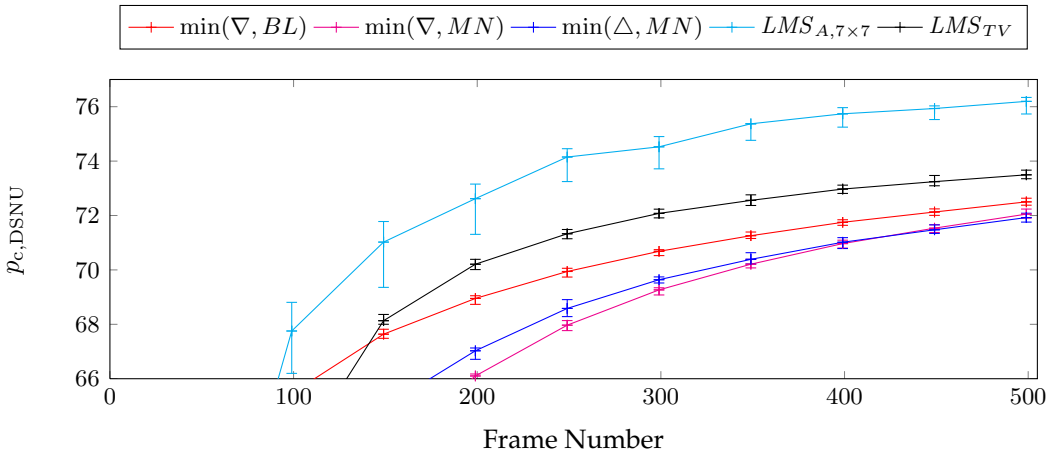
LMS_{TV} (see p. 96). The $\min(\nabla, BL)$ method reaches for the random frame sequences about about $72\,\%$ correction rate for both parameters, the optimized and the statistic. The same can be observed for the $\min(\nabla, MN)$ method, and only for the $\min(\triangle, MN)$ a decrease of about $4\,\%$ for the statistic parameters is observed.

With respect to the consecutive frame sequences, the reached correction performance cannot reach the performance of random frame input and the correction rate drops down to about $65\,\%$ for the optimized parameters. This is an expected result, as the new method explicitly demands for temporally independent feature input, and this demand is not fulfilled by consecutive frame sequences. Random frame sequences however are closer to fulfill this demand. When considering the statistic parameters, the performance drops even further and only the $\min(\triangle, MN)$ method maintains a constant performance with this parameter change. Therefore it achieves the best performance from the set of new methods with respect to the consecutive frame sequences.

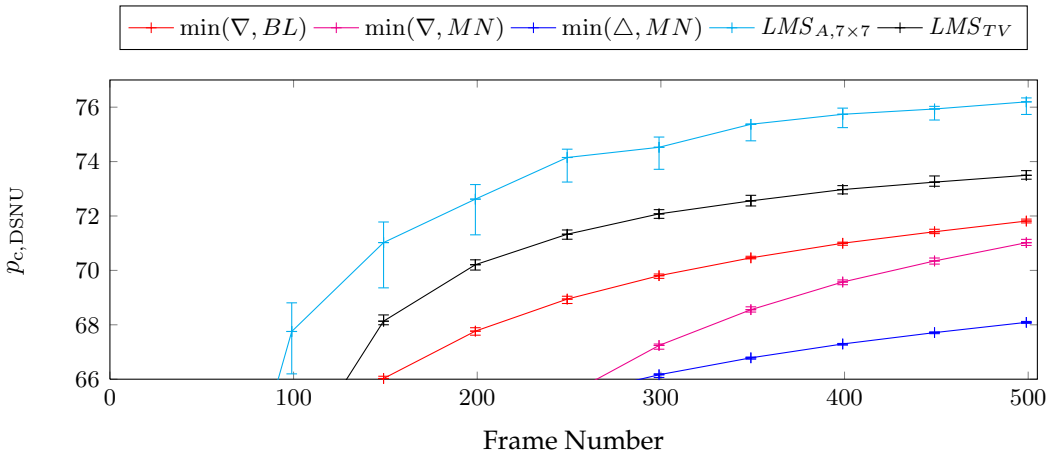
Compared to the $LMS_{7\times7}$ and $LMS_{A,7\times7}$ methods, the new methods in general do not reach these methods' performances. For the consecutive frame sequences the lack of performance can be explained as there exist a frame-to-frame dependency of the features. The random frame sequences however fulfill the demand for independent inputs as good as possible and thus further improvements will be necessary to improve the new method as described in chapter 7. A better comparison between the new methods and the literature methods will be obtained by the analysis of the temporal, spatial and spectral performances.

6.4.2. Temporal Performance Analysis

The temporal performance analysis of the new methods is conduced as described in sec. 5.3.4 (p. 98). As the new maximum likelihood method corrects for the DSNU only, this section shows the evaluations in the $p_{\mathrm{c,DSNU}}$ only. Sec. 5.3.4, already showed that this performance measure yields to the best comparability of the methods actual performance. For completeness, the other quality measures are shown in appendix D.3.



(a) Parameter set: Opt. Random (tab. 6.3)



(b) Parameter set: Stat. Params (tab. 6.1)

Figure 6.5.: Temporal evaluations with the optimized (fig. 6.5a) and the statistic (fig. 6.5b) parameter sets in the $p_{\mathrm{c,DSNU}}$ metric for random frame sequences.

Fig. 6.5 shows the results for the random frame sequences with the experimentally optimized and statistically obtained parameters, compared to the two best reference methods $LMS_{A,7\times7}$ and LMS_{TV}. Considering the optimized parameters in fig. 6.5a the performance of the new methods almost reaches the reference methods, but in the end both reference methods stay ahead by a few percent. Among the newly proposed methods, the $\min(\nabla, BL)$ method provides the best performance and outperforms the reference methods until frame 150. The $\min(\nabla, BL)$ method is followed by the $\min(\triangle, MN)$ method and last in this comparison is the $\min(\nabla, MN)$ method.

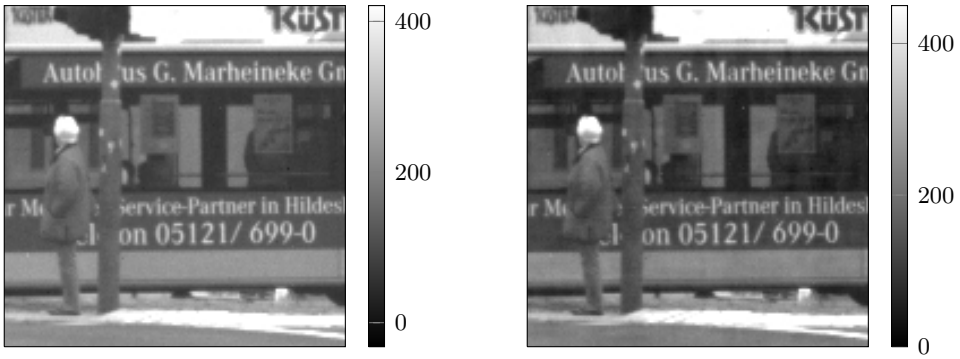
The above experiment was repeated, using the statistically obtained λ values and the results are shown in fig. 6.5b. As expected from the optimal performance analysis in fig. 6.4 the results are almost identical, except for the $\min(\triangle, MN)$ method, which drops significantly in its reached performance to $\approx 68\,\%$.

6.4.3. Spatial and Spectral Performance Analysis

The analysis of the new methods continues with the spatial performance analysis for the DSNU patterns, which has been introduced in sec. 5.3.5 (p. 102). Fig. 6.6 shows the ROI of the corrected images with almost no visual difference between the methods' corrections and the ground truth corrected image. The $LMS_{A,7\times7}$ and LMS_{TV} methods are shown as references.

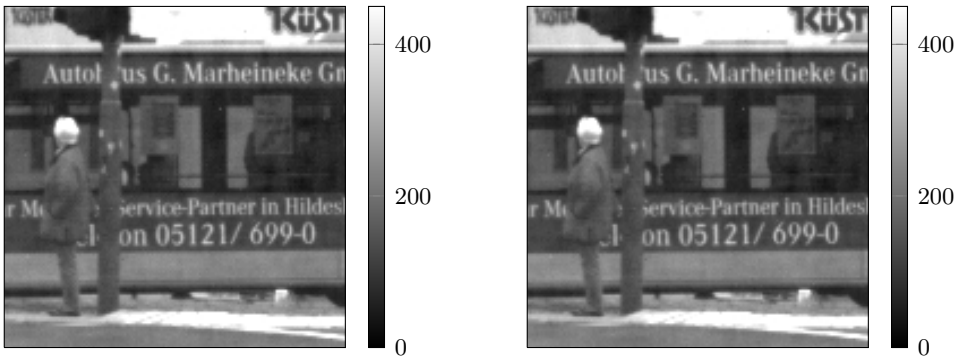
The analysis of the remaining DSNU patterns allows to investigate the methods' difference in more detail as shown in fig. 6.7. The new methods differ not much in a first visual inspection, which goes in accordance to the fact that the methods reached almost the same correction rates. Only the $\min(\nabla, MN)$ method seems to suffer from a superimposed very high spatial frequency grid, which is explained below. For the reference methods, the $LMS_{A,7\times7}$ method seems to have a lower amplitude but higher spatial frequencies in comparison to the new methods, while the LMS_{TV} method provides a similar pattern as the new methods.

The claimed spatial frequency properties in the above discussion can be better investigated with the spectral damping analysis as shown in fig. 6.8, where the actual differences between the methods become visible. Comparing the methods among themselves, the horizontal and vertical frequency components are damped less and the methods differ mainly in their ability to damp the low frequency components in the center of the spectral presentation. The $\min(\triangle, MN)$ method provides much less correction in the low frequencies compared to the other methods and for the $\min(\nabla, MN)$ method remaining high frequency components show up in the corners of the representation. This is a critical point as these frequencies deteriorate the visual image quality the most. Fig. 6.2b (p. 123) showed already the corresponding high pass filter characteristics and it is obvious that the high frequencies that are placed in the corners of the representation are allowed to pass the HP-filter of this method. As a consequence the method does not react on these frequencies, which results in the observed high frequency pattern that became



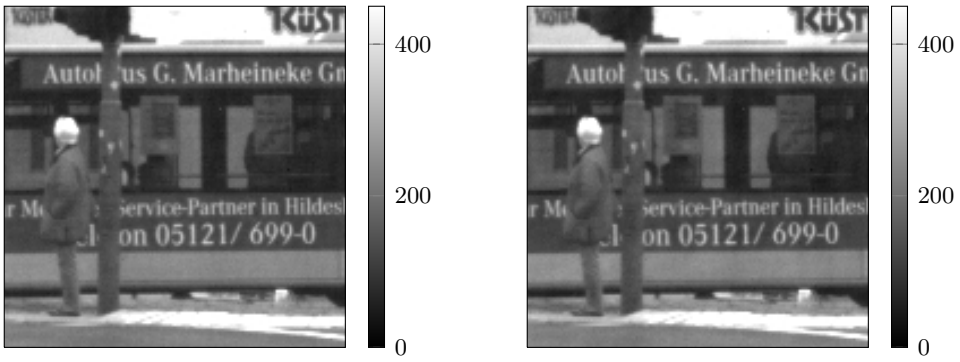
(a) Ground truth corrected

(b) Corrected with $\min(\nabla, BL)$



(c) Corrected with $\min(\nabla, MN)$

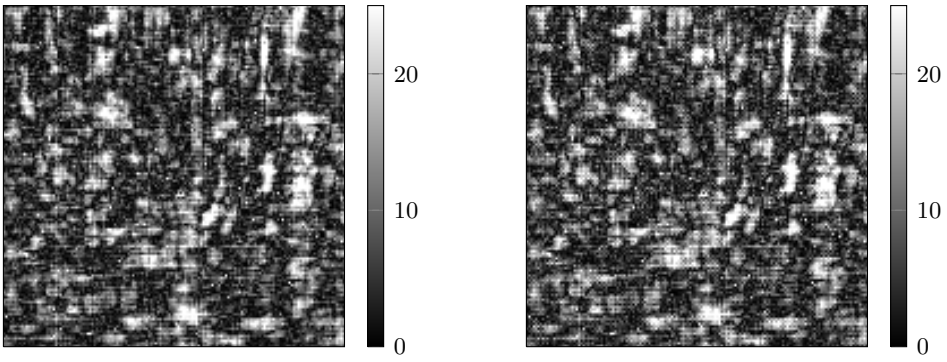
(d) Corrected with $\min(\triangle, MN)$



(e) Corrected with $LMS_{A,7\times7}$ Method

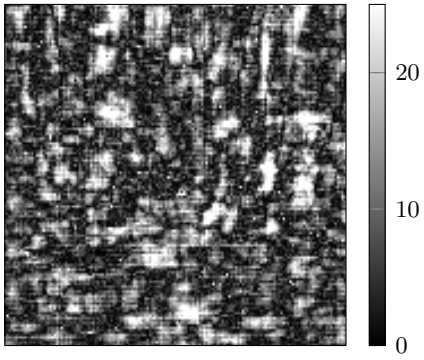
(f) Corrected with LMS_{TV} Method

Figure 6.6.: Spatial corrected images with the introduced methods and the ground truth corrected image for random frame sequences.

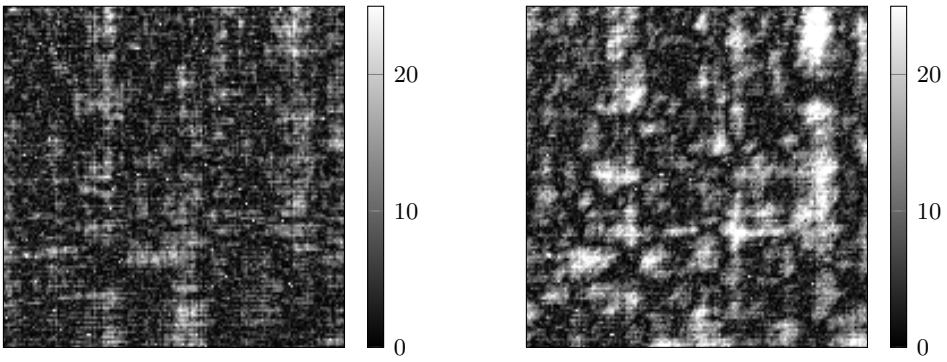


(a) $|\{b_{i,j,rem.}\}|$ for $\min(\nabla, BL)$

(b) $|\{b_{i,j,rem.}\}|$ for $\min(\nabla, MN)$



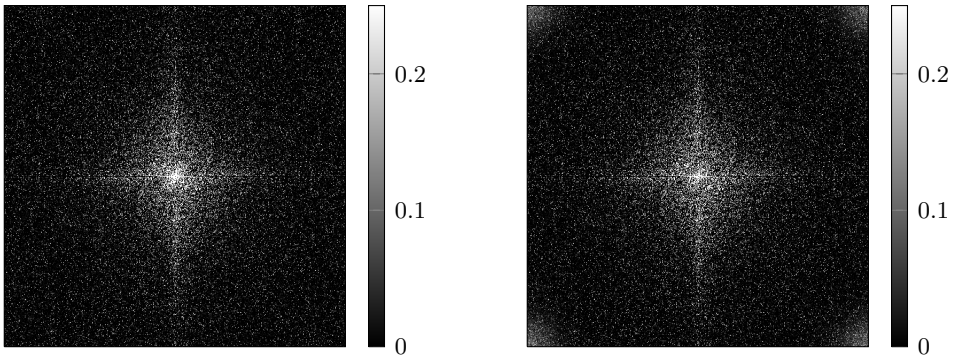
(c) $|\{b_{i,j,rem.}\}|$ for $\min(\triangle, MN)$



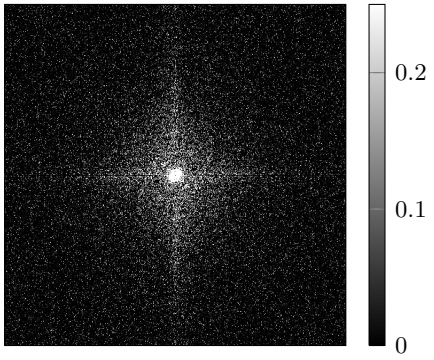
(d) $|\{b_{i,j,\mathrm{rem.}}\}|$ for $LMS_{A,7\times7}$

(e) $|\{b_{i,j,\mathrm{rem.}}\}|$ for LMS_{TV}

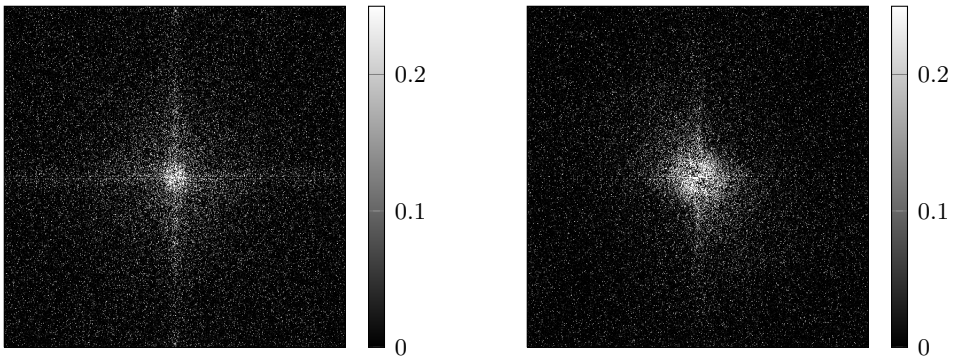
Figure 6.7.: The remaining DSNU after the correction with the introduced methods for random frame sequences.



(a) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $\min(\nabla, BL)$ (b) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $\min(\nabla, MN)$



(c) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $\min(\triangle, MN)$



(d) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $LMS_{A,7\times7}$ (e) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for LMS_{TV}

Figure 6.8.: Spectral damping factors for the DSNU correction of the introduced methods for random frame sequences.

visible in fig. 6.7. Finally, the $LMS_{A,7\times7}$ method wins the spectral damping comparison as it provides the highest damping of the central (low) frequencies, while the advantage of the LMS_{TV} method is a damping of more vertical and horizontal frequencies. This also results in a visually more pleasing appearance.

6.5. First conclusions

This chapter derived a new maximum likelihood DSNU estimation method, which is based on trivial and well-motivated assumptions. Three variants of the method have been developed and proven to function, but in a performance comparison these methods did not yet outperform the two best methods known in literature. However the performance of the new methods reaches very close to them $4\,\%$ to $6\,\%$. The main advantage of the new methods is that they reach a comparable performance with statistically derived parameters. This makes a parameter search on calibration sequences obsolete as only the statistics of the scenes need to be known. The next chapters will improve this new and promising approach in a way that it outperforms the reference methods.

7. Analysis and Improvements of the new DSNU Method

The introduced maximum likelihood DSNU estimation methods from chapter 6 are able to correct successfully for DSNU and have the advantage of motivated assumptions and defined parameters. This gives them already a certain advantage over the reference methods from chapter 5. However, the performance of the new methods does not reach up to the performance of the reference methods. This chapter identifies ways to improve the new methods in a similar way in which the $LMS_{3\times3}$ and $LMS_{7\times7}$ methods have been improved into their adaptive versions $LMS_{A,3\times3}$ and $LMS_{A,7\times7}$.

Furthermore, this chapter will introduce a modified solver for the resulting minimization that is based on cellular neural networks. Variants of the methods that consider an improved sensor model that considers the row and column wise nonuniformities are derived as well. Summarized, this chapter develops the new methods in a way that tailored solutions for common real time applications are possible and the contents of this chapter have already been partly published [11, 12].

7.1. Improvements to the Averaging Task

As a first step, the performance of the new method is increased by analyzing its input. This analysis is similar to the improvements of the LMS methods towards the adaptive LMS methods by Vera [4], who as well analyzed the input to the LMS methods.

Chapter 6 summarized the new method in a mathematical way as the task of solving eq. 6.79:

$$-2\lambda T_N(A_M - b) * HP(\hat{L}, \Phi) + 2\beta(b - \mu_b) = 0$$

Besides the methods' parameters $\lambda, T_N, \beta, \mu_b$ and $HP(\hat{L}, \Phi)$, the method also depends on the averaged sensor measurements A_M. While the parameters are fixed when the method is running, the averaging of the sensor input seems to bear a big potential to improve the method. For example a weighted sensor average would allow to preselect the sensor information and cut out disturbing elements.

As mentioned, this is similar to the improvements that lead to the adaptive least mean square methods, where the sensor input is weighted according to the spatial standard deviation (see sec. 5.2). For the new maximum likelihood approach such a weighting technique translates into a weighted average of the sensor measurements:

$$A_{\mathrm{M},i,j} = \frac{1}{T_N} \sum_{t_n} C_{\mathrm{M},i,j} \rightarrow A_{\mathrm{M}_W,i,j} = \frac{1}{\sum\limits_{t_n} w_{i,j,t_n}} \sum_{t_n} w_{i,j,t_n} C_{\mathrm{M},i,j} \tag{7.1}$$

The $A_{\mathrm{M}_W,i,j}$ stand for the weighted average of the sensor measurements and the w_{i,j,t_n} give the weight of a single light sensor's measurement. Next, an appropriate weighting rule needs to be defined.

7.1.1. Realizations of the Weighed Average

Favor of small spatial variances: A trivial approach to define a weighting rule would be to use the same weights as for the adaptive LMS method in sec. 5.32 p. 87. This gives the weights as:

$$w_{i,j,t_n} = \frac{1}{1 + s(C_{\mathrm{M},N(i,j)}(t_n))} \tag{7.2}$$

where $s(C_{\mathrm{M},N(i,j)}(t_n))$ stands for the spatial standard deviation of the measurements in a defined neighborhood N around the current position (i, j) measured at time t_n (see p. 87).

This approach sounds feasible as one may think that the lowest spatial deviation will occur for patches with uniform illumination. However, an implementation of this approach yields to performances of $< 20\%$ given the $p_{\mathrm{c,DSNU}}$-norm, which is an unsatisfying result. The reason for this poor performance seems to be a wrong selection of the sensor inputs by these weights. For a frame-to-frame based update rule only the effects of one frame are considered, and the relative DSNU effects are considered for the light sensors and the update steps. This is however not the case if the averaged weighted information is considered, as the difference between the measurements of two averaged, neighboring light sensors then depends on a combination of scene intensity and scene texture. As a consequence other differences occur than the actual DSNU would have caused, which does not allow to extract the correct DSNU. For this reason another weighting rule needs to be investigated.

Favor of Dark Regions: Considering the physical definitions of the EMVA1288 standard [1], the DSNU measurements have to take place at darkness. Of course low spatial standard deviations occur at darkness as well, which makes the adaptive LMS weights still a considerable choice, but with the above mentoined disadvantages. A better indicator for darkness is a low value of the measurements

C_{M}, which are bounded to a minimum of 0 by definition. As the C_{M} values are connected by the *HP* filter mask in the maximum likelihood algorithm, the weight should consider the average illumination in a neighborhood that corresponds to the size of the *HP* mask. The squared average of the measurements is used, to increase the importance of dark areas and the weight is defined as:

$$w_{i,j,t_n} = \frac{1}{1 + e(C^2_{\mathrm{M},N(i,j)}(t_n))} \,. \tag{7.3}$$

The $+1$ in the nominator avoids a division by zero in case of black images due to unexpected events (e.g. a temporary image sensor failure, as the DSNU usually provides values larger than 0 for the measurements). The $e(C^2_{\mathrm{M},N(i,j)}(t_n))$ represent the spatial expectation value of C^2_M in the neighborhood N around the sensor position (i, j), measured at the time t_n.

The above choice leads to an increased correction performance that outperforms the reference methods in a statistically significant way. The results of this choice will be discussed in the next sections and the improved method can be summarized as a solution to the equation:

$$0 = -2\lambda T_N(A_{\mathrm{M}_W} - b) * HP(\hat{L}, \Phi) + 2\beta(b - \mu_b) \tag{7.4}$$

$$\text{with:} \quad A_{\mathrm{M}_W,i,j} = \frac{1}{\sum\limits_{t_n} w_{i,j,t_n}} \sum_{t_n} w_{i,j,t_n} C_{\mathrm{M},i,j} \tag{7.5}$$

$$\text{and with:} \quad w_{i,j,t_n} = \frac{1}{1 + e(C^2_{\mathrm{M},N(i,j)}(t_n))} \tag{7.6}$$

7.1.2. Weighted Statistic Based Parameters

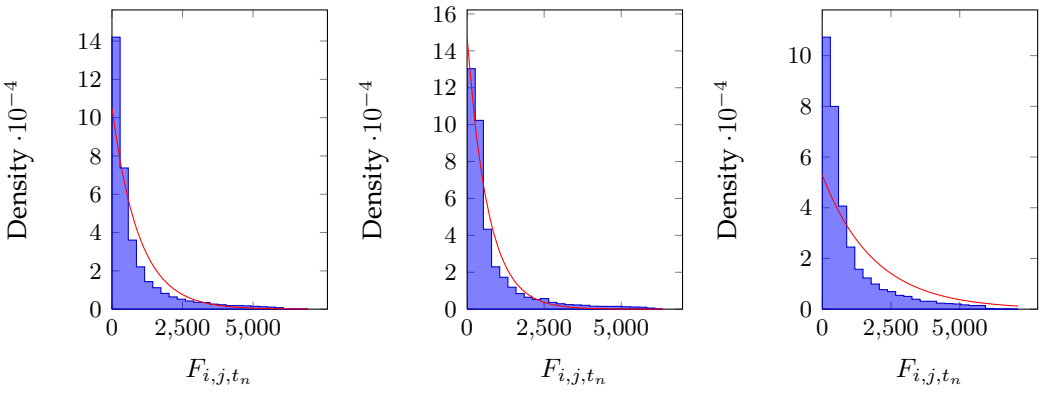
As described in sec. 6.2.3, the free parameter λ of the method is based on an exponential probability density distribution of the feature choices F_{i,j,t_n} (defined in eq. 6.45, p. 120). By evaluation of the ground truth histograms, the probability density distribution of F_{i,j,t_n} was estimated and the parameter λ could be defined from the statistics of the typical scene contents.

This method of finding the theoretically correct λ parameter changes with the newly defined weighted average. The change comes from the fact that the influence of a given feature will now contribute by its weight w_{i,j,t_n} to the temporal average of the sensor. The probability density distribution of the features F_{i,j,t_n} does however not yet consider the influence of the newly introduced weighting. A modified histogram creation, that considers the probability of an appearing feature value F_{i,j,t_n} with respect to its weight needs to be defined.

While for a usual histogram the bin count is increased by 1 if a feature value lies inside the bin borders, the bin count has to be increased by the value of the weight

instead by 1 to consider the weights. After that the overall probability is normalized according to the bin widths. To obtain the data for such a *weighted* histogram, the weight has to be sampled additionally for each feature sample.

The weighted histograms are calculated over the 25 random frame calibration sequences (see sec. 4.2) by selecting 10^6 random samples out of all the available images. Each frame is again corrected with the laboratory ground truth nonuniformity patterns before the feature measurement and fig. 7.1 shows the weighted histograms that result from the different choices of Φ and $\hat{L}$. From the visual inspection, an approximation of the probability density distribution as an exponential distribution is possible again and the maximum likelihood parameter estimation out of the collected data leads to the parameter set in tab. 7.1.



(a) F_{i,j,t_n} with $HP(\nabla, BL)$ (b) F_{i,j,t_n} with $HP(\nabla, MN)$ (c) F_{i,j,t_n} with $HP(\triangle, MN)$

Figure 7.1.: Weighted histograms of the different F_{i,j,t_n} realizations, including the maximum likelihood fit of the exponential distribution function.

Feature Choice	Resulting Parameter for weighted Averages
F_{i,j,t_n} with $HP(\nabla, BL)$	$\lambda = 0.001$
F_{i,j,t_n} with $HP(\nabla, MN)$	$\lambda = 0.0015$
F_{i,j,t_n} with $HP(\triangle, MN)$	$\lambda = 5.30 \cdot 10^{-4}$

Table 7.1.: Maximum likelihood estimated parameters for the assumed exponential distribution of the feature value distribution of F_{i,j,t_n}, given the weighted averages.

7.1.3. Experimental Results

Parameter Search and Best-Case Performance

As in all evaluations of this thesis, an optimized set of parameters is needed to guarantee a fair comparison between the methods. So far 4 sets of the parameter λ

have been extracted: Experimentally optimized parameters for random (tab. 5.2) and consecutive frames sequences (tab. 5.3), statistically obtained parameters (tab. 6.1) and newly introduced, the statistically obtained parameters that consider the influence of the weighted average (tab. 7.1)

Method	Description
$\min_{\mathrm{wA}}(\nabla, BL)$	Weighted Average (eq. 7.4 with HP as $HP(\nabla, \mathrm{BL})$)
$\min_{\mathrm{wA}}(\nabla, MN)$	Weighted Average (eq. 7.4 with HP as $HP(\nabla, \mathrm{MN})$)
$\min_{\mathrm{wA}}(\triangle, MN)$	Weighted Average (eq. 7.4 with HP as $HP(\triangle, \mathrm{MN})$)

Table 7.2.: Introduction of used abbreviations for the improved methods with weighted averaging

Method	Parameter
$\min_{\mathrm{wA}}(\nabla, BL)$	$0.0001751 \pm (4.067 \cdot 10^{-5}, 23.22\%)$
$\min_{\mathrm{wA}}(\nabla, MN)$	$0.0002039 \pm (2.63 \cdot 10^{-5}, 12.9\%)$
$\min_{\mathrm{wA}}(\triangle, MN)$	$0.08602 \pm (0.01315, 15.29\%)$

Table 7.3.: Optimized parameter set for the random frame sequences with the weighted average method

Method	Parameter
$\min_{\mathrm{wA}}(\nabla, BL)$	$2.259 \cdot 10^{-5} \pm (4.588 \cdot 10^{-5}, 203.1\%)$
$\min_{\mathrm{wA}}(\nabla, MN)$	$3.548 \cdot 10^{-5} \pm (4.605 \cdot 10^{-5}, 129.8\%)$
$\min_{\mathrm{wA}}(\triangle, MN)$	$0.007369 \pm (0.02541, 344.8\%)$

Table 7.4.: Optimized parameter set for random and consecutive sequences with the weighted average

A further parameter search for the method with the improved weighted average has been conducted for the random and consecutive frame sequences as well. These experimentally optimized parameters are shown in tab. 7.3 and tab. 7.4, while the abbreviations of the improved methods are introduced in tab. 7.2. The optimized parameters differ by more than one magnitude from the statistically obtained parameter set in tab. 7.1, and especially for the consecutive frame sequences, a range of several $100\,\%$ as standard deviation occurred.

To judge the performance of all the different parameter sets, the best-case performance analysis has been plotted. Due to the increased number of parameter sets, this analyis is plottet separately for the random frame sequences (fig. 7.2) and for the consecutive frame sequences (fig. 7.3). For each of the evaluations the performances of the optimized parameter sets of the $LMS_{A,7\times7}$ and LMS_{TV} is displayed for comparison.

Optimal performance analysis for random frame sequences: For the random frame sequences shown in fig. 7.2 all three parameter choices: optimized,

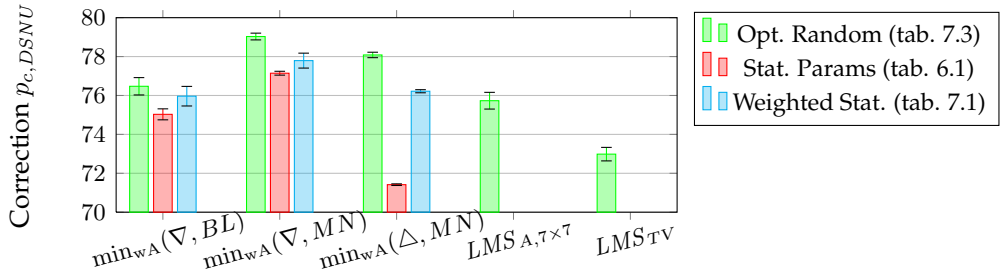


Figure 7.2.: Optimal performance analysis for the random frame sequences and the weighted average improvement.

statistic and weighted statistic, lead to very good correction performances. If the optimized parameter set is used, the reference methods are outperformed by each of the newly improved methods. The $\min_{wA}(\nabla, MN)$ method reaches a 3 % higher performance than the $LMS_{A,7\times7}$ which as well is statistically significant as the error bars are not overlapping. Only the $\min_{wA}(\nabla, BL)$ method shows a slight overlap in the error bars with $LMS_{A,7\times7}$ but the $\min_{wA}(\nabla, BL)$ method is based on a 3×3 filter mask while the $LMS_{A,7\times7}$ is based on a 7×7 mask. The LMS_{TV} method which depends as well on a 3×3 mask is outperformed significantly by all other methods.

Considering the statistic parameter set from tab. 6.1, the $\min_{wA}(\nabla, BL)$ and $\min_{wA}(\nabla, MN)$ methods still give good performance, while the $\min_{wA}(\triangle, MN)$ method decreases down to below 72 %, however this is still a slight increase over its performance without the improvement of the weighted average (see fig. 6.4 on p. 129). Last, for the weighted statistic parameters, which are the correctly deduced parameters with concern to the data statistics, all new methods still outperform the reference methods. Thus the improved maximum likelihood approach fulfills now all the demands that were discussed in the beginning of chapter 6.

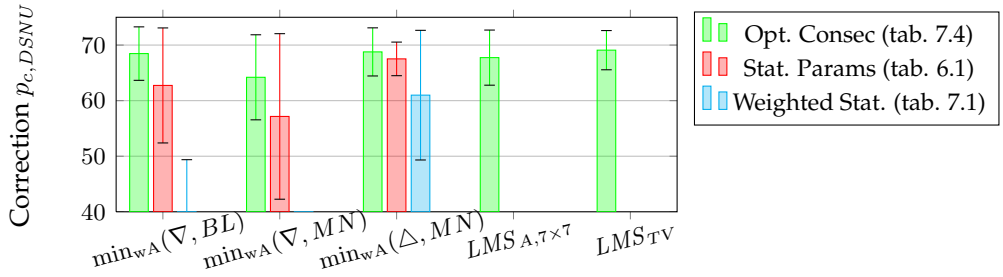


Figure 7.3.: Optimal performance analysis for the consecutive frame sequences and the weighted average improvement.

Optimal performance analysis for consecutive frame sequences: For the consecutive frame sequences the best-case results are shown in fig. 7.3 and lead to different performances for the different parameter choices. Considering the optimized parameter set, the new maximum likelihood methds and the reference methods show a significant overlap in their error bars. The general scale of the correction performance has dropped for all methods down to less than $70\,\%$. This is the general result if depended information in the input frames is considered. As all error bars overlap, a significantly best method cannot be named, but from the expectation values the LMS_{TV} method and the $\min_{\mathrm{wA}}(\triangle, MN)$ method provide a tendency to give better results. The $\min_{\mathrm{wA}}(\nabla, MN)$ method seems to have a performance drop that is larger than the one of the other new methods. For the statistical parameter set from tab. 6.1, the ∇-based methods obtain an averaged performance loss of $7\,\%$ and an increasing width of the error span is observed. The $\min_{\mathrm{wA}}(\triangle, MN)$ method gives however a stable result under the parameter change an the error bars actually shrink. Such behavior points towards instabilities for the optimized parameters and suggests to favor the statistical parameters (A further discussion on this topic is found in sec. 10.1).

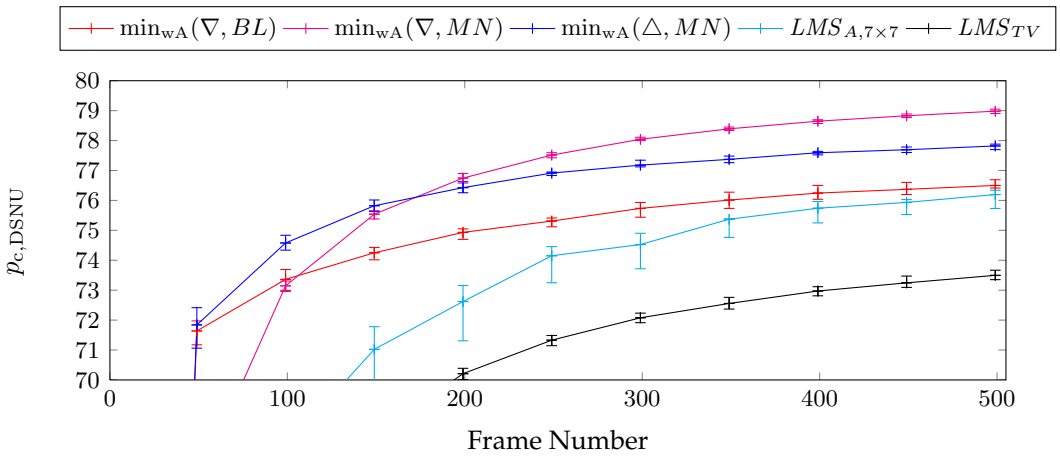
For the weighted statistic parameters the performance drop becomes even more severe and the $\min_{\mathrm{wA}}(\nabla, BL)$ and $\min_{\mathrm{wA}}(\nabla, MN)$ methods cannot reach a good performance with this parameter set. Only the $\min_{\mathrm{wA}}(\triangle, MN)$ method shows a performance that still shows up in the chosen plotting ranges. The statistical parameter set from tab. 6.1 should therefore be used for consecutive frame sequences. From a theoretical point of view the statistic parameter set can only be valid if the weighting in the averaging process can be neglected. And indeed the weighting effects should be neglected if the frames depend on each other, because if the weighted average is used for consecutive frames then similar regions in the image give similar weights for many frames. This forms a bias in the frame-to-frame dependencies, that increases the dependencies in the features' random variables. Although the demands for independent features are obviously void in the consecutive frame sequences, the $\min_{\mathrm{wA}}(\triangle, MN)$ method still gives reasonable results, while the ∇-based methods become very sensitive to parameter changes. This attests the $\min_{\mathrm{wA}}(\triangle, MN)$ method the highest reliability of the compared methods.

As the random frame sequences reached the best performances on the absolute correction rates, the further comparison is done for the random frame sequences first. The analysis for consecutive frame sequences and the image material with low valued nonuniformities is discussed in sec. 10.1 and sec. 10.2.

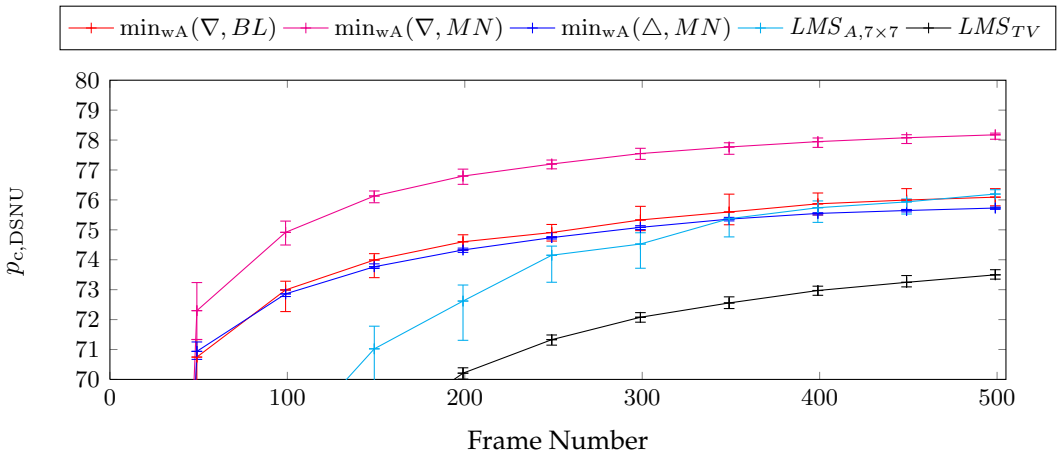
Temporal Performance Analysis

The temporal performance analysis of the improved methods is done as introduced in sec. 5.3.4, p. 98. The improvement with the weighted average of the maximum likelihood method corrects still for the DSNU only and therefore this

section again shows only the evaluations of the percentual DSNU correction performance $p_{\mathrm{c,DSNU}}$. As discussed already, this measure allows the best comparability of the methods' performances (see sec. 5.3.4). The other quality measures are shown in appendix E.1 for completeness.



(a) Opt. Random (tab. 7.3)



(b) Weighted Stat. (tab. 7.1)

Figure 7.4.: Evaluations for random frame sequences with experimentally optimized parameters (fig. 7.4a) and the parameter set from the weighted statistic histogram (fig. 7.4b) in the $p_{\mathrm{c,DSNU}}$ metric.

Fig. 7.4a shows the results of the random frame sequences with the experimentally optimized parameters in comparison to the two best reference methods $LMS_{A,7\times7}$ and LMS_{TV}. The performance of the new methods reaches already after 100 frames the mark of 72 % correction performance. The $LMS_{A,7\times7}$ method reaches this performance not until the 200 frame mark and the LMS_{TV} method not until the 300 frame mark. The new methods also keep their performance advantage until the end of the sequence. The gap between the $\min_{\mathrm{wA}}(\nabla, MN)$ method and the $LMS_{A,7\times7}$ method stays almost constant while the $\min_{\mathrm{wA}}(\nabla, BL)$ method stops

increasing its correction performance and meets up with the $LMS_{A,7\times7}$ method after 500 frames. The $\min_{\mathrm{wA}}(\triangle, MN)$ method reaches very early a high correction level ($75\,\%$ after 150 frames) and converges only very slowly from that point on, resulting in a final performance between the $\min_{\mathrm{wA}}(\nabla, BL)$ and the $\min_{\mathrm{wA}}(\nabla, MN)$ methods.

The experiment was repeated, using the new statistically obtained λ values that respect the influence of the weights (tab. 7.1). The results are shown in fig. 7.4b where the new methods perform not as good as for the optimized parameter set, which is the expected result. However, the $\min_{\mathrm{wA}}(\nabla, MN)$ method performs still better than the reference methods and the $\min_{\mathrm{wA}}(\nabla, BL)$ and $\min_{\mathrm{wA}}(\triangle, MN)$ methods give identical results as the $LMS_{A,7\times7}$ method. The LMS_{TV} method gets outperformed by all new methods and the new methods in general reach their high correction levels faster than the reference methods their optimized parameters.

Concluding, the improvement with weighted averages allows to obtain the parameter sets from the statistics of the measurements. The new methods can also be expected to converge much faster than the reference methods and just the one demand for independent input frames has to be met. This demand however leads to a generalized improved performance for all methods. Furthermore, the computational costly and theoretically unjustified parameter search with heuristic results can be substituted by a derivation of the methods' parameters by defined statistical properties of the typical input data. The analysis for consecutive frame sequences and the image material with low valued nonuniformities continues in sec. 10.1 and sec. 10.2.

Spatial and Spectral Performance Analysis

The analysis of the new methods continues with the spatial performance analysis for the DSNU patterns, which has been introduced in sec. 5.3.5 (p. 102). As reference the $LMS_{A,7\times7}$ and LMS_{TV} methods are shown as well and the optimized parameter set is used (Opt. Random (tab. 7.3)). Fig. 7.5 shows the ROI of the corrected images with almost no visual difference between the method corrected images to the ground truth corrected images.

The analysis of the remaining DSNU patterns allows to investigate the methods in detail as shown in fig. 7.6, where the remaining patterns differ significantly from each other. The $\min_{\mathrm{wA}}(\nabla, BL)$ pattern shows low magnitude but also a lot of spatial high frequencies in the horizontal and vertical direction. The $\min_{\mathrm{wA}}(\nabla, MN)$ method, which gave the best performance, does not seem to leave the before uncorrected high frequency components and thus shows an improved performance over the $\min_{\mathrm{wA}}(\nabla, BL)$ method. Finally the $\min_{\mathrm{wA}}(\triangle, MN)$ method has only very few artifacts left but therefore shows a higher amplitude in its DSNU remains.

(a) Ground Truth Corrected

(b) Corrected with $\min_{\mathrm{wA}}(\nabla, BL)$

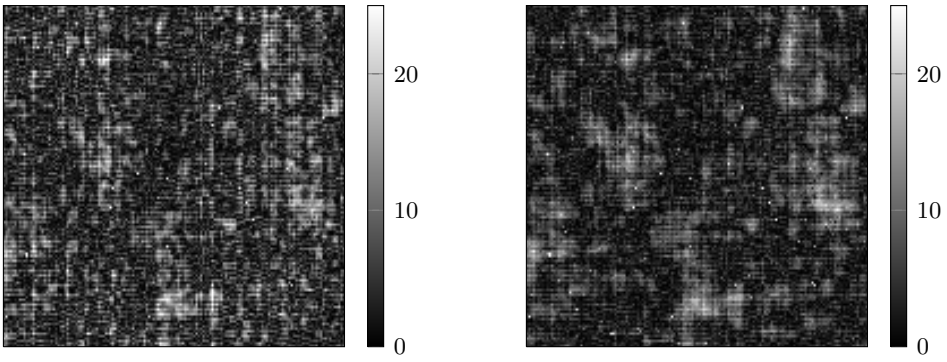
(c) Corrected with $\min_{\mathrm{wA}}(\nabla, MN)$

(d) Corrected with $\min_{\mathrm{wA}}(\triangle, MN)$

(e) Corrected with $LMS_{A,7\times7}$ Method

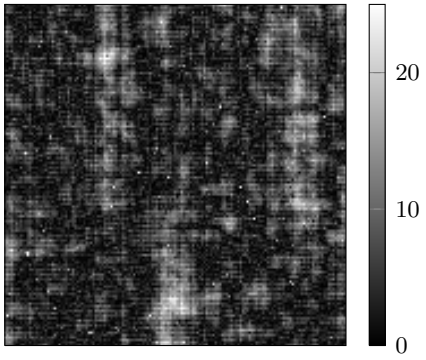
(f) Corrected with LMS_{TV} Method

Figure 7.5.: Corrected images with the ground truth correction patterns and the estimate of the different methods for random frame sequences (Parameters according to tab. 7.3).

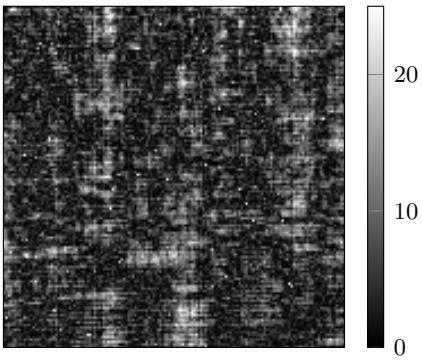


(a) $|\{b_{i,j,rem.}\}|$ for $\min_{\mathrm{wA}}(\nabla, BL)$

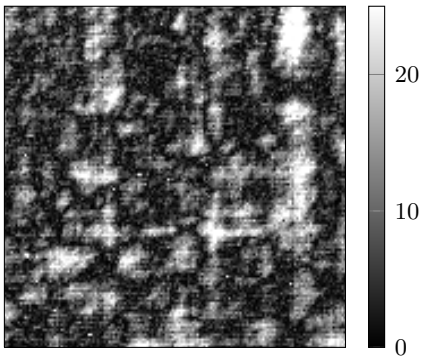
(b) $|\{b_{i,j,rem.}\}|$ for $\min_{\mathrm{wA}}(\nabla, MN)$



(c) $|\{b_{i,j,rem.}\}|$ for $\min_{\mathrm{wA}}(\triangle, MN)$



(d) $|\{b_{i,j,\mathrm{rem.}}\}|$ for $LMS_{A,7\times7}$



(e) $|\{b_{i,j,\mathrm{rem.}}\}|$ for LMS_{TV}

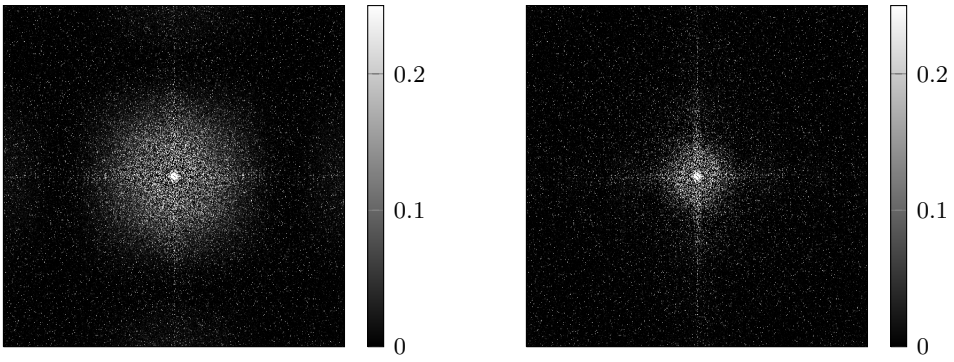
Figure 7.6.: Remaining DSNU patterns for the introduced methods and random frame sequences (Parameters according to tab. 7.3).

Compared to the new methods' without the weighted average on p. 133 in fig. 6.7, the improvements are clearly visible. This is also true for a comparison towards the reference methods, as the $LMS_{A,7\times7}$ method leaves the discussed high frequency artifacts behind, which are similar to the quality of the $\min(\nabla, BL)$ method. The LMS_{TV} method shows a rather high amplitude of the remains, which stick out in this comparison.

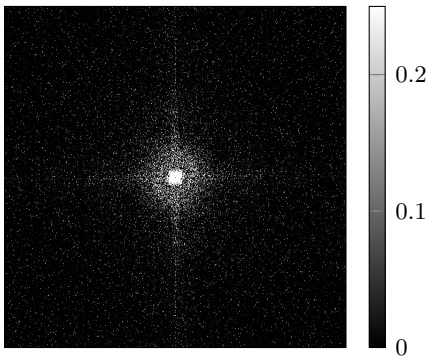
The already noticed spatial frequency properties can of course be investigated much better by the spectral analysis shown in fig. 7.7. In general the new methods correct a bit less of low frequencies than the reference methods, which is visible by the brighter dot in the middle of their spectrograms. However, the spectral remains of the $\min_{\mathrm{wA}}(\nabla, BL)$ method are much more concentrated at spatial lower frequencies which is not the case for the $LMS_{A,7\times7}$ method. Anyway, both methods result in almost the same correction performance. The $\min_{\mathrm{wA}}(\nabla, MN)$ method shows up with severely less remains in the spatial high frequencies when compared to the $\min_{\mathrm{wA}}(\nabla, BL)$ method, which is the reason for its superior performance. Only some horizontal and vertical frequencies are not well corrected.

The $\min_{\mathrm{wA}}(\triangle, MN)$ method has the same properties as the $\min_{\mathrm{wA}}(\nabla, MN)$ method, but leaves more of the low frequencies behind, which is indicated by the larger area of the white dot in the middle of the spectrogram. Comparing these representations, it is obvious that the $\min_{\mathrm{wA}}(\nabla, MN)$ method performs best. The LMS_{TV} method leaves more of the low frequencies behind and comes last in this comparison. The two reference methods still leave more random high frequency components behind, which are much better corrected by the new methods. This can be easily observed by the *darker* areas towards the boarders of the images.

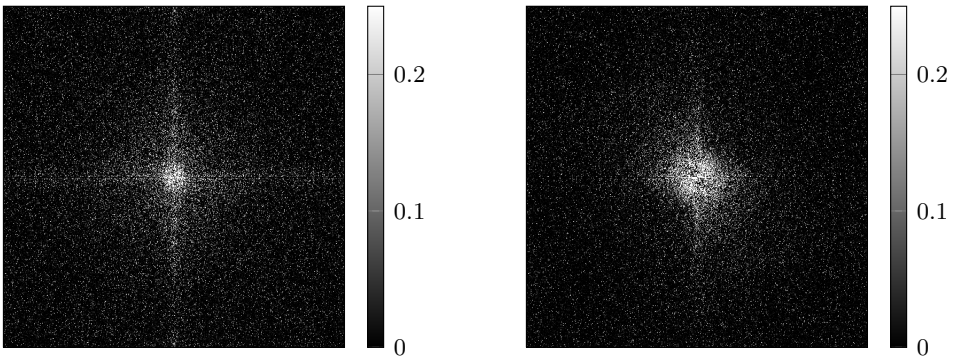
Summarizing, all methods, including the reference methods have some problems to correct horizontal and vertical frequencies to the same extend as the other frequency components. The $LMS_{A,7\times7}$ and LMS_{TV} methods show less of the lower frequencies uncorrected, but leave behind spurious high frequencies while the new methods solve this trade-off by correcting less of the low frequency components, but cut out the high frequency components much better. Consequently, the new methods behave as designed and cut away the more important high frequency components first.



(a) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $\min_{\text{wA}}(\nabla, BL)$ (b) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $\min_{\text{wA}}(\nabla, MN)$



(c) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $\min_{\text{wA}}(\triangle, MN)$



(d) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $LMS_{A,7\times7}$ (e) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for LMS_{TV}

Figure 7.7.: Spectral damping factors for the introduced methods and random frame sequences (Parameters according to tab. 7.3).

7.1.4. First Conclusions

The improved maximum likelihood approach with the weighed averaged input allows to outperform the reference methods from literature in case of the demanded random frame inputs. The results showed that a very good performance is even possible without an extended parameter search using parameters that can be extracted from the statistics of the recorded data instead. The main assumption of the method is that the input frames should be independent from each other. This assumption can be violated by using short consecutive frame high speed sequences, but even in these cases the best-case performance of the methods still overlap in their error bars.

Besides the good results, there are two main disadvantages left that I want to overcome in the following sections: First, the new methods need a computational expensive interative solver, which can be substituted by transforming the problem into a cellular neural network as explained in the next section (sec. 7.2). Second, the performance in the critical horizontal and vertical frequencies is yet not good enough and will be improved in sec. 7.3.

7.2. Approximation by Transfer into a CNN

The new DSNU correction methods have been proven to achieve a superior correction rate when compared to the reference methods. However, solving eq. 7.4 with help of an iterative solver as described in sec. 6.3 represents a computationally very expensive step. Furthermore, this evaluation has to be repeated for each point in time at which a new set of correction parameters is needed. Of course there are several trivial ideas to limit the computational demand, e.g. by selecting the latest found DSNU pattern as the start point for the new solver.

Besides those ideas, the transformation of the problem into the known Chua CNN paradigm [46] is a further possibility and limits the computational efforts while preserving the advantages of the method as shown below. We have published this transformation already at the CNNA 2012 [11]. The transformation starts by analyzing the basic energy minimization task, which was described in eq. 6.57 as:

$$\arg\min_{\{b_{i,j}\}} \{\text{const.} + E_\mathrm{P} + E_\mathrm{D}\}\,. \tag{7.7}$$

This problem was then solved with help of an iterative solver in section sec. 6.3. Most of the fast solvers use the energy's gradient, which calculates for the proposed problem as given in eq. 6.78:

$$\hat{\nabla}_{\{b_{i,j}\}}(E_\mathrm{D} + E_\mathrm{P}) = -2\lambda T_N(A_M - b) * HP(\hat{L}, \Phi) + 2\beta(b - \mu_b) \tag{7.8}$$

With a known gradient, the energy minimization can be actually transferred into a CNN. To provide a better understanding and motivation on these steps, the basics of a CNN are now revisited first.

A summary of CNNs: CNN is an abbreviation for cellular neural network, or cellular nonlinear network. The CNNs are considered in the way of the Chua CNN paradigm [46] as a network of cells that have an interconnection only within a limited neighborhood. This allows to consider the cell output of the neighboring cells as cell inputs. The different inputs to the cells can be weighted and these weights are then called synaptic weights. The dynamics of a CNN are given by temporal state equations of the cells and lead to a fast convergence of the CNN into a temporal equilibrium point after a short period of time.

If the connected neighborhood is equal in size for each cell and if the spatial synaptic weights are constant relative to the cells, then each cell's dynamics is given by the following state equation [46]:

$$\begin{aligned} \partial_t x_{k,l} &= -x_{k,l} + (y * A)_{(k,l)} + (u * B)_{(k,l)} + z_{k,l} \\ y_{m,n} &= f(x_{\mathrm{M},n})\,, \end{aligned} \tag{7.9}$$

where x is called *cell state*, y is called *output*, u is called *input* and z is called *threshold*. The filter masks A and B represent the spatial constant synaptic weights and define predominantly the CNN's behavior. It is to point out that the cell state x, cell output y and cell input u are time continuous variables.

Transforming a method into a CNN will therefore allow a time continuous input to the system and results in a time continuous cell output y. For the DSNU method, the output y should then become an approximation of the DSNU parameter set, even if the CNN is not yet fully converged. A realization of CNNs in the analogue domain is also possible and first partly analogue hardware has already been developed in this branch (eg. the ACE16k chip [47]).

7.2.1. Transferring the Energy Minimization into a CNN

When solving the energy minimization task of the new method, the iterative solvers approach their solution with increasing iterations n. As the solver itself has to be a physical system, each subsequent iteration has to be connected with an increase in the real, physical time t. Thus the iteration's n map to discrete points in the continuous time dimension t.

In an abstract view, a steepest descent solver can therefore be regarded as the quantized approximation of an yet unknown *temporal* state equation. The CNN state equation can now be regarded as this hidden state equation and the update steps of the steepest descent solvers can be investigated in this context. With the update

step of the iterative solvers in the opposite direction of the energy's gradient, eq. 6.78 can be assumed as the CNN cell state change in eq. 7.9:

$$-\nabla_b E_{tot.}(b(t)) = \nabla_t x\,. \tag{7.10}$$

With this assumption, the CNN variables need to be identified, where the following mappings are obvious:

$$\begin{aligned}
&x = b \quad \text{: cell state as } \textit{inner} \text{ DSNU}\\
&y = f(x) = \begin{cases} \mu_b - n\sigma_b & if \quad x < \mu_b - n\sigma_b \\ x & if \quad x \in (\mu_b \pm n\sigma_b) \\ \mu_b + n\sigma_b & if \quad x > \mu_b + n\sigma_b \end{cases}\\
&u = A_{\mathrm{M}_W} \quad \text{: input as weighted average of the measurements}\,.
\end{aligned} \tag{7.11}$$

The function $y = f(x)$ is used to model the physically expected boundaries of the parameters $\{b_{i,j}\}$ (e.g. by a choice of n that fits the desired outlier criteria). y is considered as the final resulting $\{b_{i,j}\}$ to correct the image with.

However, the transformation into a CNN is only possible if the synaptic weights can be chosen in a way that eq. 7.10 can be fulfilled. A calculation of the synaptic weights and the threshold can be found in the linear region of the CNN ($f(x) = x$) as shown below:

$$\partial_t x_{i,j} = -\partial_{b_{i,j}} E_{tot.} \tag{7.12}$$

$$= -\Big[-2\lambda T_N (A_M - b) * HP(\hat{L}, \Phi) + 2\beta(b - \mu_b)\Big]_{i,j} \tag{7.13}$$

$$= \Big[2\lambda T_N \Big((u * HP(\hat{L}, \Phi)) - (x * HP(\hat{L}, \Phi))\Big) - 2\beta x + 2\beta\mu_b)\Big]_{i,j} \tag{7.14}$$

$$= \Bigg[x * (-2\lambda T_N HP) - 2\beta x + u * \underbrace{(2\lambda T_N HP)}_{B} + \underbrace{2\beta\mu_b}_{z}\Bigg]_{i,j} \tag{7.15}$$

$$\begin{aligned}
&= \Bigg(\sum_{k,l\neq i,j} x_{i,j} \cdot (-2\lambda T_N HP_{k-i,l-j})\Bigg) + x_{i,j}(-2\lambda T_N HP_{0,0})\\
&\quad - 2\beta x_{i,j} + [u * B + z]_{i,j}
\end{aligned} \tag{7.16}$$

$$\begin{aligned}
&= -x_{i,j} + \sum_{k,l\neq i,j} x_{i,j} \cdot \underbrace{(-2\lambda T_N HP_{k-i,l-j})}_{A_{(k-i,l-j)\neq(0,0)}}\\
&\quad + x_{i,j} \underbrace{((-2\lambda T_N HP_{0,0}) - 2\beta + 1)}_{A_{0,0}} + u_{i,j} * B + z_{i,j}
\end{aligned} \tag{7.17}$$

$$= -x_{i,j} + (x * A)_{i,j} + (u * B)_{i,j} + z_{i,j} \tag{7.18}$$

As shown, A, B and z can be defined for the linear case where $f(x) = x$. For the cases where $f(x) \neq x$, the same definitions of A, B and z are used which then defines the CNN's state equation completely according to the mathematical considerations of eq. 7.9. All parameters that define this CNN, especially the

synaptic weight matrices A and B, are calculated directly from the parameters of the energy minimization task as:

$$\begin{aligned} A_{i,j\neq 0,0} &= -2\lambda T_N H P_{i,j} \\ A_{0,0} &= -2\lambda T_N H P_{0,0} - 2\beta + 1 \\ B &= 2\lambda H P \\ z &= 2\beta\mu_b = \frac{\mu_b}{\sigma_b^2}\,. \end{aligned} \tag{7.19}$$

Therefore the typical learning for the synaptic weights of a CNN is not needed. Considering a constant input $u = A_{\mathrm{M}_W}$, the CNN converges to a very similar solution as the maximum likelihood method. This may however not be the mathematically identical solution as the methods differ by the nonlinear CNN function $f(x)$.

With the defined CNN state equation the input becomes *time dependent* and $u(t) = A_{\mathrm{M}_W}(t)$ is used as input, which represents the averaged weighted measurements up to the current time t. The weighted average is used, as there is no reason to ignore the already achieved improvements from sec. 7.1. The system output $y(t)$ now represents the desired solution $\{b_{i,j}\}(t)$ for every point in time with a certain error if the system has not yet converged into its equilibrium point. The CNN will not converge instantly, but as the nonuniformity parameters change rather slowly with time, the CNN gives very good results even if its convergence is slowed down. The results below confirm this.

7.2.2. CNN Simulation

The derived CNN's state equation can now be directly implemented in hardware or used on actual CNN-alike hardware systems like the SCAMP, ASPA, ACE16k or QEye systems [48, 49, 47]. Most vision systems are however not fully prepared for a CNN and thus the CNN needs to be simulated (e.g. on an FPGA). A good approximation of the CNN's behavior is achieved by the following approximation of a full range CNN [50]:

$$\begin{aligned} x_{i,j,t+\epsilon_t} &= f(x_{i,j,t} + s_\epsilon \cdot \partial_t x_{i,j,t}) \\ \partial_t x_{k,l} &= -x_{k,l} + (y * A)_{(k,l)} + (u * B)_{(k,l)} + z_{k,l} \\ \text{with:} \quad & A, B, z \text{ and } f(\cdot) \text{ as defined above}\,, \end{aligned} \tag{7.20}$$

and ϵ_t as the simulated time step and s_ϵ as the corresponding convergence constant. The use of the nonlinear function $f(\cdot)$ within the update guarantees that state and output are identical for this simulation and stay within the physical boundaries of the assumed system.

As there are usually only very limited computational resources available for nonuniformity corrections, we chose ϵ_t as the time that elapses from frame to frame and s_ϵ then becomes the only free parameter of the method. The rather

large time step is justifiable as the CNN targets the DSNU correction and thus does not demand the CNN to converge on a frame to frame basis.

7.2.3. Experimental Results for the CNN approach

The CNN has 4 parameters that have to be adjusted. The first one is the HP filter mask which is chosen as before by its three realizations $HP(\nabla, BL)$, $HP(\nabla, MN)$ and $HP(\triangle, MN)$. As the CNN method is actually just another solver for the $\min_{\mathrm{wA}}(\nabla, BL)$, $\min_{\mathrm{wA}}(\nabla, MN)$ and $\min_{\mathrm{wA}}(\triangle, MN)$ method, the corresponding optimal λ parameters from tab. 7.3 for the random frame sequences and from tab. 7.4 for the consecutive frame sequences are used. For the time step ϵ_t the time between two consecutive frames is used, which effectively results in one update step of eq. 7.20 per frame in the given sequence. Finally the free parameter s_ϵ is left to be optimized and the CNN approach results in the three new methods summarized in tab. 7.5:

Parameter Search and Best-Case Performance

Unfortunately there is no theoretical way to define the free parameter s_ϵ and thus a parameter search for random frame and consecutive frame sequences was conducted which resulted in the parameter sets shown in tab. 7.6 and 7.7. The best-case performance analysis according to 5.3.1 (p. 90) was also conducted and the results are shown in fig. 7.8. It is obvious from the figure that the CNN methods outperform even the best reference methods for the random frame sequences, except for the $\mathrm{CNN}(\triangle, MN)$ method. For the consecutive frame sequences the CNN methods achieve a lower correction performance as expected but still the optimal performances overlap in their error bars. The further evaluations for the consecutive frame sequences and the image material with low valued nonuniformities is continued in sec. 10.1 and sec. 10.2. The temporal performance analysis will now give further inside into the methods' temporal performance for random frame sequences.

Method Abbreviation	Details
$\mathrm{CNN}(\nabla, BL)$	Iterations of eq. 7.20 with $HP(\nabla, BL)$
$\mathrm{CNN}(\nabla, MN)$	Iterations of eq. 7.20 with $HP(\nabla, MN)$
$\mathrm{CNN}(\triangle, MN)$	Iterations of eq. 7.20 with $HP(\triangle, MN)$

Table 7.5.: Abbreviation and details for the resulting CNN type methods.

Method	Parameter
$\mathrm{CNN}(\nabla, BL)$	$2.439 \pm (0.5219, 21.4\%)$
$\mathrm{CNN}(\nabla, MN)$	$3.445 \pm (1.955, 56.76\%)$
$\mathrm{CNN}(\triangle, MN)$	$2.034 \pm (0.2135, 10.5\%)$

Table 7.6.: Results of a parameter search to optimize CNN step size s_ϵ for random frame sequences

Method	Parameter
$\mathrm{CNN}(\nabla, BL)$	$2.501 \pm (6.526, 261\%)$
$\mathrm{CNN}(\nabla, MN)$	$8.305 \pm (23.23, 279.8\%)$
$\mathrm{CNN}(\triangle, MN)$	$5.281 \pm (6.097, 115.4\%)$

Table 7.7.: Results of a parameter search to optimize CNN step size s_ϵ for consecutive frame sequences

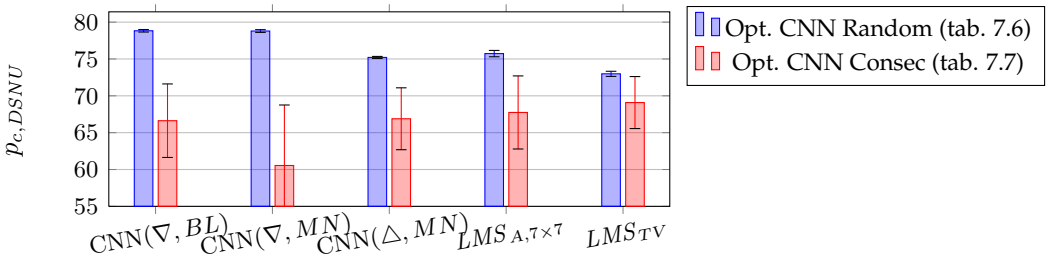


Figure 7.8.: Optimal Performance analysis with respect to the $p_{\mathrm{c,DSNU}}$ quality metric of the CNN methods.

Temporal Performance Analysis

Fig. 7.9 shows the temporal results for the random frame sequences and except for the $\mathrm{CNN}(\triangle, MN)$ method, all CNN methods outperform the reference methods. The costy iterative solver can therefore be substituted with the CNN method. The only exception is the $\mathrm{CNN}(\triangle, MN)$ method, which achieved better results with the iterative solver. The $\mathrm{CNN}(\nabla, BL)$ and $\mathrm{CNN}(\nabla, MN)$ methods actually converge towards the same result of $\approx 78\,\%$ correction rate. The only advantage of the $\mathrm{CNN}(\nabla, MN)$ method over the $\mathrm{CNN}(\nabla, BL)$ method is a slightly faster performance in the first few hundred frames. If this advantage is not needed, the additional computational costs for the larger mask of the $\mathrm{CNN}(\nabla, MN)$ method can be saved.

The temporal evaluation in the other quality measures are shown in appendix E.2 for completeness.

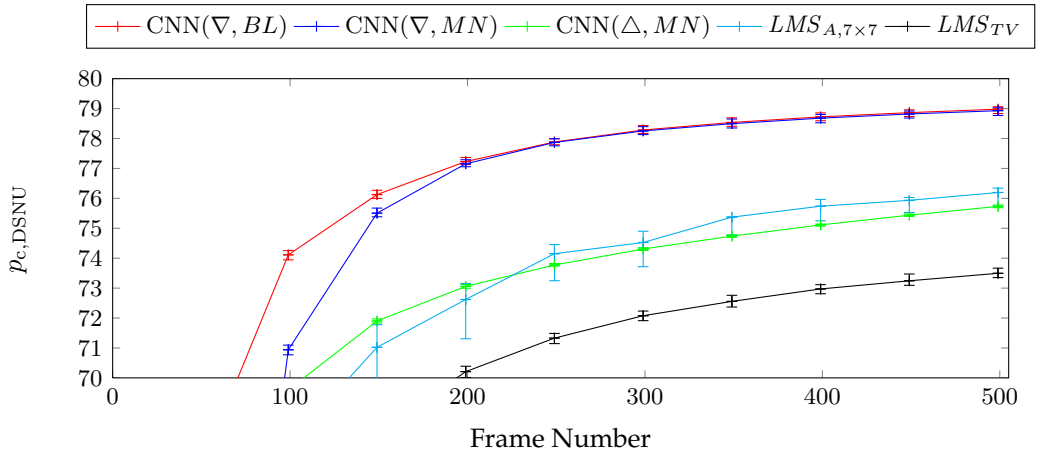


Figure 7.9.: Methods evaluations with experimental optimized parameters (tab. 7.6) for the $p_{\mathrm{c,DSNU}}$ metric and random frame sequences as input

A further interesting phenomenon is that the nonlinearity of the CNN methods results in improved performances with respect to the actual iterative solvers for the $\mathrm{CNN}(\nabla, BL)$ and $\mathrm{CNN}(\nabla, MN)$ methods. Fig. 7.10 compares the performances of the CNN methods against the corresponding energy minimizing solvers. The three best methods so far are $\mathrm{CNN}(\nabla, BL)$, $\mathrm{CNN}(\nabla, MN)$ and $\min_{\mathrm{wA}}(\nabla, MN)$ and all converge to the maximum reached performance of $78\,\%$. Only the $\mathrm{CNN}(\triangle, MN)$ method is behind the performance of its energy minimizer $\min_{\mathrm{wA}}(\triangle, MN)$ as expected from the above performance analysis. Of course other iterative solvers that limit the result space in the same way as the CNN's nonlinearity function could achieve similar results as well. However the CNN paradigm further allows to transfer the results fast into already present hardware or use other researched improvements from the CNN community.

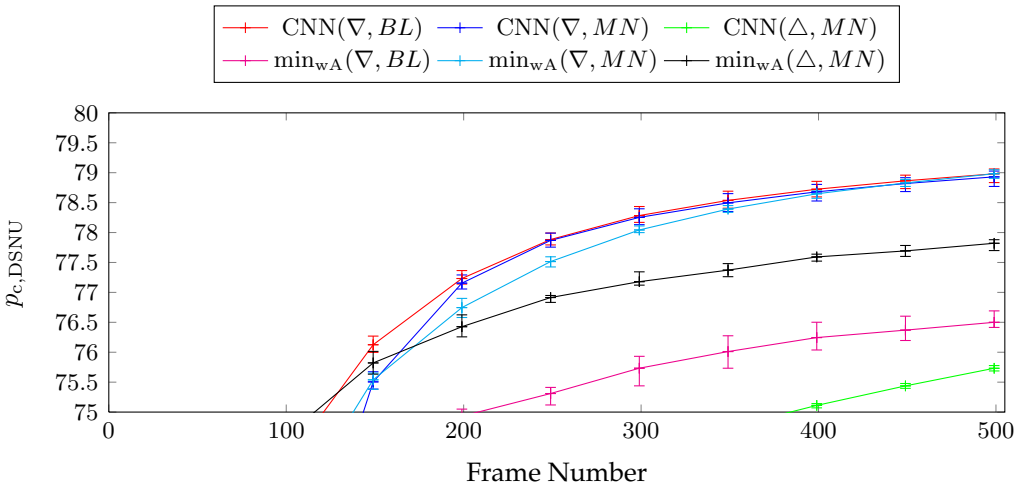
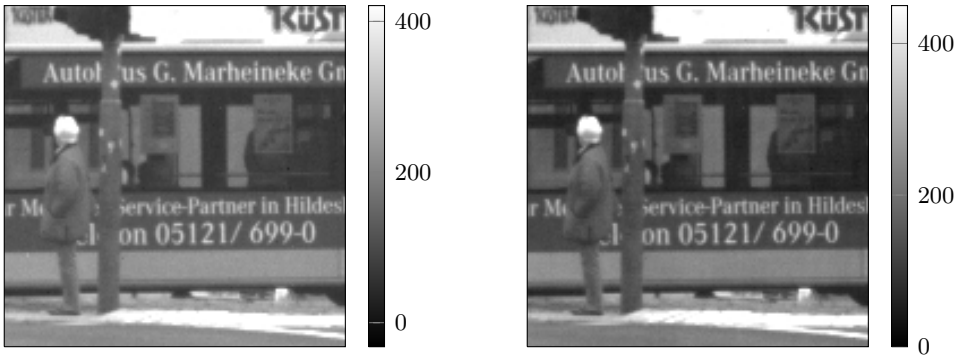


Figure 7.10.: The CNN-methods and the $\min_{wA}$-methods compared in the $p_{\mathrm{c,DSNU}}$ metric for random frame sequences. The used parameters sets are: (Opt. CNN Random (tab. 7.6)) and (Weighted Stat. (tab. 7.1)) have been used.

Spatial and Spectral Performance Analysis

The spatial and spectral performances of the CNN based methods are now evaluated for the random frame sequences. With the good correction performance, the comparison of the corrected images does not lead to any visible differences as seen in fig. 7.11. The remaining DSNU patterns are shown in fig. 7.12 and the $\mathrm{CNN}(\nabla, BL)$ and $\mathrm{CNN}(\nabla, MN)$ methods present very similar remains, which is expected as both methods reach almost identical correction performances. The $\mathrm{CNN}(\triangle, MN)$ method shows mainly horizontally remaining structures, which are similar to the ones presented by the $LMS_{A,7\times7}$ and LMS_{TV} method. Besides this quality difference, the spectral analysis gives a better overview of the methods' differences according to the spectral sensitivity of the methods. Fig. 7.13 shows the spectral damping factors and here the advantage of the CNN based methods is clearly visible. All CNN based methods show a better concentration of the remaining frequency components towards the center of the representation, which means that only spatial low frequencies are left behind. Further the horizontal and vertical frequencies are damped with slight improvments compared to the reference methods. As usual for the $\triangle$-based methods does $\mathrm{CNN}(\triangle, MN)$ damps the low frequencies not as much as the ∇-based methods. The advantage of the CNN based methods can as well be compared towards the $\min_{\mathrm{wA}}(\nabla, BL)$, $\min_{\mathrm{wA}}(\nabla, MN)$ and $\min_{\mathrm{wA}}(\triangle, MN)$ methods' spectral damping in fig. 7.7 on p. 7.7.



(a) Ground Truth Corrected

(b) Corrected with $\mathrm{CNN}(\nabla, BL)$

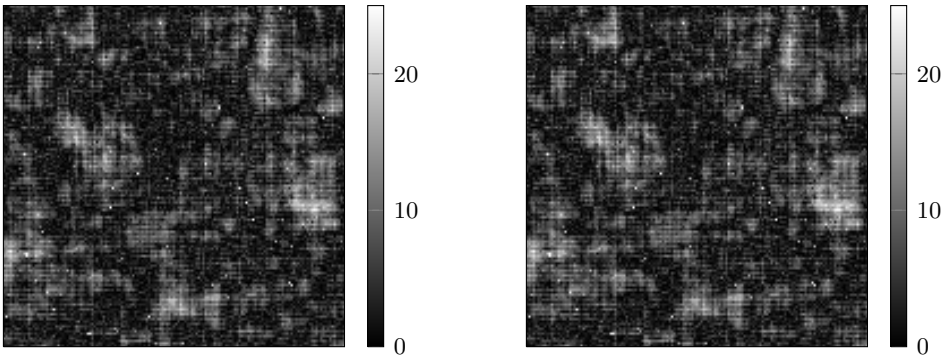
(c) Corrected with $\mathrm{CNN}(\nabla, MN)$

(d) Corrected with $\mathrm{CNN}(\triangle, MN)$

(e) Corrected with $LMS_{A,7\times7}$ Method

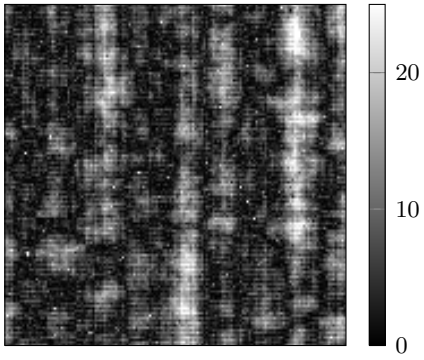
(f) Corrected with LMS_{TV} Method

Figure 7.11.: Corrected images of the introduced methods for random frame sequences (Parameters according to tab. 7.6)

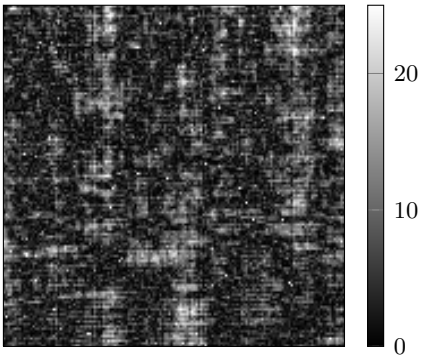


(a) $|\{b_{i,j,rem.}\}|$ for $\mathrm{CNN}(\nabla, BL)$

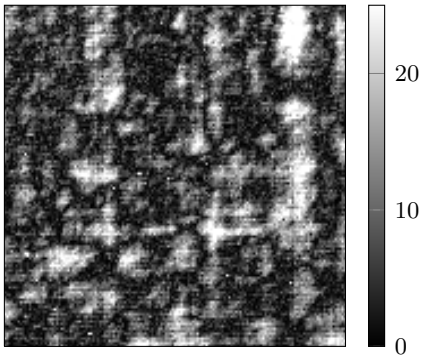
(b) $|\{b_{i,j,rem.}\}|$ for $\mathrm{CNN}(\nabla, MN)$



(c) $|\{b_{i,j,rem.}\}|$ for $\mathrm{CNN}(\triangle, MN)$

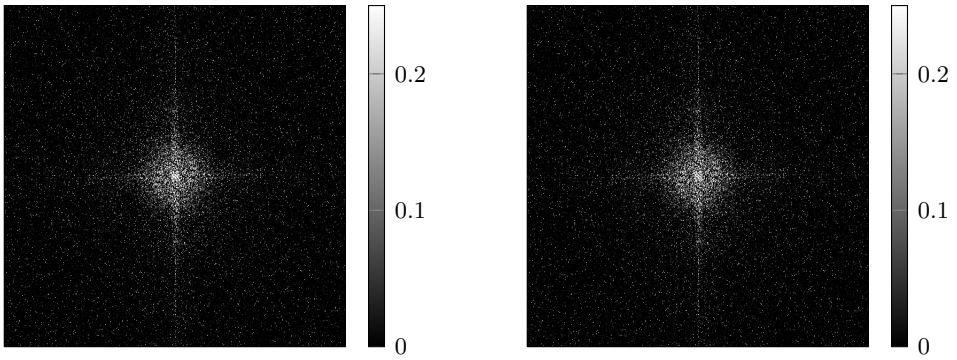


(d) $|\{b_{i,j,\mathrm{rem.}}\}|$ for $LMS_{A,7\times7}$

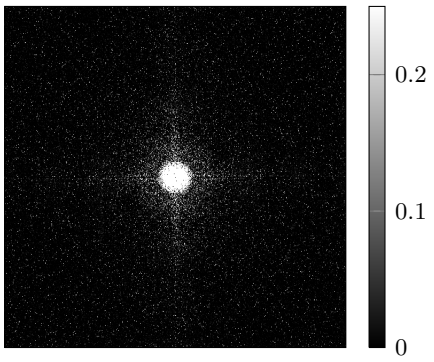


(e) $|\{b_{i,j,\mathrm{rem.}}\}|$ for LMS_{TV}

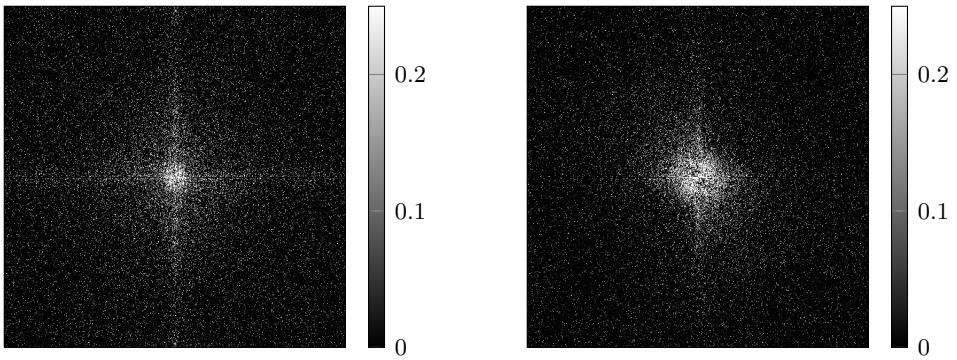
Figure 7.12.: Remaining DSNU patterns for the introduced methods and random frame sequences (Parameters according to tab. 7.6)



(a) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $\mathrm{CNN}(\nabla, BL)$ (b) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $\mathrm{CNN}(\nabla, MN)$



(c) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $\mathrm{CNN}(\triangle, MN)$



(d) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $LMS_{A,7\times7}$ (e) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for LMS_{TV}

Figure 7.13.: Spectral damping factors for the introduced methods and random frame sequences (Parameters according to tab. 7.6)

7.2.4. First Conclusions

With the introduced CNN based maximum likelihood approach the computational performance of the new methods could be severely degraded and is now in the same order as the frame-to-frame based LMS methods. Therefore next the correction of horizontal and vertical frequencies is addressed.

7.3. An Extension to Row and Column Wise Estimation

The typical characteristics of the ground truth DSNU patterns shown in sec. 4.3, showed strong horizontal and vertical spatial frequencies in the spectral analysis. This is especially of concern, as the human eye is very sensitive to those kind of patterns in combination with the fact that all introduced methods have a problem to correct for these frequency components.

7.3.1. Calculating the Gradient with an Improved Sensor Model

Given the current energy minimization task, the offset parameters $\{b_{i,j}\}$ are considered independent for each pixel. However, with the knowledge of the row and column wise dependencies, the sensor model can be extended to consider each offset as the sum of a row-, column- and pixel-wise component:

$$b_{i,j} = b_{\text{col.},i} + b_{\text{row.},j} + b_{\text{pix.},i,j}\,. \tag{7.21}$$

The resulting value $b_{i,j}$ is still the realization of the random variable $\mathbf{b}_{i,j}$ as introduced in sec. 6.2.1 and thus all the statements made about it are still valid. The dependency of the row- and column-wise components is therefore only implicitly considered. Consequently the energy function given in sec. 6.3 stays valid and the variables $b_{i,j}$ can be substituted leading to the modified energy function:

$$\begin{aligned} E_{\mathrm{D}} + E_{\mathrm{P}} = \lambda \sum_{t_n} & \sum_{(i,j)\in S} (C_{\mathrm{M},i,j} - b_{\text{col.},i} - b_{\text{row.},j} - b_{\text{pix.},i,j} - \chi_{i,j}) \\ & \sum_{(k,l)\in S} (C_{\mathrm{M},k,l} - b_{\text{col.},k} - b_{\text{row.},l} - b_{\text{pix.},k,l} - \chi_{k,l})\Psi(\cdot) \\ & + \sum_{i,j} \beta \left(b_{\text{col.},i} + b_{\text{row.},j} + b_{\text{pix.},i,j} - \mu_b\right)^2 + \text{const.} \end{aligned} \tag{7.22}$$

The target is now to minimize this energy function with respect to the new variables:

$$\arg \min_{b_{\text{col.},i}, b_{\text{row.},j}, b_{\text{pix.},i,j}} (E_{\mathrm{D}} + E_{\mathrm{P}}) \tag{7.23}$$

As before, the gradient of the energy functional in all the variables is needed, with three different kinds of variables instead of one in this new extension. For simplicity we start with the derivation for the column type offset parameters of the E_{P} energy:

$$\partial_{b_{\text{col.},m}}(E_{\mathrm{P}}) = \partial_{b_{\text{col.},m}} \sum_{i,j} \beta \left(b_{\text{col.},i} + b_{\text{row.},j} + b_{\text{pix.},i,j} - \mu_b\right)^2 \tag{7.24}$$

$$= 2\beta \sum_{j} \left(b_{\text{col.},m} + b_{\text{row.},j} + b_{\text{pix.},m,j} - \mu_b\right) \tag{7.25}$$

$$= 2\beta b_{\sum_{\text{col},m}} - \mu_b \sum_{j} 1 \tag{7.26}$$

$$\text{with:} \quad X_{\sum_{\text{col/row},m}} \quad : \text{summation of } X \text{ along the column/row at position } m \tag{7.27}$$

The newly introduced variable $b_{\sum_{\text{col}},m}$ gives the column wise summation over the *resulting* offset parameters $b_{i,j}$. In the vector notation, all the column offset derivatives can be summarized with help of the transposed nabla operator. The indicies of the vectors can then be omitted and the image sensor dimensions are considered as (M, N) with M columns and N rows.

$$\nabla^T_{b_{\text{col.}}} E_{\mathrm{P}} = 2\beta(b_{\sum_{\text{col}}} - N\mu_b) \tag{7.28}$$

$$\text{with:} \quad N\text{: Number of sensor rows}\,. \tag{7.29}$$

The row wise gradient obeys the same mathematical form and can be written as:

$$\nabla_{b_{\text{row.}}} E_{\mathrm{P}} = 2\beta(b_{\sum_{\text{row}}} - M\mu_b) \tag{7.30}$$

$$\text{with:} \quad M\text{: Number of sensor columns}\,. \tag{7.31}$$

Last, the pixel-wise part of the gradient can be expressed by the using the combined offsets $b_{i,j} = b_{\text{col.},i} + b_{\text{row.},j} + b_{\text{pix.},i,j}$ again. Using the $\hat{\nabla}$ operator it expresses as:

$$\hat{\nabla}_{b_{\text{pix.}}} E_{\mathrm{P}} = 2\beta(b - \mu_b)\,. \tag{7.32}$$

For the derivatives of the data part of the energy function, the same mathematical transformations as in sec. 6.3 are used, namely exploiting the product and chain rule in combination with the linearity of the derivation. For the partial derivatives for the column component this leads to:

$$\partial_{b_{\text{col.},m}}(E_{\text{D}}) = (\sum_{t_n}\sum_{i,j}\lambda(C_{\text{M}} - b - \chi) \otimes ((C_{\text{M}} - b - \chi) * HP(\hat{L}, \Phi)) + \text{const.}) \tag{7.33}$$

$$= \sum_{t_n}\lambda\partial_{b_{\text{col.},m}}\left(\sum_{(i,j)\in S}(C_{\text{M},i,j} - b_{i,j} - \chi_{i,j}) \right.$$

$$\left. \cdot \sum_{(k,l)\in S}(C_{\text{M},k,l} - b_{k,l} - \chi_{k,l})\Psi(|k-i|, |l-j|)\right) \tag{7.34}$$

$$= \sum_{t_n}\lambda\left((\partial_{b_{\text{col.},m}}\sum_{(i,j)\in S}(C_{\text{M},i,j} - b_{i,j} - \chi_{i,j}))\sum_{(k,l)\in S}(C_{\text{M},k,l} - b_{k,l} - \chi_{k,l})\Psi(\cdot)\right.$$

$$\left. + \sum_{(i,j)\in S}(C_{\text{M},i,j} - b_{i,j} - \chi_{i,j})(\partial_{b_{\text{col.},m}}\sum_{(k,l)\in S}(C_{\text{M},k,l} - b_{k,l} - \chi_{k,l})\Psi(\cdot))\right) \tag{7.35}$$

The actual task is now to define the derivatives of the sums and to substitute them in the above equations. With the new variables the sum simplifies to:

$$\partial_{b_{\text{col.},m}}\sum_{(i,j)\in S}(C_{\text{M},i,j} - b_{i,j} - \chi_{i,j}) \tag{7.36}$$

$$=\partial_{b_{\text{col.},m}}\sum_{(i,j)\in S}(C_{\text{M},i,j} - b_{\text{col.},i} - b_{\text{row.},j} - b_{\text{pix.},i,j} - \chi_{i,j}) \tag{7.37}$$

$$=\sum_{j}\partial_{b_{\text{col.},m}}\sum_{i}(C_{\text{M},i,j} - b_{\text{col.},i} - b_{\text{row.},j} - b_{\text{pix.},i,j} - \chi_{i,j}) \tag{7.38}$$

$$=\sum_{j}-1\,, \tag{7.39}$$

which can be directly inserted into the eq. 7.35 and allows further simplifications:

$$\partial_{b_{\text{col.},m}}(E_{\text{D}}) = \sum_{t_n}\lambda\left(-\sum_{j}\sum_{(k,l)\in S}(C_{\text{M},k,l} - b_{k,l} - \chi_{k,l})\Psi(|k-m|, |l-j|)\right. \tag{7.40}$$

$$\left. + \sum_{(i,j)\in S}(C_{\text{M},i,j} - b_{i,j} - \chi_{i,j})(-\sum_{l}\Psi(|m-i|, |l-j|))\right) \tag{7.41}$$

$$= -2\lambda\sum_{t_n}\left(\sum_{j}\sum_{(k,l)\in S}(C_{\text{M},k,l} - b_{k,l} - \chi_{k,l})\Psi(|k-m|, |l-j|)\right) \tag{7.42}$$

$$= -2\lambda \sum_{t_n} \left(\sum_{(k,l)\in S} (C_{\mathrm{M},k,l} - b_{k,l} - \chi_{k,l}) \sum_{j} \Psi(|k-m|, |l-j|) \right) . \tag{7.43}$$

7.3.2. Reduction into One Dimensional Subproblems

The sum over the Ψ reduces to a constant Ψ_{col} in each Ψ-column. This is justified by the fact that the sum over j extends far over the area where $\Psi(|k-m|, |l-j|) \neq 0$, resulting in the same summations for each column:

$$\sum_{j} \Psi(|k-m|, |l-j|) = \Psi_{\mathrm{col}}(|k-i|) \in \mathbb{R} \tag{7.44}$$

Near the borders of the image sensor the above column wise sum is valid as well if the same mirroring extension of the image is used as in sec. 6.3. The temporal expectation value and temporal average from sec. 6.3 are also used which simplifies the eq. 7.43 to:

$$- 2\lambda \sum_{t_n} T_N \frac{1}{T_N} \left(\sum_{k} (\sum_{l} C_{\mathrm{M},k,l} - \sum_{l} b_{k,l} - \sum_{l} \chi_{k,l}) \Psi_{\mathrm{col}}(|k-i|) \right) \tag{7.45}$$

$$= - 2\lambda T_N \left(\sum_{k} (\sum_{l} \sum_{t_n} \frac{1}{T_N} C_{\mathrm{M},k,l} - \sum_{l} \sum_{t_n} \frac{1}{T_N} b_{k,l} - \sum_{l} \sum_{t_n} \frac{1}{T_N} \chi_{k,l}) \Psi_{\mathrm{col}}(|k-i|) \right) \tag{7.46}$$

$$= -2\lambda T_N \left(\sum_{k} (\underbrace{\sum_{l} A_{\mathrm{M},k,l}}_{A_{\mathrm{M},\sum_{\mathrm{col}},l}} - \underbrace{\sum_{l} b_{k,l}}_{b_{\sum_{\mathrm{col}},k}}) \Psi_{\mathrm{col}}(|k-i|) \right) \tag{7.47}$$

In the last step the definition of $b_{\sum_{\mathrm{col},k}}$ from eq. 7.26 is used. The final result is a convolution again, but this time in one dimension only:

$$\nabla^T_{b_{\mathrm{col.}}} E_{\mathrm{D}} = -2\lambda T_N (A_{\mathrm{M},\sum_{\mathrm{col}}} - b_{\sum_{\mathrm{col}}}) * HP_{\mathrm{col.}}(\hat{L}, \Phi) \tag{7.48}$$

As already for the E_{P} derivatives, the results for the $b_{\mathrm{row.}}$ and $b_{\mathrm{pix.}}$ variables follow directly from the symmetry of the mathematical expressions:

$$\nabla_{b_{\mathrm{row.}}} E_{\mathrm{D}} = -2\lambda T_N (A_{\mathrm{M},\sum_{\mathrm{row}}} - b_{\sum_{\mathrm{row}}}) * HP_{\mathrm{row.}}(\hat{L}, \Phi) \tag{7.49}$$

$$\hat{\nabla}_{b_{\mathrm{pix.}}} E_{\mathrm{D}} = -2\lambda T_N (A_{\mathrm{M}} - b) * HP(\hat{L}, \Phi) \tag{7.50}$$

The filter masks $HP_{\text{row.}}(\hat{L}, \Phi)$ and $HP_{\text{col.}}(\hat{L}, \Phi)$ follow also directly from the row or column wise summation over the $HP(\hat{L}, \Phi)$ masks. Due to the symmetry of these masks they can be written directly as:

$$\begin{aligned}
HP^T_{\text{row.}}(\nabla, BL) &= \begin{pmatrix} -1 & 2 & -1 \end{pmatrix} = HP_{\text{col.}}(\nabla, BL) \\
HP^T_{\text{row.}}(\nabla, MN) &= \frac{1}{2520} \cdot \begin{pmatrix} -1 & -24 & -15 & 80 & -15 & -24 & -1 \end{pmatrix} = HP_{\text{col.}}(\nabla, MN) \\
HP^T_{\text{row.}}(\triangle, MN) &= \frac{1}{25200} \begin{pmatrix} 1 & 0 & -9 & 16 & -9 & 0 & 1 \end{pmatrix} = HP_{\text{col.}}(\triangle, MN) \quad (7.51)
\end{aligned}$$

With the knowledge of the gradient, the minimum of the energy function $E_{\text{D}} + E_{\text{P}}$ can be found by iterative gradient based solvers. This time the tree types of variables combine to a gradient consisting of the parts $\nabla^T_{b_{\text{col.}}}$, $\nabla_{b_{\text{row.}}}$ and $\hat{\nabla}_{b_{\text{pix.}}}$.

The combined gradient and the equation to solve can be expressed with help of the equations 7.28 to 7.32 for the prior energy and the equations 7.48 to 7.50 for the data energy:

$$\begin{pmatrix} \nabla^T_{b_{\text{col.}}}(E_{\text{D}} + E_{\text{P}}) \\ \nabla_{b_{\text{row.}}}(E_{\text{D}} + E_{\text{P}}) \\ \hat{\nabla}_{b_{\text{pix.}}}(E_{\text{D}} + E_{\text{P}}) \end{pmatrix} = 0 \quad (7.52)$$

$$\Leftrightarrow \begin{pmatrix} -2\lambda T_N(A_{\text{M},\sum_{\text{col}}} - b_{\sum_{\text{col}}}) * HP_{\text{col.}}(\hat{L}, \Phi) + 2\beta(b_{\sum_{\text{col}}} - N\mu_b) \\ -2\lambda T_N(A_{\text{M},\sum_{\text{row}}} - b_{\sum_{\text{row}}}) * HP_{\text{row.}}(\hat{L}, \Phi) + 2\beta(b_{\sum_{\text{row}}} - M\mu_b) \\ -2\lambda T_N(A_{\text{M}} - b) * HP(\hat{L}, \Phi) + 2\beta(b - \mu_b)\,, \end{pmatrix} = 0 \quad (7.53)$$

$$(7.54)$$

With these equations a combined row, column, and pixel-wise DSNU correction is possible. For better readability this combined form will be abbreviated as RCP-correction.

Separated Row- and Column-wise DSNU: A special case arises, if only column-wise or only row-wise offsets are present. For only column wise offsets, eq. 7.21 and 7.26 simplify to:

$$b_{i,j} = b_{\text{col.},i} \quad \text{and:} \quad b_{\sum_{\text{col},m}} = \sum_j b_{\text{col.},m} = N b_{\text{col.},m} \quad (7.55)$$

While the same goes for row wise offsets:

$$b_{i,j} = b_{\text{row.},j} \quad \text{and:} \quad b_{\sum_{\text{row},n}} = \sum_i b_{\text{row.},n} = M b_{\text{row.},n} \quad (7.56)$$

The complete gradient for these two special cases then compresses into the following one dimensional forms:

$$\nabla^{T}_{b_{\text{col.}}}(E_{\text{D}} + E_{\text{P}}) = -2\lambda T_N(A_{\text{M},\sum_{\text{col}}} - Nb_{\text{col.}}) * HP_{\text{col.}}(\hat{L}, \Phi) + 2\beta N(b_{\text{col.}} - \mu_b) \tag{7.57}$$

$$\nabla_{b_{\text{row.}}}(E_{\text{D}} + E_{\text{P}}) = -2\lambda T_N(A_{\text{M},\sum_{\text{row}}} - Mb_{\text{row.}}) * HP_{\text{row.}}(\hat{L}, \Phi) + 2\beta M(b_{\text{row.}} - \mu_b)\,. \tag{7.58}$$

This reduction to one dimensional forms has the advantage to correct for either row- or column-wise DSNU in a way that demands much less computational efforts than the pixel wise approach. Furthermore, in many real applications the row and column wise DSNU is of primary concern due to their good visibilities to the human eye. Other combinations like a combined row- and column-wise correction without the pixel-wise dependencies are possible as well and can be derived directly from the above given equations.

7.3.3. Experimental Results

First Impressions of the Performance

First results of the extensions that considers row, column and the combined RCP DSNU correction, are shown in fig. 7.14 for the choice of $HP = HP(\nabla, \text{MN})$. Thereby fig. 7.14a-7.14c show the corrected images for only row-wise, only column-wise and at last the combination of row-, column- and pixel-wise DSNU estimation (RCP). For the row-wise only and column-wise only corrections, the remaining vertical and horizontal stripes are visible in the corrected images. As this test is performed with the ground truth DSNU the row-wise dependency of the nonuniformity is not as distinct as the column-wise dependency. The RCP-combined correction shows the expected very good results with almost no visual artifacts in the remaining image, eventhough the parameter set has not yet been optimized.

Fig. 7.14d-7.14f show the remaining DSNU patterns, where the uncorrected column-wise DSNU is clearly visible in fig. 7.14d. The remains of the row-wise DSNU patterns, are not very well visible in fig. 7.14e, but they could be observed in the corrected image in fig. 7.14b. The RCP-combined correction shows a remaining DSNU pattern that differs from the usual observations, as there are no larger horizontally or vertically connected structures visible (see in fig. 7.6 on p. 147). This forms an advantage, as the human eye usually is very sensitive to these kind of structures.

The selective row- or column-wise corrections, as well as the effects of a RCP combined correction are visible as well in the spectral damping analysis shown in the figures 7.14g-7.14i. In fig. 7.14g and 7.14h only the horizontal or vertical frequency components have been corrected, which shows up as the corresponding vertical

Method Abbreviation	Description
$\min_{\mathrm{RCP}}(\nabla, BL)$	Solving of eq. 7.52 with $HP(\nabla, BL)$
$\min_{\mathrm{RCP}}(\nabla, MN)$	Solving of eq. 7.52 with $HP(\nabla, MN)$
$\min_{\mathrm{RCP}}(\triangle, MN)$	Solving of eq. 7.52 with $HP(\triangle, MN)$

Table 7.8.: Abbreviations for the resulting RCP methods.

Method	Parameter
$\min_{\mathrm{RCP}}(\nabla, BL)$	$0.0007185 \pm (0.0007276, 101.3\%)$
$\min_{\mathrm{RCP}}(\nabla, MN)$	$0.001491 \pm (0.0005025, 33.71\%)$
$\min_{\mathrm{RCP}}(\triangle, MN)$	$0.002557 \pm (0.001233, 48.2\%)$

Table 7.9.: Optimized parameter set for the RCP methods and random frame sequences.

or horizontal line of damped frequencies. In fig. 7.14i the usual image of the remaining low frequencies is visible, but this time the typical horizontal and vertical remaining structures are missing. Instead these structures are corrected to the extend that a black cross on the remaining low frequency components in the center of the representation is visible.

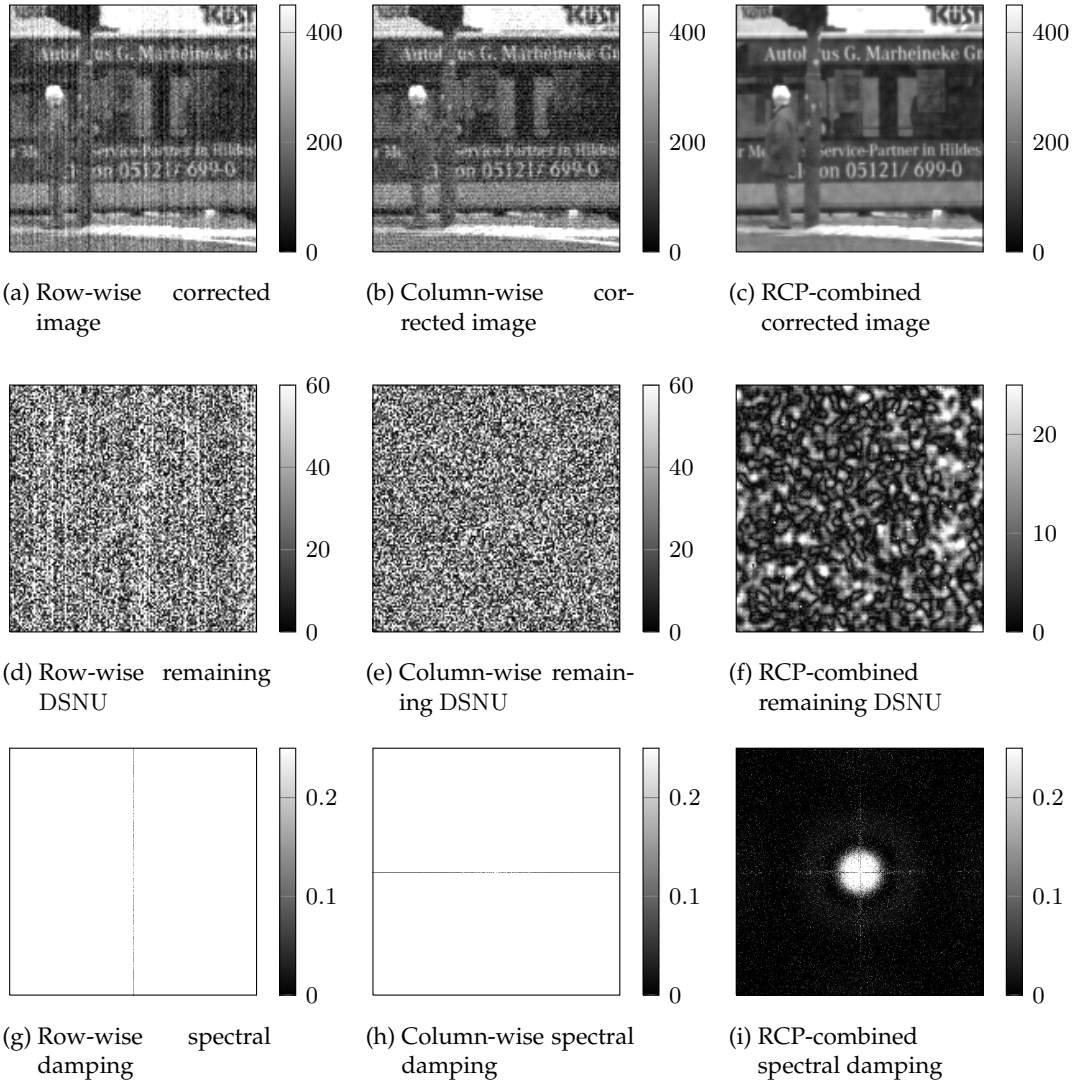
Parameter Search and Best-Case Performance

The row-only and column-only corrections are interesting for certain applications that suffer directly from these types of nonuniformities, but these modes are in general not comparable to the other methods discussed before. Therefore the combined correction of row-, column- and pixel-wise DSNU (RCP) will be further investigated. Tab. 7.8 summarizes the abbreviations for these new variants of the method and tab. 7.9 shows the resulting optimized parameters, given random frame sequences as input. The parameters for consecutive frame input are shown in tab.7.10.

The best-case performance analysis for these methods is shown in fig. 7.15 for the random frame sequences and in fig. 7.16 for the consecutive frame sequences, always in comparison to the $LMS_{A,7\times7}$ and LMS_{TV} reference methods. The two theoretically derived parameter sets, the statistic parameters from tab. 6.1 p. 121

Method	Parameter
$\min_{\mathrm{RCP}}(\nabla, BL)$	$0.000101 \pm (0.0001895, 187.7\%)$
$\min_{\mathrm{RCP}}(\nabla, MN)$	$0.000179 \pm (0.0003615, 201.9\%)$
$\min_{\mathrm{RCP}}(\triangle, MN)$	$0.0009217 \pm (0.0008477, 91.98\%)$

Table 7.10.: Optimized parameter set for the RCP methods and consecutive frame sequences.



(a) Row-wise corrected image

(b) Column-wise corrected image

(c) RCP-combined corrected image

(d) Row-wise remaining DSNU

(e) Column-wise remaining DSNU

(f) RCP-combined remaining DSNU

(g) Row-wise spectral damping

(h) Column-wise spectral damping

(i) RCP-combined spectral damping

Figure 7.14.: Proof of concept analysis of the row-, column- and row-column-pixel-wise (RCP) correction performances in the typical measures: correction of the visual image, remaining DSNU patterns, and spectral damping properties.

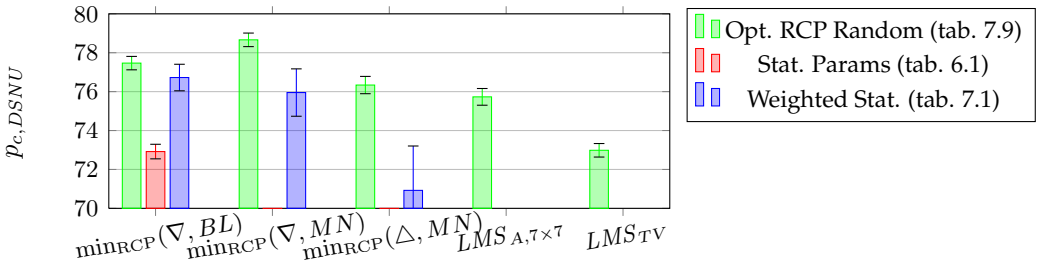


Figure 7.15.: Optimal Performance for the RCP methods with respect to the $p_{\mathrm{c,DSNU}}$ quality metric and random frame sequences

and the statistic parameter set considering the weighted average from tab. 7.1 p. 140 are investigated as well.

For the random frame sequences, the optimized parameter set still leads to a performance superior to the reference methods, but in comparison to the first improvements in fig. 7.2 (p. 142) the performance has slightly decreased for all variants of the new RCP method. For the weighted statistic parameters reasonable performances can be gained, but this performance does not keep up to the reference methods. This has however been the case for the $\min_{wA}$-methods before. With concern to the statistic parameters that did not consider the weighted averaging, only the $\min_{\mathrm{RCP}}(\nabla, BL)$ method reaches an acceptable performance.

For the consecutive frame sequences, the optimized parameter set reaches the performance of the reference methods if the error bars are considered. Compared to the non-RCP methods in fig.7.3 (p. 142) the performance has actually slightly increased for all methods. However, the weighted statistic parameters do not lead to a good performance, which was explained before, as the dependency of the image information leads to a bias in the weighted average. The pure statistic parameters from tab. 6.1 give therefore as expected reasonable good results that can almost keep up to the reference methods' performances considering the error spans. The $\min_{\mathrm{RCP}}(\triangle, MN)$ needs an extra discussion in sec. 10.1 as this method is the only method that reaches a good performance for all parameters, however for the weighted statistic parameters the error bars are extremely wide, which may result in unstable results.

Temporal Performance Analysis

As mentioned before, the optimized performance analysis gives an upper border of the performance that is not always valid if the methods are compared to those results where the parameter is fixed over the 80 evaluation sequences. Therefore the temporal performance analysis gives a better view on the methods behavior, as it represents the realistic situation.

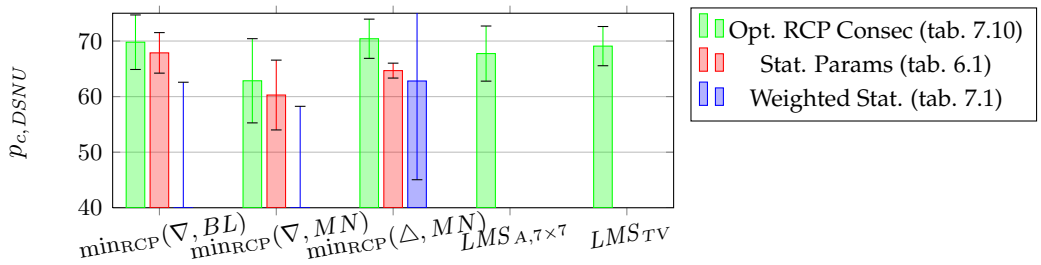
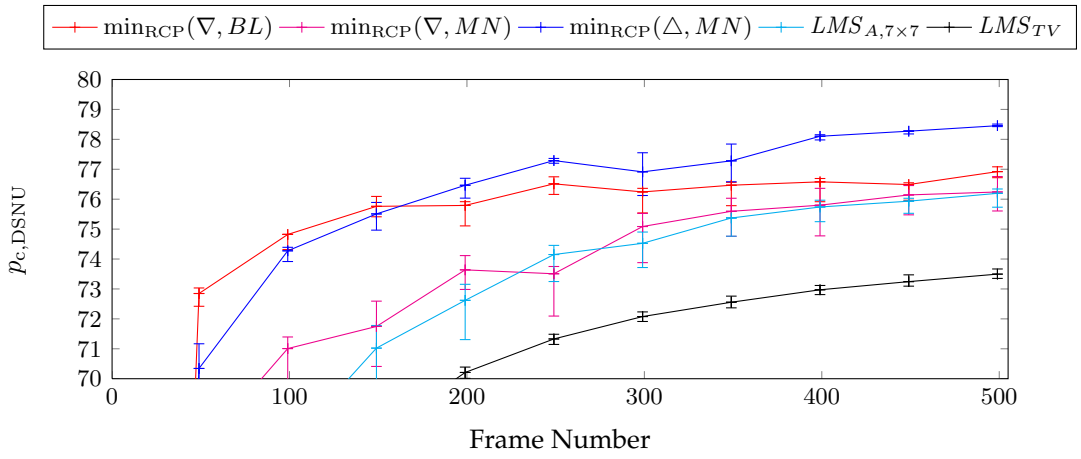


Figure 7.16.: Optimal Performance for the RCP methods with respect to the $p_{c,DSNU}$ quality metric and consecutive frame sequences

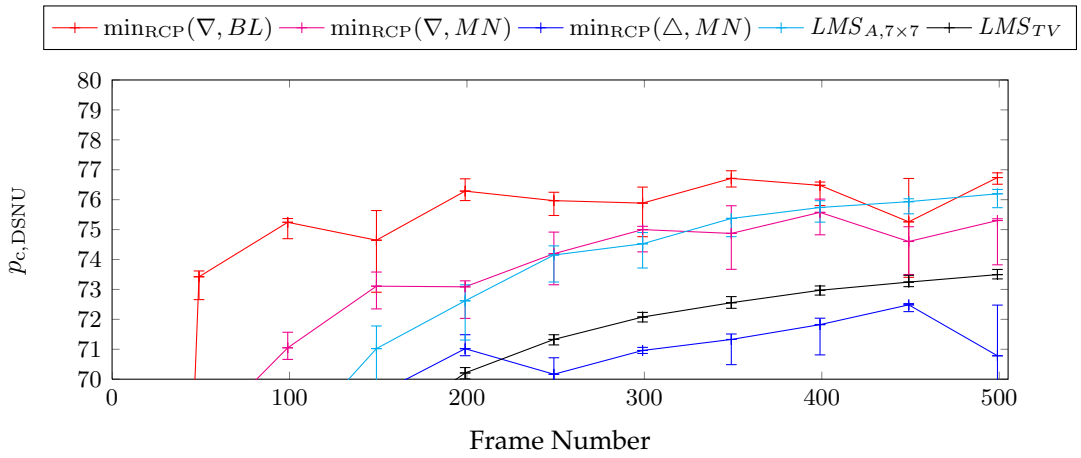
Fig. 7.17a shows temporal evaluations for the random frame sequences. While the $\min_{RCP}(\nabla, BL)$ and $\min_{RCP}(\nabla, MN)$ methods outperform the reference methods in the beginning, they converge towards the performance of the reference methods after 500 frames, however they still remain slightly above the performance of the $LMS_{A,7\times7}$ method. Surprisingly the $\min_{RCP}(\triangle, MN)$ method reaches the best performance of all methods which is founded in its stability towards parameter changes that has already been discussed. Fig. 7.17b shows a repetition of the above experiment with the statistic parameters that consider the weighted average (tab. 7.1). As expected from the optimal performance analysis, this parameter set decreases the performance of all methods and even the $\min_{RCP}(\triangle, MN)$ method cannot compensate this parameter change and results as the worst method of this comparison. The $\min_{RCP}(\nabla, BL)$ and $\min_{RCP}(\nabla, MN)$ methods however still reach the performance of the $LMS_{A,7\times7}$ method. The actual correction performances still reach the $70\,\%$ mark which means that a visually pleasing correction is expected even with this parameter choice for all methods. The temporal evaluations are also shown in the other quality norms in appendix E.3 for completeness For the consecutive frame sequences and the image material with low valued nonuniformities the experiments are presented in sec. 10.1 and sec. 10.2.

Spatial and Spectral Performance Analysis

A look at the spatial correction performances for random frame sequences, using the optimal parameters (tab. 7.9), is shown in fig. 7.18 for the visual image corrections. As expected no visual difference between the methods' corrections and the correction with the ground truth parameter set can be observed, except for the $\min_{RCP}(\nabla, MN)$ method, which shows an overlay of a high frequency noise pattern. With respect to the remaining DSNU patterns, the different qualities of the methods separate much better (see. fig. 7.19). Especially the $\min_{RCP}(\nabla, MN)$ method shows the high frequency pattern, which results from the corner frequencies that are not detected by its filter mask (compare fig. 6.2b, p. 123). The $\min_{RCP}(\triangle, MN)$ method on the other hand shows much less connected patches of high amplitude, which is the expected result from the proof of concept experiments above.



(a) Parameter set: Opt. RCP Random (tab. 7.9)



(b) Parameter set: Weighted Stat. (tab. 7.1)

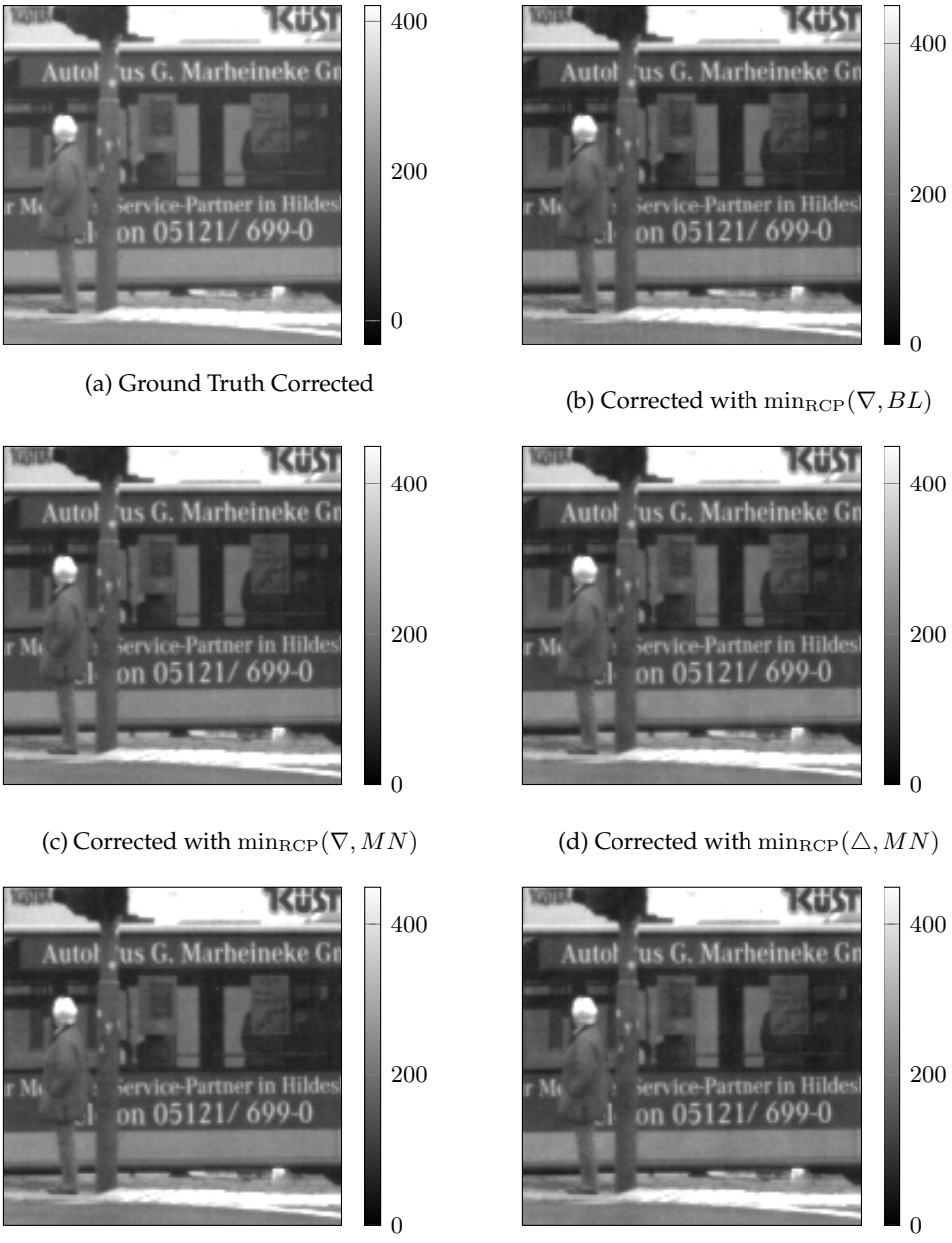
Figure 7.17.: Evaluations for random frame sequences with experimentally optimized parameters (fig. 7.17a) and the parameter set from the weighted statistic histogram (fig. 7.17b) in the $p_{\mathrm{c,DSNU}}$ metric.

In fig. 7.20, the mentioned spectral effects become visible and the separate damping of the horizontal and vertical frequencies are visible as the black cross in the middle of the representation. Compared to the $LMS_{A,7\times7}$ and LMS_{TV} methods, this effect comes very clear as the low frequency components are not damped to the extend that was reached by the former non-RCP methods. The high frequency components of the $\min_{\mathrm{RCP}}(\nabla, MN)$ method are also visible in this representation and fit the characteristics of this method's high pass filter. The $\min_{\mathrm{RCP}}(\nabla, MN)$ method is however the only method that concentrates the remaining low frequency remains on a smaller area. However the spatial high frequency components may render this method useless for practical evaluations.

7.3.4. First Conclusions

The row- and column-wise DSNU components were the last not addressed problem of the new maximum likelihood methods. The extension to a combined row-, column- and pixel-wise (RCP) maximum likelihood method shows that it is possible to correct these DSNU components successfully. Furthermore it is possible to correct separately only row- or column-wise nonuniformities with this approach, which reduces the problem into a one dimension of offset parameters. Only the appearance of remaining spatial high frequency components for the (∇, MN) choice forms a yet unsolved problem, but the other variants of the method did not show this problem.

Considering the DSNU-only correction, this chapter has provided several improvements to the bare maximum likelihood method introduced in chapter 6. With the current state of development an extension to a combined PRNU and DSNU correction method is next and follows in chapter 8.



(a) Ground Truth Corrected

(b) Corrected with $\min_{\mathrm{RCP}}(\nabla, BL)$

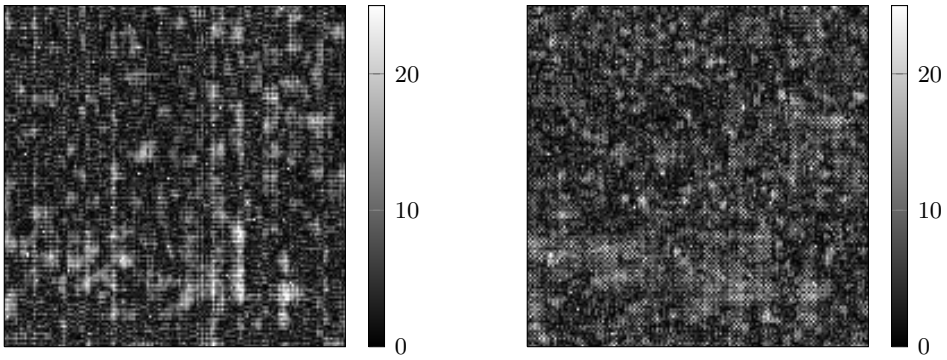
(c) Corrected with $\min_{\mathrm{RCP}}(\nabla, MN)$

(d) Corrected with $\min_{\mathrm{RCP}}(\triangle, MN)$

(e) Corrected with $LMS_{A,7\times7}$ Method

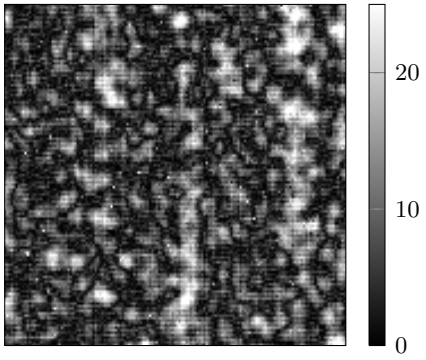
(f) Corrected with LMS_{TV} Method

Figure 7.18.: Corrected images of the introduced methods for random frame sequences (Parameters according to tab. 7.9)

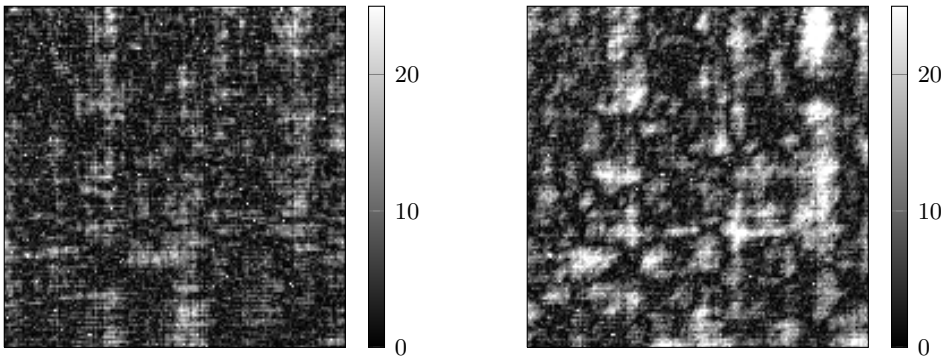


(a) $|\{b_{i,j,rem.}\}|$ for $\min_{\mathrm{RCP}}(\nabla, BL)$

(b) $|\{b_{i,j,rem.}\}|$ for $\min_{\mathrm{RCP}}(\nabla, MN)$



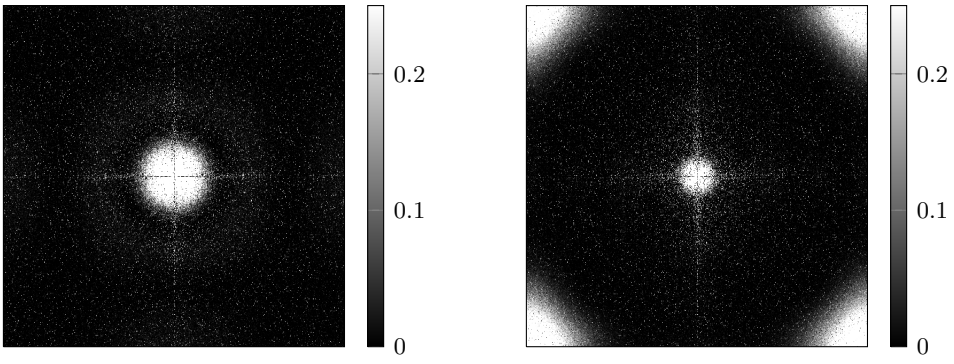
(c) $|\{b_{i,j,rem.}\}|$ for $\min_{\mathrm{RCP}}(\triangle, MN)$



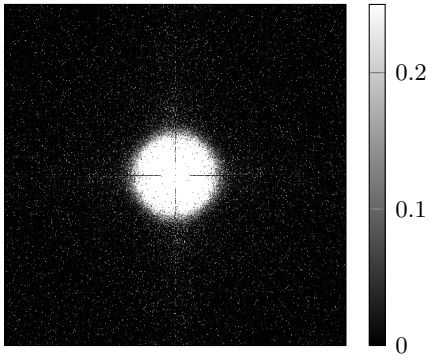
(d) $|\{b_{i,j,\mathrm{rem.}}\}|$ for $LMS_{A,7\times7}$

(e) $|\{b_{i,j,\mathrm{rem.}}\}|$ for LMS_{TV}

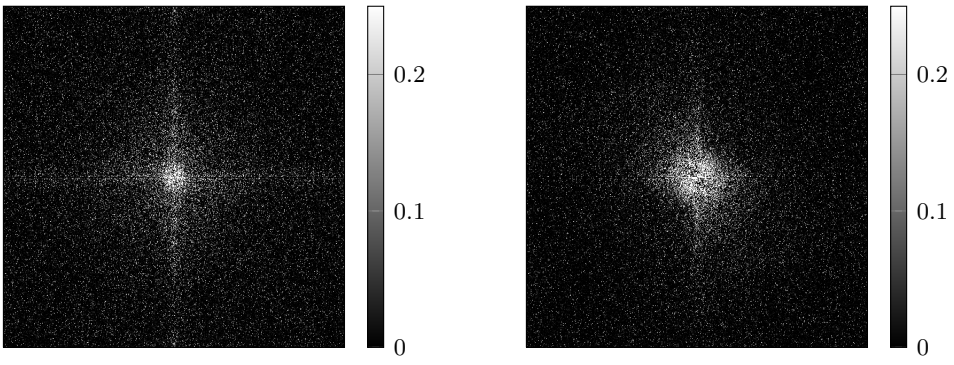
Figure 7.19.: Remaining DSNU patterns for the introduced methods and random frame sequences (Parameters according to tab. 7.9)



(a) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $\min_{\text{RCP}}(\nabla, BL)$ (b) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $\min_{\text{RCP}}(\nabla, MN)$



(c) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $\min_{\text{RCP}}(\triangle, MN)$



(d) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $LMS_{A,7\times7}$ (e) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for LMS_{TV}

Figure 7.20.: Spectral damping factors for the introduced methods and random frame sequences (Parameters according to tab. 7.9)

8. Extension to a Combined PRNU and DSNU Estimation

The current state of development for the new maximum likelihood estimation method corrects successfully the DSNU and allows to extract the free parameters from the statistics of the scene content. The next step to fully complete the new method is an extension into a combined PRNU and DSNU nonuniformity estimation as introduced in this chapter. The contents of this chapter have been partly published during the research [10, 13].

8.1. The Generalized Mathematical Ansatz

To extend the method, the approach of chapter 6 is repeated, but this time the sensor gain factors $\{a_{i,j}\}$ are included in the used sensor model. Therefore the nonuniformities are now fully described by a set of parameters that contain the gain variables $\{a_{i,j}\}$ and the offset variables $\{b_{i,j}\}$. This combined set is written as: $\{a_{i,j}, b_{i,j}\}$. Following the derivations in chapter 6, a multivariate random variable $\mathbf{NU}$ that maps all the possible realizations for the nonuniformities is introduced and has the probability density $f_{\mathbf{NU}}(\{a_{i,j}, b_{i,j}\})$.

The maximum likelihood estimation can then be rewritten to maximize the conditional probability of $\mathbf{NU}$ given the occurrence of a the spatio-temporal light signal I_{World}:

$$\arg \max_{\{a_{i,j},b_{i,j}\}} f_{\mathbf{NU}|\mathbf{I}_{\mathrm{World}}=I_{\mathrm{World}}}(\{a_{i,j}, b_{i,j}\}) \,, \tag{8.1}$$

which transfers with the Bayes' Theorem directly into:

$$\arg \max_{\{b_{i,j}\}} \frac{1}{N} \cdot f_{\mathbf{NU}}(\{a_{i,j}, b_{i,j}\}) \cdot f_{\mathbf{I}_{\mathrm{World}}|\mathbf{NU}=\{a_{i,j},b_{i,j}\}}(I_{\mathrm{World}}) \,. \tag{8.2}$$

8.2. Density Estimations of the Defined Random Variables

As next step, the random variables from eq. 8.2 need to be defined, which means, that a probability density as well as a state space are needed. These definitions follow directly the ideas of sec. 6.2.

8.2.1. The density of the random variable NU

The definition of the multivariate random variable $\mathbf{NU}$ is following the derivation of $\mathbf{DSNU}$ with the difference that the parameters $\{a_{i,j}, b_{i,j}\}$ are used instead of only $\{b_{i,j}\}$. First, the state space is given and, following the procedure in sec. 6.2.1, the state space is named Ω_{NU} and covers all the possible realizations given by the parameters $\{a_{i,j}, b_{i,j}\}$. The parameters $b_{i,j}$ are considered the realizations of their corresponding random variables $\mathbf{b}_{i,j}$ and in extension, the parameters $a_{i,j}$ are considered the realizations of their random variables $\mathbf{a}_{i,j}$. This allows to define a multivariate random variable [29]:

$$\mathbf{NU} : \Omega \to \mathbb{R}^{2m \times n} \tag{8.3}$$

$$(\mathbf{b}_{1,1}, \ldots, \mathbf{b}_{M,N}, \mathbf{a}_{1,1}, \ldots, \mathbf{a}_{M,N}) \to (b_{1,1}, \ldots, b_{m,n}, a_{1,1}, \ldots, a_{M,N})^T \tag{8.4}$$

To finish the definition of $\mathbf{NU}$, its probability density needs to be defined. According to sec. 6.2.1, the parameters $\{b_{i,j}\}$ are assumed to be independent and identically distributed (iid). This assumption is now extended onto the parameters $\{a_{i,j}\}$, which shall be independent and identically distributed as well. Furthermore, the parameters $\{a_{i,j}\}$ are considered independent from all the parameters $\{b_{i,j}\}$, which consequently leads to the independent set of random variables $\{\mathbf{a}_{i,j}, \mathbf{b}_{i,j}\}$.

From the statistical mathematics it follows that a product over the single independent densities of the variables $\mathbf{a}_{i,j}$ and $\mathbf{b}_{i,j}$ forms the combined probability density $f_{\mathbf{NU}}$:

$$f_{\mathbf{NU}}(\{a_{i,j}, b_{i,j}\}) = \prod_{i,j} f_{\mathbf{a}_{i,j}}(a_{i,j}) f_{\mathbf{b}_{i,j}}(b_{i,j})\,. \tag{8.5}$$

Again, the manufacturing process of the individual light sensors, is used to motivate the above assumptions of the iid distributions for $\{\mathbf{a}_{i,j}\}$ and $\{\mathbf{b}_{i,j}\}$.

Given the iid assumption, a random sample of the realizations out of the $\{b_{i,j}\}$ already showed successfully that the $\{\mathbf{b}_{i,j}\}$ obey a Gaussian distribution. From the measured ground truth distribution of the $\{a_{i,j}\}$ a Gaussian distribution may fit as well, but does not hold as strict as for the $b_{i,j}$. The actual histograms of the $\{a_{i,j}\}$ have been discussed in sec. 4.3. However, the mathematical advantages of a Gaussian distribution in the later log-likelihood minimization process lead to the choice of a Gaussian model, knowing that it is not the best fit. Consequently, the random variables for each given (i, j) obey a Gaussian distribution, with the same parameters μ_b and σ_b for all $\mathbf{b}_{i,j}$ and the parameters μ_a and σ_a for all $\mathbf{a}_{i,j}$:

$$\mathbf{b}_{i,j} \sim \mathcal{N}\left(\mu_b, \sigma_b^2\right) \quad \text{with} \quad f_{\mathbf{b}_{i,j}}(b_{i,j}) = \frac{1}{\sigma_b\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{b_{i,j}-\mu_b}{\sigma_b}\right)^2} \quad \forall (i,j) \tag{8.6}$$

$$\mathbf{a}_{i,j} \sim \mathcal{N}\left(\mu_a, \sigma_a^2\right) \quad \text{with} \quad f_{\mathbf{a}_{i,j}}(a_{i,j}) = \frac{1}{\sigma_a\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{a_{i,j}-\mu_a}{\sigma_a}\right)^2} \quad \forall (i,j) \tag{8.7}$$

The definition of the random variable **NU** is therefore completed.

8.2.2. The density of $\mathbf{I}_{\mathrm{World}}|\mathbf{NU}=\{a_{i,j}, b_{i,j}\}$

The conditional random variable $\mathbf{I}_{\mathrm{World}}|\mathbf{NU}=\{a_{i,j}, b_{i,j}\}$ states the probability of the world light signal I_{World} to occur, given a realization of the nonuniformities. As discussed in section 6.2.2, the realization of the world signal I_{World} can be approximated by the reconstructed signal I_{Reco}. The reconstruction formula given in eq. 2.28 leads to an I_{Reco} that depends on the PRNU and DSNU correction patterns $\{a_{i,j}\}$ and $\{b_{i,j}\}$. The rest of the argumentation goes along as in sec. 6.2.2, and leads to an approximation of the conditional random variable by:

$$I_{\mathrm{Reco}}(\{a_{i,j}, b_{i,j}\}) \approx I_{\mathrm{World}} \tag{8.8}$$

$$\mathbf{I}_{\mathrm{World}}|\mathbf{NU}=\{a_{i,j}, b_{i,j}\} \approx \mathbf{I}_{\mathrm{Reco}}|\mathbf{NU}=\{a_{i,j}, b_{i,j}\}\,. \tag{8.9}$$

The argumentation continues by defining the probability $\mathbf{I}_{\mathrm{Reco}}|\mathbf{NU}=\{a_{i,j}, b_{i,j}\}$ with help of gradient- and edge-based approximations (compare sec. 6.2.2). These features depend only on the corrected sample values C_S, and their definition is therefore identical to eq. 6.45 (p. 120) with the only difference that the C_{S} are calculated from the measurements C_M by correcting for the PRNU parameters $\{a_{i,j}\}$ as well. The definition of the random variable is then completed as:

$$f_{\mathbf{I}_{\mathrm{Reco}}|\mathbf{NU}=\{a_{i,j},b_{i,j}\}} = \prod_{i,j}\prod_{t_n} f_{\mathbf{F}_{i,j,t_n}}(F_{i,j,t_n}) \tag{8.10}$$

$$\text{with} \quad F_{i,j,t_n}(a,b) = [C_{\mathrm{S}}(a,b,t_n) \otimes (C_{\mathrm{S}}(a,b,t_n) * HP(\hat{L},\Phi)]_{i,j} \tag{8.11}$$

The definition and spectral behavior of the *HP* filter masks have already been discussed in sec. 6.2.4 and do not change. Of course, the above definition is bound to the same assumptions as in chapter 6, that states that the feature variables are assumed to be independent and identically distributed. The experimental chapters showed already that random frame input, or very slowly captured frames seem to fulfill this assumption and result in the best possible correction rates.

8.3. Solving as Log-Likelihood Energy Minimization

With all the assumptions and definitions given above, the maximum likelihood estimation is fully defined by:

$$\arg\max_{\{a_{i,j},b_{i,j}\}} \frac{1}{N} \cdot f_{\mathbf{NU}}(\{a_{i,j}, b_{i,j}\}) \cdot f_{\mathbf{I}_{\mathrm{World}}|\mathbf{NU}=\{a_{i,j},b_{i,j}\}}(I_{\mathrm{World}}) \tag{8.12}$$

$$\approx \arg\max_{\{a_{i,j},b_{i,j}\}} \frac{1}{N} \prod_{i,j} f_{\mathbf{b}_{i,j}}(b_{i,j}) f_{\mathbf{a}_{i,j}}(a_{i,j}) \cdot \prod_{i,j}\prod_{t_n} f_{\mathbf{F}_{i,j,t_n}} \tag{8.13}$$

$$\approx \arg \max_{\{a_{i,j},b_{i,j}\}} \frac{1}{N} \prod_{i,j} \frac{1}{\sigma_b\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{b_{i,j}-\mu_b}{\sigma_b}\right)^2} \frac{1}{\sigma_a\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{a_{i,j}-\mu_a}{\sigma_a}\right)^2} \cdot \prod_{i,j}\prod_{t_n} \lambda e^{-\lambda F_{i,j,t_n}} \,. \tag{8.14}$$

To simplify the maximization, eq. 8.12 is transferred into the corresponding log-likelihood energy minimization task, which consists of a prior knowledge energy term E_P and a data energy term E_D:

$$\arg\min_{\{b_{i,j}\}} \{ \underbrace{-\log\frac{1}{N}}_{\text{const.}} \underbrace{-\log f_{\mathbf{NU}}(\{a_{i,j},b_{i,j}\})}_{E_\mathrm{P}} \underbrace{-\log f_{\mathbf{I}_\mathrm{World}|\mathbf{NU}=\{a_{i,j},b_{i,j}\}}(I_\mathrm{World})}_{E_\mathrm{D}} \} \tag{8.15}$$

$$\arg\min_{\{b_{i,j}\}} \{E_\mathrm{P} + E_\mathrm{D} + \text{const.}\} \tag{8.16}$$

With the definitions of the eq. 8.5f. the energy E_P can be simplified to:

$$\begin{aligned} E_\mathrm{P} &= -\log f_{\mathbf{NU}}(\{a_{i,j},b_{i,j}\}) && (8.17)\\ &= -\log \prod_{i,j} f_{\mathbf{a}_{i,j}}(a_{i,j}) f_{\mathbf{b}_{i,j}}(b_{i,j}) && (8.18)\\ &= \sum_{i,j} \underbrace{\frac{1}{2\sigma_b^2}}_{\beta} (b_{i,j}-\mu_b)^2 + \sum_{i,j} \underbrace{\frac{1}{2\sigma_a^2}}_{\alpha} (a_{i,j}-\mu_a)^2 + \text{const.} && (8.19)\\ &= \beta \sum_{i,j} (b_{i,j}-\mu_b)^2 + \alpha \sum_{i,j} (a_{i,j}-\mu_a)^2 + \text{const.}\,, && (8.20) \end{aligned}$$

with the intermediate transformations shown in appendix F.1.

For E_D the needed transformations are identical to sec. 6.3 p. 124 if the definition of F_{i,j,t_n} is modified with a $C_{\mathrm{S},i,j}$ according to the sensor model of eq. 2.27 that includes the $\{a_{i,j}\}$ parameters:

$$E_D = \underbrace{\sum_{i,j}\sum_{t_n} -\log\lambda}_{\mathrm{C_{log}}} + \sum_{i,j}\sum_{t_n} \lambda F_{i,j,t_n} \tag{8.21}$$

$$= \lambda \sum_{t_n} \sum_{(i,j)\in S} \frac{1}{a_{i,j}} (C_{\mathrm{M},i,j} - b_{i,j} - \chi_{i,j}) \sum_{(k,l)\in S} \frac{1}{a_{k,l}} (C_{\mathrm{M},k,l} - b_{k,l} - \chi_{k,l}) \Psi(\cdot) + c_\mathrm{log} \tag{8.22}$$

Using this notation in addition with the Hadamard (element-wise) multiplication $\otimes$ and element wise division $\oslash$, the definition of E_D can be expressed in matrix notation as:

$$E_\mathrm{D} = \lambda \sum_{i,j}\sum_{t_n} [C_\mathrm{S} \otimes (C_\mathrm{S} * HP(\hat{L},\Phi)]_{i,j} + c_\mathrm{log} \tag{8.23}$$

$$= \lambda \sum_{t_n} \sum_{i,j} \left[(C_\mathrm{M} - b - \chi) \oslash a \otimes \left(((C_\mathrm{M} - b - \chi) \oslash a) * HP(\hat{L}, \Phi) \right) \right]_{i,j} + c_\mathrm{log} \tag{8.24}$$

The energies E_P and E_D are now functions of the set of variables $\{a_{i,j}, b_{i,j}\}$ and to be able to solve the energy minimization task with help of gradient based iterative solvers, the gradient in the variables $\{a_{i,j}\}$ and $\{b_{i,j}\}$ needs to be calculated. The gradients for the two sets of parameters $\{b_{i,j}\}$ and $\{a_{i,j}\}$ are concatenated below each other as given in the following definitions:

$$\begin{pmatrix} \nabla_{\{b_{i,j}\}}(\mathrm{const.} + E_\mathrm{P} + E_\mathrm{D}) \\ \nabla_{\{a_{i,j}\}}(\mathrm{const.} + E_\mathrm{P} + E_\mathrm{D}) \end{pmatrix} = \begin{pmatrix} \nabla_{\{b_{i,j}\}} E_\mathrm{P} + \nabla_{\{b_{i,j}\}} E_\mathrm{D} \\ \nabla_{\{a_{i,j}\}} E_\mathrm{P} + \nabla_{\{a_{i,j}\}} E_\mathrm{D} \end{pmatrix} \tag{8.25}$$

$$\text{with:} \quad \nabla_{\{b_{i,j}\}} = (\partial_{b_{1,1}} \dots \partial_{b_{1,N}}, \partial_{b_{2,1}} \dots \partial_{b_{m,N}})^T \tag{8.26}$$

$$\text{and with:} \quad \nabla_{\{a_{i,j}\}} = (\partial_{a_{1,1}} \dots \partial_{a_{1,N}}, \partial_{a_{2,1}} \dots \partial_{a_{m,N}})^T \tag{8.27}$$

8.3.1. The derivation of $\nabla_{\{b_{i,j}\}} E_\mathrm{P}, \nabla_{\{b_{i,j}\}} E_\mathrm{D}$ and $\nabla_{\{a_{i,j}\}} E_\mathrm{P}$

The calculation of the gradients of the different energies and variables is done separately for $\nabla_{\{b_{i,j}\}} E_\mathrm{P}, \nabla_{\{b_{i,j}\}} E_\mathrm{D}, \nabla_{\{a_{i,j}\}} E_\mathrm{P}$ and $\nabla_{\{a_{i,j}\}} E_\mathrm{D}$. For the prior energy the derivation $\nabla_{\{b_{i,j}\}} E_\mathrm{P}$ does not change in comparison to eq. 6.70, p.126, as the $\{a_{i,j}\}$ dependencies are constants with respect to the $\{b_{i,j}\}$ variables. For $\nabla_{\{a_{i,j}\}} E_\mathrm{P}$, the mathematical derivations are a direct symmetry in the mathematical expression to $\nabla_{\{b_{i,j}\}} E_\mathrm{P}$ as just the variable is exchanged. Using again the $\hat{\nabla}$ operator, which puts the derivative of the (i,j)'s variable as the (i,j)'s element of a matrix, one obtains the derivatives in the following matrix notation:

$$\hat{\nabla}_{\{b_{i,j}\}} E_\mathrm{P} = 2\beta(b - \mu_b) \tag{8.28}$$

$$\hat{\nabla}_{\{a_{i,j}\}} E_\mathrm{P} = 2\alpha(a - \mu_a) \tag{8.29}$$

The gradients of the data energy can also be calculated very similar to the mathematical strategy in sec. 6.70. For $\nabla_{\{b_{i,j}\}} E_\mathrm{P}$ the only difference is the consideration of the new (constant) parameters $\{a_{i,j}\}$. The intermediate steps are shown in the appendix F.2 and lead to:

$$\begin{aligned} \nabla_{\{b_{i,j}\}} E_\mathrm{D} &= \nabla_{\{b_{i,j}\}} \sum_{i,j,t_n} \lambda (C_\mathrm{M} - b - \chi) \oslash a \otimes \left(((C_\mathrm{M} - b - \chi) \oslash a) * HP(\hat{L}, \Phi) \right) \\ \partial_{b_{m,n}} E_\mathrm{D} &= \partial_{b_{m,n}} (\sum_{t_n} \sum_{i,j} \lambda (C_\mathrm{M} - b - \chi) \oslash a \otimes ((C_\mathrm{M} - b - \chi) \oslash a * HP(\hat{L}, \Phi))) \\ &= -2\lambda \sum_{t_n} \frac{1}{a_{m,n}} \sum_{(i,j) \in S} \frac{1}{a_{i,j}} (C_{\mathrm{M},i,j} - b_{i,j} - \chi_{i,j}) \Psi(|i-m|, |j-n|)) \end{aligned} \tag{8.30}$$

The temporal normalizing factor T_N according to eq. 6.73, allows again to express some quantities as the temporal average, or the temporal expectation value of the corresponding random variables:

$$\partial_{b_{m,n}} E_\mathrm{D} = -2\lambda T_\mathrm{N} \frac{1}{T_\mathrm{N}} \sum_{t_n} \frac{1}{a_{m,n}} \sum_{(i,j)\in S} \frac{1}{a_{i,j}} (C_{\mathrm{M},i,j} - b_{i,j} - \chi_{i,j}) \Psi(|i-m|, |j-n|)) \tag{8.31}$$

$$= -2\lambda T_\mathrm{N} \frac{1}{a_{m,n}} \sum_{(i,j)\in S} \frac{1}{a_{i,j}} (\underbrace{\frac{1}{T_\mathrm{N}} \sum_{t_n} C_{\mathrm{M},i,j}}_{A_{\mathrm{M},i,j}} - \underbrace{\frac{1}{T_\mathrm{N}} \sum_{t_n} b_{i,j}}_{b_{i,j}} - \underbrace{\frac{1}{T_\mathrm{N}} \sum_{t_n} \chi_{i,j}}_{\to 0}) \Psi(\cdot)) \tag{8.32}$$

$$= -2\lambda T_\mathrm{N} \frac{1}{a_{m,n}} \sum_{(i,j)\in S} \underbrace{\frac{1}{a_{i,j}} (A_{\mathrm{M},i,j} - b_{i,j})}_{A^{\mathrm{corr.}}_{\mathrm{M},i,j}} \Psi(|i-m|, |j-n|)) \tag{8.33}$$

The newly introduced variable $A^{\mathrm{corr.}}_{\mathrm{M},i,j}$ contains the corrected averaged measurements. The derivative again becomes the convolution with the high pass filter defined by Ψ and with help of the $\hat{\nabla}_{\{b_{i,j}\}}$, the gradient expresses in matrix notation as

$$\hat{\nabla}_{\{b_{i,j}\}} E_\mathrm{D} = -2\lambda T_\mathrm{N} \big(A^{\mathrm{corr.}}_\mathrm{M} * HP(\hat{L}, \Phi) \big) \oslash a \tag{8.34}$$

The only difference to the $\hat{\nabla}_{\{b_{i,j}\}} E_\mathrm{D}$ in the DSNU-only correction is a division of each gradient index by the corresponding $a_{i,j}$ parameter.

8.3.2. The derivation of $\nabla_{\{a_{i,j}\}} E_\mathrm{P}$

A more challenging task is the derivation of $\nabla_{\{a_{i,j}\}} E_\mathrm{P}$. It starts with the same strategy as before by splitting the sum and using the chain rule with the intermediate steps shown in appendix F.2.2:

$$\nabla_{\{a_{i,j}\}} E_\mathrm{D} = \nabla_{\{a_{i,j}\}} \sum_{t_n} \sum_{i,j} \lambda (C_\mathrm{M} - b - \chi) \oslash a \otimes \big(((C_\mathrm{M} - b - \chi) \oslash a) * HP(\hat{L}, \Phi) \big)$$

$$\partial_{a_{m,n}} E_\mathrm{D} = \partial_{a_{m,n}} \Big(\sum_{t_n} \sum_{i,j} \lambda (C_\mathrm{M} - b - \chi) \oslash a \otimes \big(((C_\mathrm{M} - b - \chi) \oslash a) * HP(\hat{L}, \Phi) \big) \Big)$$

$$= -2\lambda T_\mathrm{N} \frac{1}{T_\mathrm{N}} \sum_{t_n} \Bigg(\frac{1}{a^2_{\mathrm{M},m,n} a_{k,l}} \sum_{(k,l)\in S} (C_{\mathrm{M},m,n} - b_{m,n} - \chi_{m,n}) \tag{8.35}$$

$$\cdot (C_{\mathrm{M},k,l} - b_{k,l} - \chi_{k,l}) \Psi(|k-m|, |l-n|) \Bigg) \tag{8.36}$$

Again the temporal averaging can be used to achieve further simplifications:

$$\begin{aligned}\partial_{a_{m,n}} E_{\mathrm{D}} = -2\lambda T_{\mathrm{N}}\Big(\frac{1}{a_{\mathrm{M},n}^2 a_{k,l}}\Big(\sum_{(k,l)\in S} \underbrace{\frac{1}{T_{\mathrm{N}}}\sum_{t_n} C_{\mathrm{M},m,n} C_{\mathrm{M},k,l}}_{A^2_{\mathrm{M},m,n,k,l}} - \underbrace{\frac{1}{T_{\mathrm{N}}}\sum_{t_n} C_{\mathrm{M},m,n} b_{k,l}}_{A_{\mathrm{M},m,n} b_{k,l}} \\ - \underbrace{\frac{1}{T_{\mathrm{N}}}\sum_{t_n} C_{\mathrm{M},m,n}\chi_{k,l}}_{\to 0} - \underbrace{\frac{1}{T_{\mathrm{N}}}\sum_{t_n} b_{m,n} C_{\mathrm{M},k,l}}_{b_{m,n} A_{\mathrm{M},k,l}} + \underbrace{\frac{1}{T_{\mathrm{N}}}\sum_{t_n} b_{m,n} b_{k,l}}_{b_{m,n} b_{k,l}} \\ + \underbrace{\frac{1}{T_{\mathrm{N}}}\sum_{t_n} b_{m,n}\chi_{k,l}}_{\to 0} - \underbrace{\frac{1}{T_{\mathrm{N}}}\sum_{t_n} \chi_{m,n} C_{\mathrm{M},k,l}}_{\to 0} + \underbrace{\frac{1}{T_{\mathrm{N}}}\sum_{t_n} \chi_{m,n} b_{k,l}}_{\to 0} \\ + \underbrace{\frac{1}{T_{\mathrm{N}}}\sum_{t_n} \chi_{m,n}\chi_{k,l}}_{\to 0}\Big)\Psi(\cdot)\Big)\end{aligned} \tag{8.37}$$

For the product of $\chi_{m,n}\chi_{k,l}$, each position's random variable χ is considered independent. Then it follows that the expectation value of the product of two independent variables is the product of their expectation values. In the special case of $(m,n) = (k,l)$ the variable gets squared and the further assumption for χ is made that it's coefficient of variation is $\gamma(\chi_{i,j}) = 0$ (e.g. a symmetric probability density distribution [29]). Mathematically expressed, this yields to the following:

$$\frac{1}{T_{\mathrm{N}}}\sum_{t_n} \chi_{m,n,t}\chi_{k,l,t} \approx \mu_t(\chi_{m,n,t}\chi_{k,l,t}) = \mu_t(\chi_{m,n,t})\mu_t(\chi_{k,l,t}) = 0 \cdot 0 \tag{8.38}$$

$$\frac{1}{T_{\mathrm{N}}}\sum_{t_n} \chi_{i,j,t}\chi_{i,j,t} = \frac{1}{T_{\mathrm{N}}}\sum_{t_n} \chi^2_{i,j,t} \approx \mu_t(\chi^2_{i,j,t}) = (\mu_t(\chi_{i,j,t}))^2(1+\gamma(\chi_{i,j,t})) = 0\,. \tag{8.39}$$

The above derivative can then be further simplified, and the intermediate steps shown in appendix F.2.2, lead to:

$$\begin{aligned}\partial_{a_{m,n}} E_{\mathrm{D}} &= -2\lambda T_{\mathrm{N}}\Big(\frac{1}{a_{\mathrm{M},n}^2 a_{k,l}}\Big(\sum_{(k,l)\in S} A^2_{\mathrm{M},m,n,k,l} - A_{\mathrm{M},m,n} b_{k,l} \\ &\qquad - b_{m,n} A_{\mathrm{M},k,l} + b_{m,n} b_{k,l}\Big)\Psi(\cdot)\Big) \qquad (8.40) \\ &= -2\lambda T_{\mathrm{N}}\Big(\frac{1}{a_{\mathrm{M},n}^2 a_{k,l}}\Big(\sum_{(k,l)\in S} A^2_{\mathrm{M},m,n,k,l} - A_{\mathrm{M},m,n} A_{\mathrm{M},k,l} \\ &\qquad + ((A_{\mathrm{M},m,n} - b_{m,n})(A_{\mathrm{M},k,l} - b_{k,l}))\Psi(\cdot)\Big) \qquad (8.41)\end{aligned}$$

At this point the averaged measurements $A_{\mathrm{M},m,n}$ and $A_{\mathrm{M},k,l}$ are considered as the *temporal* expectation values $\mu_t(C_{\mathrm{M},m,n})$ and $\mu_t(C_{\mathrm{M},k,l})$. $A^2_{\mathrm{M},m,n,k,l}$ is then by definition the expectation value of the product of these two variables $A^2_{\mathrm{M},m,n,k,l} =$

$\mu_t(C_{\mathrm{M},m,n}C_{\mathrm{M},k,l})$ and combined these two parts form the definition of the covariance cov:

$$\mathrm{cov}(A,B) = \mu(AB) - \mu(A)\cdot\mu(B) \tag{8.42}$$

$$\mathrm{cov}(C_{\mathrm{M},m,n}, C_{\mathrm{M},k,l}) = \mu_t(C_{\mathrm{M},m,n}C_{\mathrm{M},k,l}) - \mu_t(C_{\mathrm{M},m,n})\cdot\mu_t(C_{\mathrm{M},k,l}) \tag{8.43}$$

$$= A^2_{\mathrm{M},m,n,k,l} - A_{\mathrm{M},m,n}\cdot A_{\mathrm{M},k,l}\,. \tag{8.44}$$

The transformation then continues with:

$$\begin{aligned}\partial_{a_{m,n}} E_{\mathrm{D}} = -2\lambda T_{\mathrm{N}}\Big(&\frac{1}{a^2_{\mathrm{M},n}a_{k,l}}\Big(\sum_{(k,l)\in S}\underbrace{A^2_{\mathrm{M},m,n,k,l} - A_{\mathrm{M},m,n}A_{\mathrm{M},k,l}}_{\mathrm{cov}(C_{\mathrm{M},m,n},C_{\mathrm{M},k,l})}\\ &+ ((A_{\mathrm{M},m,n}-b_{m,n})(A_{\mathrm{M},k,l}-b_{k,l}))\Psi(\cdot)\Big)\end{aligned} \tag{8.45}$$

$$\begin{aligned}= -2\lambda T_{\mathrm{N}}\Big(&\frac{1}{a_{m,n}}\Big(\frac{1}{a_{m,n}}\underbrace{\sum_{(k,l)\in S}\frac{1}{a_{k,l}}\mathrm{cov}(C_{\mathrm{M},m,n},C_{\mathrm{M},k,l})\Psi(|k-m|,|l-n|)}_{\Upsilon_{m,n}}\\ &+\sum_{(k,l)\in S}\underbrace{\frac{1}{a_{m,n}}(A_{\mathrm{M},m,n}-b_{m,n})}_{A^{\mathrm{corr.}}_{\mathrm{M},m,n}}\underbrace{\frac{1}{a_{k,l}}(A_{\mathrm{M},k,l}-b_{k,l})}_{A^{\mathrm{corr.}}_{\mathrm{M},k,l}}\Psi(\cdot)\Big)\end{aligned} \tag{8.46}$$

Υ is a new matrix with its elements defined by the definition of $\Upsilon_{m,n}$ given above. Finally $\nabla_{\{a_{i,j}\}}E_{\mathrm{P}}$ can be encoded with help of the $\hat{\nabla}$ operator as

$$\hat{\nabla}_{\{a_{i,j}\}}E_{\mathrm{P}} = -2\lambda T_{\mathrm{N}}((\Upsilon \oslash a) + (A^{\mathrm{corr.}}_{\mathrm{M}} \otimes A^{\mathrm{corr.}}_{\mathrm{M}} * HP(\hat{L},\Phi))) \oslash a\,, \tag{8.47}$$

and the whole gradient is given by the following expression:

$$\begin{aligned}&\begin{pmatrix}\hat{\nabla}_{\{a_{i,j}\}}(E_{\mathrm{P}}+E_{\mathrm{D}})\\ \hat{\nabla}_{\{b_{i,j}\}}(E_{\mathrm{P}}+E_{\mathrm{D}})\end{pmatrix}\\ =&\begin{pmatrix}-2\lambda T_{\mathrm{N}}((\Upsilon\oslash a) + A^{\mathrm{corr.}}_{\mathrm{M}}\otimes(A^{\mathrm{corr.}}_{\mathrm{M}} * HP(\hat{L},\Phi)))\oslash a + 2\alpha(a-\mu_a)\\ -2\lambda T_{\mathrm{N}}(A^{\mathrm{corr.}}_{\mathrm{M}} * HP(\hat{L},\Phi))\oslash a + 2\beta(b-\mu_b)\end{pmatrix}.\end{aligned} \tag{8.48}$$

The final equation that needs to be solved to find the PRNU and DSNU pattern is again given as the point where the gradient becomes 0:

$$\begin{pmatrix}-2\lambda T_{\mathrm{N}}((\Upsilon\oslash a) + (A^{\mathrm{corr.}}_{\mathrm{M}}\otimes A^{\mathrm{corr.}}_{\mathrm{M}} * HP(\hat{L},\Phi)))\oslash a + 2\alpha(a-\mu_a)\\ -2\lambda T_{\mathrm{N}}(A^{\mathrm{corr.}}_{\mathrm{M}} * HP(\hat{L},\Phi))\oslash a + 2\beta(b-\mu_b)\end{pmatrix} = 0\,. \tag{8.49}$$

Including the Weighted Average Improvement: Of course the introduced method can again be improved with help of the weighted averaged measurements

and the improved method can be summarized as a solution to the equation:

$$0 = \begin{pmatrix} -2\lambda T_{\mathrm{N}}((\Upsilon \oslash a) + (A^{\mathrm{corr.}}_{\mathrm{M}W} \otimes A^{\mathrm{corr.}}_{\mathrm{M}W} * HP(\hat{L}, \Phi))) \oslash a + 2\alpha(a - \mu_a) \\ -2\lambda T_{\mathrm{N}}(A^{\mathrm{corr.}}_{\mathrm{M}W} * HP(\hat{L}, \Phi)) \oslash a + 2\beta(b - \mu_b) \end{pmatrix} \tag{8.50}$$

$$A^{\mathrm{corr.}}_{\mathrm{M}W} = (A_{\mathrm{M}W} - b) \oslash a \tag{8.51}$$

$$A_{\mathrm{M}W,i,j} = \frac{1}{\sum\limits_{t_n} w_{i,j,t_n}} \sum_{t_n} w_{i,j,t_n} C_{\mathrm{M},i,j} \tag{8.52}$$

$$w_{i,j,t_n} = \frac{1}{1 + e(C^2_{\mathrm{M},N(i,j)}(t_n))} \tag{8.53}$$

Please be aware that for the calculation of the covariances the weighted averaged measurements can not be used.

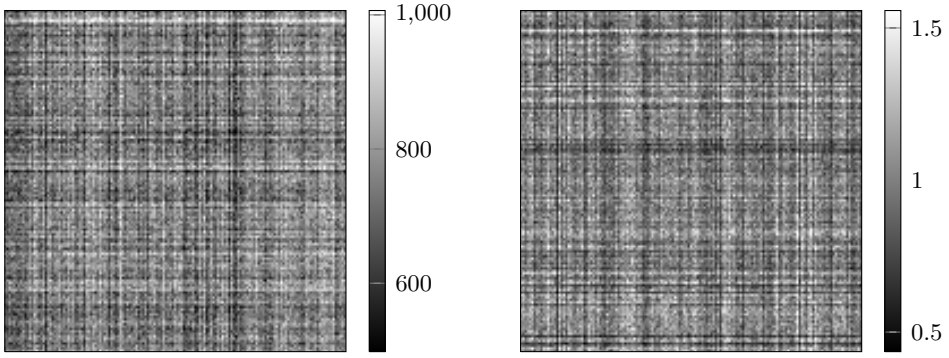
With the given gradient, this equation can be solved by a conjugate gradient method or any other iterative solver that exploits the gradient. However for this problem the starting point parameters should be close to the expectation values: $\{a_{i,j}\} = \mu_a$ and $\{b_{i,j}\} = \mu_b$. Because for too large values of $\{a_{i,j}\}$ the whole gradient converges towards 0 as well, which is a trivial and unwanted minimum. Therefore a solver with a built in save-guard that limits $\{a_{i,j}\} = \mu_a \pm n\sigma_a$ and $n = 5$ is used to avoid this problem.

8.4. Experiments and Results

8.4.1. Proof of Concept for the Combined Estimation

First results have been conducted with an artificially created nonuniformity, which has the advantage of showing PRNU image degradations as well. Such PRNU-based degradations are in general hardly visible with the ground truth PRNU patterns. For the simulation, a Gaussian row-, column-, and pixel-wise DSNU are added and a Gaussian row-, column-, and pixel-wise PRNU are multiplied. Fig. 8.1 shows the spatial representation of this artificial nonuniformities and due to the nature of its creation, it obeys perfectly to a Gaussian distribution, also for the PRNU part.

Fig. 8.2 shows the combined PRNU and DSNU corrections, with hand picked parameters and the choice of $HP = HP(\nabla, \mathrm{BL})$. The achieved DSNU correction percentages reach the combined correction $p_{\mathrm{c,Comb.}} = 49.02\,\%$ and for the PRNU and DSNU parts $p_{\mathrm{c,PRNU}} = 39.96\,\%$ and $p_{\mathrm{c,DSNU}} = 66.59\,\%$. These correction performances perform not as well as the optimized parameters, but for a proof of concept, the visual results of fig. 8.2 show a very good improvement of the image quality, for both degradations, the PRNU-caused and the DSNU-caused. The



(a) Artificial DSNU $= 85.31$ DN (b) Artificial PRNU $= 0.185\,\%$

Figure 8.1.: Spatial representation of the artificial nonuniformities.

remaining DSNU and PRNU patterns leave some horizontal and vertical frequencies behind, which is also visible in the spectral damping analysis. Especially for the PRNU it becomes visible that a damping of the remaining lower frequencies seems not to be possible without further efforts.

8.4.2. Parameter Search and Best-Case Performance

Method	Description
$\min_{\text{comb.}}(\nabla, BL)$	Solution of eq. 8.50 with HP as $HP(\nabla, \text{BL})$
$\min_{\text{comb.}}(\nabla, MN)$	Solution of eq. 8.50 with HP as $HP(\nabla, \text{MN})$
$\min_{\text{comb.}}(\triangle, MN)$	Solution of eq. 8.50 with HP as $HP(\triangle, \text{MN})$

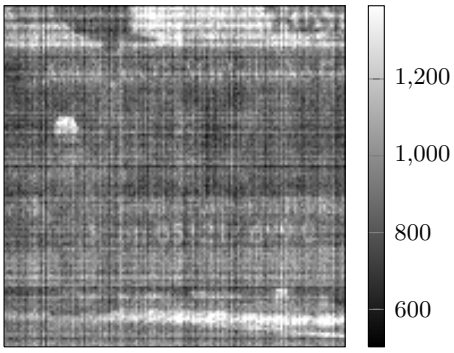
Table 8.1.: Introduction of used abbreviations for the combined DSNU and PRNU estimation methods.

As a next step the validation continues with the real nonuniformities of the car front scenes recorded with the MV1-D1312-160-CL camera described in sec. 4.1.1. The parameter search for the λ parameter is done on the 25 random and consecutive frame calibration sequences according to sec. 5.3.1.

Method	Parameter
$\min_{\text{comb.}}(\nabla, BL)$	$0.06935 \pm (0.02126, 30.65\%)$
$\min_{\text{comb.}}(\nabla, MN)$	$0.02915 \pm (0.01084, 37.17\%)$
$\min_{\text{comb.}}(\triangle, MN)$	$0.2411 \pm (0.4637, 192.4\%)$

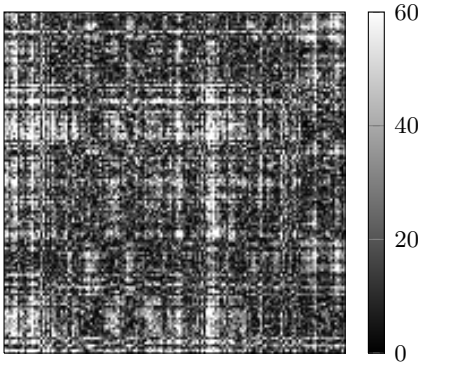
Table 8.2.: Optimized parameter set for the random frame sequences.

The three realizations of the high pass filter mask are investigated and the used the abbreviations are shown tab. 8.1 and the resulting optimal parameters are presented in tab. 8.2 for the random frame sequences and in tab. 8.3 for the consecutive frame sequences. Compared to the parameters from the statistic evalu-

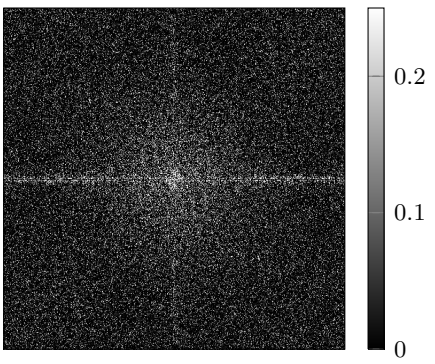


(a) Recorded Image with artifical nonuniformities

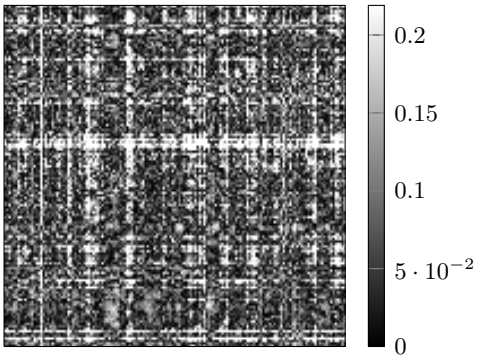
(b) Corrected with the combined method



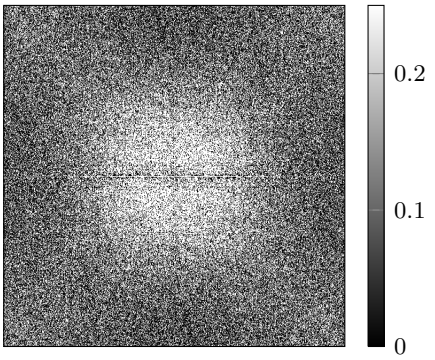
(c) Remaining DSNU pattern



(d) Spectral damping analyis for the DSNU



(e) Remaining PRNU pattern



(f) Spectral damping analyis for the PRNU

Figure 8.2.: Spatial and spectral performance analysis according to sec. 3.6 for the combined PRNU and DSNU estimation method, using $HP = HP(\nabla, \mathrm{BL})$ and the artificial nonuniformities from fig. 8.1.

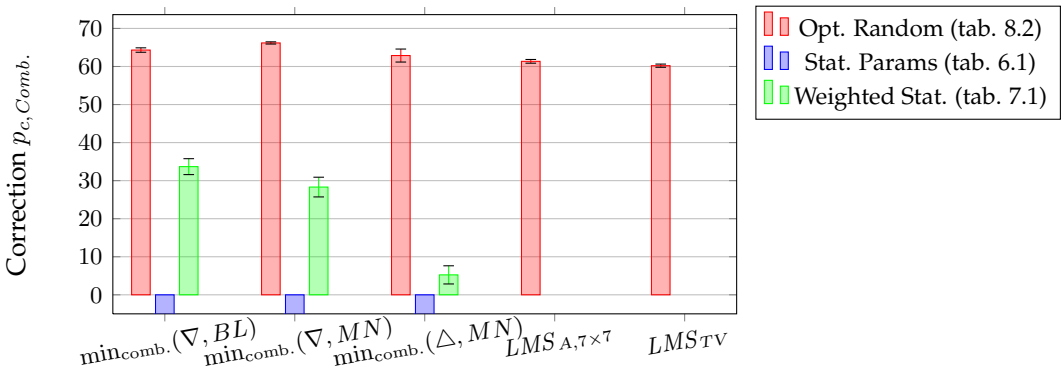
Method	Parameter
$\min_{\text{comb.}}(\nabla, BL)$	$1.244 \pm (3.585, 288.2\%)$
$\min_{\text{comb.}}(\nabla, MN)$	$0.6296 \pm (1.883, 299\%)$
$\min_{\text{comb.}}(\triangle, MN)$	$1.489 \pm (0.9741, 65.43\%)$

Table 8.3.: Optimized parameter set for the consecutive frame sequences.

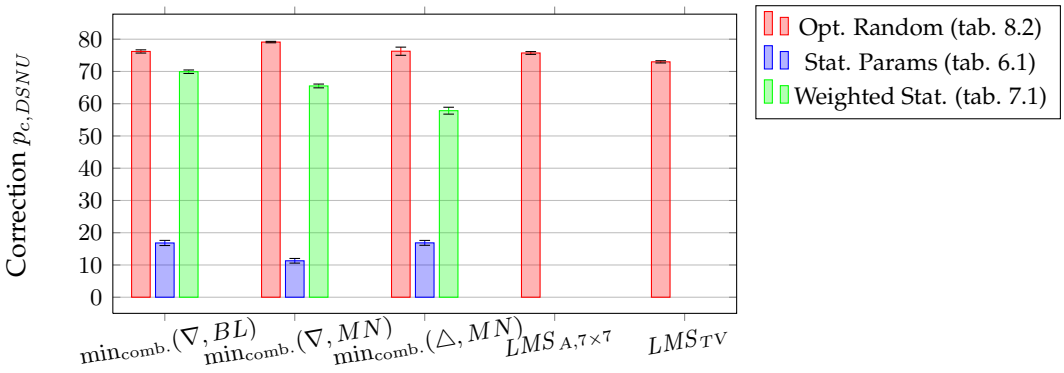
ations (tab. 6.1) or weighted statistic evaluations (tab. 7.1), the difference in the parameters is more than a factor 10. To see the influence of the different parameter choices, the best-case performance analysis is conducted as explained in sec. 5.3.2 (p. 93), including the optimized, the statistic and the weighted statistic parameters. Additionally to the prior best-case performance analysis, the separation into the $p_{\text{c,PRNU}}$, $p_{\text{c,DSNU}}$ and the $p_{\text{c,Comb.}}$ is possible now and will be discussed below.

Fig. 8.3 shows the best-case performance analysis for the random frame sequences. The combined maximum likelihood estimation method provides the best performances in the combined quality metric in fig. 8.3a , however only for the optimized parameters. For the statistic based parameter set (tab. 6.1), only a negative performance can be reached. The statistic parameter set that considers the weighted average (tab. 7.1) reaches a positive performances, but only up to $10\,\% - 35\,\%$ depending on the method. This outcome for the combined performance $p_{\text{c,Comb.}}$ can be analyzed for dependencies on its DSNU and PRNU constituents. Fig. 8.3b shows the performance of in the $p_{\text{c,DSNU}}$ measure. Here the optimized parameters give again a superior performance for the optimized parameter set with concern to the reference methods. However, in this measure the weighted statistic parameter set gives good results as well, by correcting $60\,\%$ to $70\,\%$ of the DSNU. These results are very similar to the results shown in sec. 7.2.3 for the DSNU-only correction. Considering the $p_{\text{c,PRNU}}$ performance (fig. 8.3c), the optimized parameters obtain $12\,\%$ correction rate for the $\min_{\text{comb.}}(\nabla, BL)$ method and about $5\,\%$ to $7\,\%$ for the $\min_{\text{comb.}}(\nabla, MN)$ and $\min_{\text{comb.}}(\triangle, MN)$ method. The reference methods reach only $-4\,\%$ for the $LMS_{A,7\times7}$ and actually $4\,\%$ for the LMS_{TV} method. However, for all the statistical parameter sets, the PRNU correction results in large valued negative performances, which as a consequence leads to a diminished performance of the method if the statistic parameter sets are used. A possible reason for the bad PRNU performance may be found in the assumption of Gaussian distributed $\{\mathbf{a}_{i,j}\}$ random variables, which did not fit as well as for the $\{\mathbf{b}_{i,j}\}$ variables.

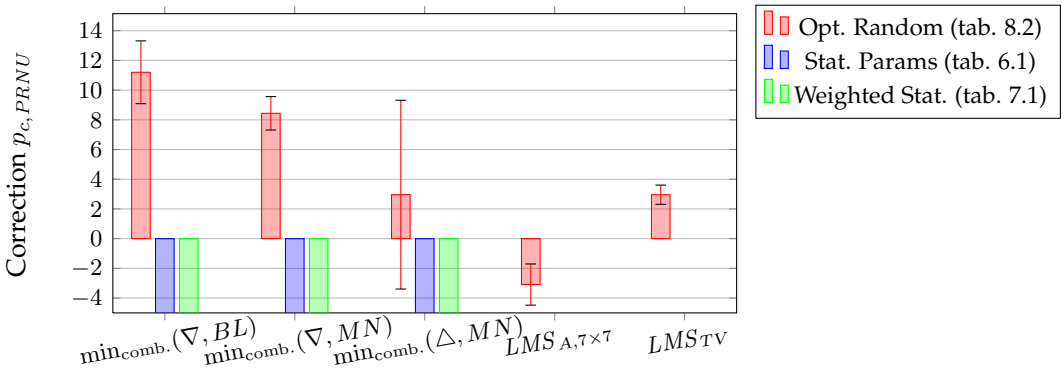
The best-case performance analysis for the consecutive frame sequences is shown in fig. 8.4. In the $p_{\text{c,Comb.}}$ measure, the new methods are not able to reach a good correction performance (see fig. 8.4a). Only the $\min_{\text{comb.}}(\triangle, MN)$ method shows again its stability and achieves $p_{\text{c,Comb.}} \approx 30\,\%$, however with large error bars. The reason for the failure of the methods can be found when separating the PRNU and DSNU correction performances. For $p_{\text{c,DSNU}}$ positive correction rates are achieved with all analyzed parameter sets (see fig. 8.4b), however the $p_{\text{c,PRNU}}$ performance



(a) Optimal performance in the $p_{\text{c,Comb.}}$ metric.

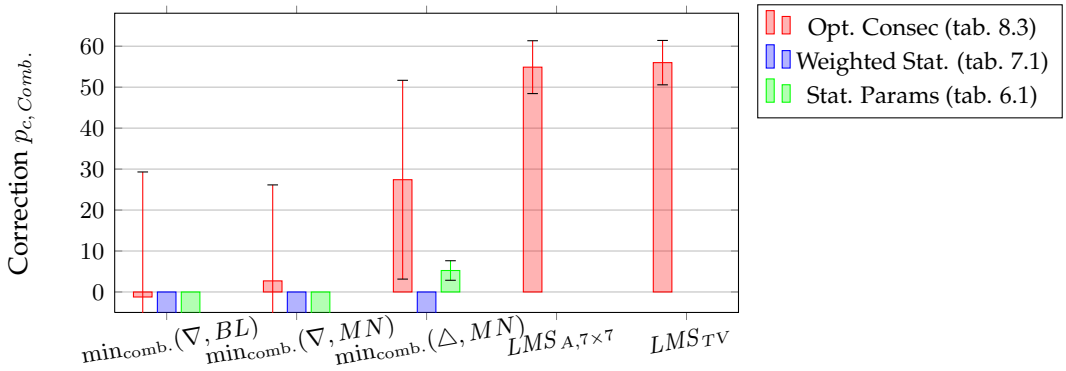


(b) Optimal performance in the $p_{\text{c,DSNU}}$ metric.

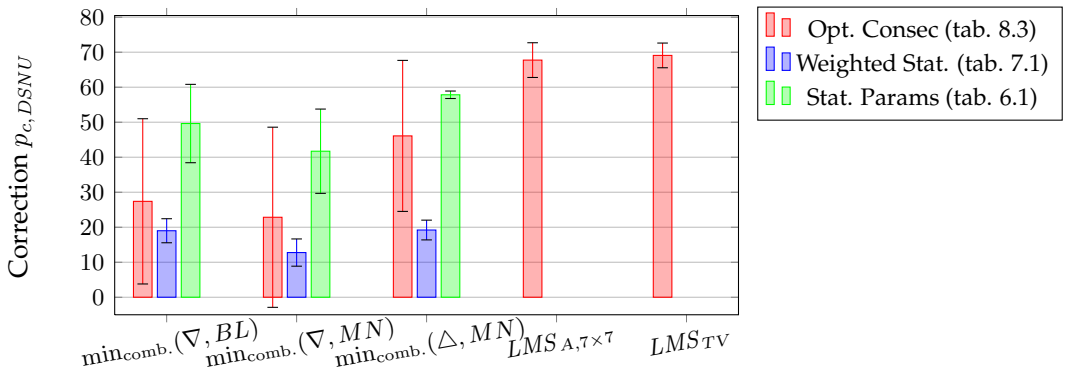


(c) Optimal performance in the $p_{\text{c,PRNU}}$ metric.

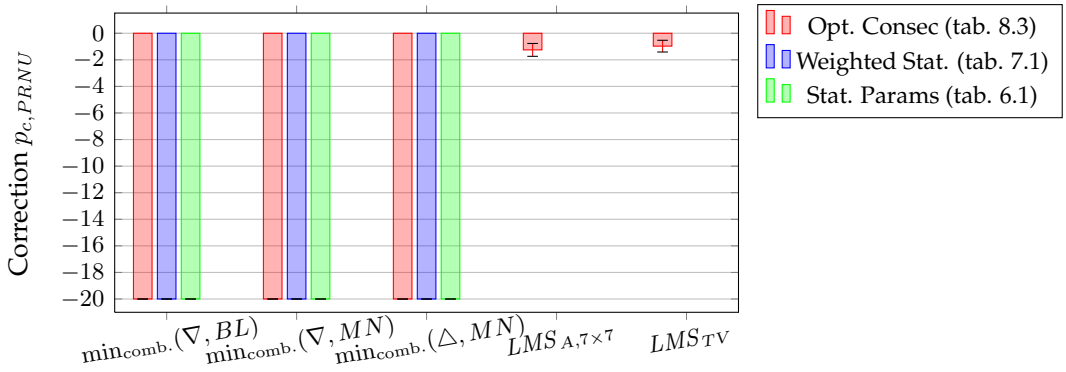
Figure 8.3.: The best-case performance analysis for the combined PRNU and DSNU estimation methods for *random* frame sequences in the different quality metrics.



(a) Optimal performance in the $p_{\text{c,Comb.}}$ metric.



(b) Optimal performance in the $p_{\text{c,DSNU}}$ metric.



(c) Optimal performance in the $p_{\text{c,PRNU}}$ metric.

Figure 8.4.: The best-case performance analysis for the combined PRNU and DSNU estimation methods for *consecutive* frame sequences in the different quality metrics.

as shown in fig. 8.4c goes beyond the $-20\,\%$ and was cut. The resulting $p_{\mathrm{c,Comb.}}$ then degrades to the extend shown above. It was already discussed that the consecutive frame sequences violate the assumption of independent features which is believed to be one of the main reasons for the observed performances. However, the severe degradation of the PRNU correction may be based on the further reasons that the PRNU correction is much more sensitive against parameter changes and that for the $\{\mathbf{a}_{i,j}\}$ the Gaussian distribution is not the best possible fit.

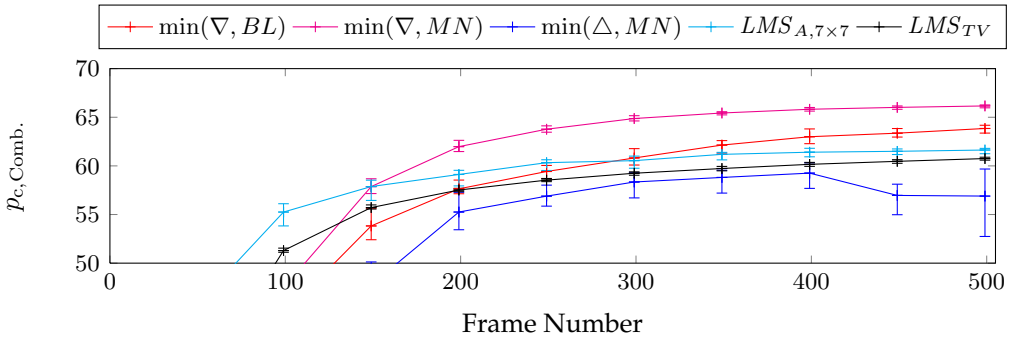
Concluding, the combined PRNU and DSNU nonuniformity correction method achieves the best possible correction if random frame sequences are used as input. For consecutive frame sequences the performance in general decreases and further discussion and analysis for these type of sequences and the image material with low valued nonuniformities is presented in sec. 10.1 and sec. 10.2.

8.4.3. Temporal Performance Analysis

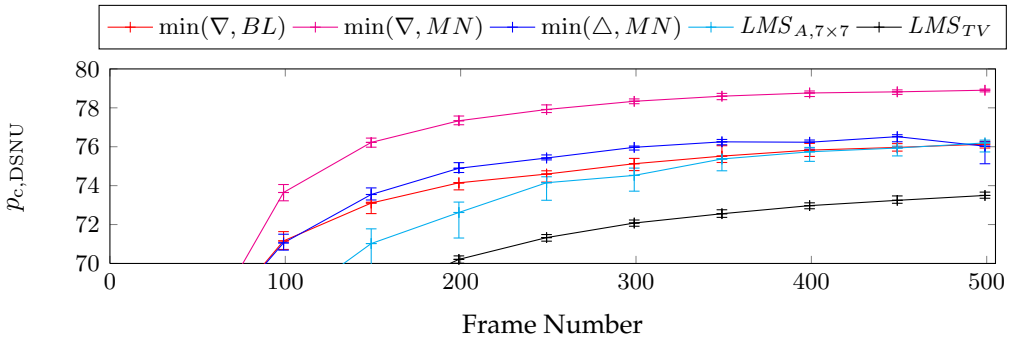
The temporal performance analysis was conducted for the methods using the optimized parameter set and fig. 8.5 shows the results for random sequence frames. Considering the $p_{\mathrm{c,Comb.}}$-measure (fig. 8.5a), the $\min_{\mathrm{comb.}}(\nabla, MN)$ method gives the best results with a statistic significant advance of more than $5\,\%$ over the best reference method $LMS_{A,7\times7}$. The $\min_{\mathrm{comb.}}(\nabla, BL)$ method gives as well a superior performance over the reference methods, only the $\min_{\mathrm{comb.}}(\triangle, MN)$ method finishes about $5\,\%$ below the reference methods. For $\min_{\mathrm{comb.}}(\triangle, MN)$, the performance drops mainly after 400 frames which is an effect from the failed PRNU correction as discussed below.

Fig. 8.5b shows the correction performance $p_{\mathrm{c,DSNU}}$. Here the $\min_{\mathrm{comb.}}(\nabla, MN)$ method gives a superior performance of $78\,\%$, while the $\min_{\mathrm{comb.}}(\nabla, BL)$ and $\min_{\mathrm{comb.}}(\triangle, MN)$ methods converge to the same results as the $LMS_{A,7\times7}$ reference method. The analysis of the $p_{\mathrm{c,PRNU}}$ is shown in fig. 8.5c and the $\min_{\mathrm{comb.}}(\nabla, BL)$ and $\min_{\mathrm{comb.}}(\nabla, MN)$ methods show the best performances. However, a better performance than the reference methods is achieved not until frame number 250 and 350, but after 500 frames the a correction rate of $p_{\mathrm{c,PRNU}}{=}9\,\%$ is reached, while the best reference method, the LMS_{TV}, reaches only $4\,\%$.

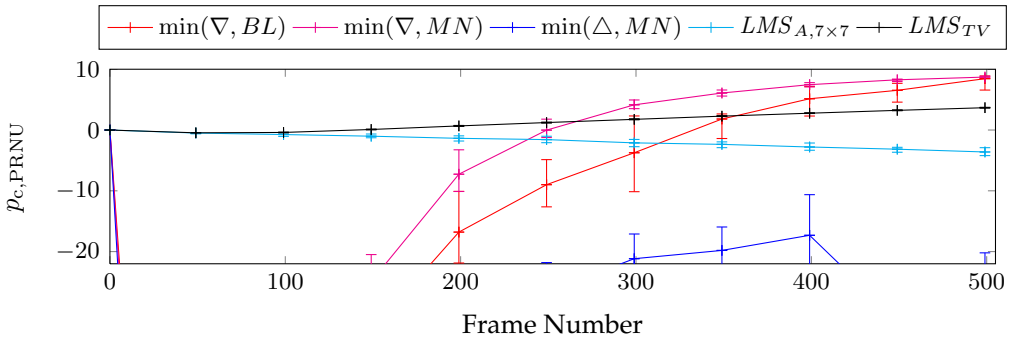
Only the $\min_{\mathrm{comb.}}(\triangle, MN)$ method does not reach up positive correction rates, which may be surprising as the method achieved about $5\,\%$ correction rate in the best-case performance analysis. The best-case performance analysis represents however an upper border of the possible performances as discussed before in sec. 5.3.2. The large error bars in the best-case analysis and the parameter search can therefore be understood as a hint for instabilities of the method against parameter changes. Fixing the parameters to the values of tab. 8.2 then leads to the general drop in the performance. As the PRNU estimate is very sensitive, the achieved performances result in large negative values.



(a) Evaluations of the $p_{\mathrm{c,Comb.}}$



(b) Evaluations of the $p_{\mathrm{c,DSNU}}$



(c) Evaluations of the $p_{\mathrm{c,PRNU}}$

Figure 8.5.: Methods evaluations with experimental optimized parameters for the random frame sequences. The used parameters are given in tab. 8.2.

8.4.4. Spatial and Spectral Performance Analysis

Fig. 8.6 shows the corrected images with the different methods. As expected from the overall high correction performances, a visual degradation of the images is not notable (see fig. 3.1c p. 46 for the uncorrected image). In fig. 8.7 however, the remaining differnces in the DSNU are visible. The temporal evaluations above showed that the correction performance for the new combined maximium likelihood method is as good as for the DSNU-only corrections and therefore the images look similar to the ones shown in fig. 7.6 on p.147. Especially for the spectral damping analysis in fig. 8.8, the similarities to the correction performance of the weighted averaging maximum likelihood method in fig. 7.7 p.149 are visible. Thus, the comparison towards to the reference methods shows as before a superiority of the new methods with concern to the DSNU correction.

With respect to the PRNU correction, fig. 8.9 shows the remaining $\{a_{i,j}\}$ coefficients, where the $\min_{\text{comb.}}(\nabla, BL)$ method gives the visually best impression and the new methods in general seem to have corrected more of the higher frequency components. A look onto the spectral damping factors of the PRNU in fig. 8.10 reveals, that the overall damping is better for the new methods, as the images have a slightly lower average gray value. Interesting is that the strongest damping has occurred for the lower frequencies, which is surprising as the method tends to correct lower frequencies not as well as the high frequencies. Compare the results of the PRNU correction in the proof of concept analysis in sec. 8.4.1. Anyway, the general better damping coefficients show that the new methods are able to correct even the low valued PRNU better than the reference methods. The only drawback against the DSNU-only methods is that the statistically derived parameters can not be used any more, especially with concern to the PRNU correction.

Concluding, the results of the temporal performance analysis confirm that the combined correction method is more unstable against parameter changes and not fulfilled demands assumptions on the input data (e.g. consecutive frame input). The discussed reasons are a misfit of the PRNU to a Gaussian distribution where a solution could be to tune the alpha parameter as well. In sec. 8.5.5, the separate tuning of the α, β and λ parameters will be the only possibility to gain reasonable corrections. But for now, as long as the input data fits the random frame input assumption, a one dimensional parameter search for λ is sufficient, to achieve superior results over the best reference methods.

(a) Ground Truth Corrected

(b) $\min_{\text{comb.}}(\nabla, BL)$ corrected Image

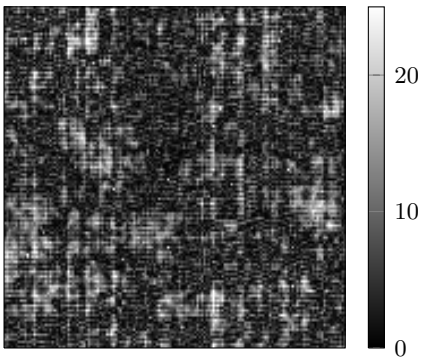
(c) $\min_{\text{comb.}}(\nabla, MN)$ corrected Image

(d) $\min_{\text{comb.}}(\triangle, MN)$ corrected Image

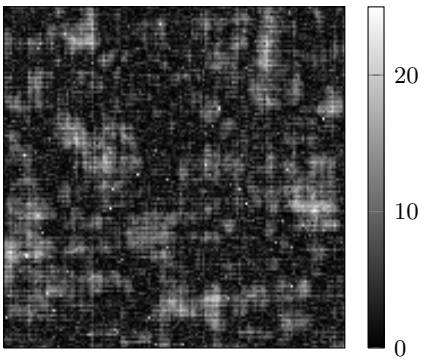
(e) Corrected with $LMS_{A,7\times7}$ Method

(f) Corrected with LMS_{TV} Method

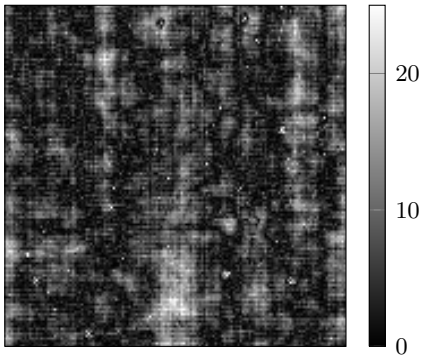
Figure 8.6.: The visual impression of the corrected images with the different methods for a combined PRNU and DSNU correction in random frame sequences.



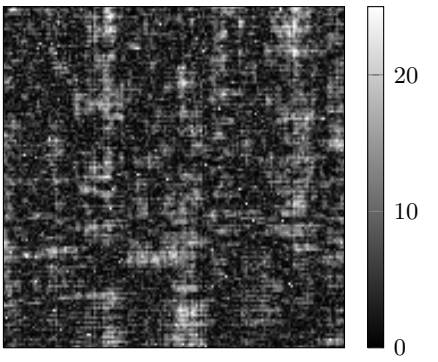
(a) $\min_{\text{comb.}}(\nabla, BL)$ remaining DSNU



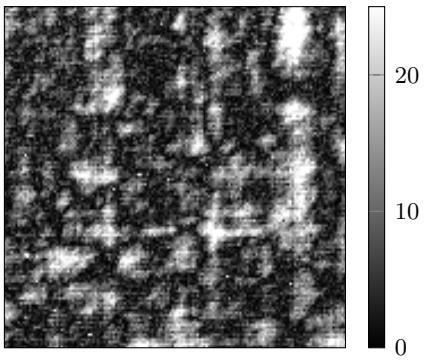
(b) $\min_{\text{comb.}}(\nabla, MN)$ remaining DSNU



(c) $\min_{\text{comb.}}(\triangle, MN)$ remaining DSNU

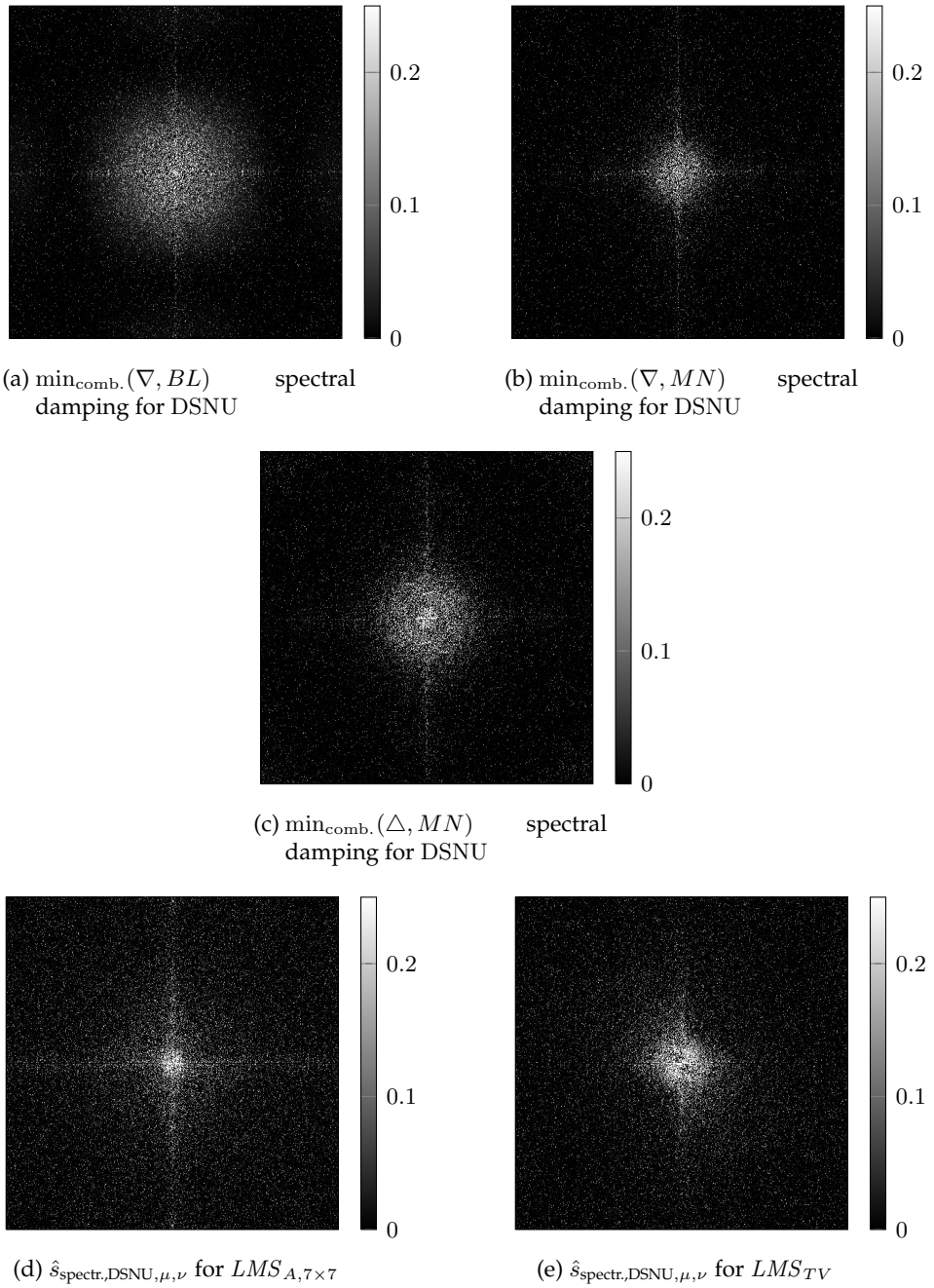


(d) $|\{b_{i,j,\text{rem.}}\}|$ for $LMS_{A,7\times7}$



(e) $|\{b_{i,j,\text{rem.}}\}|$ for LMS_{TV}

Figure 8.7.: The remaining DSNU patterns for the different methods for a combined PRNU and DSNU correction in random frame sequences.



(a) $\min_{\text{comb.}}(\nabla, BL)$ spectral damping for DSNU

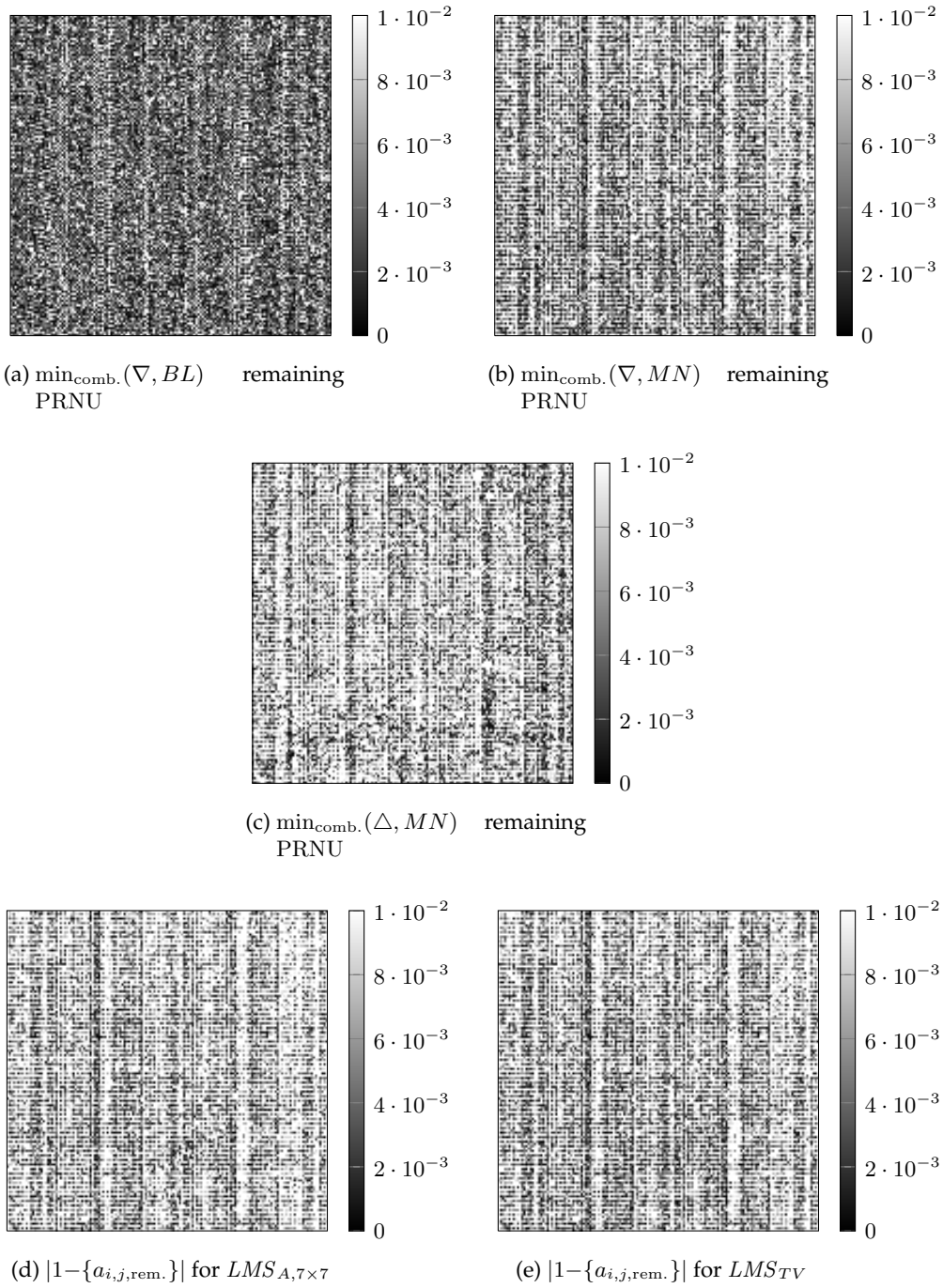
(b) $\min_{\text{comb.}}(\nabla, MN)$ spectral damping for DSNU

(c) $\min_{\text{comb.}}(\triangle, MN)$ spectral damping for DSNU

(d) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $LMS_{A,7\times7}$

(e) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for LMS_{TV}

Figure 8.8.: The spectral damping factors for the DSNU patterns and the different methods for a combined PRNU and DSNU correction in random frame sequences.



(a) $\min_{\text{comb.}}(\nabla, BL)$ remaining PRNU

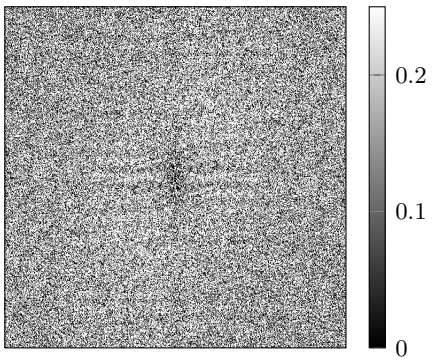
(b) $\min_{\text{comb.}}(\nabla, MN)$ remaining PRNU

(c) $\min_{\text{comb.}}(\triangle, MN)$ remaining PRNU

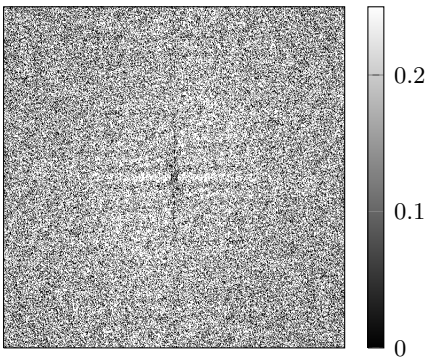
(d) $|1-\{a_{i,j,\text{rem.}}\}|$ for $LMS_{A,7\times7}$

(e) $|1-\{a_{i,j,\text{rem.}}\}|$ for LMS_{TV}

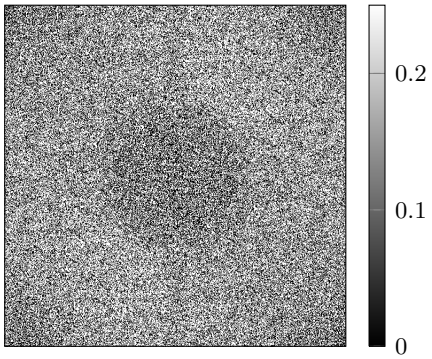
Figure 8.9.: The remaining PRNU patterns for the different methods for a combined PRNU and DSNU correction in random frame sequences.



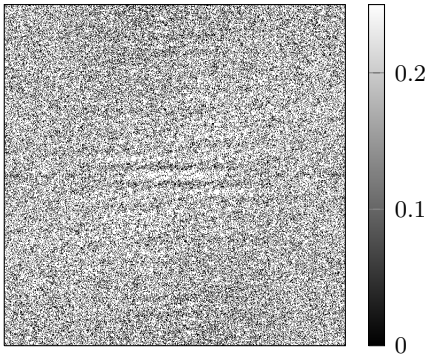
(a) $\min_{\text{comb.}}(\nabla, BL)$ spectral damping for PRNU



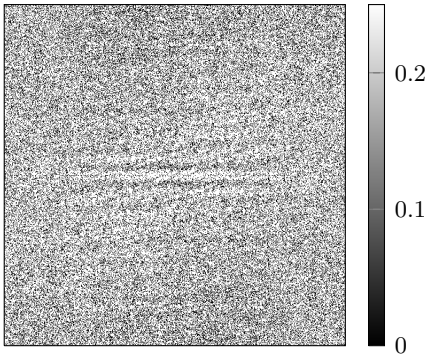
(b) $\min_{\text{comb.}}(\nabla, MN)$ spectral damping for PRNU



(c) $\min_{\text{comb.}}(\triangle, MN)$ spectral damping for PRNU



(d) $\hat{s}_{\text{spectr.,PRNU},\mu,\nu}$ for $LMS_{A,7\times7}$



(e) $\hat{s}_{\text{spectr.,PRNU},\mu,\nu}$ for LMS_{TV}

Figure 8.10.: The spectral damping factors for the PRNU patterns and the different methods for a combined PRNU and DSNU correction in random frame sequences.

8.5. An Extension to Row and Column Wise Estimation

The discussed parameter instability for the combined DSNU and PRNU estimation method in the last section inhibits the method until now from using theoretically statistically derived parameter sets. The theoretical development of the combined estimation estimation into a row-, and column-, and pixel-wise combined PRNU and DSNU estimation method is possible anyway, in equivalence to the DSNU-only extension in sec. 7.3. The motivation for this step is as before a better correction of the visually important row and column wise degradations. However, the problem with the parameter instabilities becomes even more severe with this row-, column- and pixel-wise (RCP) extension of the method and could not be solved in the research of this thesis. This section derives the mathematical extensions and provides first proof of concept experimental results.

8.5.1. Extending the Mathematical Ansatz

To derive the RCP approach for the combined estimation, the parameters $\{a_{i,j}\}$ are decomposed into row-, column-, and pixel-wise components. Due to the multiplicative nature of the gain factors, the decomposition is multiplicative for the $a_{i,j}$ variables. The whole derivation follows actually the motivations given in sec. 7.3 for the DSNU decompositions. The row-, column- and pixel-wise offset variables present as:

$$\begin{aligned} b_{i,j} &= b_{\text{col.},i} + b_{\text{row.},j} + b_{\text{pix.},i,j} \\ \text{and} \quad a_{i,j} &= a_{\text{col.},i} \cdot a_{\text{row.},j} \cdot a_{\text{pix.},i,j} \,. \end{aligned} \tag{8.54}$$

The energy function of sec. 6.3 stays valid as the resulting random variables $\{\mathbf{a}_{i,j}\}$ and $\{\mathbf{b}_{i,j}\}$ are not changed. However, the row- and column-wise dependencies are again only considered implicitly by this approach. The derivation follows the same track of mathematical transformations, considering the idea of row-, column- and pixel-wise parameters as in sec. 7.3. With the new variables the minimization task rewrites as:

$$\arg\min(E_{\text{D}} + E_{\text{P}})(\{a_{\text{col.},i}, a_{\text{row.},j}, a_{\text{pix.},i,j}, b_{\text{col.},i}, b_{\text{row.},j}, b_{\text{pix.},i,j}\}) \tag{8.55}$$

with the brackets $\{\cdot\}$ representing the set for the running indices i, j of the argument variables.

To execute the minimization, a gradient based iterative solver will be used and therefore the gradient of the energy function needs to be calculated. By concate-

nating the gradient vectors of the different variable types below each other, the gradient can be expressed as:

$$\begin{pmatrix} \nabla_{\{a_{\text{col.},i}\}}(\text{const.} + E_\text{P} + E_\text{D}) \\ \nabla_{\{a_{\text{row.},j}\}}(\text{const.} + E_\text{P} + E_\text{D}) \\ \nabla_{\{a_{\text{pix.},i,j}\}}(\text{const.} + E_\text{P} + E_\text{D}) \\ \nabla_{\{b_{\text{col.},i}\}}(\text{const.} + E_\text{P} + E_\text{D}) \\ \nabla_{\{b_{\text{row.},j}\}}(\text{const.} + E_\text{P} + E_\text{D}) \\ \nabla_{\{b_{\text{pix.},i,j}\}}(\text{const.} + E_\text{P} + E_\text{D}) \end{pmatrix} = \begin{pmatrix} \nabla_{\{a_{\text{col.},i}\}}E_\text{P} + \nabla_{\{a_{\text{col.},i}\}}E_\text{D} \\ \nabla_{\{a_{\text{row.},j}\}}E_\text{P} + \nabla_{\{a_{\text{row.},j}\}}E_\text{D} \\ \nabla_{\{a_{\text{pix.},i,j}\}}E_\text{P} + \nabla_{\{a_{\text{pix.},i,j}\}}E_\text{D} \\ \nabla_{\{b_{\text{col.},i}\}}E_\text{P} + \nabla_{\{b_{\text{col.},i}\}}E_\text{D} \\ \nabla_{\{b_{\text{row.},j}\}}E_\text{P} + \nabla_{\{b_{\text{row.},j}\}}E_\text{D} \\ \nabla_{\{b_{\text{pix.},i,j}\}}E_\text{P} + \nabla_{\{b_{\text{pix.},i,j}\}}E_\text{D} \end{pmatrix} \tag{8.56}$$

$$\text{with:} \quad \nabla_{\{a_{\text{col.},i}\}} = (\partial_{a_{\text{col.},1}} \dots \partial_{a_{\text{col.},M}})^T \tag{8.57}$$

$$\nabla_{\{a_{\text{row.},j}\}} = (\partial_{a_{\text{row.},1}} \dots \partial_{a_{\text{row.},N}})^T \tag{8.58}$$

$$\nabla_{\{a_{\text{pix.},i,j}\}} = (\partial_{a_{\text{pix.},1,1}} \dots \partial_{a_{\text{pix.},1,N}}, \partial_{a_{\text{pix.},2,1}} \dots \partial_{a_{\text{pix.},M,N}})^T \tag{8.59}$$

$$\nabla_{\{b_{\text{col.},i}\}} = (\partial_{b_{\text{col.},1}} \dots \partial_{b_{\text{col.},M}})^T \tag{8.60}$$

$$\nabla_{\{b_{\text{row.},j}\}} = (\partial_{b_{\text{row.},1}} \dots \partial_{b_{\text{row.},N}})^T \tag{8.61}$$

$$\nabla_{\{b_{\text{pix.},i,j}\}} = (\partial_{a_{\text{pix.},1,1}} \dots \partial_{a_{\text{pix.},1,N}}, \partial_{a_{\text{pix.},2,1}} \dots \partial_{a_{\text{pix.},M,N}})^T \tag{8.62}$$

The gradients for the different energy and variable sets can be calculated separately. In total there are the 12 parts of the derivatives: $\nabla_{\{a_{\text{col.},i}\}}E_\text{P}$, $\nabla_{\{a_{\text{col.},i}\}}E_\text{D}$,, $\nabla_{\{a_{\text{row.},j}\}}E_\text{P}$, $\nabla_{\{a_{\text{row.},j}\}}E_\text{D}$, $\nabla_{\{a_{\text{pix.},i,j}\}}E_\text{P}$, $\nabla_{\{a_{\text{pix.},i,j}\}}E_\text{D}$, $\nabla_{\{b_{\text{col.},i}\}}E_\text{P}$, $\nabla_{\{b_{\text{col.},i}\}}E_\text{D}$, $\nabla_{\{b_{\text{row.},j}\}}E_\text{P}$, $\nabla_{\{b_{\text{row.},j}\}}E_\text{D}$, $\nabla_{\{b_{\text{pix.},i,j}\}}E_\text{P}$ and $\nabla_{\{b_{\text{pix.},i,j}\}}E_\text{D}$.

8.5.2. Calculating the E_P Gradients

For simplicity we start with the derivation for the column type offset parameters $b_{\text{col.},m}$ and $a_{\text{col.},m}$ of the E_P part of the energy function:

$$\partial_{b_{\text{col.},m}}(E_\text{P}) = \partial_{b_{\text{col.},m}} \Bigg(\sum_{i,j} \beta \left(b_{\text{col.},i} + b_{\text{row.},j} + b_{\text{pix.},i,j} - \mu_b\right)^2 + \sum_{i,j} \alpha \left(a_{\text{col.},i} \cdot a_{\text{row.},j} \cdot a_{\text{pix.},i,j} - \mu_a\right)^2 \Bigg) \tag{8.63}$$

$$= 2\beta \sum_j \left(b_{\text{col.},m} + b_{\text{row.},j} + b_{\text{pix.},m,j} - \mu_b\right) \tag{8.64}$$

$$= 2\beta \sum_j \left(b_{i,j} - \mu_b\right) \tag{8.65}$$

$$\partial_{a_{\text{col.},m}}(E_\text{P}) = \partial_{a_{\text{col.},m}} \Bigg(\sum_{i,j} \beta \left(b_{\text{col.},i} + b_{\text{row.},j} + b_{\text{pix.},i,j} - \mu_b\right)^2 + \sum_{i,j} \alpha \left(a_{\text{col.},i} \cdot a_{\text{row.},j} \cdot a_{\text{pix.},i,j} - \mu_a\right)^2 \Bigg) \tag{8.66}$$

$$= 2\alpha \frac{1}{a_{\text{col.},m}} \sum_j \left(a_{\text{col.},m} \cdot a_{\text{row.},j} \cdot a_{\text{pix.},m,j} - \mu_a\right) a_{\text{col.},m} \cdot a_{\text{row.},j} \cdot a_{\text{pix.},m,j} \tag{8.67}$$

$$= 2\alpha \frac{1}{a_{\text{col.},m}} \sum_j \left(a_{i,j} - \mu_a\right) a_{i,j} \tag{8.68}$$

In the vector notation, all the column offset derivatives can be summarized with help of the transposed nabla operator. The indices of the vectors can then be omitted:

$$\nabla^T_{a_{\text{col.}}} E_{\text{P}} = \left(\sum_{Col.wise} 2\alpha(a - \mu_a) \otimes a \right) \oslash a_{\text{col.}} \tag{8.69}$$

$$\nabla^T_{b_{\text{col.}}} E_{\text{P}} = \sum_{Col.wise} 2\beta(b - \mu_b)\,. \tag{8.70}$$

For the row-wise gradient the same mathematics apply leading to:

$$\nabla_{a_{\text{row.}}} E_{\text{P}} = 2\alpha \left(\sum_{rowwise} (a - \mu_a) \otimes a \right) \oslash a_{\text{row.}} \tag{8.71}$$

$$\nabla_{b_{\text{row.}}} E_{\text{P}} = 2\beta \sum_{rowwise} (b - \mu_b)\,. \tag{8.72}$$

The pixel-wise part of the gradient can as always be expressed over the combined offsets $b_{i,j}$ and gains $a_{i,j}$ with the $\hat{\nabla}$ operator as

$$\hat{\nabla}_{a_{\text{pix.}}} E_{\text{P}} = 2(\alpha(a - \mu_a) \otimes a) \oslash a_{pix} \tag{8.73}$$

$$\hat{\nabla}_{b_{\text{pix.}}} E_{\text{P}} = 2\beta(b - \mu_b)\,. \tag{8.74}$$

8.5.3. Calculating the $\nabla_{\{b_{i,j}\}} E_{\text{D}}$ Gradients

The E_{P} part of the gradient is now defined and the next step is the E_{D} part of the energy. As there are many similarities to the DSNU-only version, we start with $\partial_{b_{\text{col.},m}}(E_{\text{D}})$:

$$= \partial_{b_{\text{col.},m}} \Big(\sum_{t_n} \sum_{i,j} \lambda(C_{\text{M}} - b - \chi) \oslash a \otimes ((C_{\text{M}} - b - \chi) \oslash a * HP(\hat{L}, \Phi)) + \text{const.} \Big) \tag{8.75}$$

$$= \sum_{t_n} \lambda \partial_{b_{\text{col.},m}} \left(\sum_{(i,j) \in S} \frac{1}{a_{i,j}} (C_{\text{M},i,j} - b_{i,j} - \chi_{i,j}) \right.$$
$$\left. \cdot \sum_{(k,l) \in S} \frac{1}{a_{k,l}} (C_{\text{M},k,l} - b_{k,l} - \chi_{k,l}) \Psi(|k - i|, |l - j|) \right) \tag{8.76}$$

$$= \sum_{t_n} \lambda \left((\partial_{b_{\text{col.},m}} \sum_{(i,j)\in S} \frac{1}{a_{i,j}} (C_{\text{M},i,j} - b_{i,j} - \chi_{i,j})) \sum_{(k,l)\in S} \frac{1}{a_{k,l}} (C_{\text{M},k,l} - b_{k,l} - \chi_{k,l}) \Psi(\cdot) \right.$$
$$\left. + \sum_{(i,j)\in S} \frac{1}{a_{i,j}} (C_{\text{M},i,j} - b_{i,j} - \chi_{i,j}) (\partial_{b_{\text{col.},m}} \sum_{(k,l)\in S} \frac{1}{a_{k,l}} (C_{\text{M},k,l} - b_{k,l} - \chi_{k,l}) \Psi(\cdot)) \right) \tag{8.77}$$

The actual task is now to define the derivatives of the sums, which simplifies to:

$$\partial_{b_{\text{col.},m}} \sum_{(i,j)\in S} \frac{1}{a_{i,j}} (C_{\text{M},i,j} - b_{i,j} - \chi_{i,j}) \tag{8.78}$$

$$= \partial_{b_{\text{col.},m}} \sum_{(i,j)\in S} \frac{1}{a_{\text{col.},i} \cdot a_{\text{row.},j} \cdot a_{\text{pix.},i,j}} (C_{\text{M},i,j} - b_{\text{col.},i} - b_{\text{row.},j} - b_{\text{pix.},i,j} - \chi_{i,j}) \tag{8.79}$$

$$= \sum_j \frac{-1}{a_{\text{col.},m} \cdot a_{\text{row.},j} \cdot a_{\text{pix.},m,j}} = \sum_j \frac{-1}{a_{m,j}}, \tag{8.80}$$

and can directly be inserted into the derivative for $\partial_{b_{\text{col.},m}}$:

$$\partial_{b_{\text{col.},m}} (E_{\text{D}}) \tag{8.81}$$

$$= \sum_{t_n} \lambda \left(\sum_j \frac{-1}{a_{m,j}} \sum_{(k,l)\in S} \frac{1}{a_{k,l}} (C_{\text{M},k,l} - b_{k,l} - \chi_{k,l}) \Psi(|k-m|, |l-j|) \right.$$
$$\left. + \sum_{(i,j)\in S} (C_{\text{M},i,j} - b_{i,j} - \chi_{i,j}) \sum_l \frac{-1}{a_{m,l}} \Psi(|m-i|, |l-j|)) \right) \tag{8.82}$$

$$= 2\lambda \sum_{t_n} \sum_n \frac{-1}{a_{m,n}} \sum_{(i,j)\in S} \frac{1}{a_{i,j}} (C_{\text{M},i,j} - b_{i,j} - \chi_{i,j}) \Psi(|i-m|, |j-n|)) \tag{8.83}$$

$$= -2\lambda T_{\text{N}} \frac{1}{T_{\text{N}}} \sum_{t_n} \sum_n \frac{1}{a_{m,n}} \sum_{(i,j)\in S} \frac{1}{a_{i,j}} (C_{\text{M},i,j} - b_{i,j} - \chi_{i,j}) \Psi(|i-m|, |j-n|)) \tag{8.84}$$

$$= -2\lambda T_{\text{N}} \sum_n \frac{1}{a_{m,n}} \sum_{(i,j)\in S} \underbrace{\frac{1}{a_{i,j}} (A_{\text{M},i,j} - b_{i,j})}_{A^{\text{corr.}}_{\text{M},i,j}} \Psi(|i-m|, |j-n|)). \tag{8.85}$$

For the row- and pixel-wise notations, the results are obtained following the same mathematics, which in are written in matrix notation like:

$$\nabla^T_{b_{\text{col.}}} E_{\text{D}} = -2\lambda T_{\text{N}} \sum_{\text{column-wise}} (A^{\text{corr.}}_{\text{M}} * HP(\hat{L}, \Phi)) \oslash a \tag{8.86}$$

$$\nabla_{b_{\text{row.}}} E_{\text{D}} = -2\lambda T_{\text{N}} \sum_{\text{row-wise}} (A^{\text{corr.}}_{\text{M}} * HP(\hat{L}, \Phi)) \oslash a \tag{8.87}$$

$$\hat{\nabla}_{b_{\text{pix.}}} E_{\text{D}} = -2\lambda T_{\text{N}} (A_{\text{M}}^{\text{corr.}} * HP(\hat{L}, \Phi)) \oslash a \,. \tag{8.88}$$

Separated Row- and Column-wise Nonuniformities: As for the DSNU-only case further simplifications are possible if only column-wise or only row-wise DSNU is assumed. Eq. 8.54 then compresses to

$$b_{i,j} = b_{\text{col.},i} \quad \text{and} \quad a_{i,j} = a_{\text{col.},i} \tag{8.89}$$

$$\text{or} \quad b_{i,j} = b_{\text{row.},j} \quad \text{and} \quad a_{i,j} = a_{\text{row.},j} \,, \tag{8.90}$$

and the gradients from eq. 8.86 transform into the following one dimensional form (with the definitions of M and N as the sensor dimensions, see p. 161):

$$\nabla_{b_{\text{col.}}}^{T} E_{\text{D}} = -2\lambda T_{\text{N}} (A_{\text{M,col}} - M b_{\text{col.}} * HP_{\text{col.}}(\hat{L}, \Phi) \oslash (M a_{\text{col.},m}) \tag{8.91}$$

$$\nabla_{b_{\text{row.}}} E_{\text{D}} = -2\lambda T_{\text{N}} (A_{\text{M,row}} - N b_{\text{row.}} * HP_{\text{row.}}(\hat{L}, \Phi) \oslash (N a_{\text{row.},n}) \,. \tag{8.92}$$

8.5.4. Calculating the $\nabla_{\{a_{i,j}\}} E_{\text{D}}$ Gradients

The final step is the derivation against the gain parameters $\{a_{i,j}\}$. To start $\partial_{a_{\text{col.},m}}(E_{\text{D}})$ is considered, which has exactly the same mathematical form as eq. 8.35 (p. 182). Therefore the derivation can be followed until the eq. F.17 in the appendix F.2.2 (p. 351). The wanted derivative can then be expressed by substituting $\partial_{a_{m,n}} \to \partial_{a_{\text{col.},m}}$ which results in

$$\begin{aligned} \partial_{a_{\text{col.},m}}(E_{\text{D}}) = 2\lambda \sum_{t_n} \Bigg(& (\partial_{a_{m,n}} \sum_{(i,j)\in S} \frac{1}{a_{i,j}} (C_{\text{M},i,j} - b_{i,j} - \chi_{i,j})) \\ & \cdot \sum_{(k,l)\in S} \frac{1}{a_{k,l}} (C_{\text{M},k,l} - b_{k,l} - \chi_{k,l}) \Psi(|k-i|, |l-j|) \Bigg) \,. \end{aligned} \tag{8.93}$$

The evaluation of the partial derivative on the sum leads to:

$$\partial_{a_{\text{col.},m}} \sum_{(i,j)\in S} \frac{1}{a_{i,j}} (C_{\text{M},i,j} - b_{i,j} - \chi_{i,j}) \tag{8.94}$$

$$= \partial_{a_{\text{col.},m}} \sum_{(i,j)\in S} \frac{1}{a_{\text{col.},i} \cdot a_{\text{row.},j} \cdot a_{\text{pix.},i,j}} (C_{\text{M},i,j} - b_{\text{col.},i} - b_{\text{row.},j} - b_{\text{pix.},i,j} - \chi_{i,j}) \tag{8.95}$$

$$= \sum_{j} \frac{-a_{\text{row.},j} \cdot a_{\text{pix.},m,j}}{(a_{\text{col.},m} \cdot a_{\text{row.},j} \cdot a_{\text{pix.},m,j})^2} (C_{\text{M},m,j} - b_{\text{col.},m} - b_{\text{row.},j} - b_{\text{pix.},m,j} - \chi_{m,j}) \tag{8.96}$$

$$= \frac{-1}{a_{\text{col.},m}} \sum_{j} \frac{1}{a_{m,j}} (C_{\text{M},m,j} - b_{m,j} - \chi_{m,j}) \,, \tag{8.97}$$

inserted in the above equations it allows the further transformations:

$$\begin{aligned}\partial_{a_{\mathrm{col.},m}}(E_{\mathrm{D}}) =& 2\lambda \sum_{t_n}\Big(\frac{-1}{a_{\mathrm{col.},m}}\sum_{j}\frac{1}{a_{m,j}}(C_{\mathrm{M},m,j}-b_{m,j}-\chi_{m,j})\Big) \\ & \cdot \sum_{(k,l)\in S}\frac{1}{a_{k,l}}(C_{\mathrm{M},k,l}-b_{k,l}-\chi_{k,l})\Psi(|k-m|,|l-j|) \qquad (8.98)\\ =& \frac{-1}{a_{\mathrm{col.},m}}\sum_{j}-2\lambda\sum_{t_n}\Big(\frac{1}{a_{m,j}}(C_{\mathrm{M},m,j}-b_{m,j}-\chi_{m,j})\Big) \\ & \cdot \sum_{(k,l)\in S}\frac{1}{a_{k,l}}(C_{\mathrm{M},k,l}-b_{k,l}-\chi_{k,l})\Psi(|k-m|,|l-j|)\,. \qquad (8.99)\end{aligned}$$

Except for the sum $\sum_j$ this equation has the identical mathematical form as eq. F.17 and therefore all the transformations until eq. 8.46 can be applied using the substitution $\partial_{a_{m,n}} \rightarrow \partial_{a_{\mathrm{col.},m}}$ and by considering the additional sum:

$$\begin{aligned}\partial_{a_{\mathrm{col.},m}}(E_{\mathrm{D}}) =& \frac{-2\lambda T_{\mathrm{N}}}{a_{\mathrm{col.},m}}\sum_{j}\Big(\Big(\frac{1}{a_{m,j}}\underbrace{\sum_{(k,l)\in S}\frac{1}{a_{k,l}}\mathrm{cov}(C_{\mathrm{M},m,j},C_{\mathrm{M},k,l})\Psi(\cdot)}_{\Upsilon_{m,j}}\Big) \\ & + \sum_{(k,l)\in S}\underbrace{\frac{1}{a_{m,j}}(A_{\mathrm{M},m,j}-b_{m,j})}_{A^{\mathrm{corr.}}_{\mathrm{M},m,j}}\underbrace{\frac{1}{a_{k,l}}(A_{\mathrm{M},k,l}-b_{k,l})}_{A^{\mathrm{corr.}}_{\mathrm{M},k,l}}\Psi(\cdot) \qquad (8.100)\end{aligned}$$

Given the matrix notation and the mathematical symmetries, the gradients with respect to the different gain variables result in the expressions:

$$\nabla^{T}_{a_{\mathrm{col.}}}E_{\mathrm{D}} = -2\lambda T_{\mathrm{N}}\left[\sum_{\mathrm{col.-wise}}((\Upsilon \oslash a) + (A^{\mathrm{corr.}}_{\mathrm{M}} \otimes A^{\mathrm{corr.}}_{\mathrm{M}} * HP(\hat{L},\Phi)))\right] \oslash a_{\mathrm{col}} \qquad (8.101)$$

$$\nabla_{a_{\mathrm{row.}}}E_{\mathrm{D}} = -2\lambda T_{\mathrm{N}}\left[\sum_{\mathrm{row-wise}}((\Upsilon \oslash a) + (A^{\mathrm{corr.}}_{\mathrm{M}} \otimes A^{\mathrm{corr.}}_{\mathrm{M}} * HP(\hat{L},\Phi)))\right] \oslash a_{\mathrm{row}} \qquad (8.102)$$

$$\hat{\nabla}_{a_{\mathrm{pix.}}}E_{\mathrm{D}} = -2\lambda T_{\mathrm{N}}((\Upsilon \oslash a) + (A^{\mathrm{corr.}}_{\mathrm{M}} \otimes A^{\mathrm{corr.}}_{\mathrm{M}} * HP(\hat{L},\Phi))) \oslash a_{\mathrm{pix}} \qquad (8.103)$$

Separated Row- and Column-wise Nonuniformities: Given the special cases when the gain and offset factors have either only row- or only column-wise components, the gradient could be simplified into a one dimensional form. In the current case for the the $\{a_{i,j}\}$ gradients, the sum's in the below equation are not

collapsing into a one dimensional problem at once, due to the direct coupling of the gain factors to the measurements A_M:

$$\text{given:} \quad b_{i,j} = b_{\text{col.},i} \quad \text{and} \quad a_{i,j} = a_{\text{col.},i} \tag{8.104}$$

$$\begin{aligned}\partial_{a_{\text{col.},m}}(E_{\text{D}}) = \sum_j -2\lambda T_{\text{N}}\Big(&\frac{1}{a_{\text{col.},m}}\Big(\frac{1}{a_{\text{col.},m}}\underbrace{\sum_{(k,l)\in S}\frac{1}{a_{\text{col.},k}}\text{cov}(C_{\text{M},m,j}, C_{\text{M},k,l})\Psi(\cdot)}_{\Upsilon_{k,l}}\Big)\\ &+\sum_{(k,l)\in S}\underbrace{\frac{1}{a_{\text{col.},m}}(A_{\text{M},m,j} - b_{\text{col.},m})}_{A^{\text{corr.}}_{\text{M},m,j}}\underbrace{\frac{1}{a_{\text{col.},k}}(A_{\text{M},k,l} - b_{\text{col.},m})}_{A^{\text{corr.}}_{\text{M},k,l}}\Psi(\cdot)\end{aligned} \tag{8.105}$$

$$\begin{aligned}= -2\lambda T_{\text{N}}\Big(&\frac{1}{a_{\text{col.},m}}\Big(\frac{1}{a_{\text{col.},m}}\underbrace{\sum_{(k,l)\in S}\frac{1}{a_{\text{col.},k}}\sum_j\text{cov}(C_{\text{M},m,j}, C_{\text{M},k,l})\Psi(\cdot)}_{\Upsilon_{k,l}}\Big)\\ &+\sum_j\sum_{(k,l)\in S}\underbrace{\frac{1}{a_{\text{col.},m}}(A_{\text{M},m,j} - b_{\text{col.},m})}_{A^{\text{corr.}}_{\text{M},m,j}}\underbrace{\frac{1}{a_{\text{col.},k}}(A_{\text{M},k,l} - b_{\text{col.},m})}_{A^{\text{corr.}}_{\text{M},k,l}}\Psi(\cdot)\end{aligned} \tag{8.106}$$

The main reason is the multiplicative coupling of the measurements, instead of an additive coupling. Of course careful assumptions in special cases may allow to move the $\sum_j$ to the Ψ function, yielding to the same notations as in eq. 8.91 for the DSNU. However, a generalized way to reduce the problem into a one dimensional convolution is not possible at the moment.

8.5.5. Solution and a Proof of Concept

The final equations to solve a combined PRNU and DSNU maximum likelihood estimation with the extended row-, column-, and pixel-wise nonuniformity dependencies is expressed as:

$$0 = \begin{pmatrix} \nabla_{\{a_{\text{col.},i}\}}E_{\text{P}} + \nabla_{\{a_{\text{col.},i}\}}E_{\text{D}} \\ \nabla_{\{a_{\text{row.},j}\}}E_{\text{P}} + \nabla_{\{a_{\text{row.},j}\}}E_{\text{D}} \\ \nabla_{\{a_{\text{pix.},i,j}\}}E_{\text{P}} + \nabla_{\{a_{\text{pix.},i,j}\}}E_{\text{D}} \\ \nabla_{\{b_{\text{col.},i}\}}E_{\text{P}} + \nabla_{\{b_{\text{col.},i}\}}E_{\text{D}} \\ \nabla_{\{b_{\text{row.},j}\}}E_{\text{P}} + \nabla_{\{b_{\text{row.},j}\}}E_{\text{D}} \\ \nabla_{\{b_{\text{pix.},i,j}\}}E_{\text{P}} + \nabla_{\{b_{\text{pix.},i,j}\}}E_{\text{D}} \end{pmatrix}, \tag{8.107}$$

where the partial gradients are given by the equations 8.69-8.74, 8.86-8.88 and 8.101-8.103. A typical iterative gradient decent solver is capable of solving this problem [45], but as start value the offset variables should be close to μ_b and the gain variables should be close to μ_a. As for the combined approach large values of

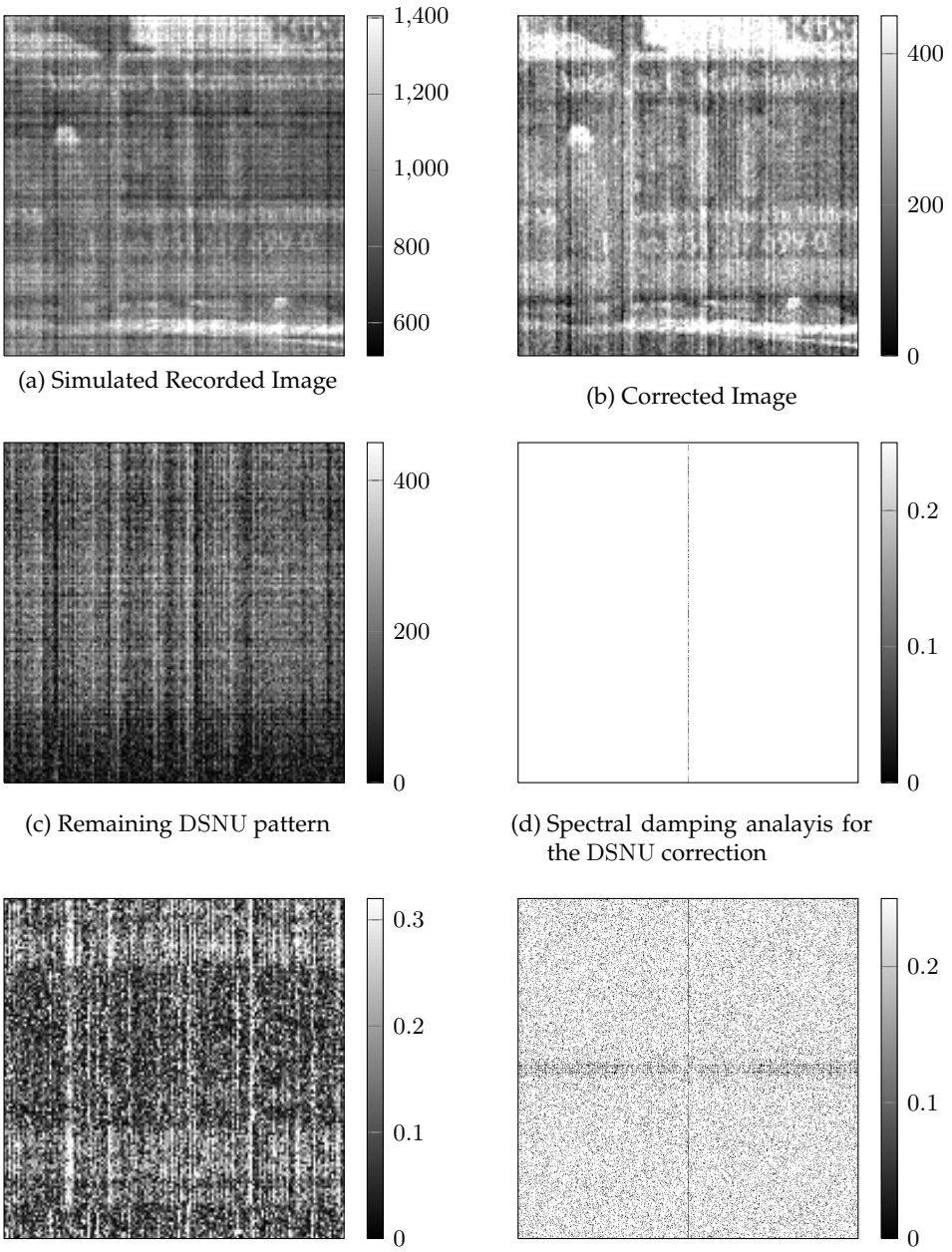
$\{a_{i,j}\}$ lead to a converging solution towards 0, which is a trivial and wrong minimum of the energy function. The gain variables have therefore been safeguarded in the solver to avoid unnaturally large values. The methods that arise from eq. 8.107 with the different choices of the high pass filter HP can now be evaluated and the improvements by substituting $A_M \rightarrow A_{\mathrm{M}_W}$ are of course applied as well.

First experimental results that engaged in finding an optimal parameter set as done before did however not succeed in reasonable correction rates. Anyway, a proof of concept of the new method was possible with the help of the artificial nonuniformities from sec. 8.4.1. These nonuniformities lead to a severe visual degradation by each component of the nonuniformity parameters: row-, column- and pixel-wise offset and gain parameters. In addition to the severe degradation, not only λ, but also α and β had to be tuned to achieve the results presented below. This effectively ignores the fixed statistics and the previous knowledge assumptions.

The results of the best found parameter set for the choice of $HP = HP(\nabla, \mathrm{BL})$ are shown in fig. 8.11 for the row-wise only correction and in fig. 8.12 for the column-wise only correction. The remaining nonuniformities with their column- and row-wise preference are clearly visible in the corrected images, and due to the high amplitude of the nonuniformity, the correction is visible in the remaining patterns as well. The spectral damping shows the expected line of one pixel width in the horizontal and vertical direction in fig. 8.11 and in fig. 8.12 for the DSNU correction. For the PRNU correction an additional damping of a wider stripe of the frequencies of the orthogonal direction is visible as well. This effect may be a result of the multiplicative nature of the PRNU and the consequent effects in the frequency domain. The specific mechanism is to be investigated in further research and can not yet be given.

Fig. 8.13 finally shows the combined row-, column-, and pixel-wise corrections where a well corrected image is displayed. The remaining DSNU pattern has still many high frequencies left, but in the spectral damping analysis, the preference for a correction of spatial high frequencies is clearly visible. Further, the effects of the row-, and column-wise specific corrections show up as the typical cross in the area of remaining low frequency components. The PRNU components as well left behind many artifacts and here as well the low frequencies are damped less in the spectral analysis, which is the expected result. On the other hand, much more of the higher spatial frequency components are not damped and the effects of the horizontal and vertical corrections are not clearly visible.

Summarizing, the mathematics expressions for the combined DSNU and PRNU RCP method have been derived, however a solution for the mentioned parameter instabilities needs to be found before this version of the method can be used in a reliable way.



(a) Simulated Recorded Image

(b) Corrected Image

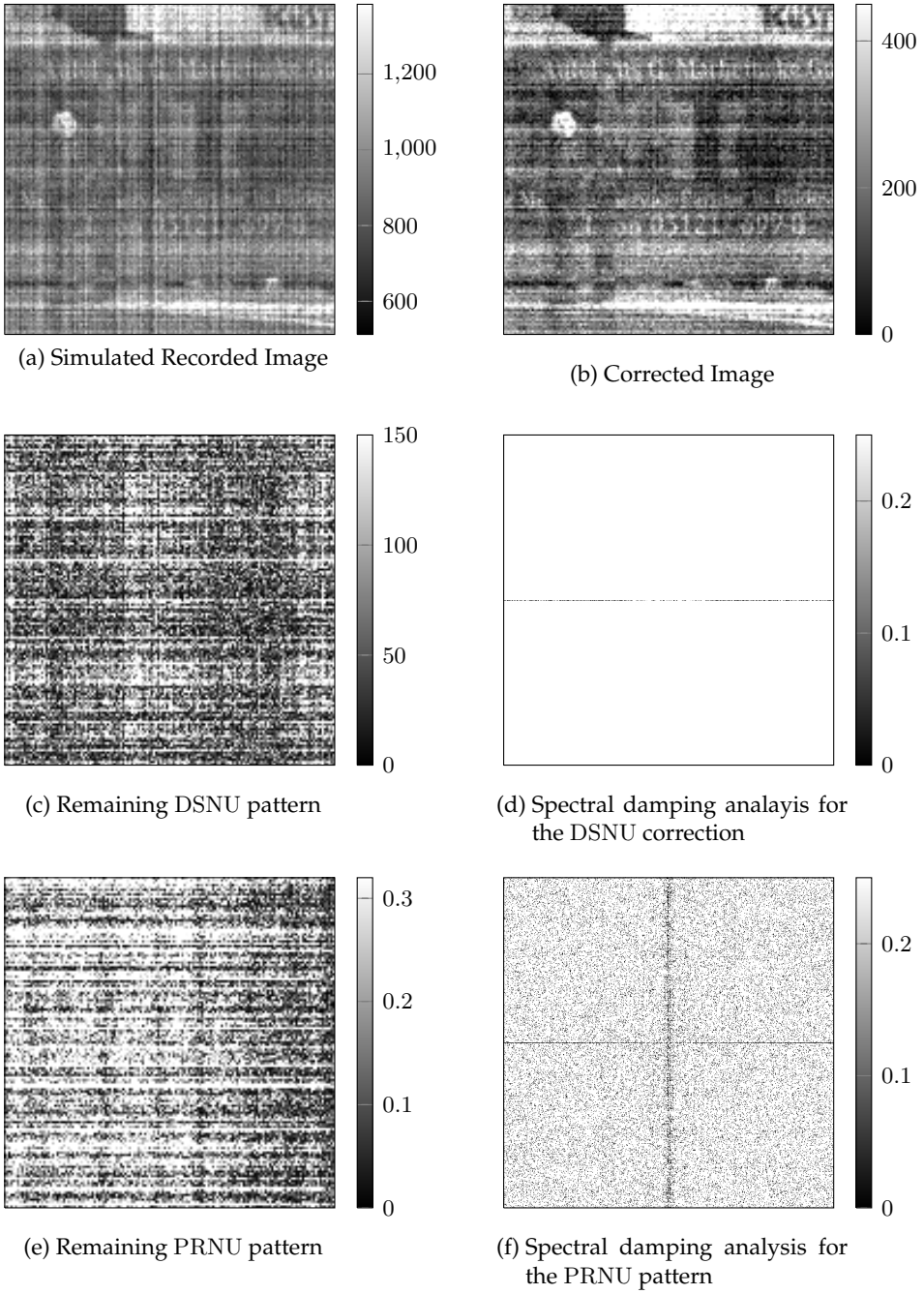
(c) Remaining DSNU pattern

(d) Spectral damping analayis for the DSNU correction

(e) Remaining PRNU pattern

(f) Spectral damping analysis for the PRNU pattern

Figure 8.11.: Proof of concept for the row-wise combined PRNU and DSNU correction



(a) Simulated Recorded Image

(b) Corrected Image

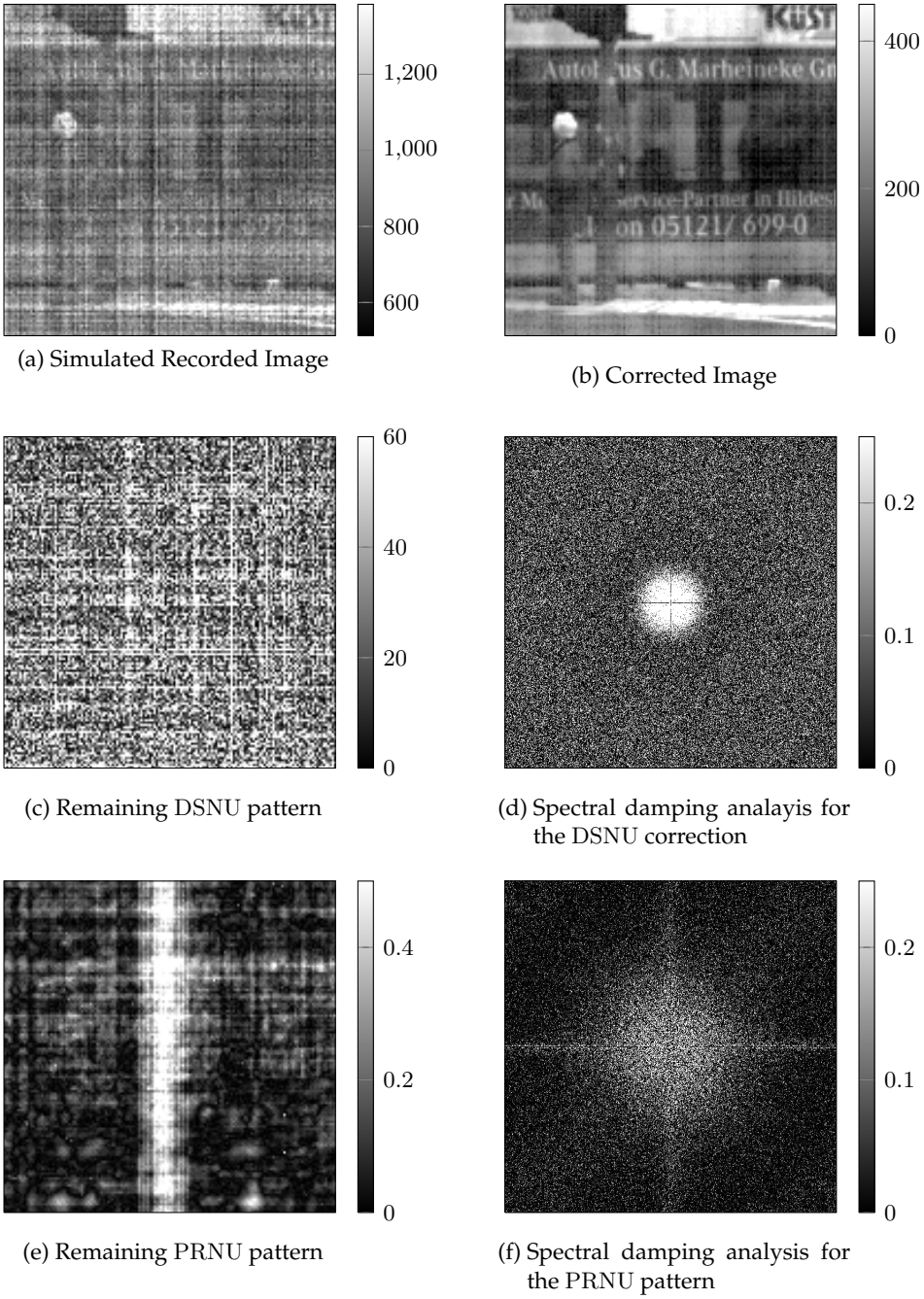
(c) Remaining DSNU pattern

(d) Spectral damping analayis for the DSNU correction

(e) Remaining PRNU pattern

(f) Spectral damping analysis for the PRNU pattern

Figure 8.12.: Proof of concept for the column-wise combined PRNU and DSNU correction.



(a) Simulated Recorded Image

(b) Corrected Image

(c) Remaining DSNU pattern

(d) Spectral damping analayis for the DSNU correction

(e) Remaining PRNU pattern

(f) Spectral damping analysis for the PRNU pattern

Figure 8.13.: Proof of concept for the combined combined row-, column-, and pixel-wise PRNU and DSNU correction.

9. A Mathematical Link to the Reference Methods

As already mentioned in the chapters 5, 6 and 8, the new new maximum likelihood method has some common properties to the LMS method and constant statistic methods like the dependency on a high pass filter and the usage of temporal averages. This chapter shows how the newly developed methods can be linked in a mathematical correct way to the LMS and constant statistic methods. The thereby used assumptions give a deeper understanding of the methods and help to explain the further results in chapter 10. The contents of this chapter have been partly published during the research [13].

To repeat, the maximum likelihood estimation method was given in eq. 6.79 as the solution of the following equation:

$$\begin{pmatrix} -2\lambda T_N((\Upsilon \oslash a) + (A_{\mathrm{M}}^{\mathrm{corr.}} \otimes A_{\mathrm{M}}^{\mathrm{corr.}} * HP(\hat{L}, \Phi))) \oslash a + 2\alpha(a - \mu_a) \\ -2\lambda T_N(A_{\mathrm{M}}^{\mathrm{corr.}} * HP(\hat{L}, \Phi)) \oslash a + 2\beta(b - \mu_b) \end{pmatrix} = 0 \quad (9.1)$$

9.1. A Link to the Constant Statistic Method

The link to the constant statistic method can be formed by two assumptions. First, a simplified probability distribution of the parameters $\{a_{i,j}, b_{i,j}\}$ is used and second, a simplified HP filter mask is used.

9.1.1. Uniform Probability Densities for DSNU and PRNU

As a first approximation, another model for the occurrence probabilities of the $\{a_{i,j}, b_{i,j}\}$ parameters is used. In sec. 6.2.1 and sec. 8.2 the probability densities of the $\{a_{i,j}\}$ and $\{b_{i,j}\}$ parameters were given as Gaussian probability distributions, which was motivated by measurements of the sensor's ground truth data.

If the ground truth knowledge is not available, the more trivial assumption of a unity probability distribution can still be made:

$$\mathbf{b}_{i,j} \sim \mathcal{U}(b_{\mathrm{M}in}, b_{\mathrm{M}ax}) \tag{9.2}$$

$$\text{with} \quad f_{\mathbf{b}_{i,j}}(b_{i,j}) = \begin{cases} \frac{1}{b_{\mathrm{M}ax} - b_{\mathrm{M}in}} & \text{for} \quad b_{i,j} \in [b_{\mathrm{M}in}, b_{\mathrm{M}ax}] \\ 0 & \text{else} \end{cases} \quad \forall (i,j) \tag{9.3}$$

$$\text{and} \quad \mathbf{a}_{i,j} \sim \mathcal{U}(a_{\mathrm{M}in}, a_{\mathrm{M}ax}) \tag{9.4}$$

$$\text{with} \quad f_{\mathbf{a}_{i,j}}(a_{i,j}) = \begin{cases} \frac{1}{a_{\mathrm{M}ax} - a_{\mathrm{M}in}} & \text{for} \quad a_{i,j} \in [a_{\mathrm{M}in}, a_{\mathrm{M}ax}] \\ 0 & \text{else} \end{cases} \quad \forall (i,j) \tag{9.5}$$

Following the log-likelihood derivation of sec. 6.3, the probability density just gives a constant value for parameters inside the boundaries $[a_{\mathrm{M}in}, a_{\mathrm{M}ax}]$ and $[b_{\mathrm{M}in}, b_{\mathrm{M}ax}]$. The derivative of these constants results as 0. Outside the boundaries, the negative logarithm converges towards $+\infty$, forming a box energy potential. When deriving the gradient this jump from a constant to $+\infty$ actually binds the solution to the given intervals $[a_{\mathrm{M}in}, a_{\mathrm{M}ax}]$ and $[b_{\mathrm{M}in}, b_{\mathrm{M}ax}]$. This of course can be implemented in the iterative solver, or embedded into the nonlinear CNN function from sec. 7.2.

Using this assumption, eq. 9.1 can be rewritten as:

$$\text{solve:} \quad \begin{pmatrix} -2\lambda T_N((\Upsilon \oslash a) + (A_{\mathrm{M}}^{\mathrm{corr.}} \otimes A_{\mathrm{M}}^{\mathrm{corr.}} * HP(\hat{L}, \Phi))) \oslash a \\ -2\lambda T_N(A_{\mathrm{M}}^{\mathrm{corr.}} * HP(\hat{L}, \Phi)) \oslash a \end{pmatrix} = 0 \tag{9.6}$$

$$\text{subject to:} \quad b_{i,j} \in [b_{\mathrm{M}in}, b_{\mathrm{M}ax}] \quad \text{and} \quad a_{i,j} \in [a_{\mathrm{M}in}, a_{\mathrm{M}ax}] \quad \forall i,j$$

9.1.2. A Trivial Filter Mask

As a further approximation, the filter mask HP is modified. Especially the choice of a smaller filter mask may be considered due to large advantages in the resulting computational efforts. The most trivial approximation of the filter mask is given as the 1×1 mask

$$HP = 1\,. \tag{9.7}$$

As a consequence simplifications occur for the components that are connected by the HP-mask. For the Υ matrix one obtains:

$$\sum_{(k,l)\in S} \frac{1}{a_{k,l}} \mathrm{cov}(C_{\mathrm{M},m,n}, C_{\mathrm{M},k,l}) \Psi(|m-k|, |n-l|)) \tag{9.8}$$

$$= \frac{1}{a_{\mathrm{M},n}} \mathrm{cov}(C_{\mathrm{M},m,n}, C_{\mathrm{M},m,n}) \tag{9.9}$$

$$= \frac{1}{a_{\mathrm{M},n}} \sigma^2(C_{\mathrm{M},m,n})\,, \tag{9.10}$$

with σ^2 as the temporal variance of the measurements. This and the substitution of $HP = 1$ can be included into eq. 9.6, which gives the following simplifications:

$$\begin{pmatrix} -2\lambda T_N((\sigma^2(C_\mathrm{M}) \oslash a) + (A_\mathrm{M}^{\mathrm{corr.}} \otimes A_\mathrm{M}^{\mathrm{corr.}})) \oslash a \\ -2\lambda T_N A_\mathrm{M}^{\mathrm{corr.}} \oslash a \end{pmatrix} = 0 \tag{9.11}$$

$$\begin{pmatrix} ((\sigma^2(C_\mathrm{M}) \oslash a) + (A_\mathrm{M}^{\mathrm{corr.}} \otimes A_\mathrm{M}^{\mathrm{corr.}})) \\ A_\mathrm{M}^{\mathrm{corr.}} \end{pmatrix} = 0 \tag{9.12}$$

$$\begin{pmatrix} ((\sigma^2(C_\mathrm{M}) \oslash a) + (A_\mathrm{M}^{\mathrm{corr.}} \otimes A_\mathrm{M}^{\mathrm{corr.}})) \\ \frac{1}{a}(A_\mathrm{M} - b) \end{pmatrix} = 0 \tag{9.13}$$

$$\begin{pmatrix} ((\sigma^2(C_\mathrm{M}) \oslash a) + (A_\mathrm{M}^{\mathrm{corr.}} \otimes A_\mathrm{M}^{\mathrm{corr.}})) \\ A_\mathrm{M} \end{pmatrix} = \begin{pmatrix} 0 \\ b \end{pmatrix} \tag{9.14}$$

$$\begin{pmatrix} \sigma^2(C_\mathrm{M}) \oslash a \\ A_\mathrm{M} \end{pmatrix} = \begin{pmatrix} 0 \\ b \end{pmatrix} \tag{9.15}$$

$$\text{subject to:} \quad b_{i,j} \in [b_{\mathrm{M}in}, b_{\mathrm{M}ax}] \text{and} \quad a_{i,j} \in [a_{\mathrm{M}in}, a_{\mathrm{M}ax}] \quad \forall i,j \tag{9.16}$$

The last equation $\sigma^2(C_\mathrm{M}) \oslash a = 0$ does not have a solution, but if we approximate its solution for a small ϵ instead of 0 we obtain:

$$\begin{pmatrix} \sigma^2(C_\mathrm{M}) \oslash a \\ A_\mathrm{M} \end{pmatrix} = \begin{pmatrix} \epsilon \\ b \end{pmatrix} \tag{9.17}$$

$$\begin{pmatrix} \sigma^2(C_\mathrm{M}) \\ \mu(C_\mathrm{M}) \end{pmatrix} = \begin{pmatrix} a\epsilon \\ b \end{pmatrix} \tag{9.18}$$

This means that the offset parameters $\{b_{i,j}\}$ are approximated by the temporal averaged sensor measurements, which is exactly what eq. 5.1 on p. 80 states for the constant statistic method. The estimation of the gain factors depends linearly on the temporal variance of the measurements. Given the additional assumption of the CS method that the spatial mean of the gain factors should be one ($e(\{a_{i,j}\}) = 1$), then this result is similar to estimation rule of the constant statistics method. The only difference is that the gain factors a had been identified with either the standard deviation or the absolute deviation of the measurements in the final CS algorithms.

With the assumptions from sec. 9.1.1 and 9.1.2 it is shown that the constant statistic methods can be regarded as the approximation of a maximum likelihood approach under the approximations of neglected pixel-to-pixel interactions (HP-mask becomes 1) and a uniform distribution of the parameters. Of course these approximations neglect a huge amount of the actually available information. However, this way of deriving the constant statistic method forms an alternative access to the CS method and allows to explain its weaknesses from another point of view. This insight might lead to an improvement of the CS method in further investigations. For example the weighted averages A_{M_W} could be considered for the estimation of b.

9.2. Approximations towards the LMS Algorithms

Besides the above introduced approximation of the new maximum likelihood approach into the constant statistic method, the approximation into the LMS type methods is possible as well. Therefore the assumptions of uniform distributed parameter sets $\{a_{i,j}\}$ and $\{b_{i,j}\}$ from sec. 9.1.1 are used in addition to the here introduced further assumptions.

9.2.1. Neglecting the Explicit Spatial Sensor-Dependencies

The approximation of the HP-mask as a scalar 1 leads to a total loss of the information that is stored in the dependencies of the neighboring light sensors. In the end, the poor performance of the CS method made clear that this approximation simplifies too much to achieve the desired correction performance.

A modification to the above approximation can be found by just considering the individual sensor responses as independent in the transformations made during the derivation. Of course this is not a good approximation, because the point spread function of the optics and the sensor responses of typical scenes usually correlate the responses of neighboring sensors. However, it is possible to specifically neglect only the correlation in the definition of Υ from eq. 8.41, which will preserve most of the neighbor dependencies:

$$\begin{aligned}\partial_{a_{\mathrm{M},n}} E_D = -2\lambda T_N \Big(\frac{1}{a_{\mathrm{M},n}^2 a_{k,l}} \Big(\sum_{(k,l)\in S} \underbrace{A^2_{\mathrm{M},m,n,k,l} - A_{\mathrm{M},m,n} A_{\mathrm{M},k,l}}_{=0 \quad \text{see below.}} \\ + \left((A_{\mathrm{M},m,n} - b_{\mathrm{M},n})(A_{\mathrm{M},k,l} - b_{k,l})\right)\Psi(\cdot)\Big)\end{aligned} \tag{9.19}$$

we approximate $C_{\mathrm{M},m,n}$ and $C_{\mathrm{M},k,l}$ as independent for this equation $\forall (m,n),(k,l)$:

$$\Rightarrow A_{\mathrm{M},m,n} A_{\mathrm{M},k,l} = A^2_{\mathrm{M},m,n,k,l} \tag{9.20}$$

$$\begin{aligned}&\text{with:} \quad \mu_t(C_{\mathrm{M},m,n}) = A_{\mathrm{M},m,n} \quad \text{and} \quad \mu_t(C_{\mathrm{M},k,l}) = A_{\mathrm{M},k,l} \\ &\text{because then:} \quad \mu_t(C_{\mathrm{M},m,n}, C_{\mathrm{M},k,l}) = \mu_t(C_{\mathrm{M},k,l}) \cdot \mu_t(C_{\mathrm{M},m,n})\end{aligned} \tag{9.21}$$

As a result the covariance term disappears from the equation and together with the uniform distribution assumption from eq. 9.6 the equation to solve develops to:

$$\begin{pmatrix} ((A_{\mathrm{M}}^{\mathrm{corr.}} \otimes A_{\mathrm{M}}^{\mathrm{corr.}} * HP(\hat{L},\Phi)) \\ (A_{\mathrm{M}}^{\mathrm{corr.}} * HP(\hat{L},\Phi)) \end{pmatrix} = 0\,. \tag{9.22}$$

Next, the corresponding gradient in the gain and offset variables can be rewritten as:

$$\begin{pmatrix} \hat{\nabla}_{\{a_{i,j}\}}(E_{\text{Appr.}}) \\ \hat{\nabla}_{\{b_{i,j}\}}(E_{\text{Appr.}}) \end{pmatrix} = \begin{pmatrix} -2\lambda T_N(((A_{\text{M}}^{\text{corr.}} \otimes A_{\text{M}}^{\text{corr.}} * HP(\hat{L}, \Phi))) \oslash a \\ -2\lambda T_N(A_{\text{M}}^{\text{corr.}} * HP(\hat{L}, \Phi)) \oslash a \end{pmatrix} \tag{9.23}$$

9.2.2. Approximating the Steepest Descent Solver

Following the maximum likelihood approach of sec. 6, then eq. 9.23 has to be solved with help of an iterative gradient based solver. Therefore the minimization of the corresponding energy function $E_{\text{Appr.}}(\{a_{i,j}, b_{i,j}\})$ should be considered to find further possibilities for simplification:

$$\{a_{i,j}, b_{i,j}\} = \arg \min_{\{a_{i,j}, b_{i,j}\}} \{E_{\text{Appr.}}\} \tag{9.24}$$

As a trivial iterative solver, a steepest descent solver can fulfill this task and it can be expressed as:

$$\{a_{i,j}, b_{i,j}\} = \arg \min_{\{a_{i,j}, b_{i,j}\}} \{E_{\text{Appr.}}\} \tag{9.25}$$

$$\Rightarrow \{a_{i,j}, b_{i,j}\}_{n+1} = \{a_{i,j}, b_{i,j}\}_n - \epsilon_n \cdot \nabla E_{\text{Appr.}}(\{a_{i,j}, b_{i,j}\}_n) \quad n = 0 \ldots \infty \tag{9.26}$$

$$= \{a_{i,j}, b_{i,j}\}_n - \epsilon \cdot \nabla E_{\text{Appr.}}(\{a_{i,j}, b_{i,j}\}_n) \quad n = 0 \ldots \infty \,. \tag{9.27}$$

The last equation contains the simplification of $\epsilon_n \rightarrow \epsilon$. While in the typical solver the step size ϵ_n is optimized for each step, a suitable solution can be also be achieved with a fixed and small step size ϵ [45], if a slow convergence is accepted.

The gradient $\nabla E_{\text{Appr.}}$ was given in eq. 9.23 and the temporal averaging can be moved in front to allow further transformations:

$$\begin{pmatrix} \hat{\nabla}_{\{a_{i,j}\}}(E_{\text{Appr.}}) \\ \hat{\nabla}_{\{b_{i,j}\}}(E_{\text{Appr.}}) \end{pmatrix} = \begin{pmatrix} \sum_{t_n} -2\lambda(((C_{\text{M}}^{\text{corr.}} \otimes C_{\text{M}}^{\text{corr.}} * HP(\hat{L}, \Phi))) \oslash a \\ \sum_{t_n} -2\lambda(C_{\text{M}}^{\text{corr.}} * HP(\hat{L}, \Phi)) \oslash a \end{pmatrix} \tag{9.28}$$

$$= \begin{pmatrix} \sum_{t_n} \hat{\nabla}_{\{a_{i,j}\}}(E_{\text{Appr.},t_n}) \\ \sum_{t_n} \hat{\nabla}_{\{b_{i,j}\}}(E_{\text{Appr.},t_n}) \end{pmatrix} , \tag{9.29}$$

which allows to rewrite the whole gradient as a sum over its temporal sample points:

$$\nabla E_{\text{Appr.}}(\{a_{i,j}, b_{i,j}\}_n) = \sum_{t_n} \nabla E_{\text{Appr.},t_n} \,. \tag{9.30}$$

The steepest descent update rule of eq. 9.26 is then expressed by

$$\{a_{i,j}, b_{i,j}\}_{n+1} = \{a_{i,j}, b_{i,j}\}_n - \epsilon \sum_{t_n} \nabla E_{\text{Appr.},t_n}(\{a_{i,j}, b_{i,j}\}_n) \tag{9.31}$$

$$= \underbrace{\{a_{i,j}, b_{i,j}\}_n - \epsilon \cdot \nabla E_{\text{Appr.},t_1}(\{a_{i,j}, b_{i,j}\}_n) - \epsilon \cdot \nabla E_{\text{Appr.},t_2}(\{a_{i,j}, b_{i,j}\}_n) \ldots}_{\approx \{a_{i,j}, b_{i,j}\}_{n^*}}$$

$$- \epsilon \cdot \nabla E_{\text{Appr.},t_n}(\{a_{i,j}, b_{i,j}\}_n) \tag{9.32}$$

$$= \{a_{i,j}, b_{i,j}\}_{n^*} - \epsilon \cdot \nabla E_{\text{Appr.},t_n}(\{a_{i,j}, b_{i,j}\}_n)\,, \tag{9.33}$$

where the defined $\{a_{i,j}, b_{i,j}\}_{n^*}$ is the result of the temporally split gradient updates for the *current* point $\{a_{i,j}, b_{i,j}\}_n$. It is now possible to approximate $\{a_{i,j}, b_{i,j}\}_{n^*} \approx \{a_{i,j}, b_{i,j}\}_n$, if one assumes that the gradient parts from the previous time steps t_n are small in comparison to the gradient part $E_{\text{Appr.},t_n}(\{a_{i,j}, b_{i,j}\}_n)$. This approximation suggests that most of the new information comes from the actual time step, which in general is not true, as discussed below. However, such simplified resulting modified steepest descent solver is the first step to approximate the LMS methods:

$$\{a_{i,j}, b_{i,j}\} = \arg \min_{\{a_{i,j}, b_{i,j}\}} \{E_{\text{Appr.}}\} \tag{9.34}$$

$$\Rightarrow \{a_{i,j}, b_{i,j}\}_{n+1} = \{a_{i,j}, b_{i,j}\}_{n^*} - \epsilon \cdot \nabla E_{\text{Appr.},t_n}(\{a_{i,j}, b_{i,j}\}_n) \tag{9.35}$$

An advantage of this approximation is that only the gradient $\nabla E_{\text{Appr.},t_n}(\{a_{i,j}, b_{i,j}\}_n)$ of the actual frame input is needed which can be calculated just from the actual frame as shown in the next section.

As already mentioned, this approximation is obviously wrong, as the information of each frame will influence the gradient and not only the last frame. On the other hand, for high speed recorded image data, the influence of the highly correlated older frames can be neglected as it does not provide additional information. The LMS method will neglect this redundant information which is actually the main reason why the LMS methods will have their advantage in those applications (see chap. 10).

9.2.3. Extracting the LMS Update Rules from the Approximation

To finally extract the LMS update rules, eq. 9.29 together with eq. 9.35 gives the definition of the gradient which results in the following update rules in the variables a and b:

$$a_{n+1} = a_n + 2\epsilon\lambda((C_{\text{M}}^{\text{corr.}} \otimes C_{\text{M}}^{\text{corr.}} * HP(\hat{L}, \Phi)) \oslash a \tag{9.36}$$

$$b_{n+1} = b_n + 2\lambda(C_{\text{M}}^{\text{corr.}} * HP(\hat{L}, \Phi)) \oslash a\,. \tag{9.37}$$

The definition of $C_{\mathrm{M}}^{\mathrm{corr.}\cdot}$ correspond to the C_{est} of eq. 5.28 and thus only the variable naming and the $\oslash a$ factors differ. The LMS method uses the variables

$$w_{i,j} = \frac{1}{a_{i,j,\mathrm{est.}}} \tag{9.38}$$

$$o_{i,j} = -\frac{b_{i,j,\mathrm{est.}}}{a_{i,j,\mathrm{est.}}} \tag{9.39}$$

in their update rules and ignores their dependencies on the physical parameters a and b. This approximation is only valid in a linear region of $a_{i,j,\mathrm{est.}} \approx 1$ (as mentioned on p. 84f.). The $\oslash a$ terms can therefore be approximated as 1, leading to identical update rules for the gain and offset correction variables as presented by the LMS method (compare eq. 5.21 on p. 85 and consider the variable transformations):

$$a_{n+1} = a_n + 2\underbrace{\epsilon\lambda}_{\epsilon^*} C_{\mathrm{M}}^{\mathrm{corr.}\cdot} \otimes (C_{\mathrm{M}}^{\mathrm{corr.}\cdot} * HP(\hat{L}, \Phi)) \tag{9.40}$$

$$b_{n+1} = b_n + 2\underbrace{\epsilon\lambda}_{\epsilon^*} C_{\mathrm{M}}^{\mathrm{corr.}\cdot} * HP(\hat{L}, \Phi) \tag{9.41}$$

$$\text{subject to:} \quad b_{i,j} \in [b_{\mathrm{M}in}, b_{\mathrm{M}ax}] \quad \text{and} \quad a_{i,j} \in [a_{\mathrm{M}in}, a_{\mathrm{M}ax}] \quad \forall i,j \tag{9.42}$$

In the current approximation the additional *subject to* constrains have to be considered, which are not mentioned explicitly in the derivation of the LMS method. However, these subject to constrains however prevent the method's parameters to leave their considered scopes, which is an important feature.

Summarizing, the LMS update rules are an approximation of the maximum likelihood estimation if the following assumptions are made: Uniform probability distribution of the parameters $\{a_{i,j}, b_{i,j}\}$, neglecting the sensor-wise correlations concerning the measurements, neglecting gradient parts that arise from *older* information and finally, by approximating the gain solution around $a_{i,j,\mathrm{est.}} = 1$. While the LMS methods provide their HP filter masks from ad-hoc assumptions, the newly derived filter masks can be used as well as a substitution in further research.

9.3. Approximation of the Fast Adaptive LMS Algorithms

The improvement of the LMS update rules towards the *adaptive* LMS update rules differ only by the definition of the step size. For the adaptive LMS method, the step size changes depending on the sensor measurements: $\epsilon^* \rightarrow \epsilon^*(C_{\mathrm{M}})$. In eq. 9.29 the temporal averaging was exploited to extract the LMS update rules. This temporal averaging was improved in sec. 7.1 with a weighted averaging,

motivated by the adaptive LMS methods' adaptive step size. Using a weighted temporal averaging in the above deliberations allows to rewrite the gradient to:

$$\begin{pmatrix} \hat{\nabla}_{\{a_{i,j}\}}(E_{\text{Appr.}}) \\ \hat{\nabla}_{\{b_{i,j}\}}(E_{\text{Appr.}}) \end{pmatrix} = \begin{pmatrix} \sum\limits_{t_n} w_{t_n} \hat{\nabla}_{\{a_{i,j}\}}(E_{\text{Appr.}}) \oslash (\sum\limits_{t_n} w_{t_n}) \\ \sum\limits_{t_n} w_{t_n} \hat{\nabla}_{\{b_{i,j}\}}(E_{\text{Appr.}}) \oslash (\sum\limits_{t_n} w_{t_n}) \end{pmatrix}. \tag{9.43}$$

Here w_{t_n} is the matrix that contains the individual weights for the light sensors at the point in time t_n.

Eq. 9.35 can then be changed with exactly the same motivations as before and the weight for the actual time step t_n remains in the update step, producing an adaptive step $\epsilon_{adapt.}$:

$$a_{n+1} = a_n + \underbrace{2\,\epsilon\lambda w_{t_n} \oslash (\sum_{t_n} w_{t_n})}_{\epsilon_{adapt.}} \otimes C_{\text{M}}^{\text{corr.}} \otimes (C_{\text{M}}^{\text{corr.}} * HP(\hat{L}, \Phi)) \tag{9.44}$$

$$b_{n+1} = b_n + \underbrace{2\,\epsilon\lambda w_{t_n} \oslash (\sum_{t_n} w_{t_n})}_{\epsilon_{adapt.}} \otimes C_{\text{M}}^{\text{corr.}} * HP(\hat{L}, \Phi) \tag{9.45}$$

$$\text{subject to:} \quad b_{i,j} \in [b_{\text{M}in}, b_{\text{M}ax}] \quad \text{and} \quad a_{i,j} \in [a_{\text{M}in}, a_{\text{M}ax}] \quad \forall i, j \tag{9.46}$$

The weight factor w_{i,j,t_n} is by now the same as used for the averaging process in sec. 7.1. However it was mentioned that for actually averaged signals, another weighing factor should be used than for frame-to-frame based updates. Therefore the original weights of the adaptive LMS methods from eq. 5.32 p. 87 can be used. This makes the above eq. 5.35 identical to the original adaptive LMS eq. 5.35 on p. 87 and the mathematical link is completed.

9.4. Summary

This chapter showed the different assumptions that were implicitly made by the constant statistic and least mean square reference methods. These assumptions lead to methods that differ from the optimal solution in the newly derived maximum likelihood approach. With the assumptions it can be explained which part of the information is neglected by the reference methods, which ultimately leads to the superior results of the new methods. However, it is also possible now to explain why the LMS methods reach the good performances for consecutive frame sequences in the next chapter.

In cases of high speed sequences, the user has now the choice to learn with the LMS methods a first correction pattern from the correlated frames and use the maximum likelihood method at a later stage if many uncorrelated frames could be averaged.

10. Further Experimental Results and Limitation

In the recent chapters, the newly developed maximum likelihood methods have been extended and improved. The results on random frame sequences showed that the new methods could achieve a superior performance when compared to all reference methods. However, the additional image material of short, high speed consecutive frame sequences and the image material recorded by the cameras with low valued nonuniformities has not yet been fully discussed. Especially the high speed sequences violate the assumptions made in the derivation of the new methods and the expected performance loss has already been discussed in the best-case analysis in the respective sections. These further image materials form a limitation of the superiority of the newly developed methods discussed below. Given the link between the new methods and the reference methods from chapter 9, the further insight in the methods' functionalities can be discussed on the the experimental results.

This chapter discusses at first the performances for the consecutive frame sequences, followed by the evaluations for the low valued nonuniformity image material.

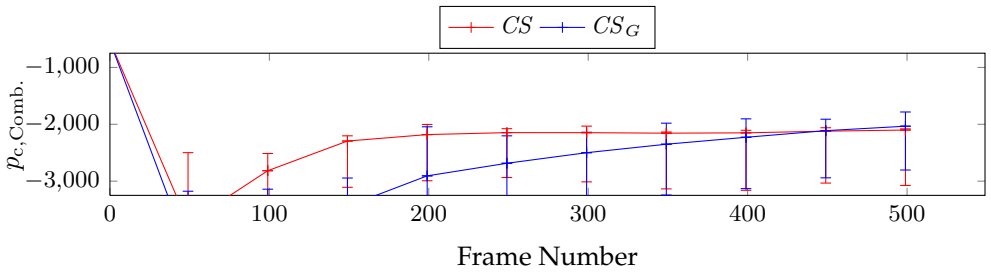
10.1. Results for the Consecutive Frame Sequences

The consecutive sequences that were introduced in sec. 4.1.1 are recorded with the photonfocus MV1-D1312-160-CL camera that has the high valued DSNU of $\approx 50\,\mathrm{DN}$ at 12 bit resolution. The recorded sequence lengths were 500 frames with and the recording rate was $100\,\mathrm{fps}$. The following analysis will show that the new maximum likelihood methods still give a reasonable correction performance despite the fact that the demand for independent input frames is not fulfilled. The reference methods have however the advantage of neglecting the dependencies of the input data by the frame-to-frame update strategy, which results in improved performances.

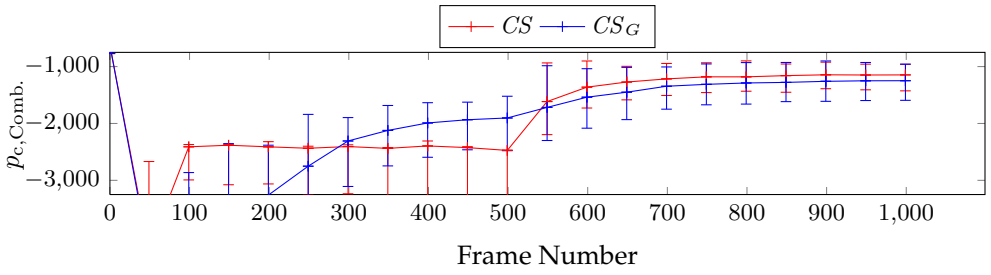
10.1.1. Analysis of the Reference Methods

Resuming the best-case performance analysis of sec. 5.3.2 (p.93) it is already shown that the new methods decrease in their correction performances for the

consecutive frame sequences. The temporal correction performance gives however a better impression as the parameters are fixed, like in the real applications.



(a) CS and CS_G evaluations in the $p_{\mathrm{c,Comb.}}$ metric for 500 frame consecutive frames



(b) CS and CS_G evaluations in the $p_{\mathrm{c,Comb.}}$ metric for 1000 consecutive frames

Figure 10.1.: Temporal performance analysis for consecutive frame sequences of 500 and 1000 frame length of the CS and CS_G methods.

Fig. 10.1 shows the correction performance for the constant statistics methods, CS and CS_G, in the $p_{\mathrm{c,Comb.}}$ quality measure. As expected from their previous performances, the CS and CS_G methods are not able to reach a positive correction rate for neither 500 nor 1000 frame sequences. Similar to the random frame sequences, the performance increases drastically for the longer 1000 frame sequences, where both methods reach similar (negative) performances. This similar performance proves also that the gating threshold does not lead to significant improvements with the reasons discussed in sec. 5.1. The same arguments also repeat for the effect of the gated adaptive LMS methods, which do not show any significant difference to their non-gated equivalent methods (see fig. 10.2).

For the other LMS type methods the behavior is shown in fig. 10.3a-10.3c where the correction rate is measured in the combined quality norm $p_{\mathrm{c,Comb.}}$ and the effect separated quality norms $p_{\mathrm{c,DSNU}}$ and $p_{\mathrm{c,PRNU}}$. In the $p_{\mathrm{c,Comb.}}$-norm the LMS_{TV} method shows the best performance of the compared methods and the $LMS_{A,7\times7}$ method follows directly. The same behavior is observed in the $p_{\mathrm{c,DSNU}}$ and the $p_{\mathrm{c,PRNU}}$ norm. However a positive correction rate for the PRNU is reached by no method. In general, the overall reached performance for the consecutive frame sequences is 5% behind the performance that has been reached with random frame sequences, especially in the important $p_{\mathrm{c,DSNU}}$ norm.

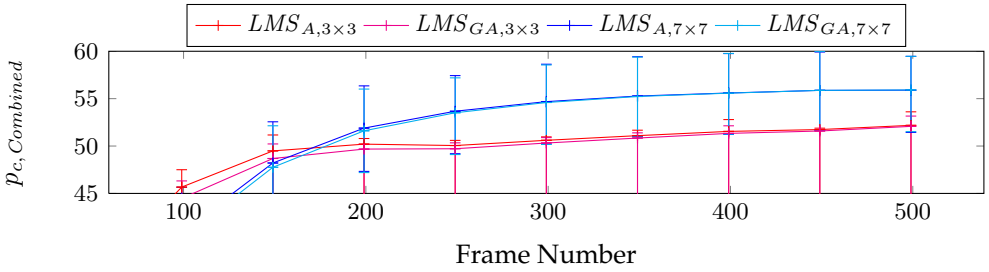
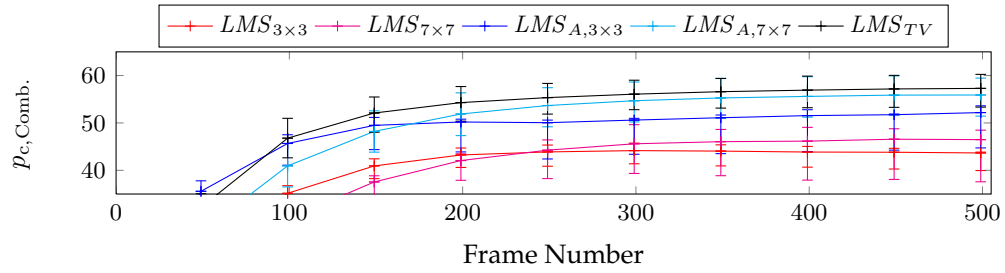
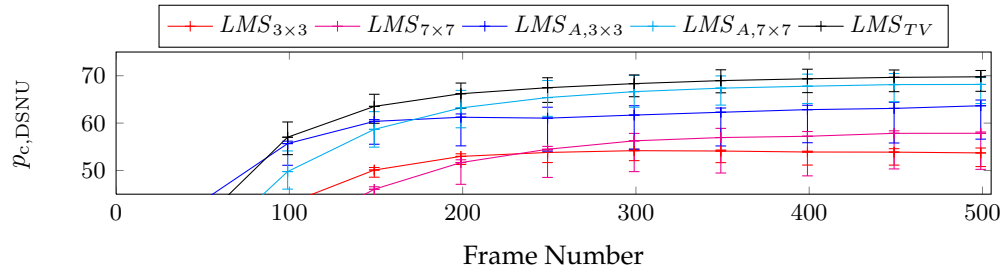


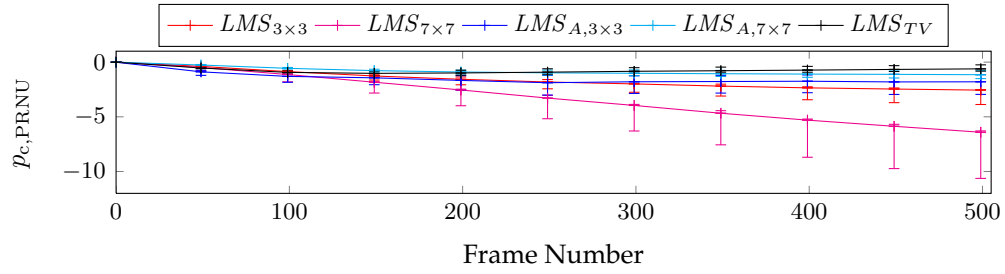
Figure 10.2.: Comparison between the gated adaptive and the adaptive LMS methods in the $p_{c,Combined}$ quality metric for consecutive frame sequences. The differences between the two methods is not notable, as only 2 plot lines are separable.



(a) Evaluation for the $p_{c,\mathrm{Comb.}}$ metric.



(b) Evaluation for the $p_{c,\mathrm{DSNU}}$ metric.



(c) Evaluation for the $p_{c,\mathrm{PRNU}}$ metric.

Figure 10.3.: The LMS methods' evaluations for the different metrics with consecutive frame sequences as input. Parameter set according to tab. 5.3.

For completeness the performances with the literature presented parameter set is shown in fig. 10.4 and results in a decreased performance for consecutive frame sequences, as expected. Only LMS_{TV} reaches a good performance with the literature parameters, but for this method an optimized parameter set is demanded and the method therefore serves only as a reference in the plot.

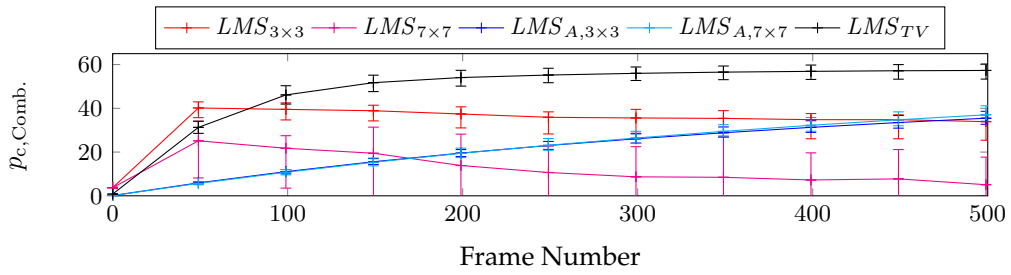


Figure 10.4.: The LMS methods evaluations for the $p_{c,\mathrm{Comb.}}$ metric with the **literature** parameter set from tab. 5.4 and consecutive frame sequences as input.

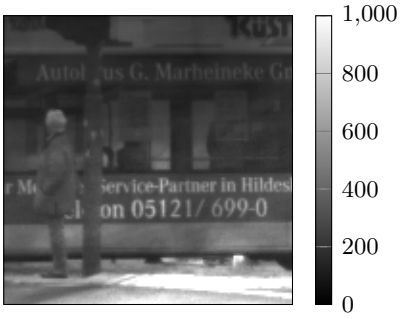
Spatial and Spectral Performance Analysis

The visual quality of the corrections is shown in fig. 10.5 while the ground truth corrected image is shown in fig. 5.16c on p.102 or in the following comparative evaluations. The visual corrections for the CS and CS_G method show a severe deterioration of the images, which did not occur for random frame inputs.

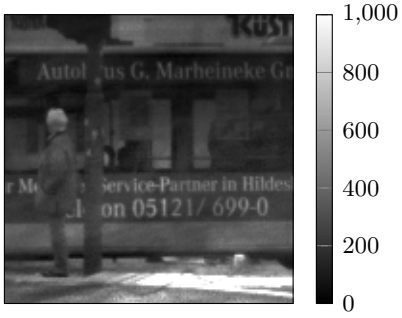
In the group of the LMS-type methods, $LMS_{3\times3}$ and $LMS_{7\times7}$ show high frequency degradations while the adaptive methods $LMS_{A,3\times3}$ and $LMS_{A,7\times7}$ show only one immediate visible artifact, which is located on the street below the feet of the woman. The $LMS_{GA,7\times7}$ method is also shown and as expected and discussed, this method gives no difference to the $LMS_{A,7\times7}$ method. Finally the LMS_{TV} method gives a visually similar impression as the adaptive LMS methods, which is not surprising as this method uses as well a high pass filter with an adaptive step size, however with different realizations.

Fig. 10.6 shows the remaining PRNU factors, and CS and CS_G leave remains with high amplitudes. The LMS-type methods on the other hand show almost no corrections and leave remains with spatial high frequency components. The remaining DSNU patterns are shown in 10.7 and for the CS and CS_G methods the pattern shows huge amplitudes, while the LMS-type algorithms give the typical low amplitude remains. However, the pure LMS methods, $LMS_{3\times3}$ and $LMS_{7\times7}$, did not succeed in correcting the spectral high frequency components completely, while their adaptive equivalents, including the LMS_{TV} method, succeed in this task.

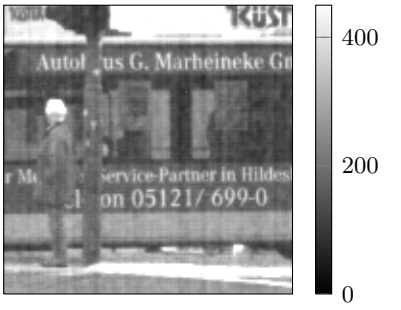
The spectral damping properties of the methods are shown in fig. 10.8 for the PRNU and in fig. 10.9 for the DSNU. For the PRNU there is as expected no special-



(a) Corrected with CS Method



(b) Corrected with CS_G Method



(c) Corrected with $LMS_{3\times3}$ Method

(d) Corrected with $LMS_{7\times7}$ Method

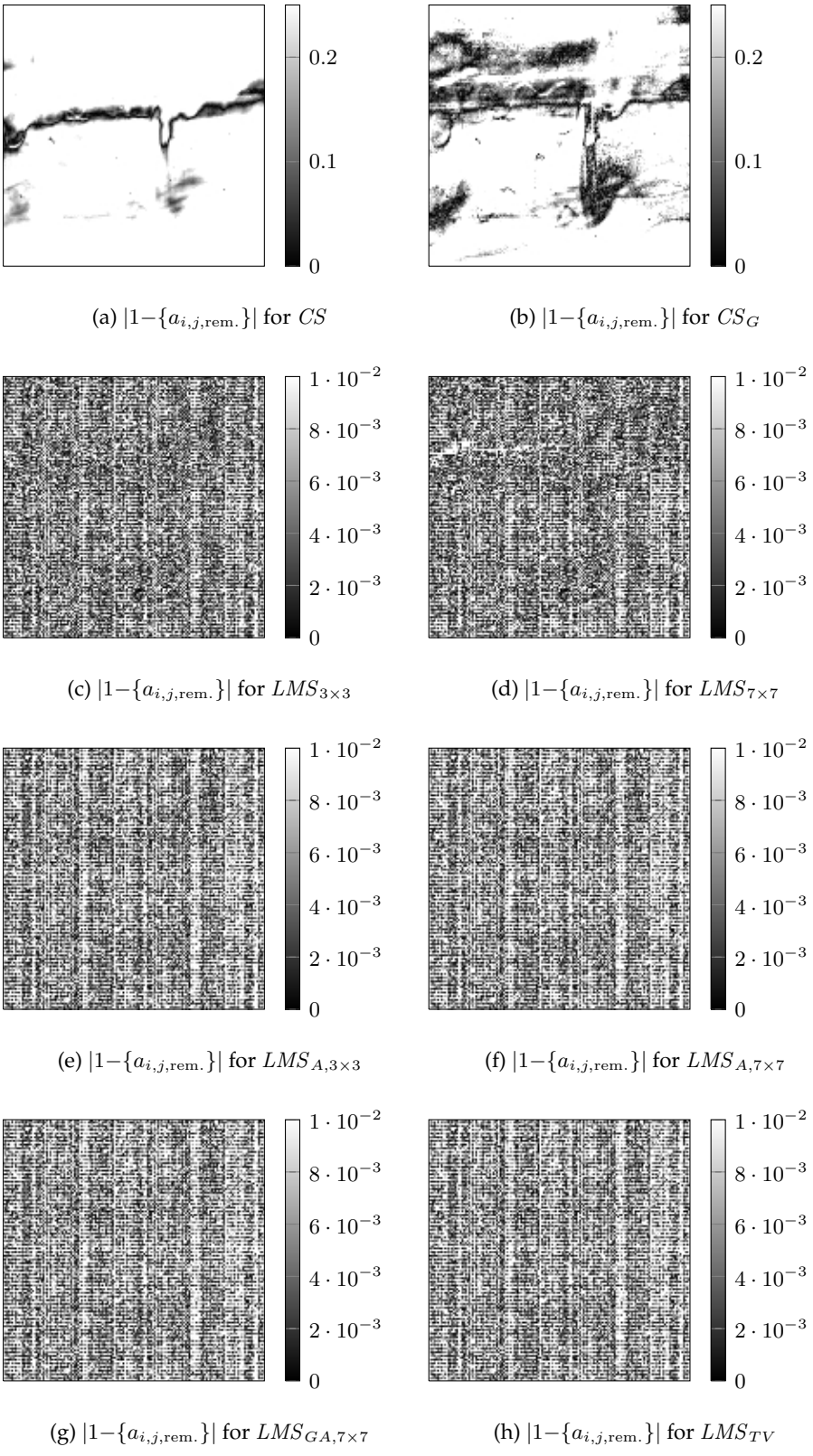
(e) Corrected with $LMS_{A,3\times3}$ Method

(f) Corrected with $LMS_{A,7\times7}$ Method

(g) Corrected with $LMS_{GA,7\times7}$ Method

(h) Corrected with LMS_{TV} Method

Figure 10.5.: Visual correction performance of the introduced methods for consecutive frame sequences of 500 frames length, using the parameter set of tab. 5.3.



(a) $|1-\{a_{i,j,\mathrm{rem.}}\}|$ for CS

(b) $|1-\{a_{i,j,\mathrm{rem.}}\}|$ for CS_G

(c) $|1-\{a_{i,j,\mathrm{rem.}}\}|$ for $LMS_{3\times3}$

(d) $|1-\{a_{i,j,\mathrm{rem.}}\}|$ for $LMS_{7\times7}$

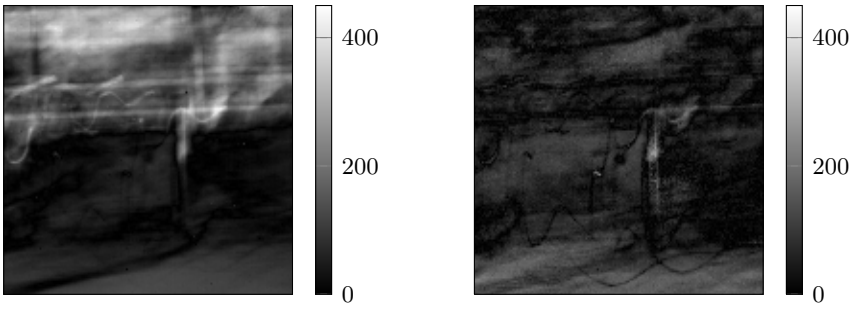
(e) $|1-\{a_{i,j,\mathrm{rem.}}\}|$ for $LMS_{A,3\times3}$

(f) $|1-\{a_{i,j,\mathrm{rem.}}\}|$ for $LMS_{A,7\times7}$

(g) $|1-\{a_{i,j,\mathrm{rem.}}\}|$ for $LMS_{GA,7\times7}$

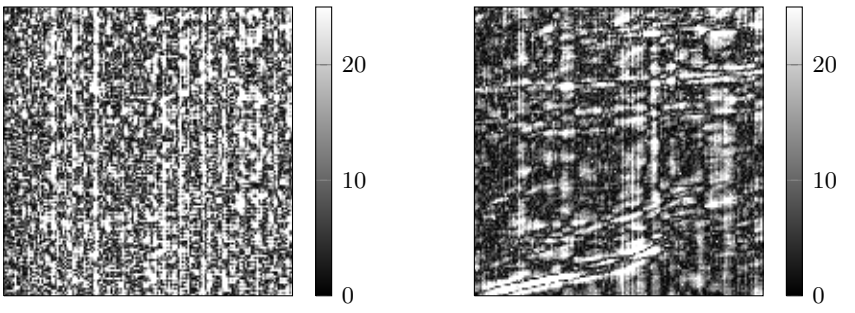
(h) $|1-\{a_{i,j,\mathrm{rem.}}\}|$ for LMS_{TV}

Figure 10.6.: Remaining PRNU patterns for the introduced methods for consecutive frame sequences of 500 frames length, using the parameter set of tab. 5.3.



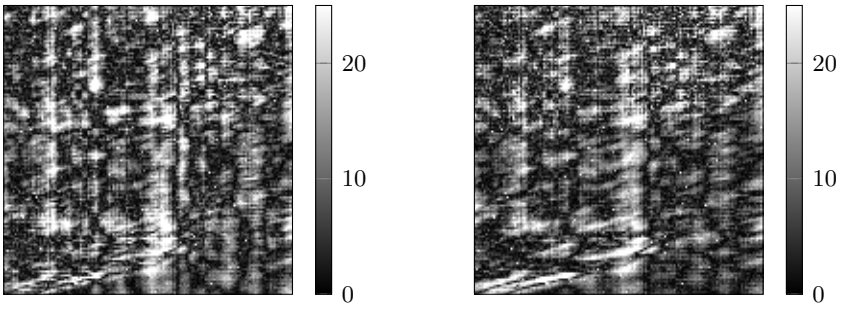
(a) $|\{b_{i,j,\mathrm{rem.}}\}|$ for CS

(b) $|\{b_{i,j,\mathrm{rem.}}\}|$ for CS_G



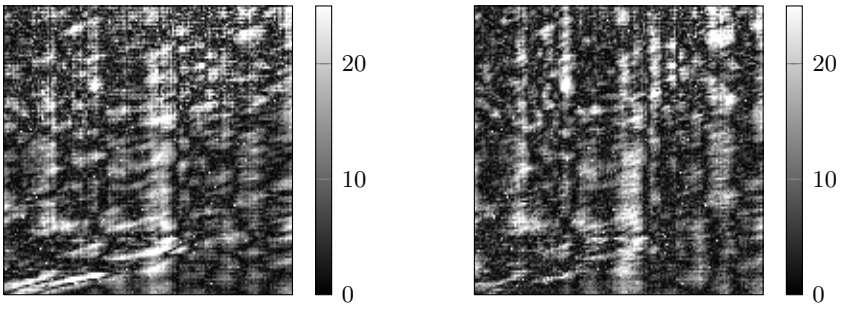
(c) $|\{b_{i,j,\mathrm{rem.}}\}|$ for $LMS_{3\times3}$

(d) $|\{b_{i,j,\mathrm{rem.}}\}|$ for $LMS_{7\times7}$



(e) $|\{b_{i,j,\mathrm{rem.}}\}|$ for $LMS_{A,3\times3}$

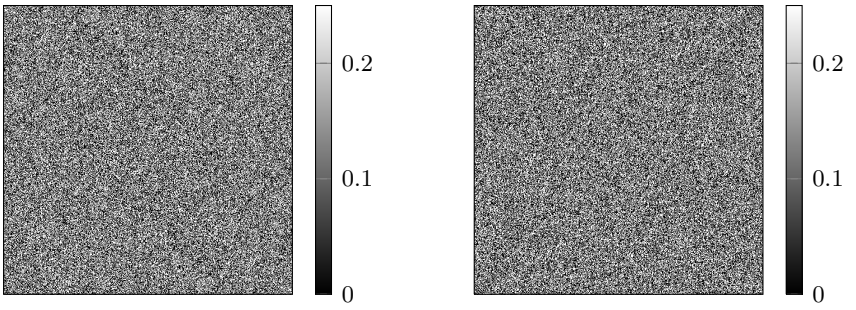
(f) $|\{b_{i,j,\mathrm{rem.}}\}|$ for $LMS_{A,7\times7}$



(g) $|\{b_{i,j,\mathrm{rem.}}\}|$ for $LMS_{GA,7\times7}$

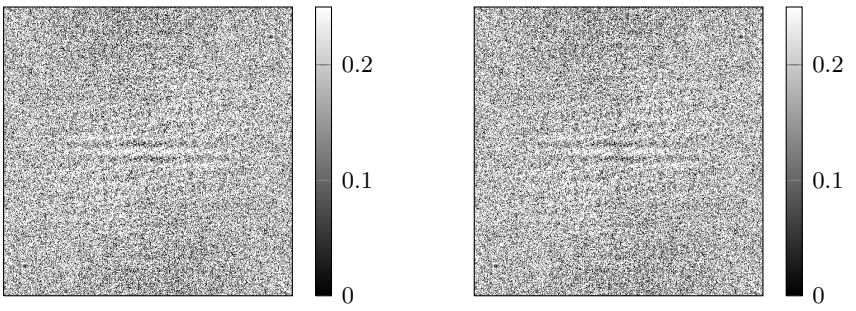
(h) $|\{b_{i,j,\mathrm{rem.}}\}|$ for LMS_{TV}

Figure 10.7.: Remaining DSNU patterns for the introduced methodsfor consecutive frame sequences of 500 frames length, using the parameter set of tab. 5.3.



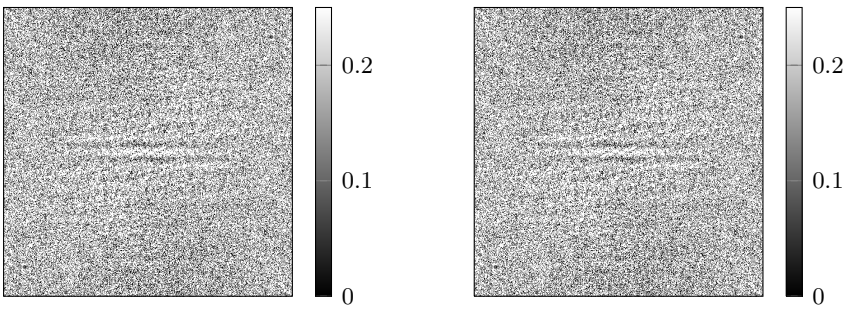
(a) $\hat{s}_{\text{spectr.,PRNU},\mu,\nu}$ for CS

(b) $\hat{s}_{\text{spectr.,PRNU},\mu,\nu}$ for CS_G



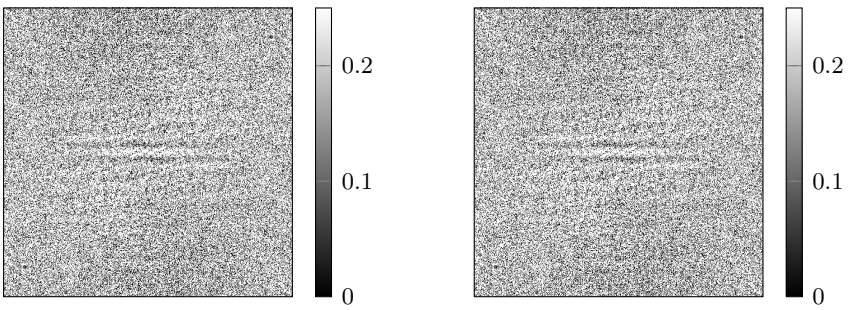
(c) $\hat{s}_{\text{spectr.,PRNU},\mu,\nu}$ for $LMS_{3\times3}$

(d) $\hat{s}_{\text{spectr.,PRNU},\mu,\nu}$ for $LMS_{7\times7}$



(e) $\hat{s}_{\text{spectr.,PRNU},\mu,\nu}$ for $LMS_{A,3\times3}$

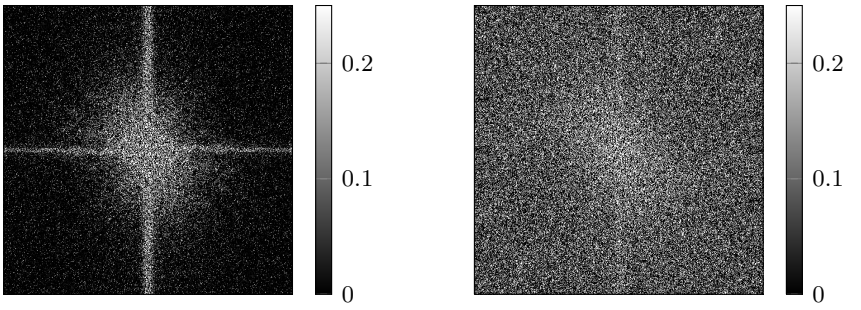
(f) $\hat{s}_{\text{spectr.,PRNU},\mu,\nu}$ for $LMS_{A,7\times7}$



(g) $\hat{s}_{\text{spectr.,PRNU},\mu,\nu}$ for $LMS_{GA,7\times7}$

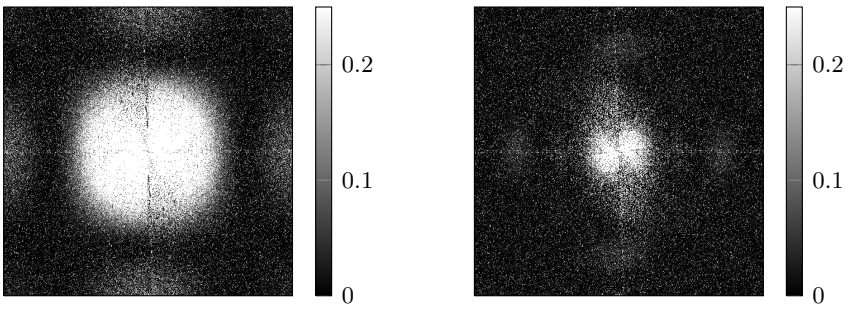
(h) $\hat{s}_{\text{spectr.,PRNU},\mu,\nu}$ for LMS_{TV}

Figure 10.8.: Analysis of the spectral damping factors for consecutive frame sequences of 500 frames length, using the parameter set of tab.5.3.



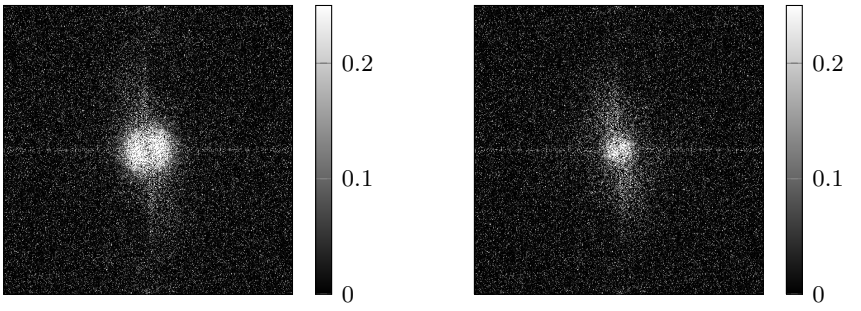
(a) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for CS

(b) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for CS_G



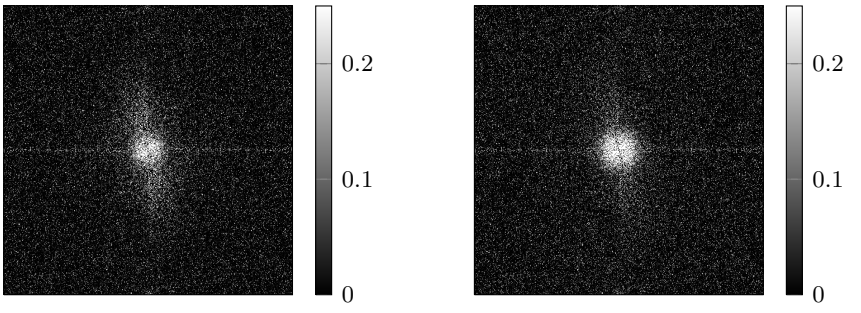
(c) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $LMS_{3\times3}$

(d) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $LMS_{7\times7}$



(e) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $LMS_{A,3\times3}$

(f) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $LMS_{A,7\times7}$



(g) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $LMS_{GA,7\times7}$

(h) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for LMS_{TV}

Figure 10.9.: Analysis of the spectral damping factors for consecutive frame sequences of 500 frames length, using the parameter set of tab. 5.3.

ized spectral damping visible, and the differences in the methods show up mainly in the DSNU analysis. Here the CS method provides an actually good damping in the spatial high frequencies while the improved CS_G method does reach such a large spectral damping. However, given the corrected images both methods cannot keep up with the LMS methods. The pure LMS methods, especially the $LMS_{3\times3}$ method, cannot correct the spatial high frequency components, while the more advanced methods like the $LMS_{A,3\times3}$ and $LMS_{A,7\times7}$ correct these high frequency components much better. Especially the $LMS_{A,7\times7}$ and LMS_{TV} methods provide the best correction performances for these spectral high frequencies. The $LMS_{GA,7\times7}$ method does not show a difference to its $LMS_{A,7\times7}$ equivalent, as expected.

Given these evaluations, the $LMS_{A,7\times7}$ and the LMS_{TV} methods are still the best reference methods.

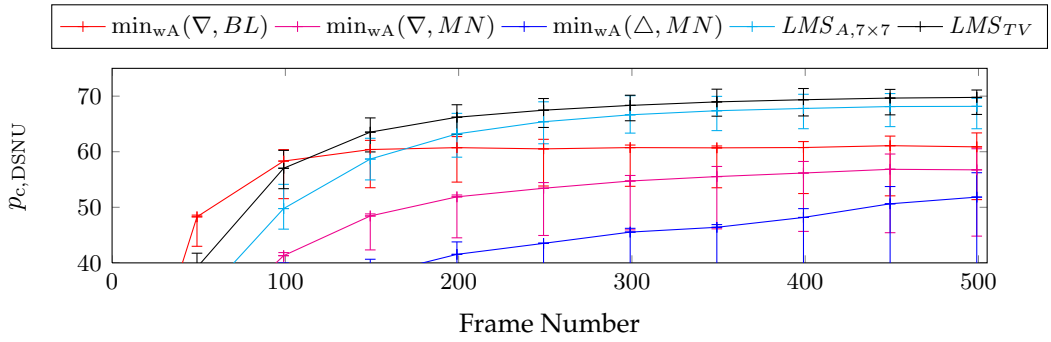
10.1.2. Analysis for the Weighted Average Improvement

With consideration to the new maximum likelihood methods, it has already been discussed in sec. 6.4.1 that the basic method does not keep up to the reference methods. The results of this intermediate step in the development are therefore skipped. The improved methods that consider the weighted average of the measurements as input have already been discussed in sec. 7.1.3 and showed performance decreases for the consecutive frame sequences. However, in the best-case performance analysis the methods still keep up with the reference methods. For the parameters that are extracted from the statistics, the non-weighted statistics from tab. 6.1 showed the best results in the case of consecutive frame sequences, while the parameters that extract from the weighted statistics lead to significantly worse performances. This result has been discussed as a logical consequence of the strong data dependencies in case of consecutive frame sequences. The weighted averaging will lead in these scenarios to an unwanted bias of the weighted statistics.

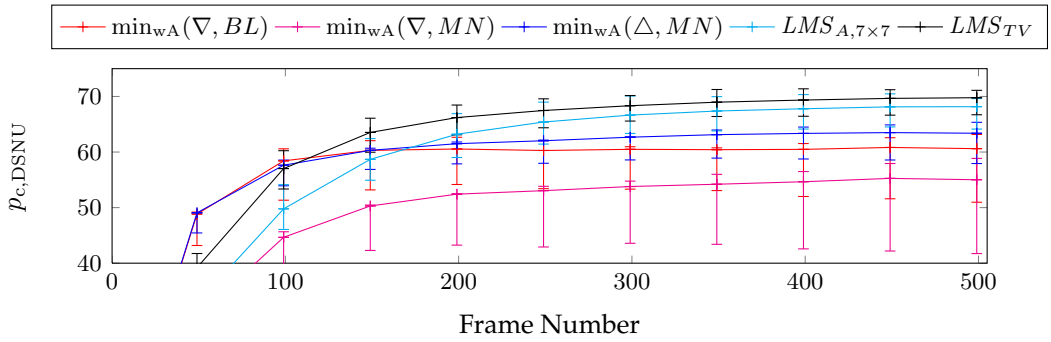
To analyze the performance of the improved methods, the optimized best-case parameter set and the statistic parameter set are used in a temporal evaluation of the methods (see fig. 10.10). The figure shows that the new methods can not keep up to the reference methods when their parameters are fixed, however this is possible in the best-case analysis (see fig. 7.3, p. 142). The figure also shows that the $\min_{\mathrm{wA}}(\nabla, BL)$ method outperforms the reference methods on the first few frames, but after 500 frames all the new methods do not even reach the reference methods with their upper error bars (realized as the 70% percentil). Surprisingly for the statistic parameter set, the $\min_{\mathrm{wA}}(\triangle, MN)$ method performs better than for the best-case parameter set (see fig. 10.10a). The $\min_{\mathrm{wA}}(\nabla, BL)$ and $\min_{\mathrm{wA}}(\nabla, MN)$ methods however perform slightly worse, which is the expected behavior (see fig. 10.10b). The reason for the untypical behavior of the $\min_{\mathrm{wA}}(\triangle, MN)$ method is a parameter instability of this method for the very optimized parameters. If in the

underlying optimization the best possible parameters per sequence are not stable in their performances then another non-optimal parameter can result in better performances, even if it is fixed. In case of the $\min_{\mathrm{wA}}(\triangle, MN)$ method, this more stable parameter is the one extracted from the statistics.

This result on the one hand shows that the statistic parameters and the $\min_{\mathrm{wA}}(\triangle, MN)$ method are both more stable with respect to a violation of the independent feature constraint. On the other hand, none of the new methods can reach an overlap of the error bars with the reference methods.



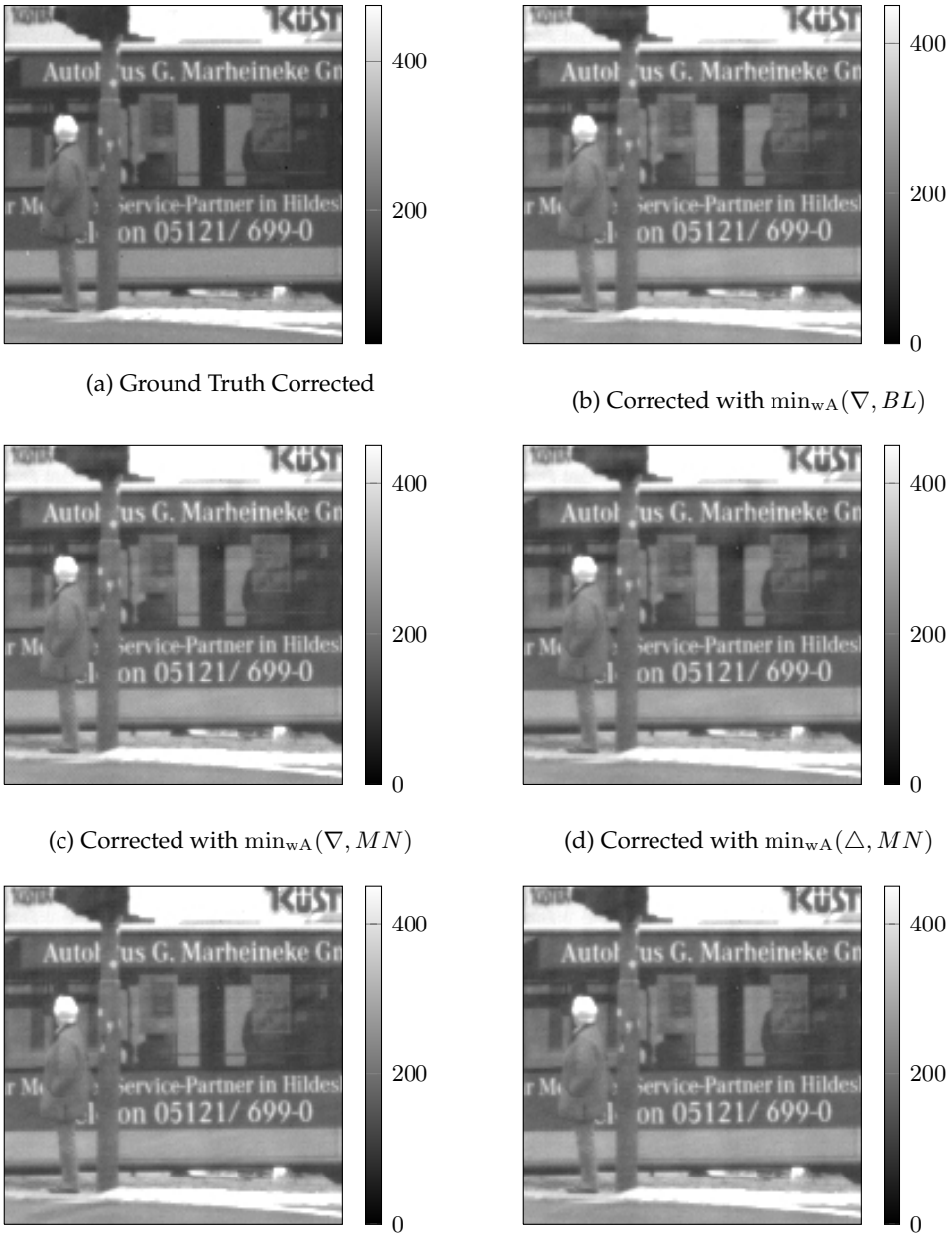
(a) Opt. Consec (tab. 7.4)



(b) Stat. Params (tab. 6.1)

Figure 10.10.: Evaluations for consecutive frame sequences with experimentally optimized parameters (fig. 7.4a) and the parameter set from the statistic histogram (fig. 7.4b) in the $p_{\mathrm{c,DSNU}}$ metric.

Fig. 10.11 shows the visual impression of the corrected images by the different methods in comparison to the ground truth corrected version. A visual difference between the new methods is not detectable, but all of them show the same artifact in the window of the bus, which presents as slight horizontal stripes. The reference methods, show this artifact as well but with much less amplitude. While this artifact is visible in all methods, the already discussed artifact below the feet of the woman in front of the bus is not visible in the new developed methods. As the analyzed sequence is only one of many possible choices, the only generalized conclusion that can be made is that the reached correction rates of 70% or less may not be sufficient to suppress artifacts in the current setup.



(a) Ground Truth Corrected

(b) Corrected with $\min_{\mathrm{wA}}(\nabla, BL)$

(c) Corrected with $\min_{\mathrm{wA}}(\nabla, MN)$

(d) Corrected with $\min_{\mathrm{wA}}(\triangle, MN)$

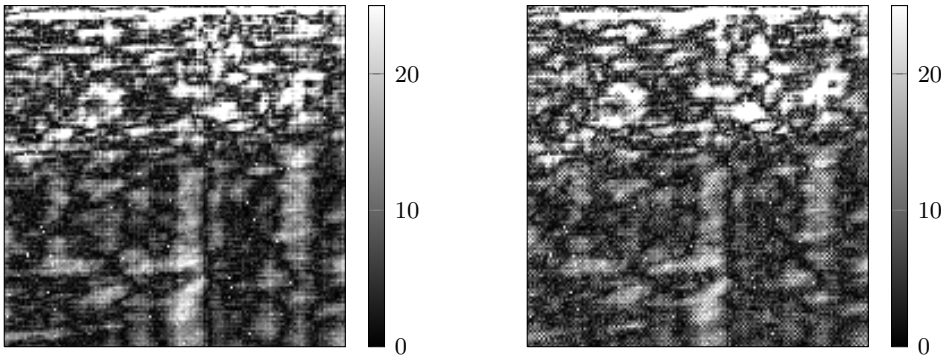
(e) Corrected with $LMS_{A,7\times7}$ Method

(f) Corrected with LMS_{TV} Method

Figure 10.11.: Corrected images with the ground truth correction patterns and the estimate of the different methods for consecutive frame sequences (Parameters according to tab. 7.3).

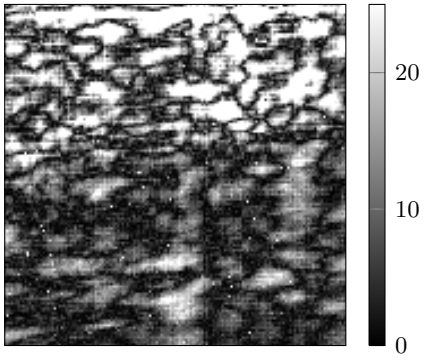
Fig. 10.12 and fig. 10.13 show the remaining DSNU patterns and the according spectral damping factors. The remaining DSNU shows in general a low amplitude and spatial low frequency remainders. The origin of the stripe-like artifact is also visible as an area of a higher amplitude in the upper parts of the sensor. The remaining patterns reaches here the scale limit of $20\,\mathrm{DN}$ in these areas. This observation is also true for the $LMS_{A,7\times7}$ method and the artifact in the lower left corner of the images, while for the LMS_{TV} method the artifact is not that high in amplitude, but still visible. The spectral damping shows the expected behavior of a good damping of the high frequency components, however the $\min_{\mathrm{wA}}(\nabla, MN)$ method has the known problem with the corner frequency components. This phenomenon was already observed for the RCP-type extension in sec. 7.3.3 and arises due to the design of the filter mask of this method (compare fig. 6.2b, p. 123). The spectral damping anaylsis summarizes by the observations that the new methods provide in general a better damping in the lower frequencies, but the reference methods are able to damp the important higher frequency components better.

Concluding, as the disadvantages and advantages of the methods have been discussed, the best reference methods seem to have a slight advance over the improved maximum likelihood methods. This advantage arises from the fact that these methods perform their updates on a frame-to-frame basis, that does not consider the information of older frames. The new methods however include the knowledge of the old frames which leads to a biased consideration of the repeated information. On the other hand, for longer sequences or random frame input, this consideration of all available information leads to the advantage of the new methods in the discussed scenarios.

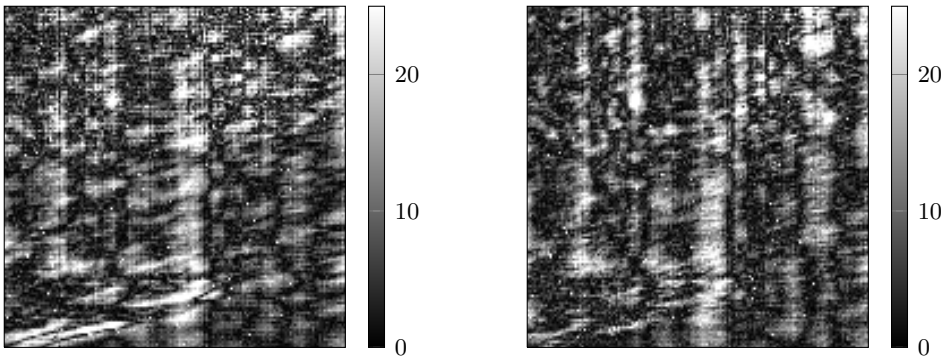


(a) $|\{b_{i,j,rem.}\}|$ for $\min_{\mathrm{wA}}(\nabla, BL)$

(b) $|\{b_{i,j,rem.}\}|$ for $\min_{\mathrm{wA}}(\nabla, MN)$



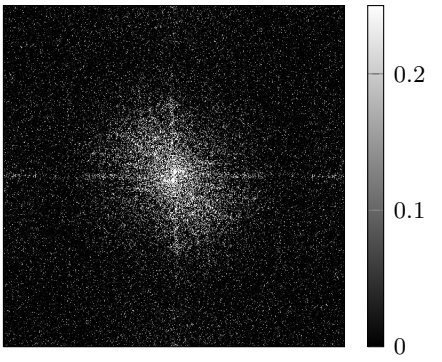
(c) $|\{b_{i,j,rem.}\}|$ for $\min_{\mathrm{wA}}(\triangle, MN)$



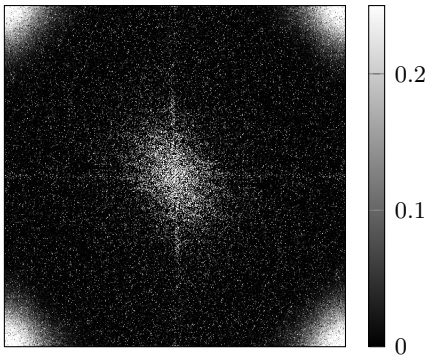
(d) $|\{b_{i,j,\mathrm{rem.}}\}|$ for $LMS_{A,7\times7}$

(e) $|\{b_{i,j,\mathrm{rem.}}\}|$ for LMS_{TV}

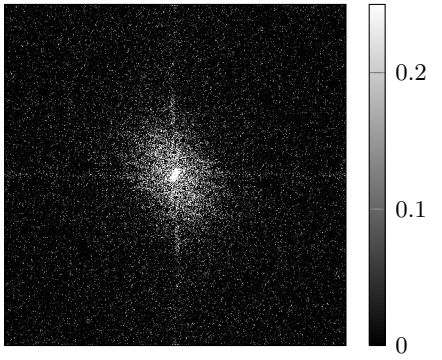
Figure 10.12.: Remaining DSNU patterns for the introduced methods and consecutive frame sequences (Parameters according to tab. 7.3).



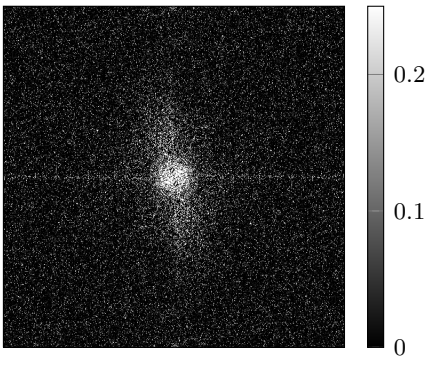
(a) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $\min_{\text{wA}}(\nabla, BL)$



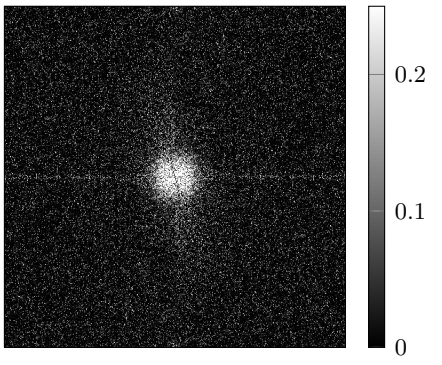
(b) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $\min_{\text{wA}}(\nabla, MN)$



(c) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $\min_{\text{wA}}(\triangle, MN)$



(d) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $LMS_{A,7\times7}$



(e) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for LMS_{TV}

Figure 10.13.: Spectral damping factors for the introduced methods and consecutive frame sequences (Parameters according to tab. 7.3).

10.1.3. Analysis of the CNN-type Methods

For the CNN-type maximum likelihood methods the best-case performances for consecutive frame sequences have already been discussed in sec. 7.2.3 p. 154 and showed again a performance decrease in comparison to the random frame sequences. However, this best-case analysis showed in fig. 7.8 that the CNN-type methods overlap significantly in their error bars with the reference methods, even for consecutive frame input.

Given the optimized parameter set for the CNN methods and consecutive frames, fig. 10.14 shows that the reference methods outperform the CNN methods. Thus the nonlinear function $f(\cdot)$ of the CNN does not help to improve the performance as much as observed in the case of random frame input. Anyway, the $\mathrm{CNN}(\triangle, MN)$ method still gains a performance increase that leads to a significant overlap of the error bars of of the $\mathrm{CNN}(\triangle, MN)$ method and the reference methods. As discussed before, the CNN methods rely on the energy minimization methods and for the random frame input, the CNN methods did outperform their energy minimizer equivalents. Fig. 10.15 shows that for the consecutive frame sequences this is only the case for the $\mathrm{CNN}(\triangle, MN)$ method. This proves again that the choice of $\hat{L} = \triangle$ is more stable against the violation of the demanded independent frame input than the ∇ choice.

Given the visual impression of the corrected images, fig. 10.16 shows that the CNN methods again corrects successfully the artifact below the feet of the woman. The artifacts in the window of the bus are still visible, but seem a bit more damped when compared to fig. 10.11. This observation can be confirmed by fig. 10.17 which shows the remaining DSNU patterns. In general less spatial high frequency components are visible but the strong patterns in the upper part of the sensor are still there (see fig. 10.12 to compare). However, $\mathrm{CNN}(\triangle, MN)$ shows less amplitude in its remains and $\mathrm{CNN}(\nabla, MN)$ does not show the remaining corner frequency components. These improvements are also visible for the spectral damping factors as depicted in fig. 10.18 where much more of the higher frequency com-

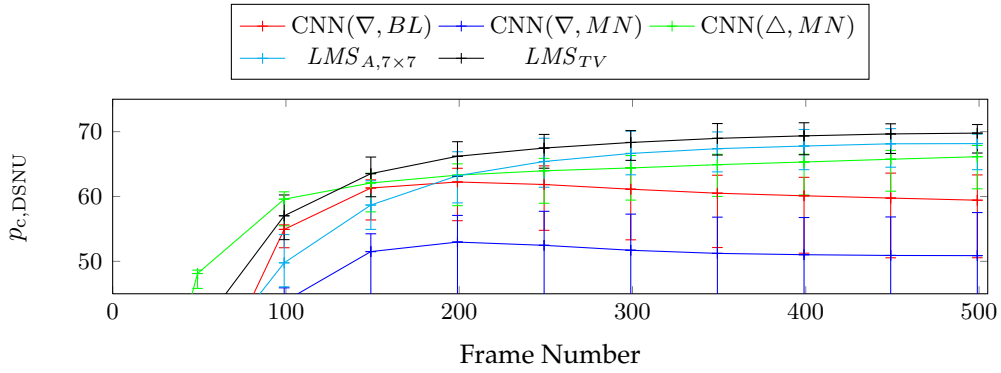


Figure 10.14.: Methods evaluations with experimental optimized parameters (tab. 7.6) for the $p_{\mathrm{c,DSNU}}$ metric and consecutive frame sequences as input

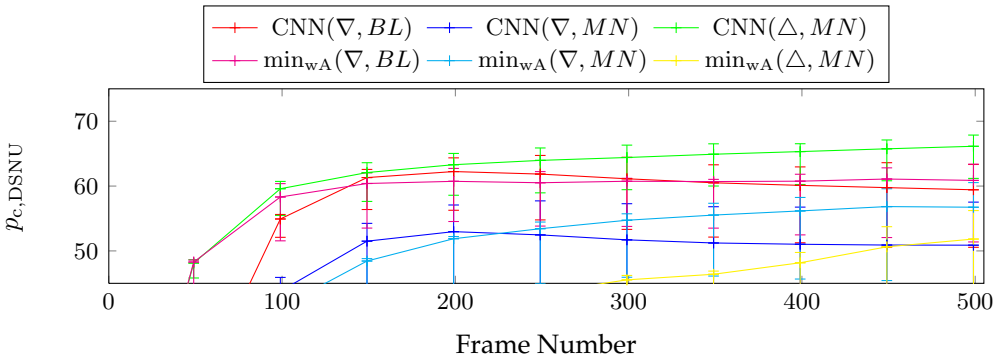
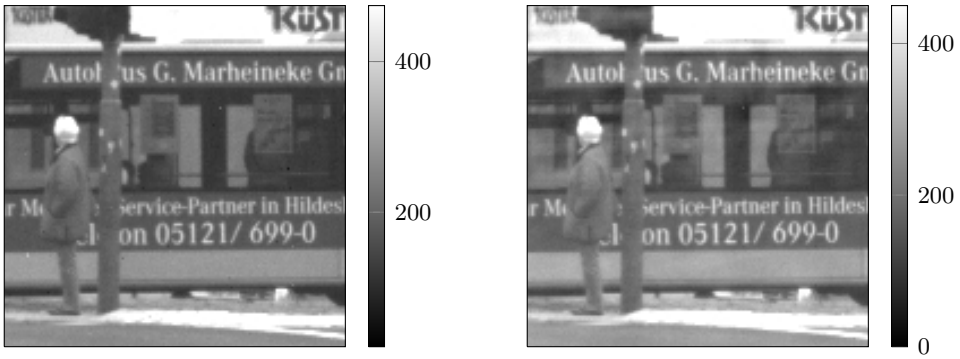


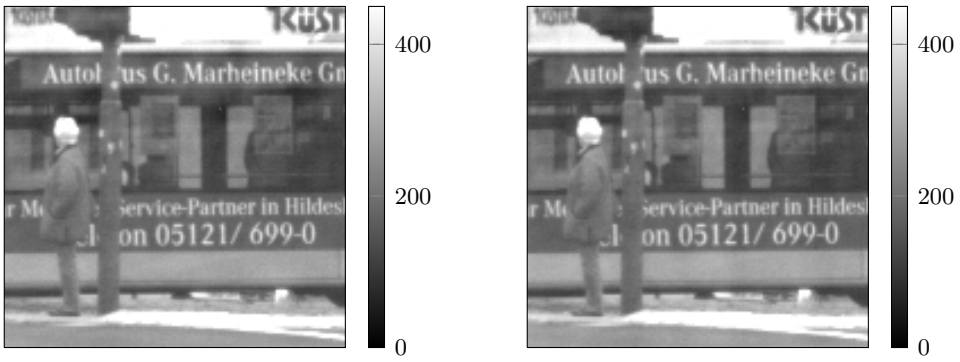
Figure 10.15.: Comparison between the CNN variants and the improved maximum likelihood methods in the $p_{\mathrm{c,DSNU}}$ metric for consecutive frame sequences. The optimized parameters for the CNN method (Opt. CNN Consec (tab. 7.7)) and the improved energy minimizers (Weighted Stat. (tab. 7.1)) have been used.

ponents are damped in comparison to fig. 10.13. However, the reference methods are still able to damp slightly more of the spatial high frequency components. Concluding, the CNN methods correct the DSNU to some extend and especially the $\mathrm{CNN}(\triangle, MN)$ method showed a performance that is close to one of the reference methods.



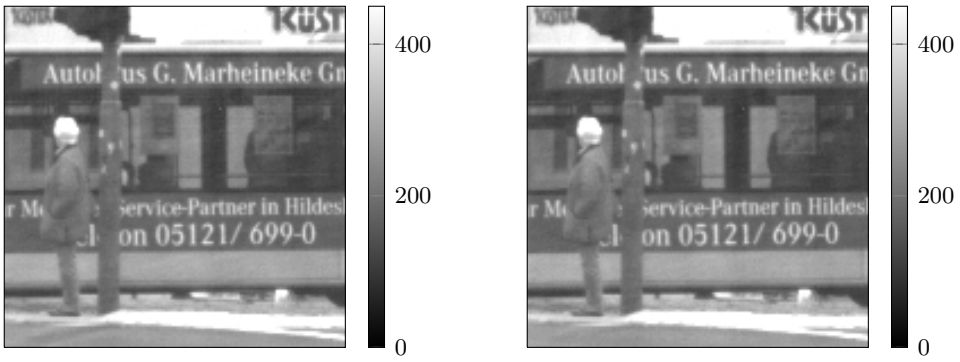
(a) Ground Truth Corrected

(b) Corrected with CNN(∇, BL)



(c) Corrected with CNN(∇, MN)

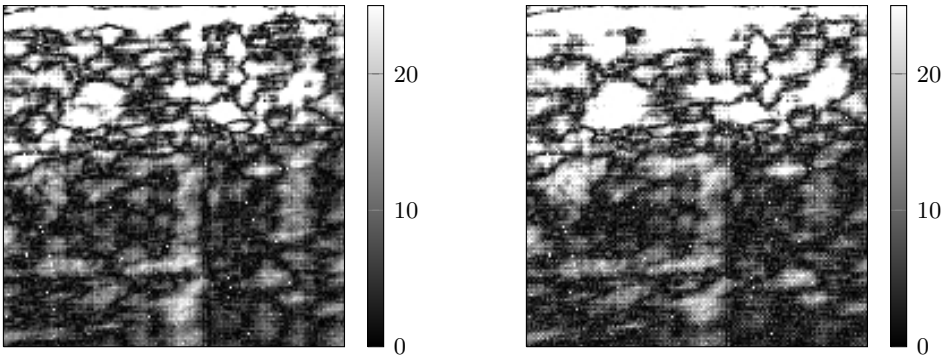
(d) Corrected with CNN($\triangle, MN$)



(e) Corrected with $LMS_{A,7\times7}$ Method

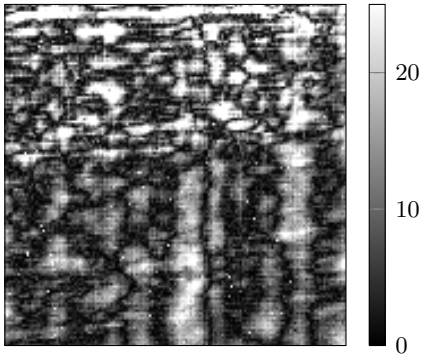
(f) Corrected with LMS_{TV} Method

Figure 10.16.: Corrected images of the introduced methods for consecutive frame sequences (Parameters according to tab. 7.6)

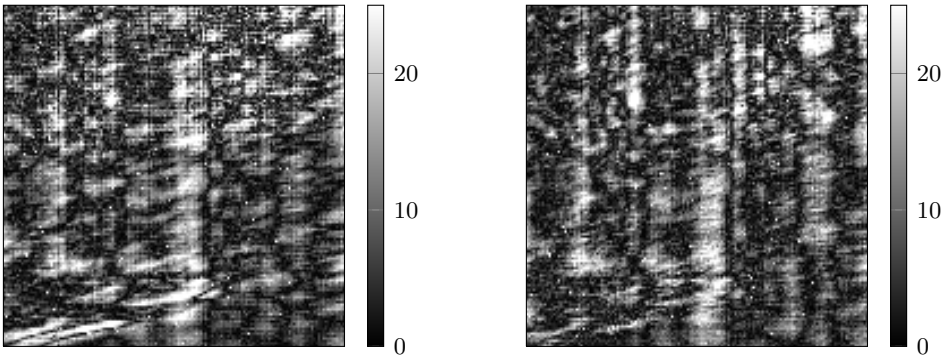


(a) $|\{b_{i,j,rem.}\}|$ for $\mathrm{CNN}(\nabla, BL)$

(b) $|\{b_{i,j,rem.}\}|$ for $\mathrm{CNN}(\nabla, MN)$



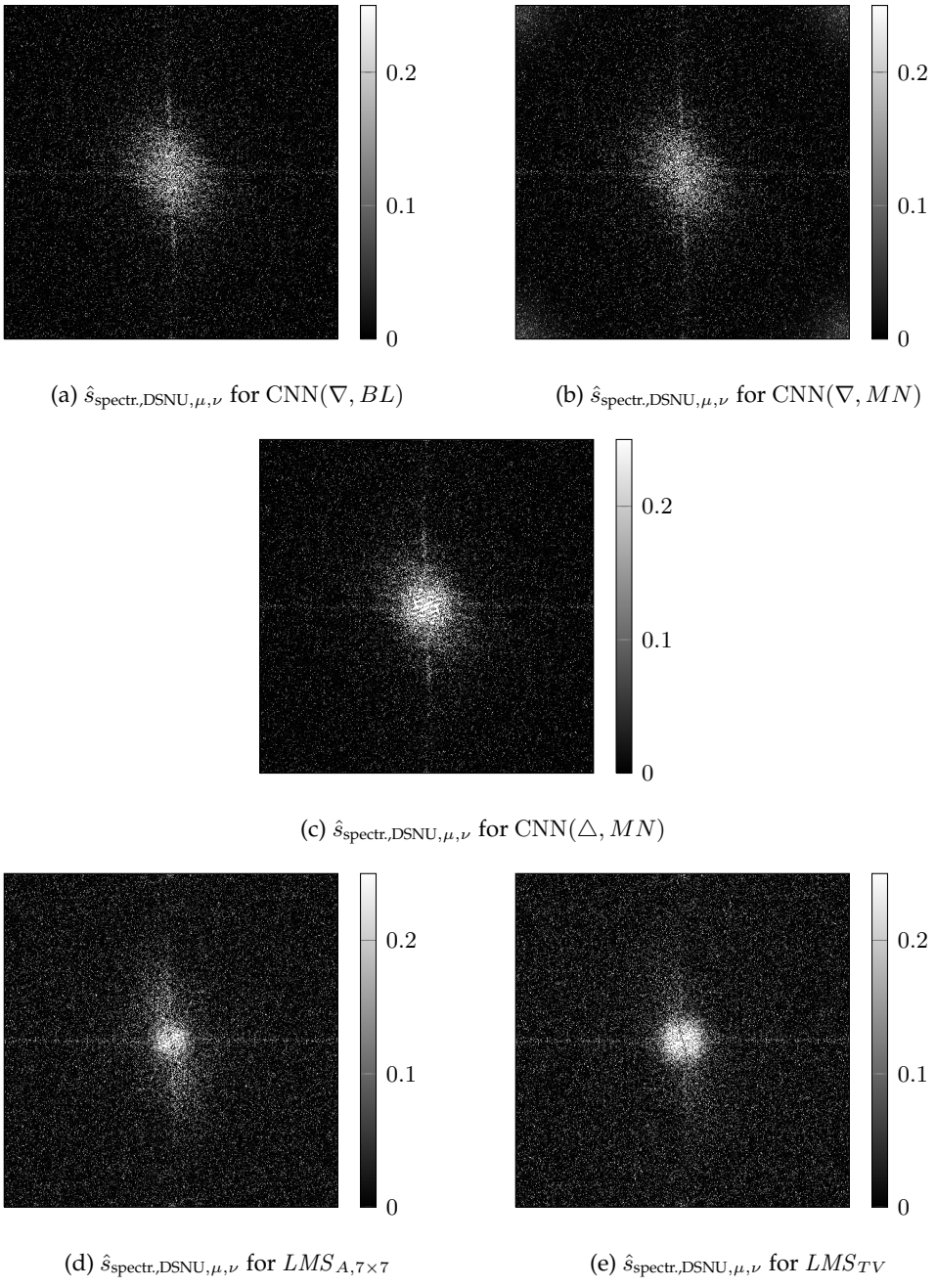
(c) $|\{b_{i,j,rem.}\}|$ for $\mathrm{CNN}(\triangle, MN)$



(d) $|\{b_{i,j,\mathrm{rem.}}\}|$ for $LMS_{A,7\times7}$

(e) $|\{b_{i,j,\mathrm{rem.}}\}|$ for LMS_{TV}

Figure 10.17.: Remaining DSNU patterns for the introduced methods and consecutive frame sequences (Parameters according to tab. 7.6)



(a) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $\mathrm{CNN}(\nabla, BL)$

(b) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $\mathrm{CNN}(\nabla, MN)$

(c) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $\mathrm{CNN}(\triangle, MN)$

(d) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $LMS_{A,7\times7}$

(e) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for LMS_{TV}

Figure 10.18.: Spectral damping factors for the introduced methods and consecutive frame sequences (Parameters according to tab. 7.6)

10.1.4. Analysis of the Row-, Column- and Pixel-Wise Extensions

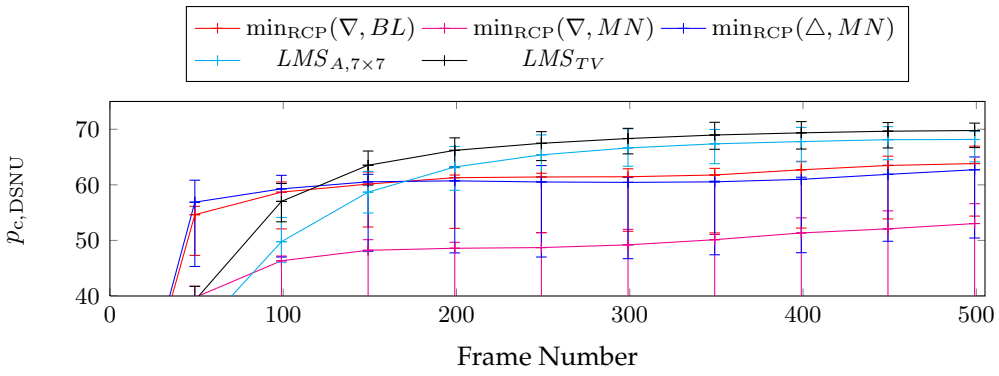
Given the row-, column- and pixel-wise (RCP) extension of the maximum likelihood approach their best-case performance for consecutive frame input has already been discussed in sec. 7.3.3. The general observed performance decrease for consecutive frame sequences, the RCP-methods were able to keep up to the reference methods in the best-case evaluations. This is common to the other analyzed methods in this chapter. However, the weighted statistic parameter set showed again as not suitable for consecutive frame sequences and the non-weighted statistic parameter set has to be used. The reasons for this have already been discussed in sec. 10.1.2 and 7.1.3.

The evaluations for the fixed parameter sets and consecutive frame input sequences give results that do not reach the performances of the reference methods and show a similar performance as the improved maximum likelihood methods in sec. 10.1.2 above (see fig. 10.19a for the optimized parameter results and fig. 10.19b for the statistic parameter results of the RCP methods). While the performances of the RCP methods are similar in their averages, the statistic parameters produce much smaller error bars. This confirms the statement made above in sec. 10.1.2 that the theoretical parameters provide more stable results than the average of the *per sequence* optimized parameters.

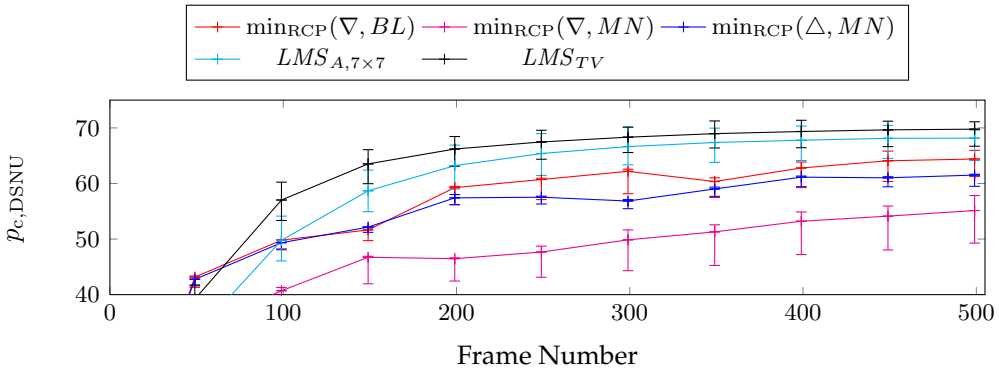
The RCP methods provide the corrected images shown in fig. 10.20, that show a further improved correction of the artifact in the window of the bus, while the artifact below the feed of the woman is not visible either. However for the $\min_{\mathrm{RCP}}(\nabla, BL)$ and the $\min_{\mathrm{RCP}}(\nabla, MN)$ methods a new vertical low frequency artifact shows up around the position where the number 699 is written on the bus. This artifact is not visible for the $\min_{\mathrm{RCP}}(\triangle, MN)$ method, which therefore expresses several high amplitude low-frequency artifacts over the whole sensor area. For the $\min_{\mathrm{RCP}}(\nabla, MN)$ method one can further observe a high amplitude spatial high frequency pattern in the corrected image.

All these effects are of course more clearly visible in the remaining DSNU patterns in fig. 10.21, where especially the high amplitude vertical artifact in the $\min_{\mathrm{RCP}}(\nabla, BL)$ and the $\min_{\mathrm{RCP}}(\nabla, MN)$ methods and the spatial high frequency overlay of the $\min_{\mathrm{RCP}}(\nabla, MN)$ are the most remarkable properties. For the $\min_{\mathrm{RCP}}(\triangle, MN)$ method only unconnected patches with spatial low frequency components are visible, which shows that the additional row- and column-wise correction have the most effects for this method. The mentioned effects are visible as well in the spectral damping analysis in fig. 10.22. Here the white corners for the $\min_{\mathrm{RCP}}(\nabla, MN)$ method's high frequency components show up clearly, which are the result of the high pass filter design as already discussed in sec. 10.1.2 and 7.3.3. For the $\min_{\mathrm{RCP}}(\nabla, MN)$ method a good correction of the row- and column-wise components is visible as the dark incisions in the white area of not well damped low frequency components. In general the remaining spatial frequency components cover a much wider area as those of the reference methods. Concluding, for consecutive frame input, only the $\min_{\mathrm{RCP}}(\triangle, MN)$ method can be

recommended, and only if the higher amplitude in the the remaining spatial low frequency pattern is of no concern. Again the choice of $\hat{L}{=}\triangle$ gives the best results and is therefore the most reliable method for consecutive frame sequences.



(a) Parameter set: Opt. RCP Consec (tab. 7.10)



(b) Parameter set: Stat. Params (tab. 6.1)

Figure 10.19.: Evaluations for consecutive frame sequences with experimentally optimized parameters (fig. 7.17a) and the parameter set from the non-weighted histogram (fig. 7.17b) in the $p_{\mathrm{c,DSNU}}$ metric.

(a) Ground Truth Corrected

(b) Corrected with $\min_{\mathrm{RCP}}(\nabla, BL)$

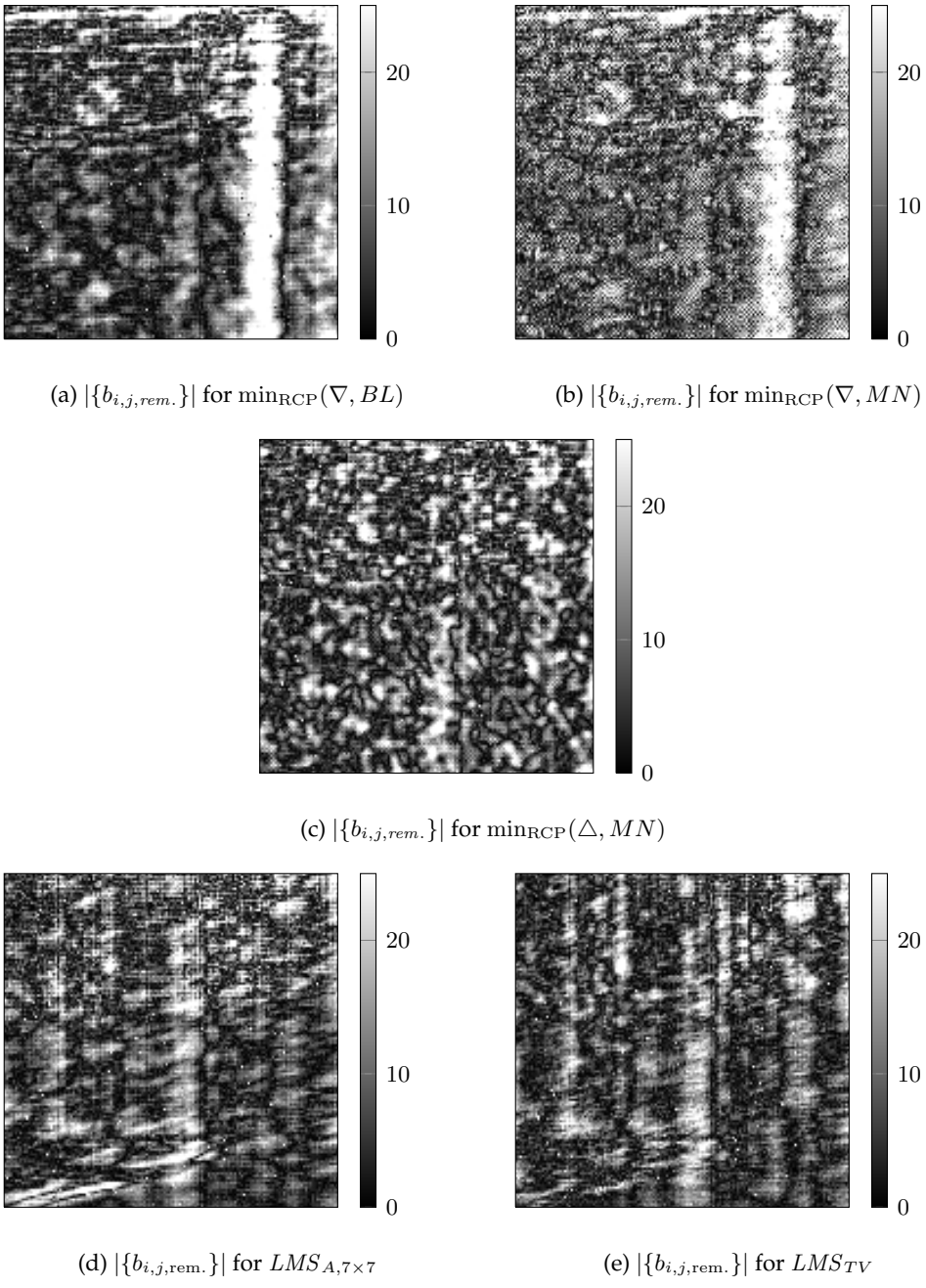
(c) Corrected with $\min_{\mathrm{RCP}}(\nabla, MN)$

(d) Corrected with $\min_{\mathrm{RCP}}(\triangle, MN)$

(e) Corrected with $LMS_{A,7\times7}$ Method

(f) Corrected with LMS_{TV} Method

Figure 10.20.: Corrected images of the introduced methods for consecutive frame sequences (Parameters according to tab. 7.7)



(a) $|\{b_{i,j,rem.}\}|$ for $\min_{\mathrm{RCP}}(\nabla, BL)$

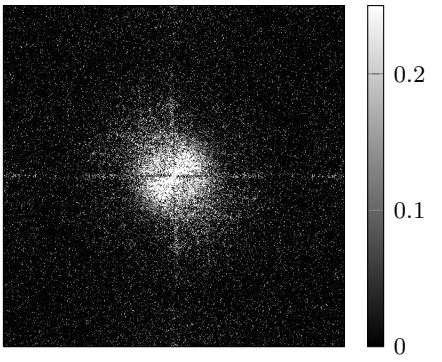
(b) $|\{b_{i,j,rem.}\}|$ for $\min_{\mathrm{RCP}}(\nabla, MN)$

(c) $|\{b_{i,j,rem.}\}|$ for $\min_{\mathrm{RCP}}(\triangle, MN)$

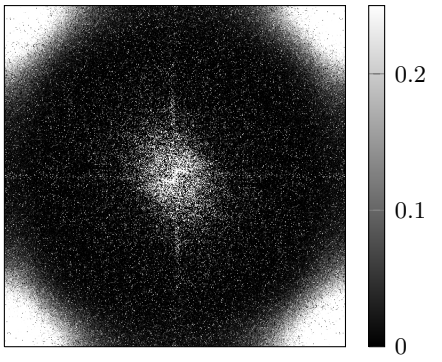
(d) $|\{b_{i,j,\mathrm{rem.}}\}|$ for $LMS_{A,7\times7}$

(e) $|\{b_{i,j,\mathrm{rem.}}\}|$ for LMS_{TV}

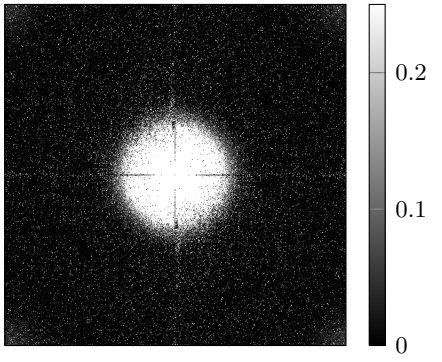
Figure 10.21.: Remaining DSNU patterns for the introduced methods and consecutive frame sequences (Parameters according to tab. 7.7)



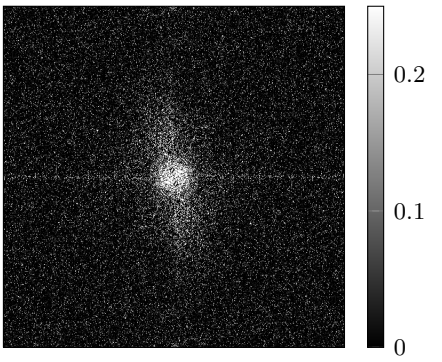
(a) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $\min_{\text{RCP}}(\nabla, BL)$



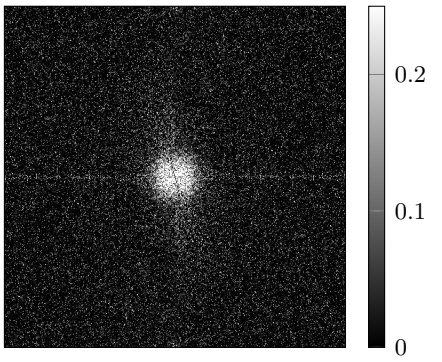
(b) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $\min_{\text{RCP}}(\nabla, MN)$



(c) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $\min_{\text{RCP}}(\triangle, MN)$



(d) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $LMS_{A,7\times7}$



(e) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for LMS_{TV}

Figure 10.22.: Spectral damping factors for the introduced methods and consecutive frame sequences (Parameters according to tab. 7.7)

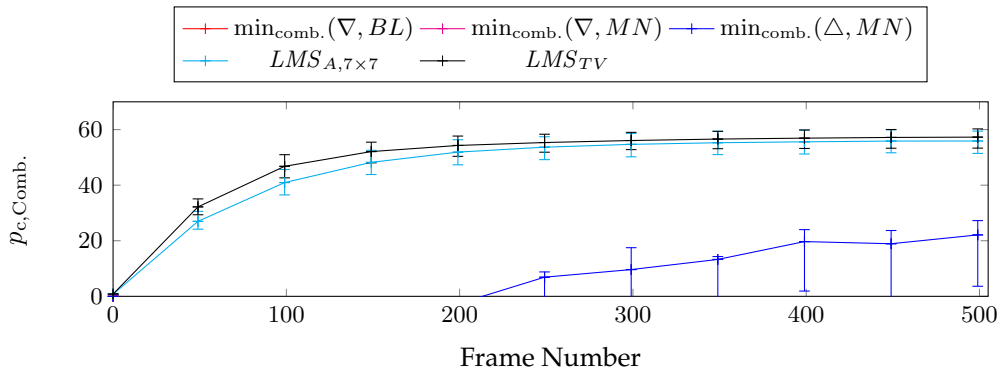
10.1.5. Analysis of the Combined PRNU and DSNU Method

The last method to discuss is the combined PRNU and DSNU estimation method as introduced in chapter 8. The discussion of the best-case performance has been done in sec. 8.4.2 and it has been shown in fig. 8.4 that the new methods do not perform well for consecutive frame input because of a failure in the PRNU part of the method. Only the $\min_{\text{comb.}}(\triangle, MN)$-method results in a significant positive correction rate in the $p_{\text{c,Comb.}}$ measure. The temporal analysis with the optimized parameters for consecutive frame input is now shown in fig. 10.23 and confirms the results of the best-case analysis. Only $\min_{\text{comb.}}(\triangle, MN)$ achieves a positive correction rate for the $p_{\text{c,Comb.}}$ (The same is observed for the $p_{\text{c,DSNU}}$-norm). With respect to $p_{\text{c,PRNU}}$, the $\min_{\text{comb.}}(\nabla, BL)$ method achieves a similar performance as the $\min_{\text{comb.}}(\triangle, MN)$ method of $\approx -100\%$, while the $\min_{\text{comb.}}(\nabla, MN)$ method does not show on the scale due to even worse correction rates.

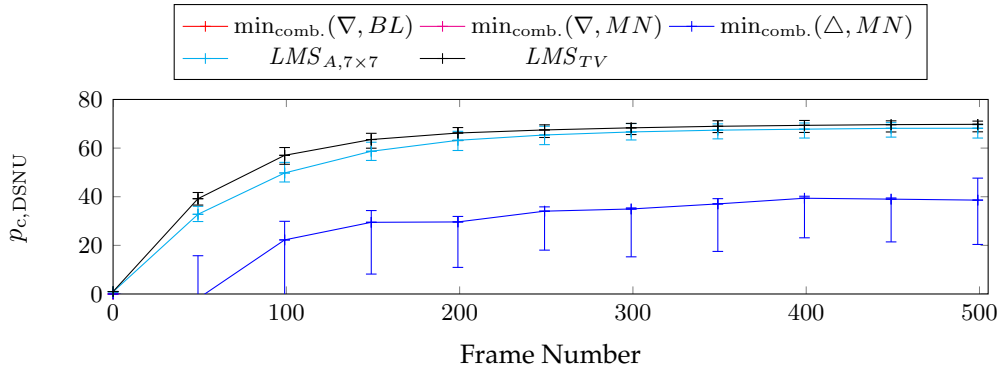
The DSNU and PRNU performances can of course not be regarded as independent any more and errors in the PRNU estimation are covered by adapted DSNU parameters in the final visual impressions. This can be observed in fig. 10.24 where a visually pleasing correction of the combined method shows up and only the typical artifact in the window of the bus can be seen. However, given the remaining DSNU pattern in fig. 10.25, the $\min_{\text{comb.}}(\nabla, BL)$ and $\min_{\text{comb.}}(\nabla, MN)$ methods show a high amplitude in their remains. Only the $\min_{\text{comb.}}(\triangle, MN)$ method shows a remaining pattern that is comparable to the reference methods, but even this still shows a significantly higher amplitude, which explains why this method did not reach the reference methods' performances. The analysis of the spectral damping factors in fig. 10.26 then reveals that all of the new combined methods were able to correct the spatial high frequencies, which is the reason why a visually pleasant corrected image is achieved. The remaining frequency components are however not limited on spatial low frequencies and further show a strong vertical and noticeable horizontal remainder.

The remaining PRNU patterns do not give any insight of the methods functionality and mainly reveal the failure to correct successfully for the PRNU (see fig. 10.27), the same applies for the spectral damping analysis which shows a slight advantage over the reverence methods by an increased damping of the spatial low frequencies (see fig. 10.28).

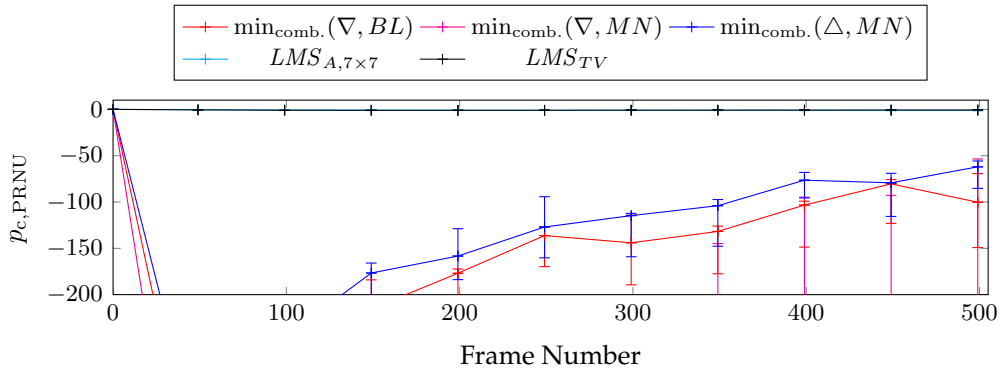
Summarized, the combined estimation method struggles more than the DSNU-only methods to correct for consecutive frame input. However, it still leads to a visually pleasant correction, as discussed.



(a) $p_{\text{c,Comb.}}$



(b) $p_{\text{c,DSNU}}$



(c) $p_{\text{c,PRNU}}$

Figure 10.23.: Methods evaluations with the optimized parameters for the combined PRNU and DSNU estimation method in the quality norms $p_{\text{c,Comb.}}$, $p_{\text{c,DSNU}}$ and $p_{\text{c,PRNU}}$ using the optimized parameters from tab. 8.3.

(a) Ground Truth Corrected

(b) $\min_{\text{comb.}}(\nabla, BL)$ corrected Image

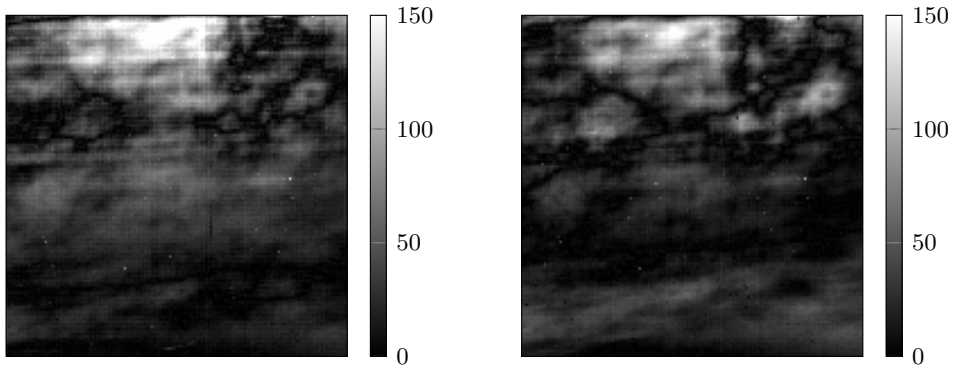
(c) $\min_{\text{comb.}}(\nabla, MN)$ corrected Image

(d) $\min_{\text{comb.}}(\triangle, MN)$ corrected Image

(e) Corrected with $LMS_{A,7\times7}$ Method

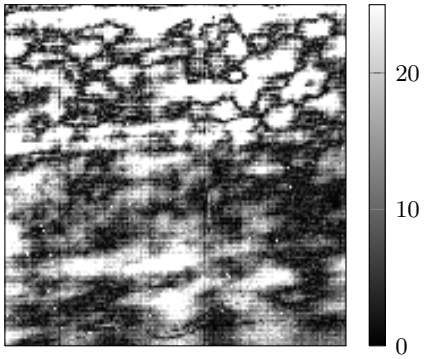
(f) Corrected with LMS_{TV} Method

Figure 10.24.: The visual impression of the corrected images with the different methods for a combined PRNU and DSNU correction in random frame sequences.

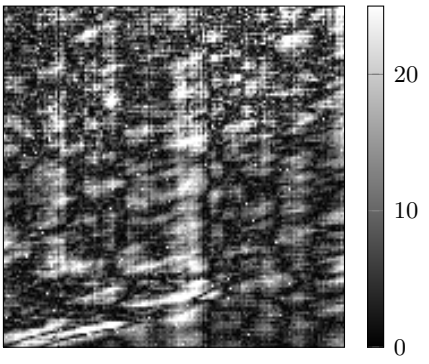


(a) $\min_{\text{comb.}}(\nabla, BL)$ remaining DSNU

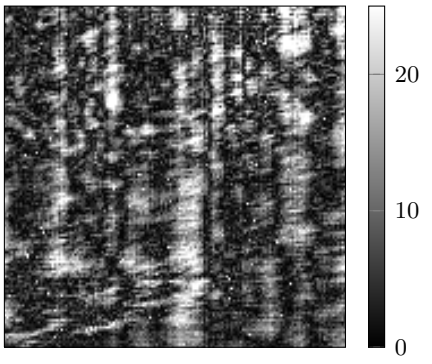
(b) $\min_{\text{comb.}}(\nabla, MN)$ remaining DSNU



(c) $\min_{\text{comb.}}(\triangle, MN)$ remaining DSNU

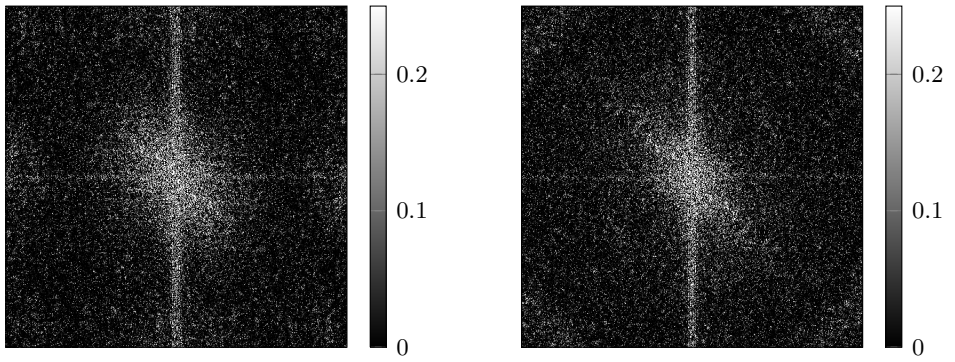


(d) $|\{b_{i,j,\text{rem.}}\}|$ for $LMS_{A,7\times7}$



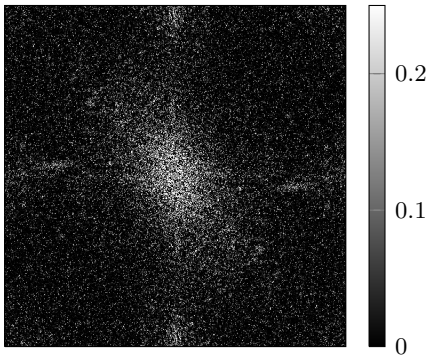
(e) $|\{b_{i,j,\text{rem.}}\}|$ for LMS_{TV}

Figure 10.25.: The remaining DSNU patterns for the different methods for a combined PRNU and DSNU correction in random frame sequences.

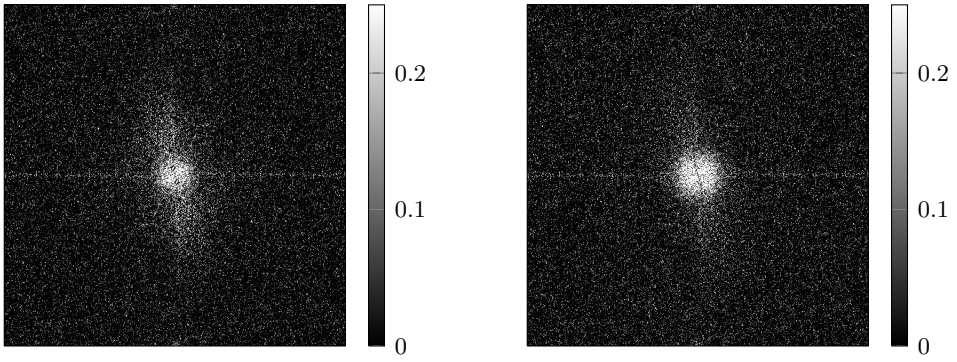


(a) $\min_{\text{comb.}}(\nabla, BL)$ spectral damping for DSNU

(b) $\min_{\text{comb.}}(\nabla, MN)$ spectral damping for DSNU



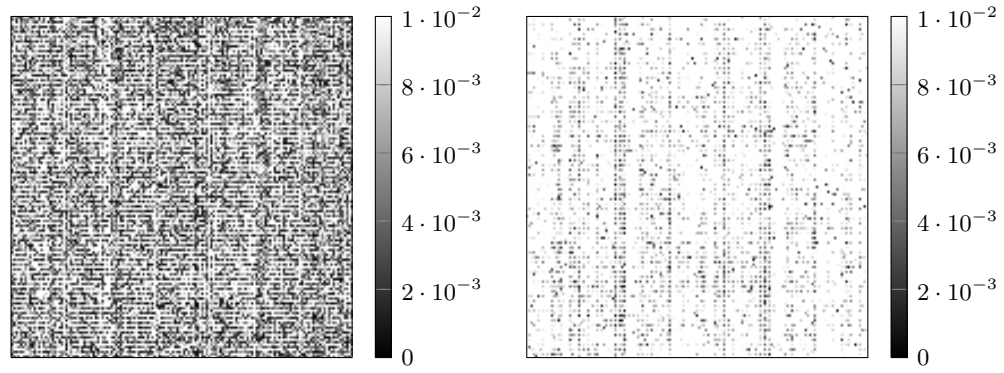
(c) $\min_{\text{comb.}}(\triangle, MN)$ spectral damping for DSNU



(d) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $LMS_{A,7\times7}$

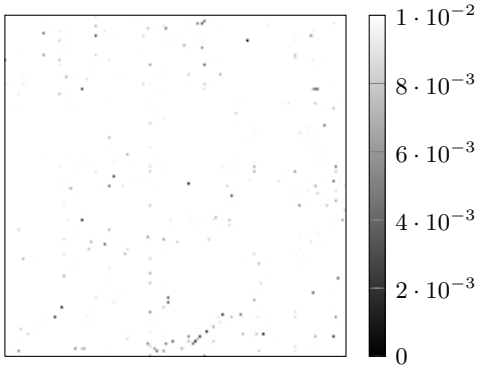
(e) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for LMS_{TV}

Figure 10.26.: The spectral damping factors for the DSNU patterns and the different methods for a combined PRNU and DSNU correction in random frame sequences.

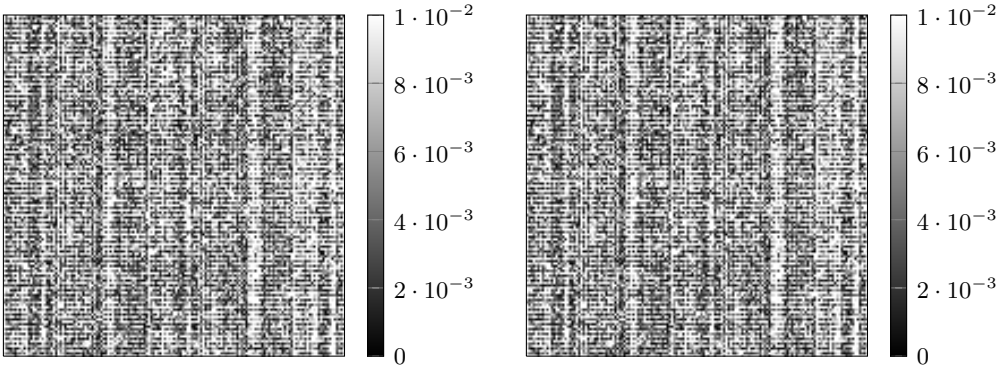


(a) $\min_{\text{comb.}}(\nabla, BL)$ remaining PRNU

(b) $\min_{\text{comb.}}(\nabla, MN)$ remaining PRNU



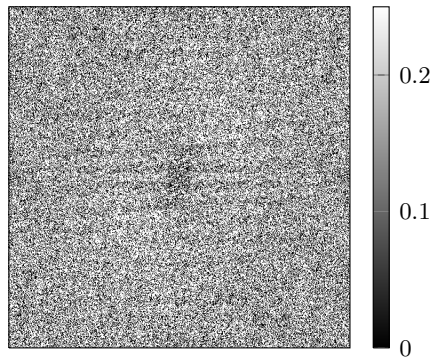
(c) $\min_{\text{comb.}}(\triangle, MN)$ remaining PRNU



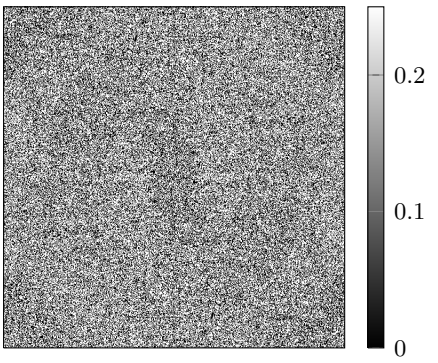
(d) $|1-\{a_{i,j,\text{rem.}}\}|$ for $LMS_{A,7\times7}$

(e) $|1-\{a_{i,j,\text{rem.}}\}|$ for LMS_{TV}

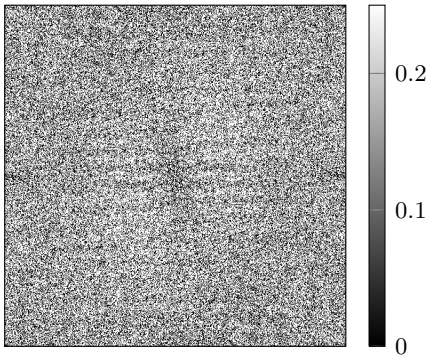
Figure 10.27.: The remaining PRNU patterns for the different methods for a combined PRNU and DSNU correction in random frame sequences.



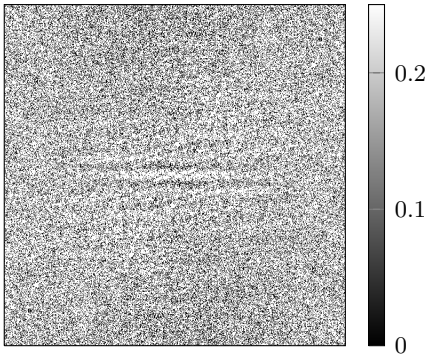
(a) $\min_{\text{comb.}}(\nabla, BL)$ spectral damping for PRNU



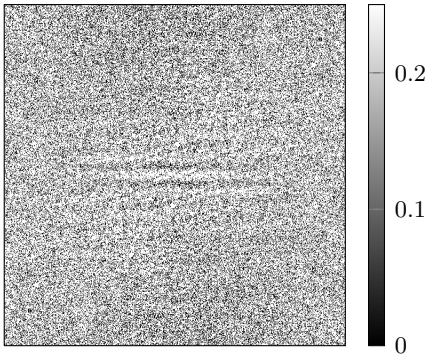
(b) $\min_{\text{comb.}}(\nabla, MN)$ spectral damping for PRNU



(c) $\min_{\text{comb.}}(\triangle, MN)$ spectral damping for PRNU



(d) $\hat{s}_{\text{spectr.,PRNU},\mu,\nu}$ for $LMS_{A,7\times7}$



(e) $\hat{s}_{\text{spectr.,PRNU},\mu,\nu}$ for LMS_{TV}

Figure 10.28.: The spectral damping factors for the PRNU patterns and the different methods for a combined PRNU and DSNU correction in random frame sequences.

10.1.6. Intermediate Conclusion

The above experimental results showed that the violation of the independent frame input results in a performance decrease for the new methods compared to the best reference methods. For most realizations of the maximum likelihood methods the theoretically obtained parameters still lead to a visually comparable correction quality, combined with a good stability of these results, especially for the parameters that result from the histogram extraction.

To mention is as well that the used consecutive frame sequences pose a very special case of being short (500 frames only) and recorded with a high speed frame rate of $100\,\mathrm{Hz}$. For longer sequences the dependency of the averaged measured information will play a less important role, as already discussed in sec. 6.2.2 and the new proposed methods might then be able to outperform the reference methods.

10.2. Behavior For Low Valued Nonuniformities

The above used results have been obtained by using image material that was recorded with the MV1-D1312 camera, which has a DSNU of $\approx 50\,\mathrm{DN}$. This of course raises the question, how the methods perform if they are presented with images that were recorded with the other camera types of sec. 4.3, which in general show a low valued nonuniformity (e.g. a DSNU of about $1\,\mathrm{DN}\,\mathrm{to}\,2\,\mathrm{DN}$ only). The recorded sequences with the other cameras were recorded by camera panning with a slow frame rate of $20\,\mathrm{Hz}$ of the scene *L600a in HD*, and form an intermediate stage between the above used high speed sequences and the random frame sequences. Furthermore the sequences have been recorded with different temperatures and exposure time setups which will show that the algorithms work under a change of these parameters as well.

10.2.1. Analysis of the Reference Methods

To start, the analysis of the reference methods is needed and first results showed that the previously found parameter sets for random and consecutive frame sequences lead to a degradation of the image quality instead of an improvement (see fig. 10.29-10.31). This is not surprising as the frame rate and therefore as well the correlation between two consecutive frames is different in these sequences. To see if the methods can contribute at all to a correction at this low degradations, a further parameter search was conducted which included only the sequences of the *L600a in HD* scene and excluded the recordings of the MV1-D1312 camera. This found *low nonuniformity* optimized parameter set is shown in tab. 10.1 and is a compromise for all the different cameras, temperatures and exposure times mentioned in sec. 4.

Method	Parameter
CS_G	$86.25 \pm (124.1, 143.9\%)$
$LMS_{3\times3}$	$0.001866 \pm (0.001894, 101.5\%)$
$LMS_{7\times7}$	$0.000945 \pm (0.0009802, 103.7\%)$
$LMS_{A,3\times3}$	$0.02543 \pm (0.01218, 47.9\%)$
$LMS_{A,7\times7}$	$0.02513 \pm (0.0122, 48.54\%)$
$LMS_{GA,3\times3}$	$84.3 \pm (306.3, 363.3\%)$
$LMS_{GA,7\times7}$	$38.87 \pm (139.9, 359.9\%)$
LMS_{TV}	$0.04622 \pm (0.02533, 54.81\%)$

Table 10.1.: Results of a parameter search to optimize the literature methods for low valued nonuniformities

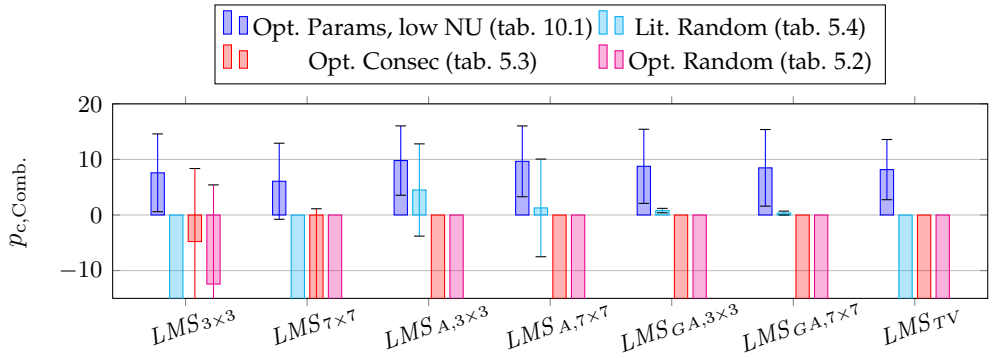


Figure 10.29.: Best-case Performance with respect to the $p_{c,Comb.}$ quality metric for cameras with low nonuniformities

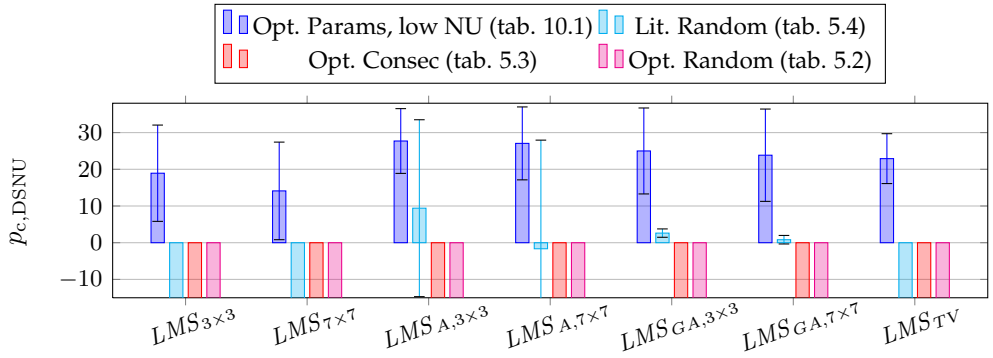


Figure 10.30.: Best-case Performance with respect to the $p_{c,DSNU}$ quality metric for cameras with low nonuniformities

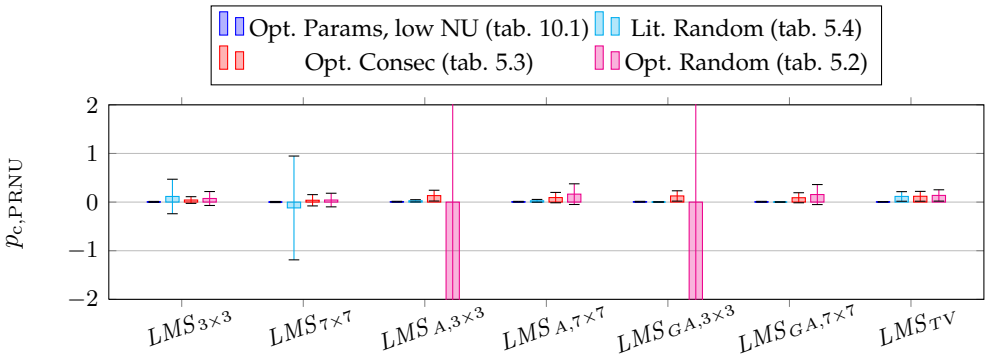


Figure 10.31.: Best-case Performance with respect to the $p_{c,PRNU}$ quality metric for cameras with low nonuniformities

Fig. 10.29 shows the best-case performance analysis for this low nonuniformity sequences for the $p_{c,Comb.}$-norn. It can be seen that a positive correction rate is possible with the optimized parameter set. The proposed parameter set from literature is able to prevent a degradation for the $LMS_{A,3\times3}$ and $LMS_{A,7\times7}$ methods. These results already show that the choice of the correct parameter seems to be a very critical point for this evaluation and slightly wrong parameters immediately lead to a severe loss of performance or even a degradation of the image quality.

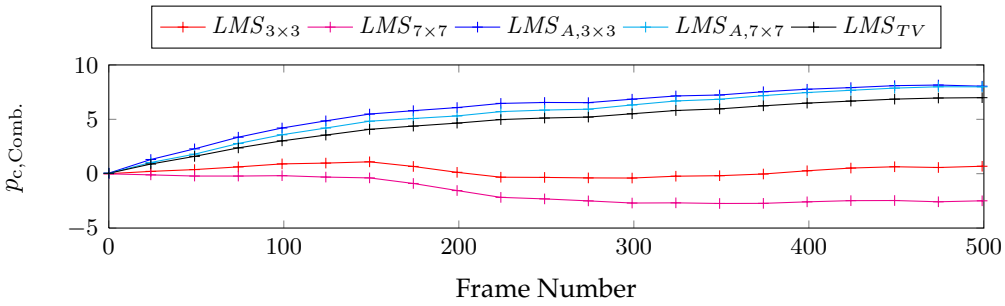


Figure 10.32.: Evaluation for the Basler A602f with $t_{exp.}$=1ms @ 22°C

The temporal analysis gives more insight into this phenomenon and the specificity of the parameter set. As there are many evaluations with different exposure times and temperatures, only the most extreme sets are shown here. Fig. 10.32 and 10.33 show the evaluation for the Basler $A602f$ camera for $t_{exp.} = 1\text{ms} \,@\, 22°\text{C}$ and $t_{exp.} = 10\text{ms} \,@\, 50°\text{C}$. At the lower temperatures the $LMS_{3\times3}$ and $LMS_{7\times7}$ method achieve slightly negative correction rates, while the other methods do actually not reach a performance of 10%. For the higher temperatures, correction rates of $10\,\%$ to $15\,\%$ are possible.

Fig. 10.34 and 10.35 show the temporal evaluation for the MV-640-66-CL camera for $t_{exp.} = 5\text{ms} \,@\, 25°\text{C}$ and $t_{exp.} = 10\text{ms} \,@\, 50°\text{C}$. At the low temperatures the parameters seem to be a misfit to the specific sequence or camera type. Only the $LMS_{3\times3}$ and $LMS_{7\times7}$ methods achieve a stable, positive correction rate. At higher

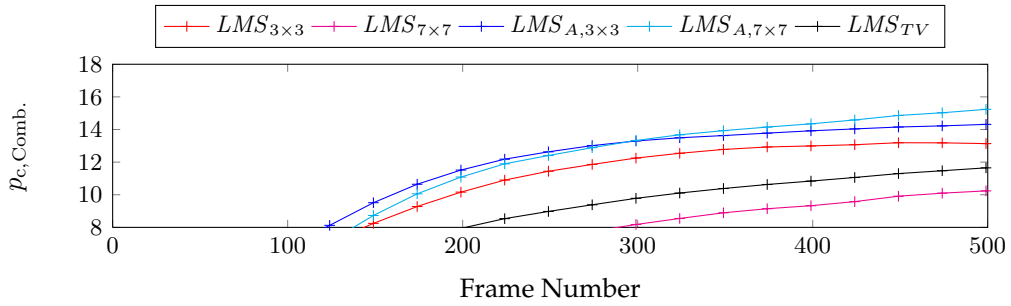


Figure 10.33.: Temporal evaluation for the Basler A602f with $t_{\mathrm{exp.}}$=10ms @ 50°C

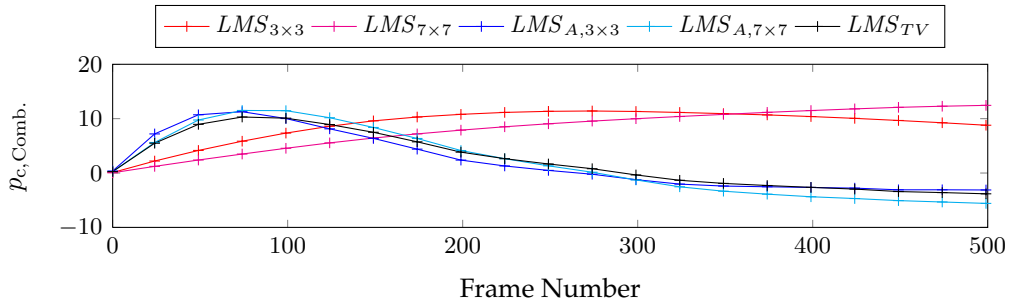


Figure 10.34.: Evaluation for the MV-640-66-CL with $t_{\mathrm{exp.}}$=5ms @ 25°C

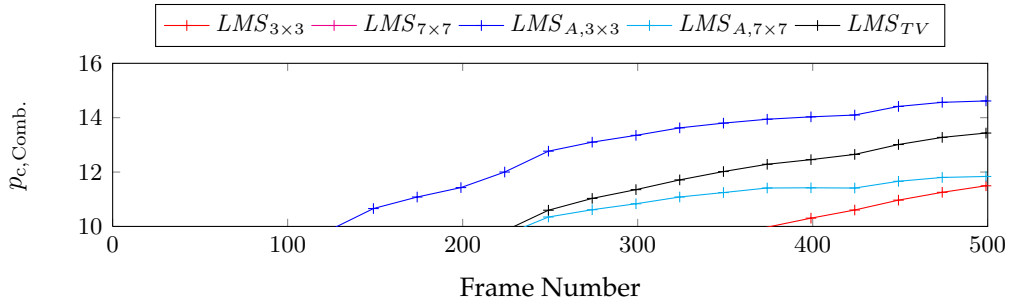


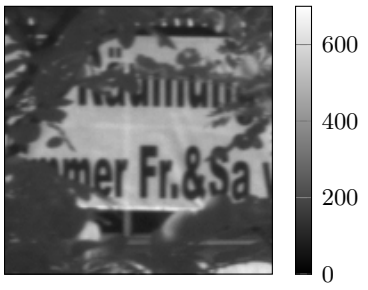
Figure 10.35.: Evaluation for the MV-640-66-CL with $t_{\mathrm{exp.}}$=10ms @ 50°C

temperatures, the results resemble those of the $A602f$ camera for the higher temperatures of $t_{\text{exp.}}$=10ms @ 50°C. The other evaluated temperatures and and exposure times showed a similar behavior that suggests a very parameter specific performance. Their evaluations are shown in appendix G.1 and some methods reach higher correction rates than shown in these sample evaluations.

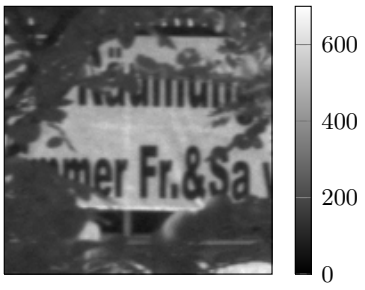
To select a set of comparing methods, the spatial and spectral performance analysis are necessary. A severe visual degradation of the images is not expected as the nonuniformities have effects of only 1 DN to 2 DN. Fig. 10.36 shows an example of the corrected images using the optimized parameters and the sequence A602f-6a from tab.4.4 (p.70) A visual difference between the corrected images is not recognizable. The remaining PRNU patterns in fig. 10.37 shows a partly good correction towards the center of the sensor, except for the CS and CS_G methods which provide high amplitude remains over the whole image. Further, all methods seem to leave spatial high frequency components uncorrected. For the remaining DSNU patterns, the CS and CS_G methods again provide high amplitude remains, while the LMS-type methods give the impression that some of the spatial high frequencies have been corrected (see fig. 10.38).

The above results can be confirmed by the analysis of the spectral damping factors. Fig. 10.39 shows the analysis for the PRNU correction, where the LMS-type methods correct spatial low frequencies around the center and the corrections hardly differ between the methods. The DSNU spectral damping is presented by fig. 10.40 and here the differences between the methods can be observed. Especially the $LMS_{A,3\times3}$ method gives a good damping in the spatial high frequencies. The $LMS_{A,7\times7}$ method damps even further low frequency components, however it does not reach the damping in the higher frequencies the $LMS_{A,3\times3}$ method does. The gated adaptive version $LMS_{GA,7\times7}$ however fails to provide an improved correction, which has the same reasons as discussed before for this type of method.

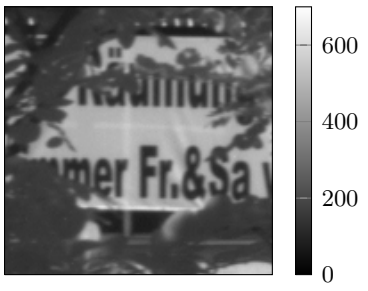
Of course these results are specific for the one selected sequence, but in general the observations are valid for the other methods as well, if a positive correction rate is achieved.



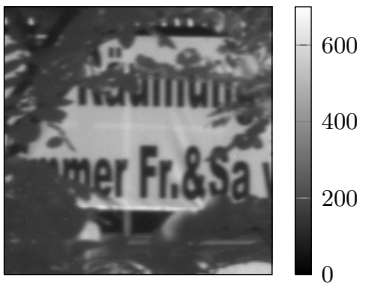
(a) Corrected with CS Method



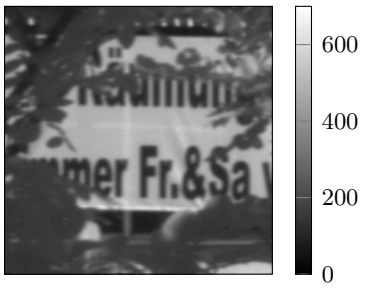
(b) Corrected with CS_G Method



(c) Corrected with $LMS_{3\times3}$ Method

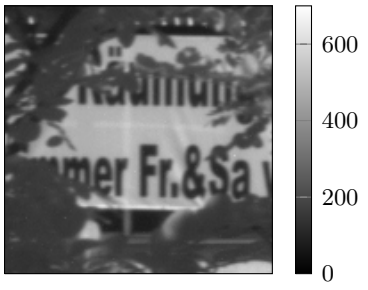


(d) Corrected with $LMS_{7\times7}$ Method



(e) Corrected with $LMS_{A,3\times3}$ Method

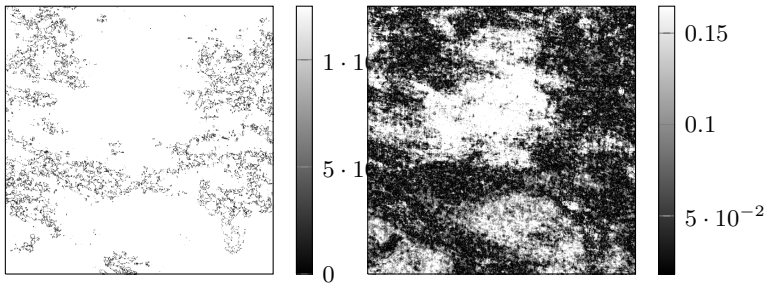
(f) Corrected with $LMS_{A,7\times7}$ Method



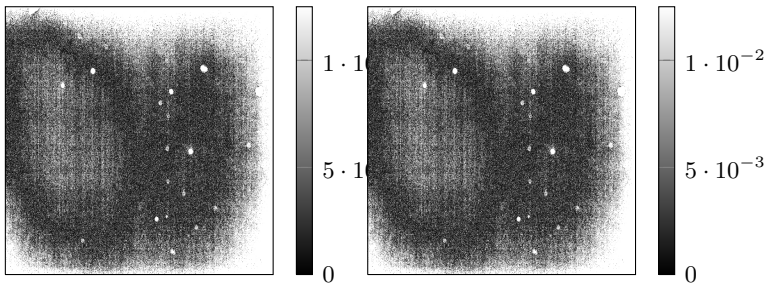
(g) Corrected with $LMS_{GA,7\times7}$ Method

(h) Corrected with LMS_{TV} Method

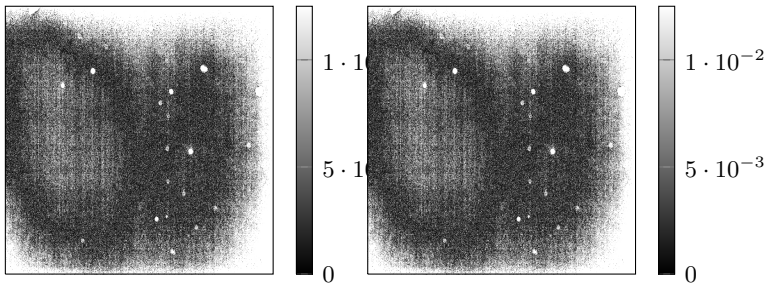
Figure 10.36.: Corrected images by the reference methods for sequences with low valued nonuniformities (Parameter set according to tab. 10.1).



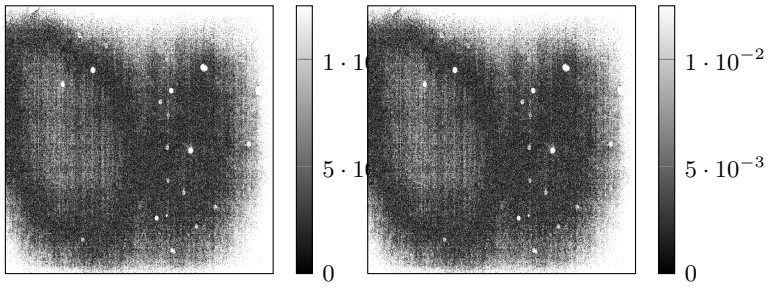
(a) $|1 - \{a_{i,j,rem.}\}|$ for CS (b) $|1 - \{a_{i,j,rem.}\}|$ for CS_G



(c) $|1 - \{a_{i,j,rem.}\}|$ for $LMS_{3\times3}$ (d) $|1 - \{a_{i,j,rem.}\}|$ for $LMS_{7\times7}$

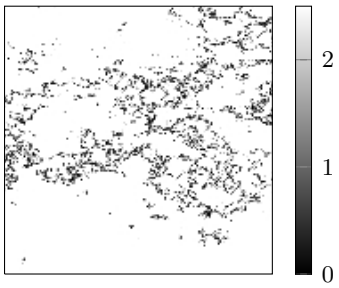


(e) $|1 - \{a_{i,j,rem.}\}|$ for $LMS_{A,3\times3}$ (f) $|1 - \{a_{i,j,rem.}\}|$ for $LMS_{A,7\times7}$

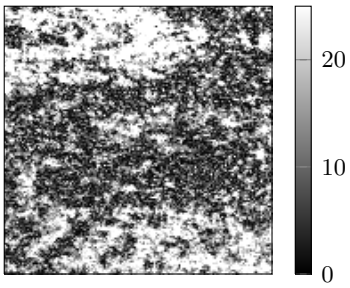


(g) $|1 - \{a_{i,j,rem.}\}|$ for $LMS_{GA,7\times7}$ (h) $|1 - \{a_{i,j,rem.}\}|$ for LMS_{TV}

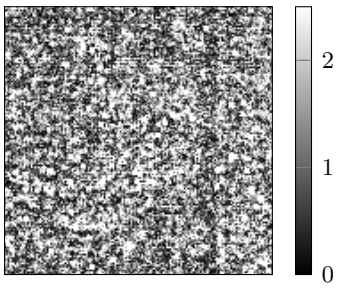
Figure 10.37.: Remaining PRNU pattern for the reference methods for sequences with low valued nonuniformities (Parameter set according to tab. 10.1).



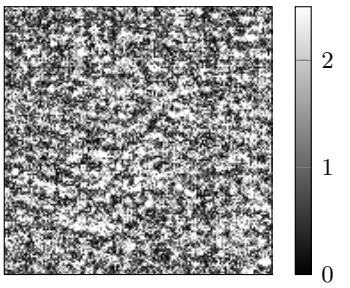
(a) $|\{b_{i,j,rem.}\}|$ for CS



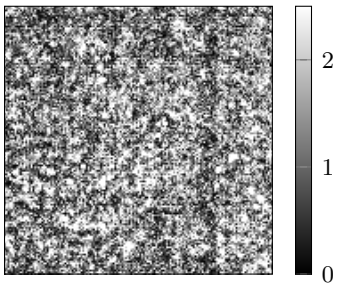
(b) $|\{b_{i,j,rem.}\}|$ for CS_G



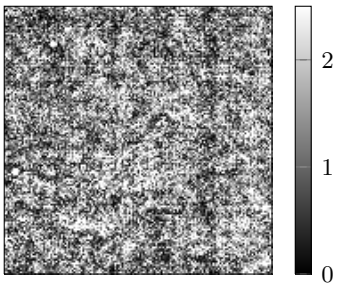
(c) $|\{b_{i,j,rem.}\}|$ for $LMS_{3\times3}$



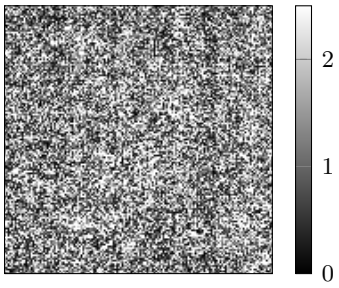
(d) $|\{b_{i,j,rem.}\}|$ for $LMS_{7\times7}$



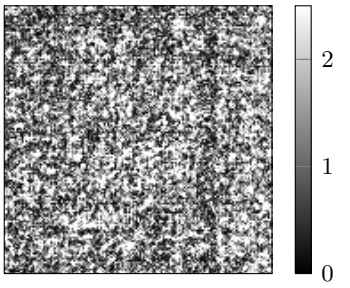
(e) $|\{b_{i,j,rem.}\}|$ for $LMS_{A,3\times3}$



(f) $|\{b_{i,j,rem.}\}|$ for $LMS_{A,7\times7}$

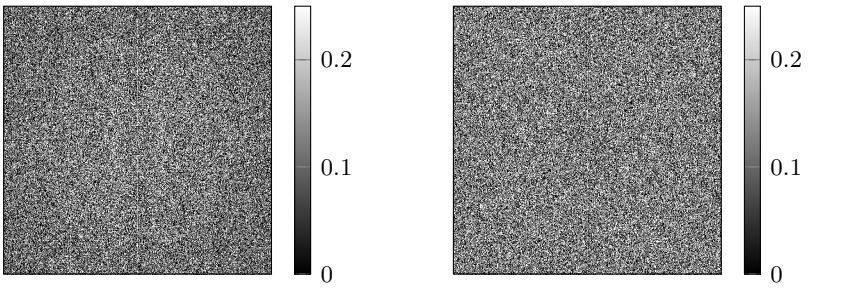


(g) $|\{b_{i,j,rem.}\}|$ for $LMS_{GA,7\times7}$



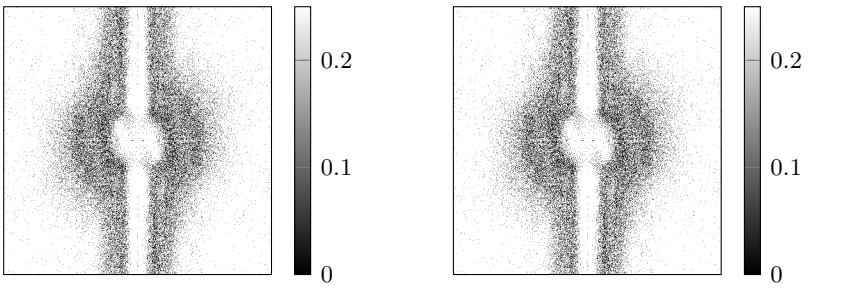
(h) $|\{b_{i,j,rem.}\}|$ for LMS_{TV}

Figure 10.38.: Remaining DSNU pattern for the reference methods for sequences with low valued nonuniformities (Parameter set according to tab. 10.1).



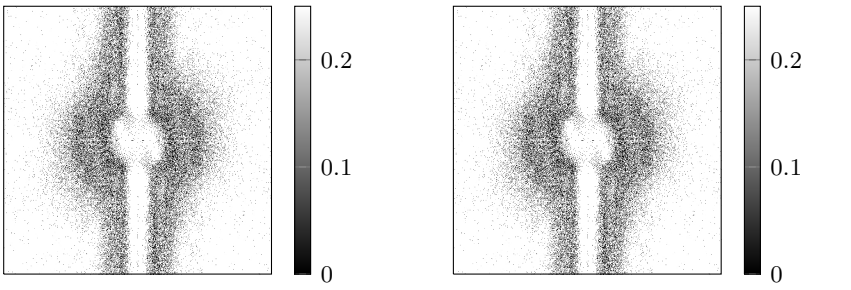
(a) $\hat{s}_{\text{spectr.,PRNU},\mu,\nu}$ for CS

(b) $\hat{s}_{\text{spectr.,PRNU},\mu,\nu}$ for CS_G



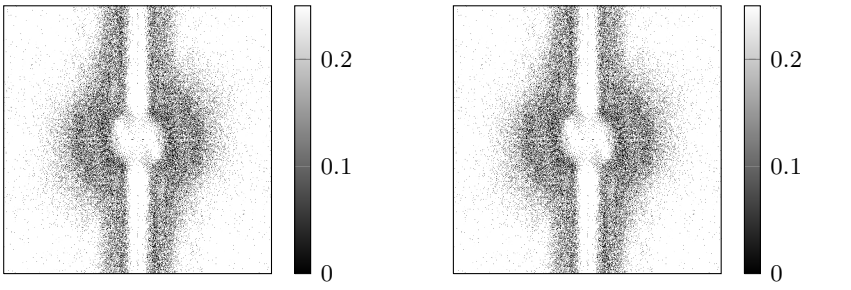
(c) $\hat{s}_{\text{spectr.,PRNU},\mu,\nu}$ for $LMS_{3\times3}$

(d) $\hat{s}_{\text{spectr.,PRNU},\mu,\nu}$ for $LMS_{7\times7}$



(e) $\hat{s}_{\text{spectr.,PRNU},\mu,\nu}$ for $LMS_{A,3\times3}$

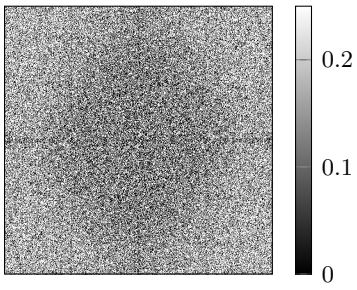
(f) $\hat{s}_{\text{spectr.,PRNU},\mu,\nu}$ for $LMS_{A,7\times7}$



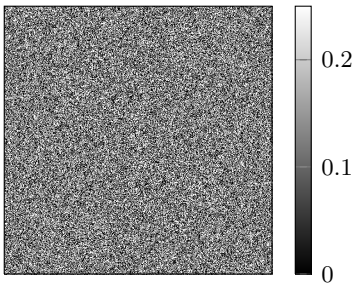
(g) $\hat{s}_{\text{spectr.,PRNU},\mu,\nu}$ for $LMS_{GA,7\times7}$

(h) $\hat{s}_{\text{spectr.,PRNU},\mu,\nu}$ for LMS_{TV}

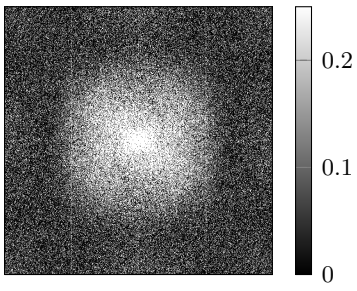
Figure 10.39.: Spectral damping coefficients for PRNU and the reference methods for sequences with low valued nonuniformities (Parameter set according to tab. 10.1).



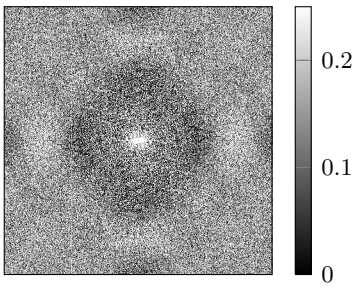
(a) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for CS



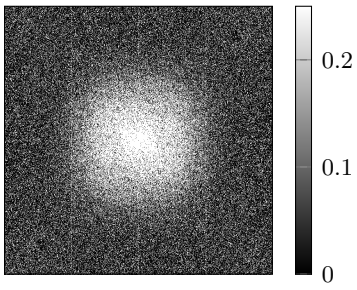
(b) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for CS_G



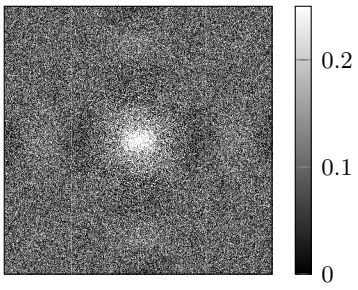
(c) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $LMS_{3\times3}$



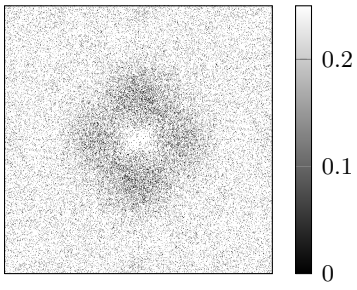
(d) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $LMS_{7\times7}$



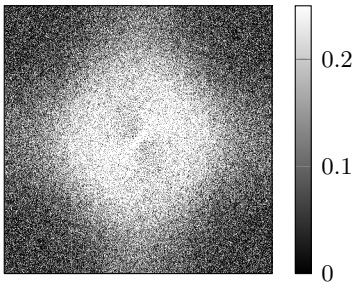
(e) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $LMS_{A,3\times3}$



(f) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $LMS_{A,7\times7}$



(g) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $LMS_{GA,7\times7}$

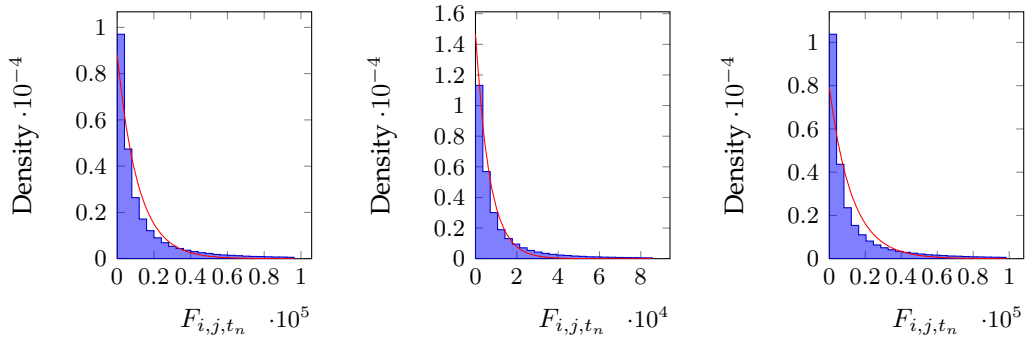


(h) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for LMS_{TV}

Figure 10.40.: Spectral damping coefficients for DSNU and the reference methods for sequences with low valued nonuniformities (Parameter set according to tab. 10.1).

10.2.2. Analysis of the DSNU Maximum Likelihood Method

With concern to the maximum likelihood methods, a new parameter search for these methods is conducted and the resulting optimized parameter set for the low valued nonuniformities and is presented in tab. 10.3. The extraction of the statistic parameters as already introduced in sec. 6.2.3 leads to new histograms for the new image material which are shown in fig. 10.41. The obtained parameter set is presented in tab. 10.2 and differs from the set obtained for the MV1-1312 camera statistics (see tab. 6.1 on page p. 121).



(a) with $HP(\nabla, \mathrm{BL})$ (b) with $HP(\nabla, \mathrm{MN})$ (c) with $HP(\triangle, \mathrm{MN})$

Figure 10.41.: Histograms of the differnent F_{i,j,t_n} realizations for the image material with the low valued nonuniformities.

Method	Parameter
$\min(\nabla, BL)$	$8.831 \cdot 10^{-5}$
$\min(\nabla, MN)$	0.0001466
$\min(\triangle, MN)$	$7.928 \cdot 10^{-5}$

Table 10.2.: Results of a parameter evaluations from the statistics for low valued nonuniformities

Method	Parameter
$\min(\nabla, BL)$	$0.000343 \pm (0.0004985, 145.3\%)$
$\min(\nabla, MN)$	$0.0006452 \pm (0.0009839, 152.5\%)$
$\min(\triangle, MN)$	$0.0003371 \pm (0.0005028, 149.2\%)$

Table 10.3.: Results of a parameter search to optimize the new methods for low valued nonuniformities

The best-case performances for all the available parameter sets is shown in fig. 10.42. To be able to compare all the methods' performances against the reference methods, the performance of the $LMS_{A,7\times7}$ and LMS_{TV} methods are included into the plot and the new DSNU correction methods perform around $5\,\%$ to $10\,\%$ lower than the reference methods. At least the new methods are able to produce

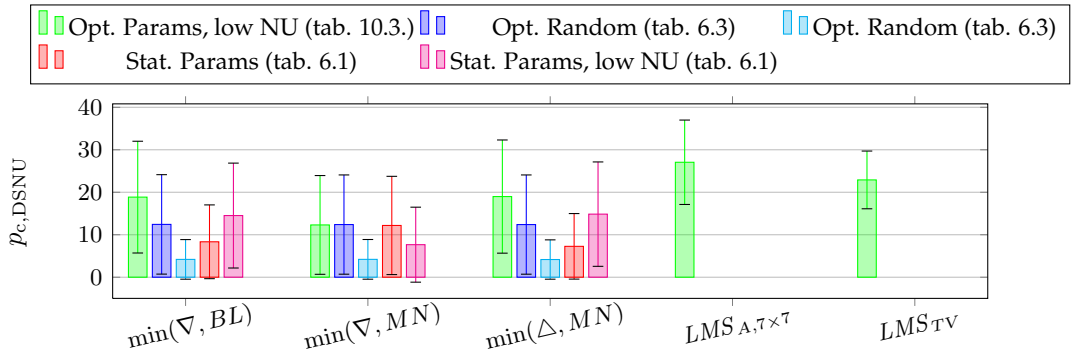
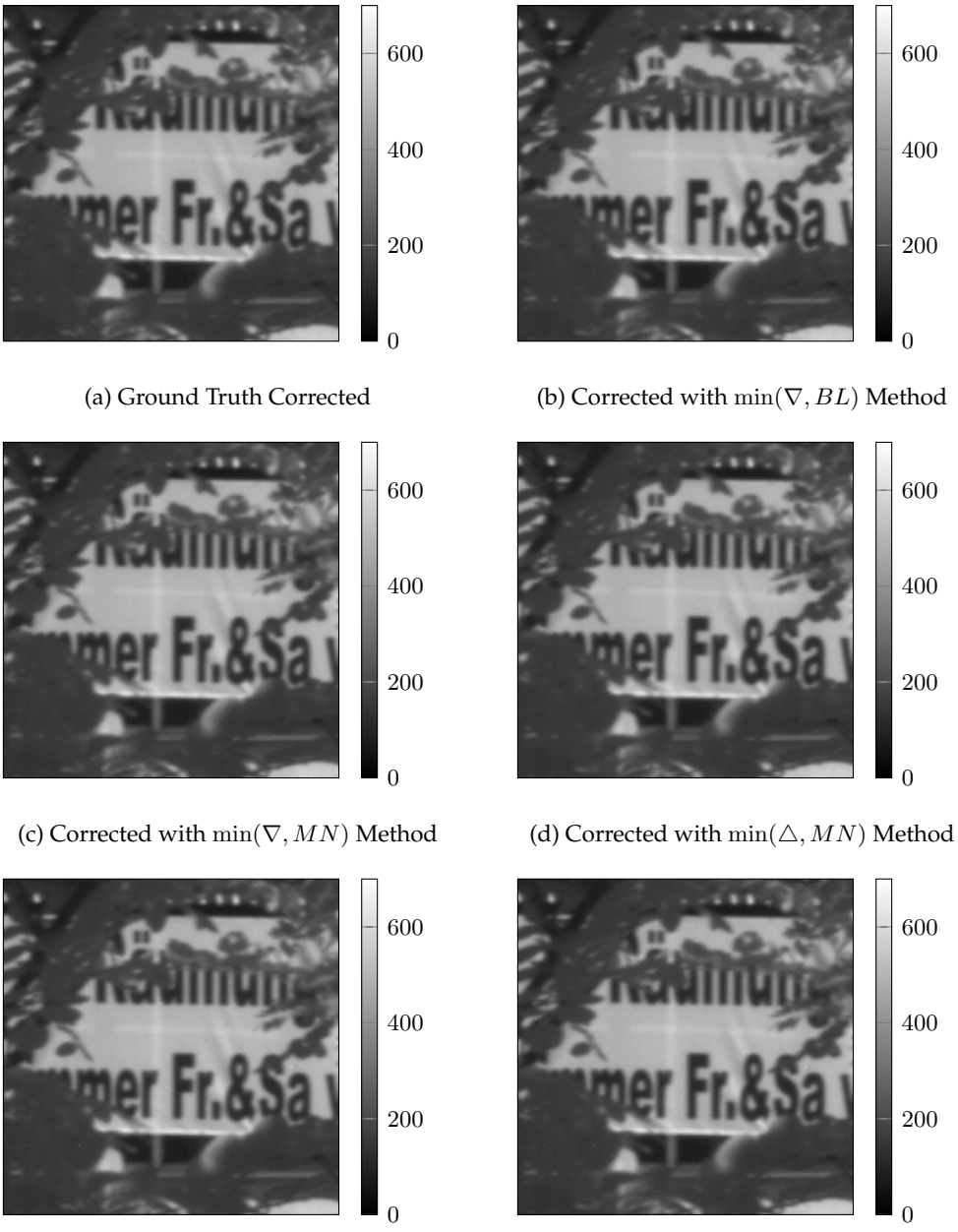


Figure 10.42.: Best-case performance with respect to the $p_{c,DSNU}$ quality metric for cameras with low nonuniformities

a positive correction rate for the non-optimal parameter sets, while the reference methods tend to achieve only negative correction rates if the parameters do not fit (see in fig. 10.30). The temporal evaluations for this very first maximum likelihood approach are skipped, as already the best case performances could not reach the reference methods.

For the visual correction performance, a degradation of the corrected images is not expected and fig. 10.43 confirms this. Actually there is no perception of a visual difference between the ground truth corrected images and the corrected images by the different methods possible. The remaining DSNU patterns in fig. 10.44 show however that spatial high frequencies have been corrected successfully and the spectral damping coefficients in fig. 10.45 confirm these results. Each of the presented method has its own damping characteristics and for the newly introduced methods these characteristics follow the HP filter masks. The most unfortunate remains are found for the $\min(\nabla, MN)$ method which are the uncorrected spatial high frequency components in the corners as discussed in sec. 10.1 already several times. Considering the average spectral remains, all compared methods perform equally well for this sequence and I would recommend to select a method according to the desired spectral correction demands.



(a) Ground Truth Corrected

(b) Corrected with $\min(\nabla, BL)$ Method

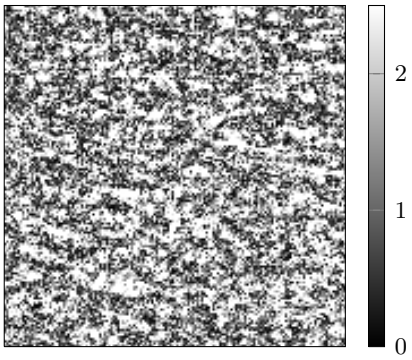
(c) Corrected with $\min(\nabla, MN)$ Method

(d) Corrected with $\min(\triangle, MN)$ Method

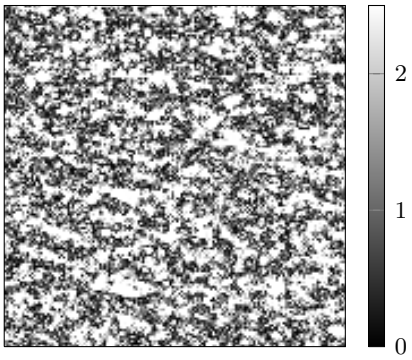
(e) Corrected with $LMS_{A,3\times3}$ Method

(f) Corrected with $LMS_{A,7\times7}$ Method

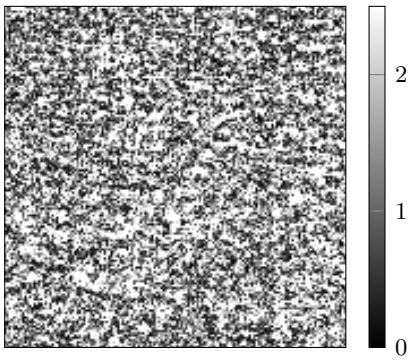
Figure 10.43.: Corrected images by the introduced methods for sequences with low valued nonuniformities (Parameter set according to tab. 10.42).



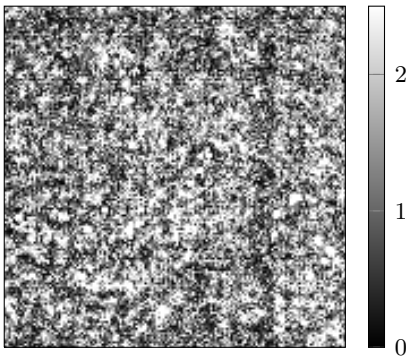
(a) $|\{b_{i,j,rem.}\}|$ for $\min(\nabla, BL)$



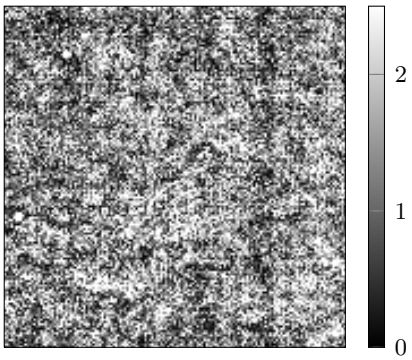
(b) $|\{b_{i,j,rem.}\}|$ for $\min(\nabla, MN)$



(c) $|\{b_{i,j,rem.}\}|$ for $\min(\triangle, MN)$

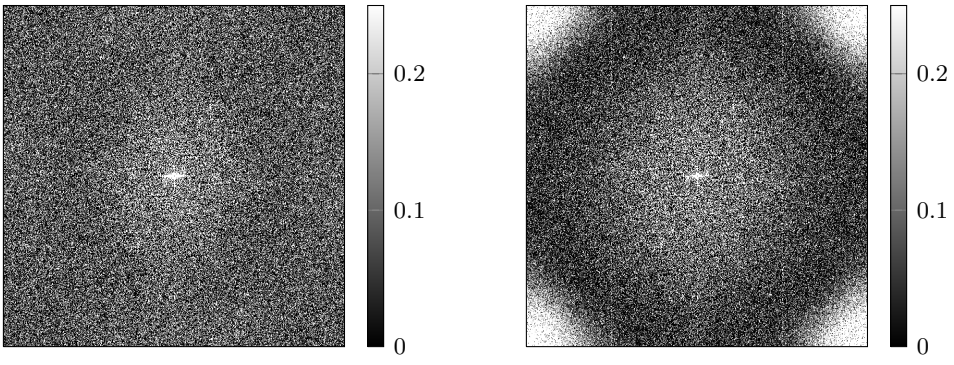


(d) $|\{b_{i,j,rem.}\}|$ for $LMS_{A,3\times3}$

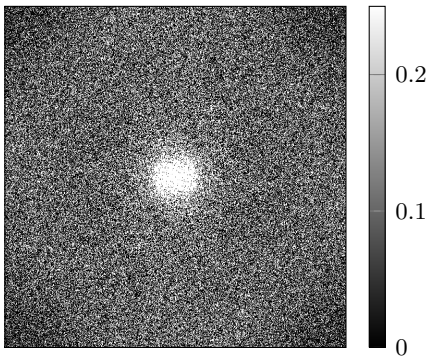


(e) $|\{b_{i,j,rem.}\}|$ for $LMS_{A,7\times7}$

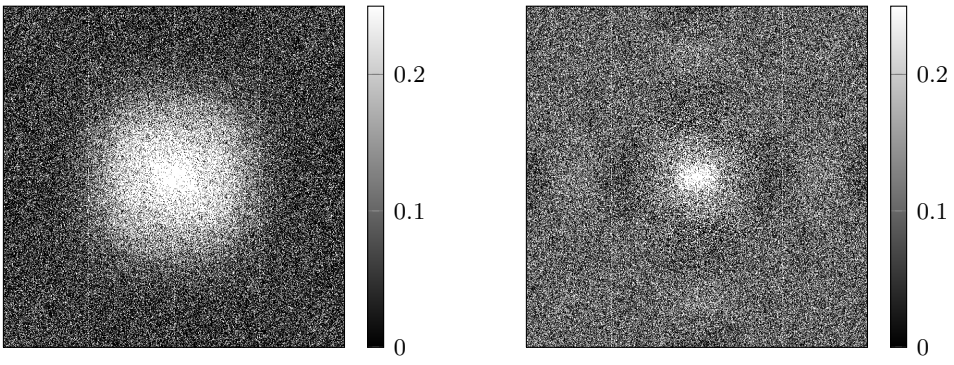
Figure 10.44.: Remaining DSNU patterns for the introduced methods and sequences with low valued nonuniformities (Parameter set according to tab. 10.42).



(a) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $\min(\nabla, BL)$ (b) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $\min(\nabla, MN)$



(c) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $\min(\triangle, MN)$

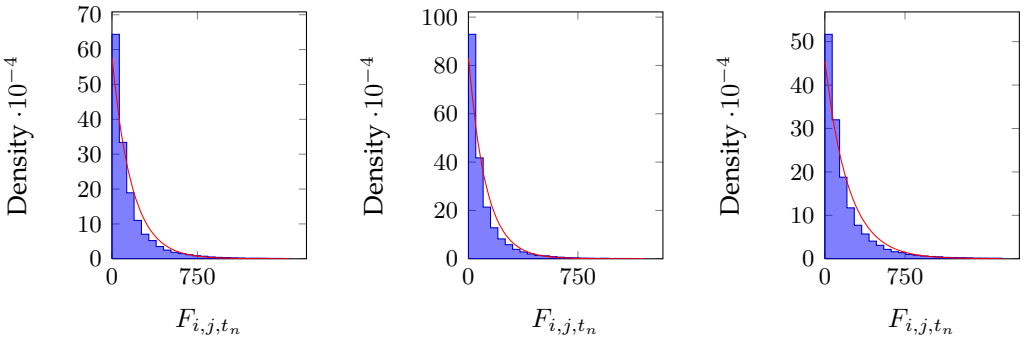


(d) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $LMS_{A,3\times3}$ (e) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $LMS_{A,7\times7}$

Figure 10.45.: Spectral damping coefficients for the DSNU for the introduced methods and sequences with low valued nonuniformities (Parameter set according to tab. 10.42).

10.2.3. Analysis for the Weighted Average Improvement

Next the methods that come from the variations of the improved maximum likelihood approach are investigated. The methods depend on the weighted average of the sensor measurements and the modified statistical evaluations give a further parameter set (evaluations of the statistics according to sec. 7.1.2). The histograms are shown in fig. 10.46 and the obtained parameter set is summarized in tab. 10.4. Further, a parameter search led to the best-case optimized parameter set shown in tab. 10.5.



(a) with $HP(\nabla, \mathrm{BL})$ (b) with $HP(\nabla, \mathrm{MN})$ (c) with $HP(\triangle, \mathrm{MN})$

Figure 10.46.: Histograms of the differnent $M_{\hat{L},t_n}$ realizations

Method	Parameter
$\min_{\mathrm{wA}}(\nabla, BL)$	0.00577
$\min_{\mathrm{wA}}(\nabla, MN)$	0.00831
$\min_{\mathrm{wA}}(\triangle, MN)$	0.00455

Table 10.4.: Results of a parameter evaluations from the statistics for low valued nonuniformities

Method	Parameter
$\min_{\mathrm{wA}}(\nabla, BL)$	$0.0009358 \pm (0.0006629, 70.84\%)$
$\min_{\mathrm{wA}}(\nabla, MN)$	$0.001776 \pm (0.001248, 70.26\%)$
$\min_{\mathrm{wA}}(\triangle, MN)$	$0.0009475 \pm (0.0007426, 78.37\%)$

Table 10.5.: Results of a parameter search to optimize the new methods for low valued nonuniformities

To see the different performances of all these parameter sets, the best-case performance analysis for all the available parameter sets is shown in fig. 10.42. As expected, the best results are obtained for the experimentally optimized parameter set. Compared to the performances for the reference methods $LMS_{A,7\times7}$ and LMS_{TV}, the optimized parameter set results in an equal performance with respect to the estimated error bars. Surprising is however that none of the low valued

nonuniformity statistic parameter sets leads to a good performance. This points towards a severe parameter instability of the methods that yields from this input material and will now be analyzed in the temporal performance analysis.

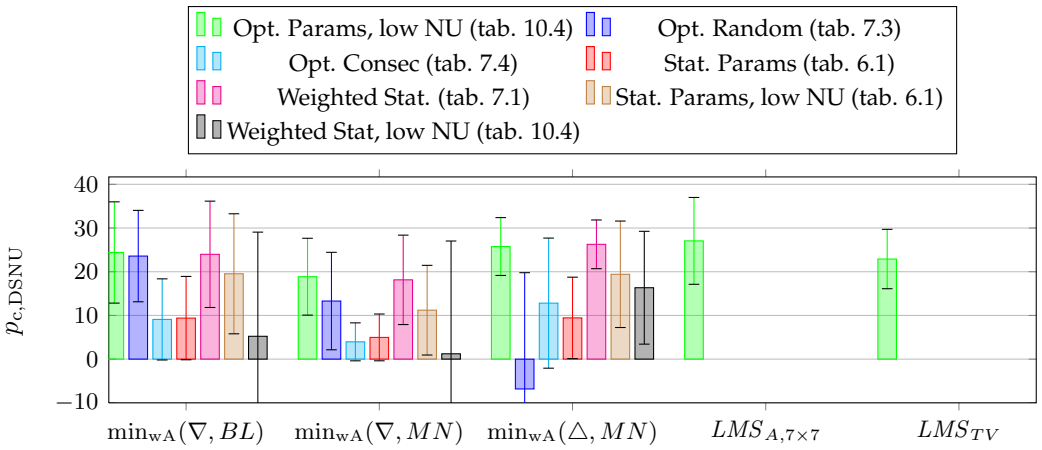


Figure 10.47.: Best-case Performance with respect to the $p_{\mathrm{c,DSNU}}$ quality metric for cameras with low valued nonuniformities

The temporal evaluations show a direct comparison between the improved new methods and the two best reference methods $LMS_{A,7\times7}$ and LMS_{TV}. As there are many evaluations with different exposure times and temperatures, only the sets and cameras with the most diverting experimental parameters are shown. Fig. 10.48 and 10.49 show the temporal evaluations for the Basler $A602f$ camera for $t_{\mathrm{exp.}}$=1ms @ 22°C and $t_{\mathrm{exp.}}$=10ms @ 50°C. At the low temperature the $\min_{\mathrm{wA}}(\nabla, BL)$ performs almost $10\,\%$ better then the best reference method $LMS_{A,7\times7}$ and the $\min_{\mathrm{wA}}(\triangle, MN)$ method also outperforms this method. Only the $\min_{\mathrm{wA}}(\nabla, MN)$ method does not reach the other methods, which again results from the high frequency problems for consecutive frame sequences. For the higher temperatures, the new methods' parameters seem to mismatch, as they reach up to a $30\,\%$ correction rate, but then fall back to $10\,\%$ to $20\,\%$.

Fig. 10.50 and 10.51 show the temporal evaluation for the MV-640-66-CL camera for $t_{\mathrm{exp.}}$=5ms @ 25°C and $t_{\mathrm{exp.}}$=10ms @ 48°C with the optimized parameter set. At low temperatures the parameters fit well and the $\min_{\mathrm{wA}}(\nabla, BL)$ and $\min_{\mathrm{wA}}(\triangle, MN)$ method perform best, while the $\min_{\mathrm{wA}}(\nabla, MN)$ method again does not reach a good performance. The reference methods seem however to have mismatched parameters, but instead of declining their performance towards $0\,\%$, they actually deteriorate the image by creating a negative correction rate. For the higher temperatures the $\min_{\mathrm{wA}}(\triangle, MN)$ and $\min_{\mathrm{wA}}(\nabla, BL)$ methods outperform the reference methods, where the $\min_{\mathrm{wA}}(\triangle, MN)$ is the overall best method, confirming again its stability. An overall result on which method tends to be the best can not be given by this comparison. However it is obvious that each method is able to correct the remaining nonuniformities to some extend and that

the choice of wrong parameters leads to a decrease of the corrections, even to negative correction rates.

The evaluations of the other experimental setups are shown in appendix G.2 and show a similar behavior as discussed here, that mainly depends on the chosen sequence. However, some higher correction rates up to $40\,\%$ are reached in the shown plots.

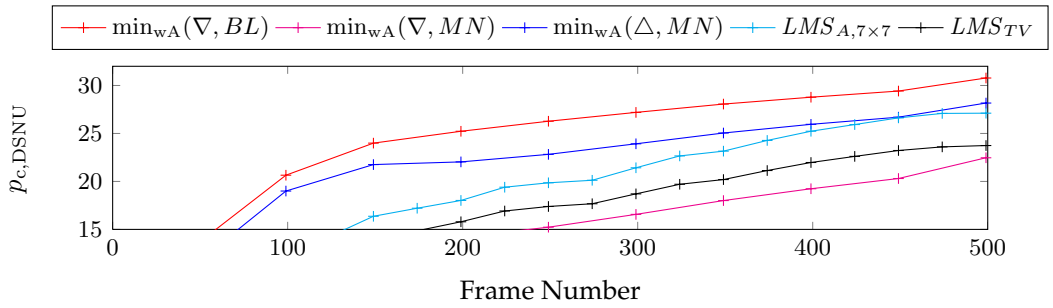


Figure 10.48.: Temporal evaluation for the Basler A602f with $t_{\mathrm{exp.}}$=1ms @ 22°C

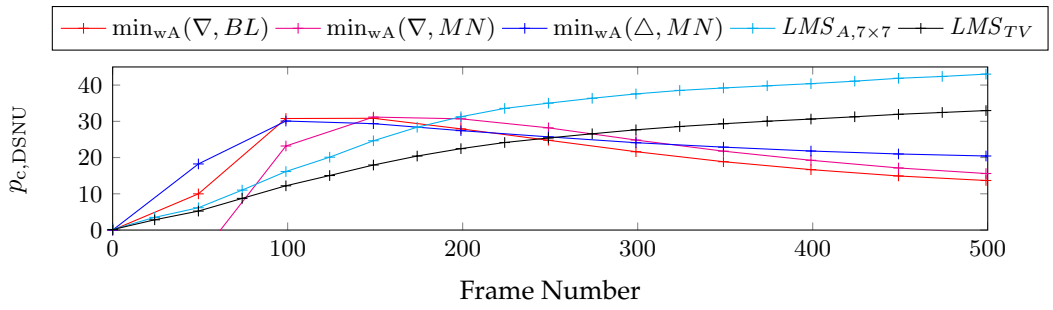


Figure 10.49.: Temporal evaluation for the Basler A602f with $t_{\mathrm{exp.}}$=10ms @ 50°C

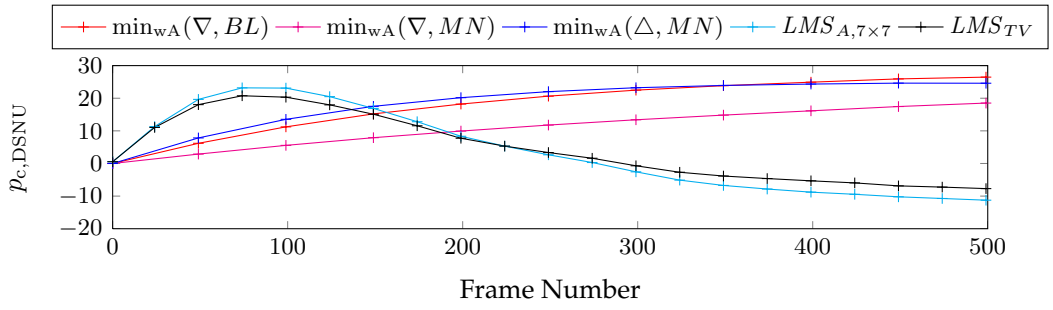


Figure 10.50.: Temporal evaluation for the MV-640-66-CL with $t_{\mathrm{exp.}}$=5ms @ 25°C

The analysis of the corrected images of the sequence A602f-6a in fig. 10.52 shows again no significant difference for the method-wise corrections compared to the ground corrections. Given the remaining DSNU patterns in fig. 10.53, a correction of the spatial high frequency components can be observed and this behavior is as always seen best in the spectral damping analysis in fig. 10.54. However

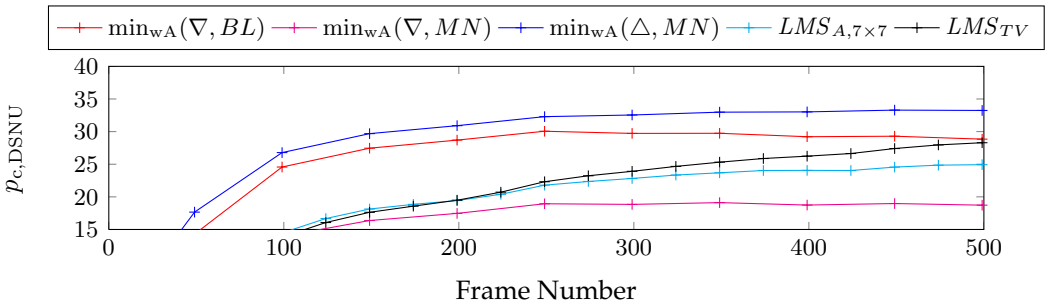
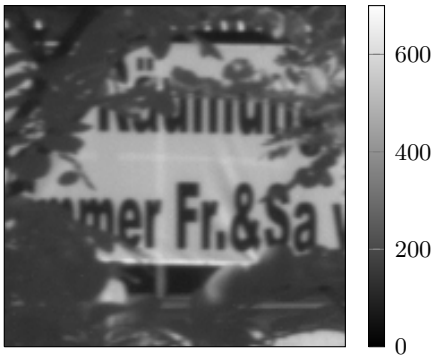


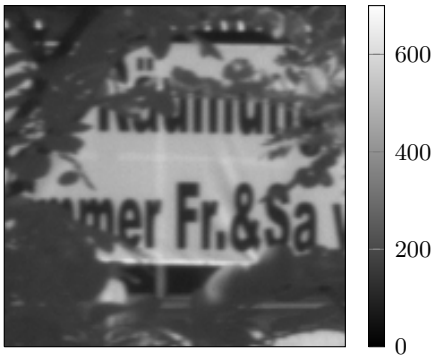
Figure 10.51.: Temporal evaluation for the MV-640-66-CL with $t_{\mathrm{exp.}}$=10ms @ 50°C

an improved correction performance compared to the non-weighted method in fig. 10.45 can not be seen directly for this sequence. The spectral damping follows tightly the characteristics of the high pass filter masks and the new methods merely outperform the reference methods by a good damping of the spatial low frequency components while the reference methods provide a better damping for the spatial higher frequency components.

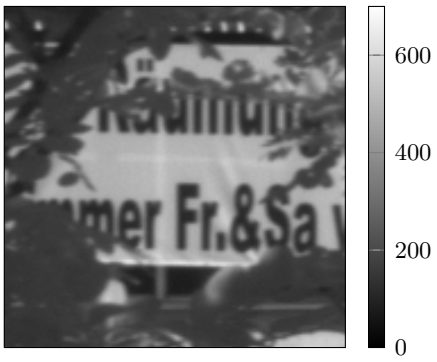
Concluding, the results above suggest that a parameter search is mandatory for the low valued nonuniformities. However, there is still a chance that the results of the statistic parameter extraction failed because not enough image material was present and therefore the several different cameras and experimental setups had to be mixed. As the temporal results show, the correct parameters may also be influenced by external parameters like type of camera, temperature or exposure time.



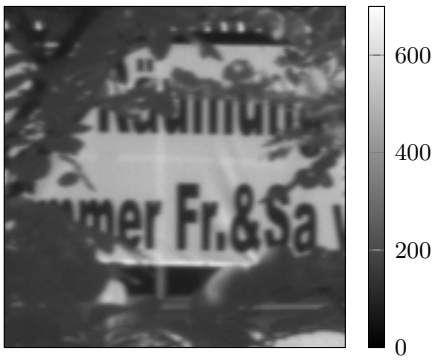
(a) Ground Truth Corrected



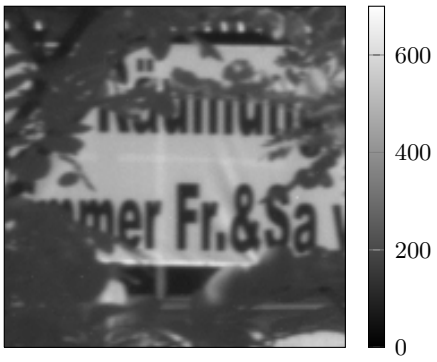
(b) Corrected with $\min(\nabla, BL)$ Method



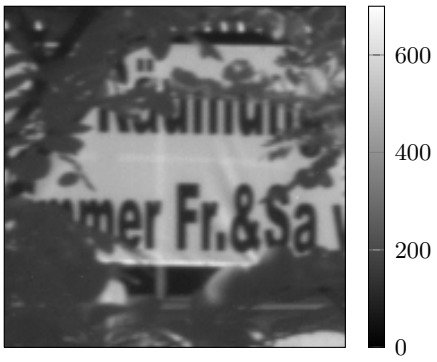
(c) Corrected with $\min(\nabla, MN)$ Method



(d) Corrected with $\min(\triangle, MN)$ Method

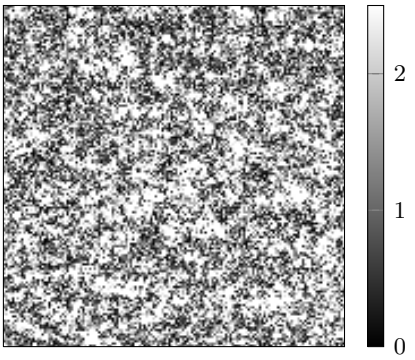


(e) Corrected with $LMS_{A,3\times3}$ Method

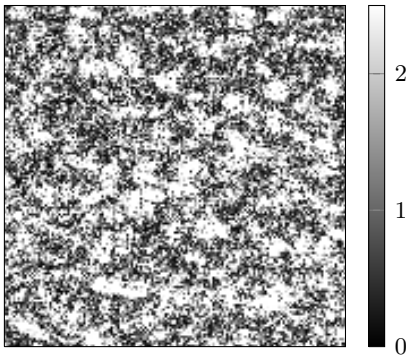


(f) Corrected with $LMS_{A,7\times7}$ Method

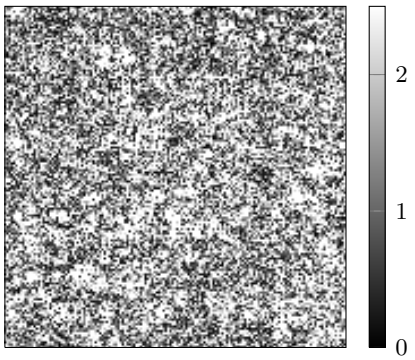
Figure 10.52.: Corrected images by the introduced methods for sequences with low valued nonuniformities (Parameter set according to tab. 10.47).



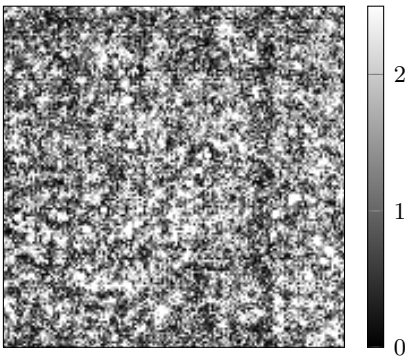
(a) $|\{b_{i,j,rem.}\}|$ for $\min(\nabla, BL)$



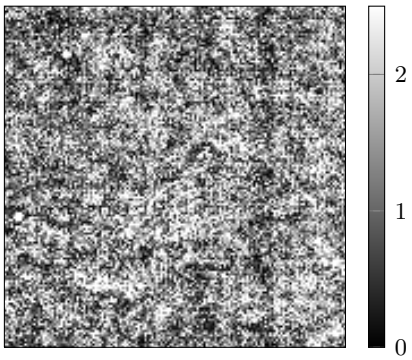
(b) $|\{b_{i,j,rem.}\}|$ for $\min(\nabla, MN)$



(c) $|\{b_{i,j,rem.}\}|$ for $\min(\triangle, MN)$

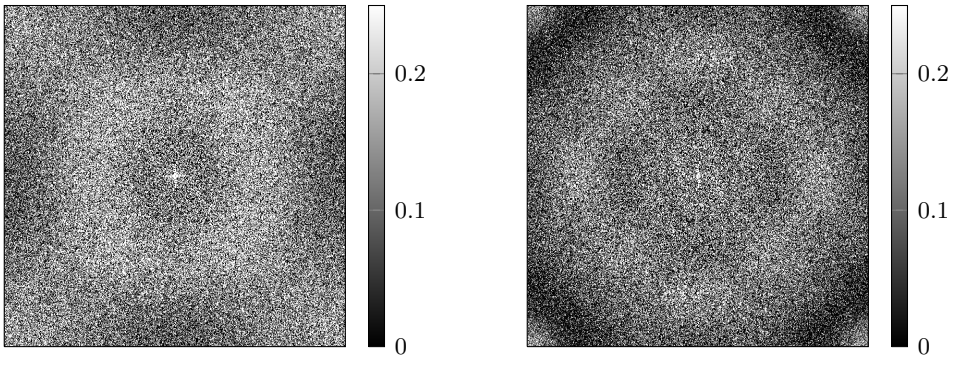


(d) $|\{b_{i,j,rem.}\}|$ for $LMS_{A,3\times3}$

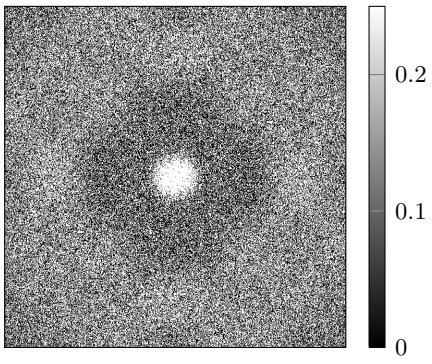


(e) $|\{b_{i,j,rem.}\}|$ for $LMS_{A,7\times7}$

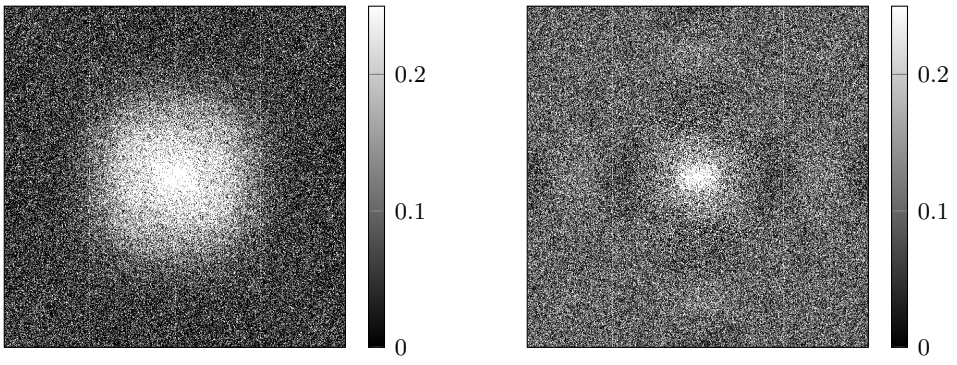
Figure 10.53.: Remaining DSNU patterns for the introduced methods and sequences with low valued nonuniformities (Parameter set according to tab. 10.47).



(a) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $\min(\nabla, BL)$ (b) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $\min(\nabla, MN)$



(c) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $\min(\triangle, MN)$



(d) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $LMS_{A,3\times3}$ (e) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $LMS_{A,7\times7}$

Figure 10.54.: Spectral damping coefficients for the DSNU for the introduced methods and sequences with low valued nonuniformities (Parameter set according to tab. 10.47).

10.2.4. Analysis for the CNN-Type Methods

The CNN type methods can also be analyzed for their behavior on low valued nonuniformities. Therefore the optimized λ parameters from tab.10.5 are used and a new parameter search was conducted for the s_ϵ parameter (see tab. 10.6) The step sizes are actually in the order of 10^{-1} and the expected DSNU is around $1\,\mathrm{DN}$ to $2\,\mathrm{DN}$. For comparison, the step sizes for the $MV1312$ camera resulted in values of $2-8$ for the given DSNU of about $50\,DN$. The difference in the step sizes go by the same factor the DSNU values differ, which may give a hint for future parameter searches.

Fig. 10.55 shows the corresponding best-case performance analysis, where the CNN methods show a better performance than the reference methods, except for the $\mathrm{CNN}(\nabla, MN)$. In fact, the reference methods have much wider error bars than the CNN methods, which lets them appear less stable.

Method	Parameter
$\mathrm{CNN}(\nabla, BL)$	$0.1398 \pm (0.1649, 117.9\%)$
$\mathrm{CNN}(\nabla, MN)$	$0.5633 \pm (0.5069, 89.99\%)$
$\mathrm{CNN}(\triangle, MN)$	$0.2158 \pm (0.1784, 82.66\%)$

Table 10.6.: Results of a parameter search to optimize CNN step size s_ϵ for low nonuniformities.

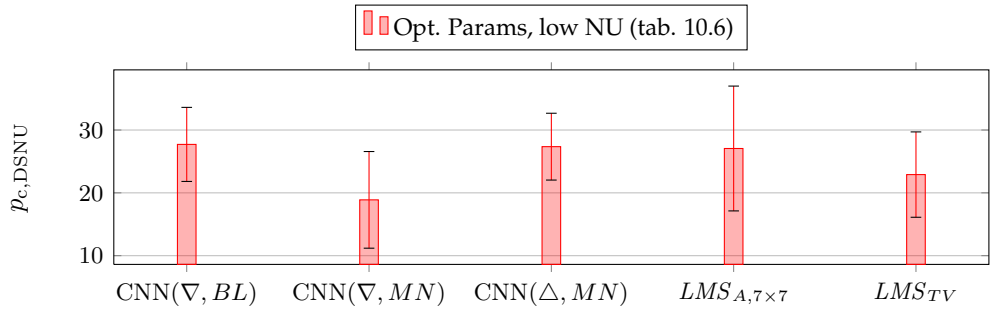


Figure 10.55.: CNN best-case Performance with respect to the $p_{\mathrm{c,DSNU}}$ quality metric for cameras with low nonuniformities

For the temporal evaluation the same set of experiments is shown as before. Fig. 10.56 and 10.57 show the temporal evaluation for the Basler $A602f$ camera for $t_{\mathrm{exp.}} = 1\mathrm{ms}\,@\,22°\mathrm{C}$ and $t_{\mathrm{exp.}} = 10\mathrm{ms}\,@\,50°\mathrm{C}$. At low temperatures the $\mathrm{CNN}(\nabla, BL)$ and $\mathrm{CNN}(\triangle, MN)$ method outperform the reference methods. Only the $\mathrm{CNN}(\nabla, MN)$ method stays below the reference methods, which was expected, as it did not perform well in the best-case performance analysis either. For the high temperatures, the new methods' parameters seem to mismatch, but they do not lead to negative correction rates.

Fig. 10.58 and 10.59 show the temporal evaluation for the MV-640-66-CL camera for $t_{\mathrm{exp.}} = 5\mathrm{ms}\,@\,25°\mathrm{C}$ and $t_{\mathrm{exp.}} = 10\mathrm{ms}\,@\,50°\mathrm{C}$ with the optimized parame-

ter set. At low temperatures the parameters fit well and the CNN(∇, BL) and CNN$(\triangle, MN)$ method perform best, while the CNN(∇, MN) reaches negative correction rates. The reference methods seem actually to deteriorate the image with their negative correction rates, as mentioned already in the earlier discussions. For the higher temperatures the CNN$(\triangle, MN)$ and CNN(∇, BL) outperform the reference methods, where the CNN$(\triangle, MN)$ is again the best method and the LMS_{TV} the best of the reference methods. The CNN(∇, MN) method again does not reach up to the reference methods, which may be founded in the not achieved correction of the spatial high frequencies discussed for this method.

The evaluations of the other experimental setups are shown in appendix G.3 and show a similar behaviour as discussed here, that mainly depends on the chosen sequence. However, some higher correction rates up to $40\,\%$ are reached in the shown plots.

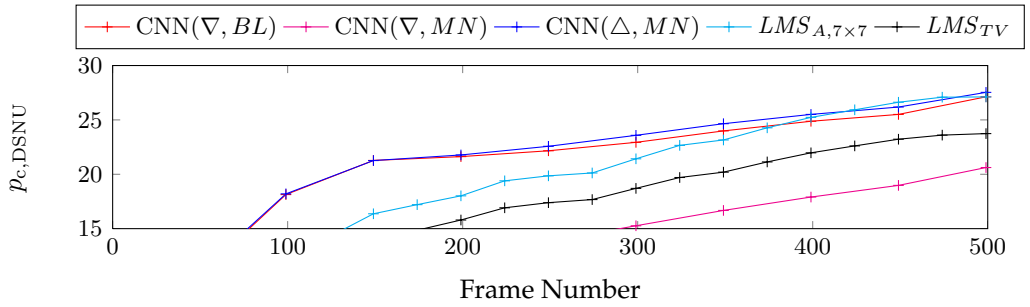


Figure 10.56.: Temporal evaluation for the Basler A602f with $t_{\text{exp.}}$=1ms @ 22°C

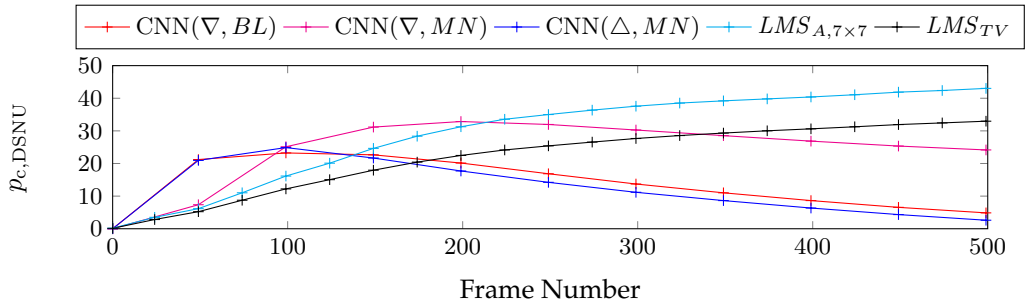


Figure 10.57.: Temporal evaluation for the Basler A602f with $t_{\text{exp.}}$=10ms @ 50°C

The impression of the visual correction is as always without any noticeable differences for the low valued nonuniformities (see fig. 10.60). In the remaining DSNU, the differences in the methods are visible. Compared to the above evaluations (see fig. 10.61), the CNN methods tend to leave a low frequency noise pattern with a higher amplitude behind. In the analysis of the spectral damping factors in fig. 10.62, the methods show as before a good damping that reaches also the spatial low frequency components. However, the reference methods seem to damp the spatial high frequencies better. As the results represent only one sequence one has to consider further experiments for all methods to find the best suitable method

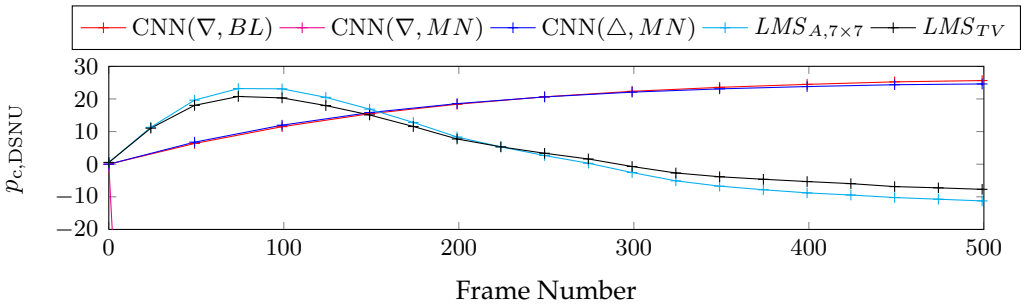


Figure 10.58.: Temporal evaluation for the MV-640-66-CL with $t_{\text{exp.}}$=5ms @ 25°C

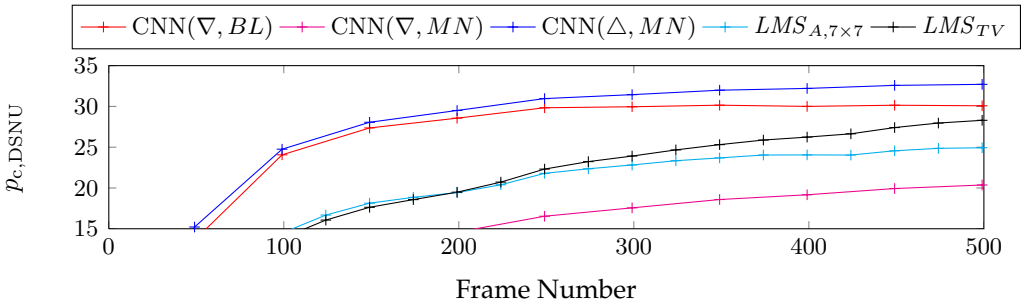
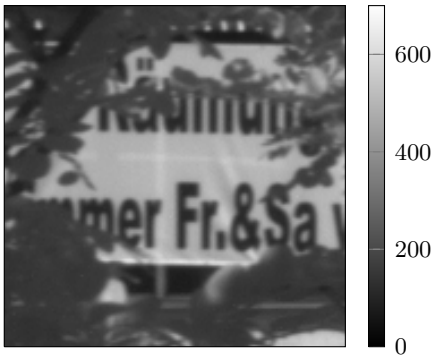
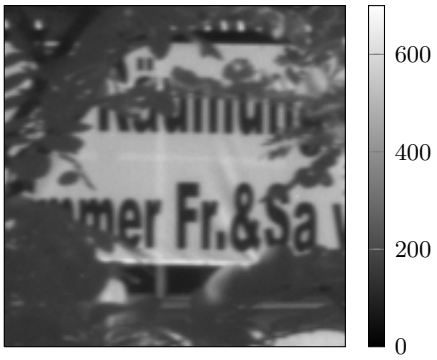


Figure 10.59.: Temporal evaluation for the MV-640-66-CL with $t_{\text{exp.}}$=10ms @ 50°C

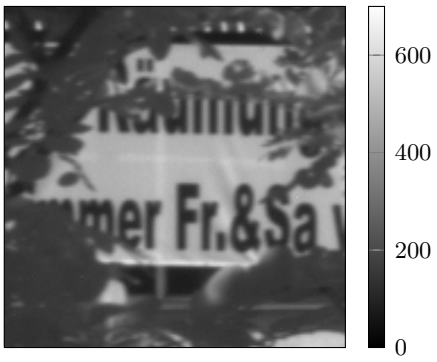
for the desired application. However, given the best-case scenarios, the CNN-methods outperform the reference methods considering the expectation values.



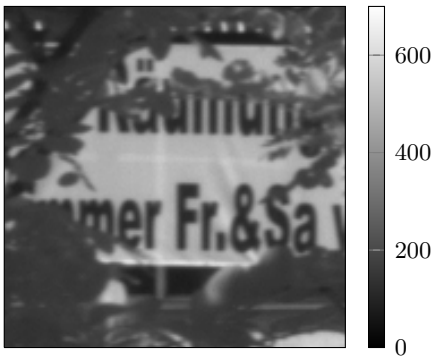
(a) Ground Truth Corrected



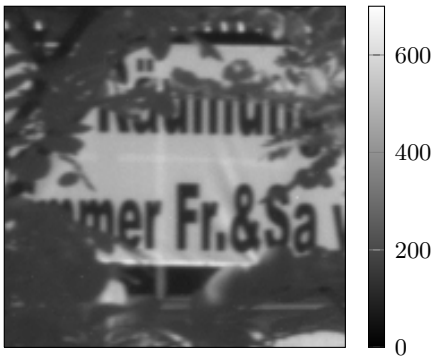
(b) Corrected with $\min(\nabla, BL)$ Method



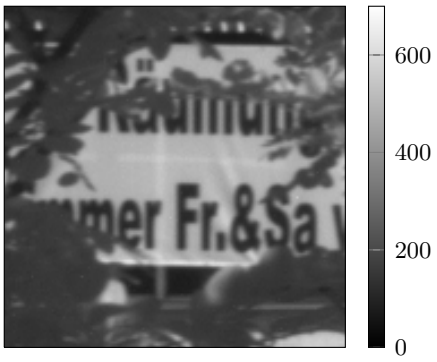
(c) Corrected with $\min(\nabla, MN)$ Method



(d) Corrected with $\min(\triangle, MN)$ Method

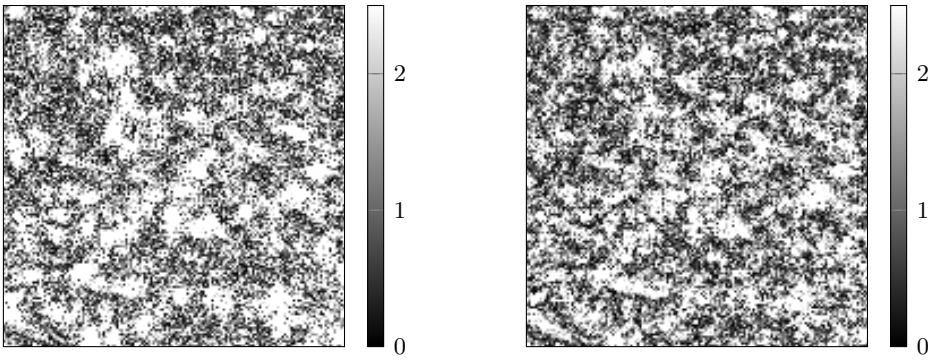


(e) Corrected with $LMS_{A,3\times3}$ Method



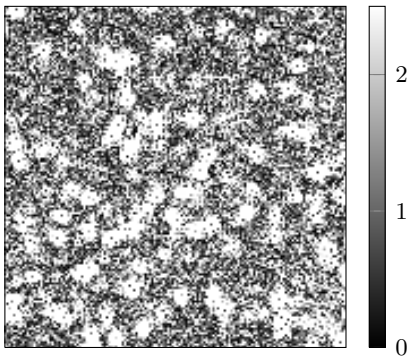
(f) Corrected with $LMS_{A,7\times7}$ Method

Figure 10.60.: Corrected images by the introduced methods for sequences with low valued nonuniformities (Parameter set according to tab. 10.6).

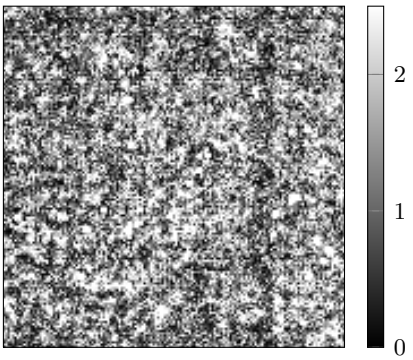


(a) $|\{b_{i,j,rem.}\}|$ for $\min(\nabla, BL)$

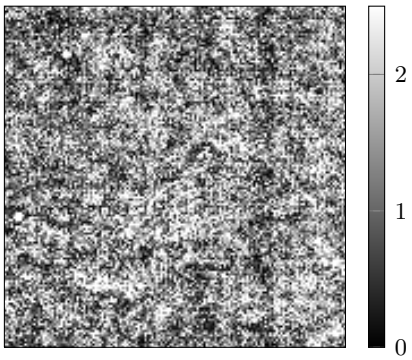
(b) $|\{b_{i,j,rem.}\}|$ for $\min(\nabla, MN)$



(c) $|\{b_{i,j,rem.}\}|$ for $\min(\triangle, MN)$

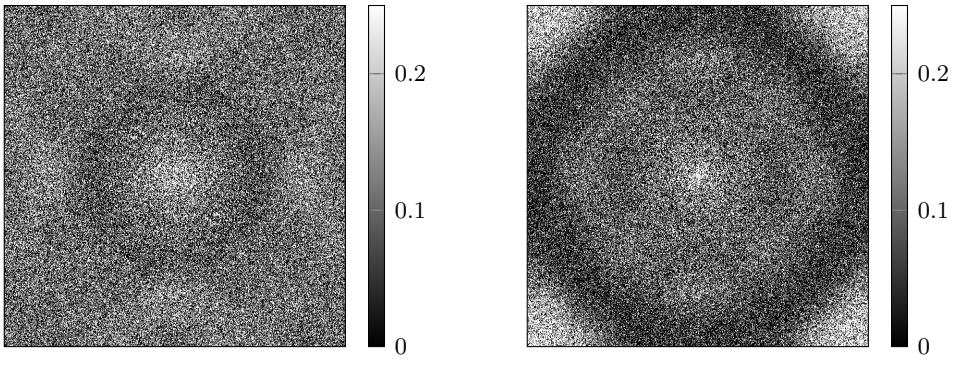


(d) $|\{b_{i,j,rem.}\}|$ for $LMS_{A,3\times3}$

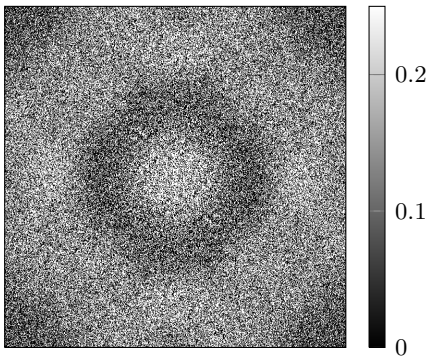


(e) $|\{b_{i,j,rem.}\}|$ for $LMS_{A,7\times7}$

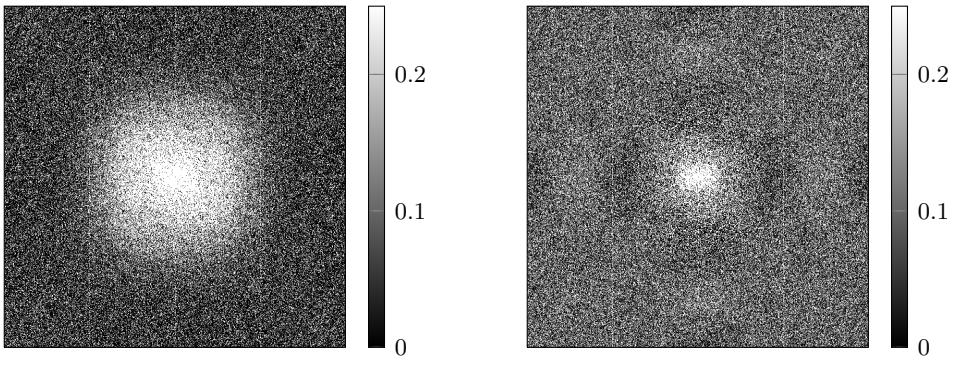
Figure 10.61.: Remaining DSNU patterns for the introduced methods and sequences with low valued nonuniformities (Parameter set according to tab. 10.6).



(a) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $\min(\nabla, BL)$ (b) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $\min(\nabla, MN)$



(c) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $\min(\triangle, MN)$



(d) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $LMS_{A,3\times3}$ (e) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $LMS_{A,7\times7}$

Figure 10.62.: Spectral damping coefficients for the DSNU for the introduced methods and sequences with low valued nonuniformities (Parameter set according to tab. 10.6).

10.2.5. Analysis of the Row-, Column- and Pixel-Wise Extensions

For the row-, column- and pixel-wise (RCP) combined methods, the parameter search yielded to the parameters shown in tab. 10.7. Fig. 10.63 shows the best-case performance analysis and besides large error bars, the $\min_{\mathrm{RCP}}(\nabla, BL)$ and the $\min_{\mathrm{RCP}}(\triangle, MN)$ methods achieve the best yet observed performances for low valued nonuniformities.

Method	Parameter
$\min_{\mathrm{RCP}}(\nabla, BL)$	$0.001299 \pm (0.001484, 114.3\%)$
$\min_{\mathrm{RCP}}(\nabla, MN)$	$0.001687 \pm (0.001039, 61.58\%)$
$\min_{\mathrm{RCP}}(\triangle, MN)$	$0.003864 \pm (0.006174, 159.8\%)$

Table 10.7.: Optimized parameter set for the RCP methods for low valued nonuniformities

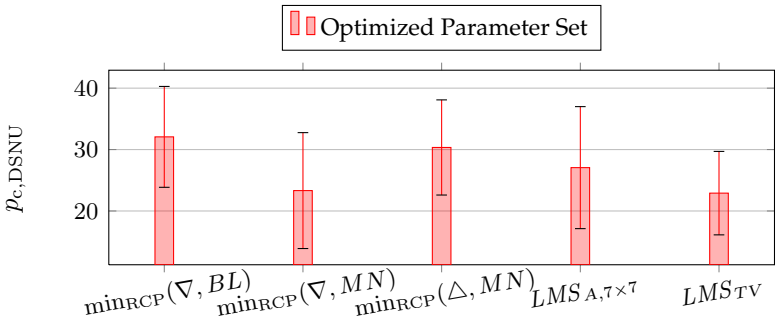


Figure 10.63.: RCP best-case performance with respect to the $p_{\mathrm{c,DSNU}}$ quality metric for cameras with low nonuniformities

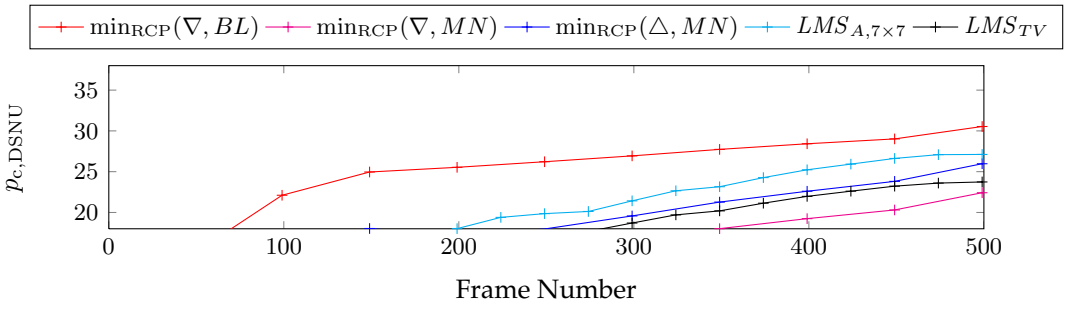


Figure 10.64.: Temporal evaluation for the Basler A602f with $t_{\mathrm{exp.}}$=1ms @ 22°C

The temporal evaluation of the usual set of experiments is shown in fig. 10.64 and 10.65 for the Basler $A602f$ camera for $t_{\mathrm{exp.}}$=1ms @ 22°C and $t_{\mathrm{exp.}}$=10ms @ 50°C. At low temperatures the $\min_{\mathrm{RCP}}(\nabla, BL)$ outperforms the reference methods, and the $\min_{\mathrm{RCP}}(\triangle, MN)$ reaches the reference methods in the very last frames. The $\min_{\mathrm{RCP}}(\nabla, MN)$ method does not gain such a good performance, as a result of the missed corner frequencies (see below). For the higher temperatures, the new

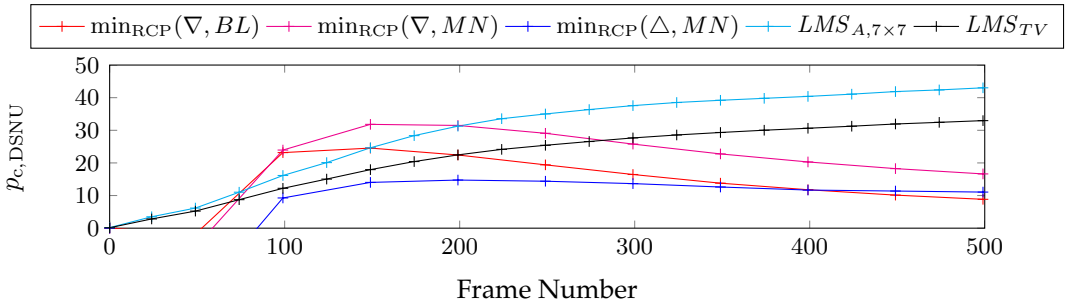


Figure 10.65.: Temporal evaluation for the Basler A602f with $t_{\mathrm{exp.}}$=10ms @ 50°C

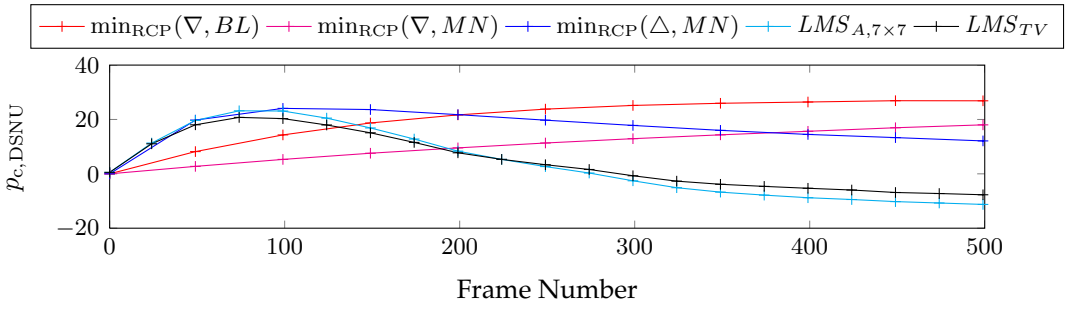


Figure 10.66.: Temporal evaluation for the MV-640-66-CL with $t_{\mathrm{exp.}}$=5ms @ 25°C

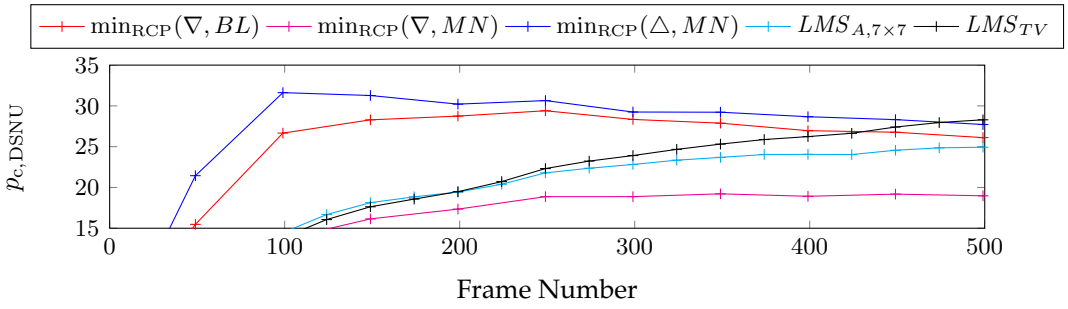


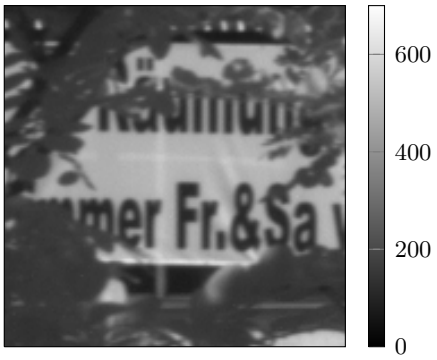
Figure 10.67.: Temporal evaluation for the MV-640-66-CL with $t_{\mathrm{exp.}}$=10ms @ 50°C

methods' parameters seem to be a mismatch as always in this comparison. However they do not degrade the image quality although they do not reach a good correction performance.

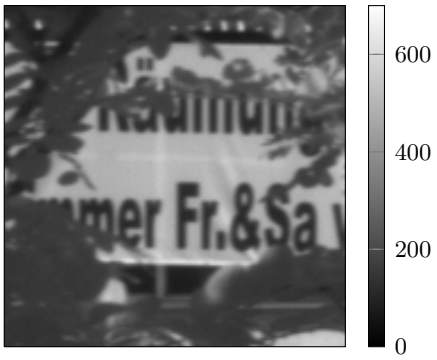
Fig. 10.66 and 10.67 show the temporal evaluation for the MV-640-66-CL camera for $t_{\text{exp.}}$=5ms @ 25°C and $t_{\text{exp.}}$=10ms @ 50°C with the optimized parameter set. At low temperatures the parameters fit well and the $\min_{\text{RCP}}(\nabla, BL)$ and $\min_{\text{RCP}}(\nabla, MN)$ methods perform best, while the $\min_{\text{RCP}}(\triangle, MN)$ method decreases slightly in its performance after the first 100 frames. The reference methods seem to deteriorate the image by creating negative correction rates, as already mentioned in the earlier discussions. For higher temperatures $\min_{\text{RCP}}(\triangle, MN)$ and $\min_{\text{RCP}}(\nabla, BL)$ outperform the reference methods again, but decrease in their performance after a few 100 frames and only the $\min_{\text{RCP}}(\triangle, MN)$ method stays at the same level as the reference methods after the complete 500 frames, which again confirms the $\hat{L} = \triangle$ choice as more stable than the $\hat{L} = \nabla$.

The evaluations of the other experimental setups are shown in appendix G.4 and show a similar behaviour as discussed here, that mainly depends on the chosen sequence. However, some higher correction rates up to $40\,\%$ are reached in the shown plots.

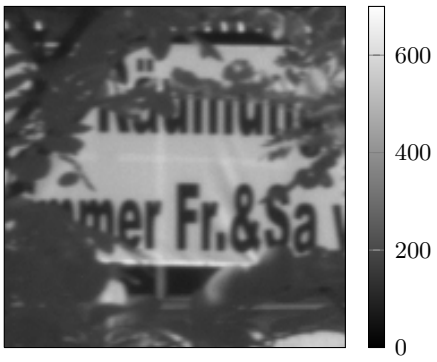
The impressions of the visual correction goes as before without any noticeable differences in the corrected images (see fig. 10.68). In the remaining DSNU, the differences become visible and the RCP methods tend to leave a low frequency noise pattern with a higher amplitude behind, compared to the above evaluations (see fig. 10.69). The main difference in the methods can be observed in the spectral analysis of fig. 10.70, where the RCP methods tend to leave more of the spatial low frequencies behind than their precedent developments. The other qualities stay the same and the reference method result again in a slightly better correction performance towards the spatial high frequency components. A distinct correction of vertical and horizontal nonuniformity structures can also not be observed. Still, the best case performance was superior to the reference methods.



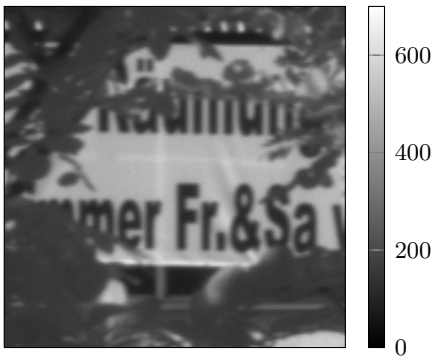
(a) Ground Truth Corrected



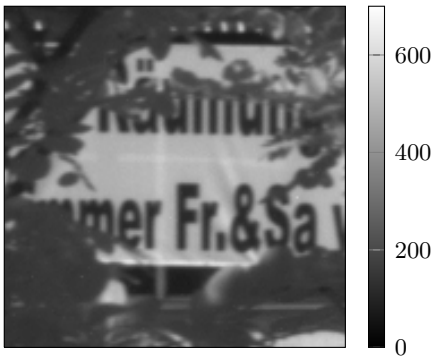
(b) Corrected with $\min(\nabla, BL)$ Method



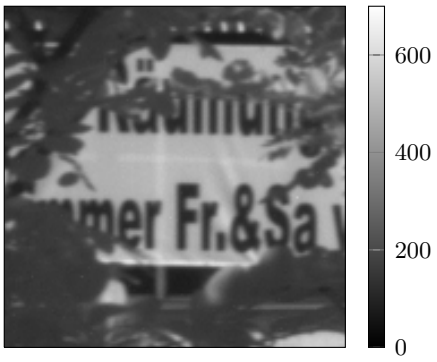
(c) Corrected with $\min(\nabla, MN)$ Method



(d) Corrected with $\min(\triangle, MN)$ Method

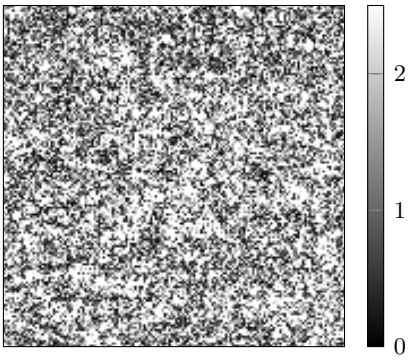


(e) Corrected with $LMS_{A,3\times3}$ Method

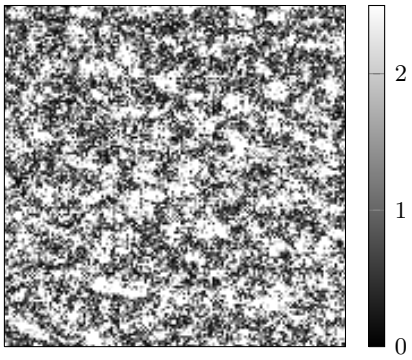


(f) Corrected with $LMS_{A,7\times7}$ Method

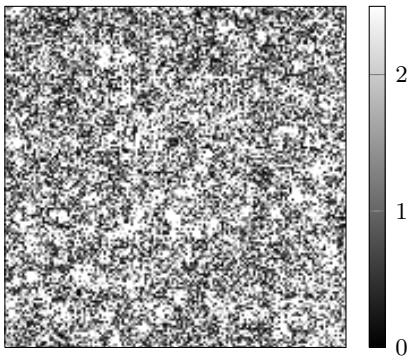
Figure 10.68.: Corrected images by the introduced methods for sequences with low valued nonuniformities (Parameter set according to tab. 10.7).



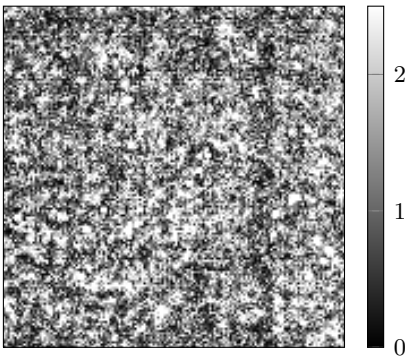
(a) $|\{b_{i,j,rem.}\}|$ for $\min(\nabla, BL)$



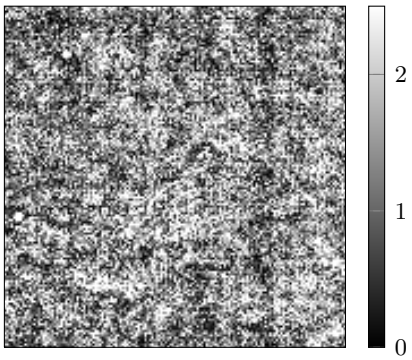
(b) $|\{b_{i,j,rem.}\}|$ for $\min(\nabla, MN)$



(c) $|\{b_{i,j,rem.}\}|$ for $\min(\triangle, MN)$

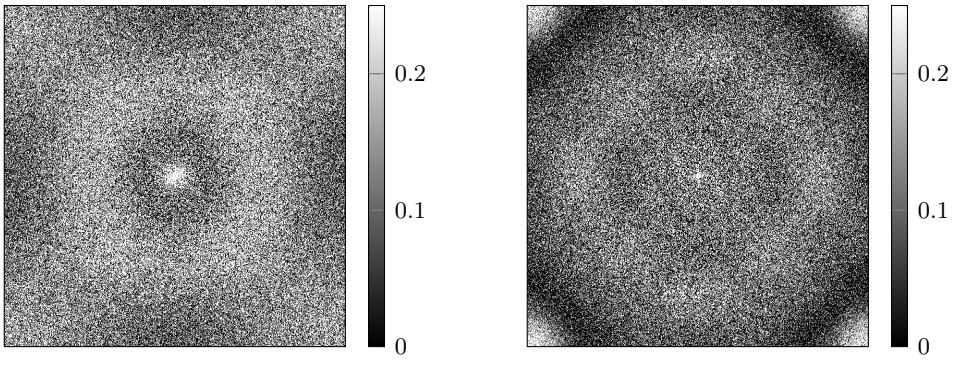


(d) $|\{b_{i,j,rem.}\}|$ for $LMS_{A,3\times3}$

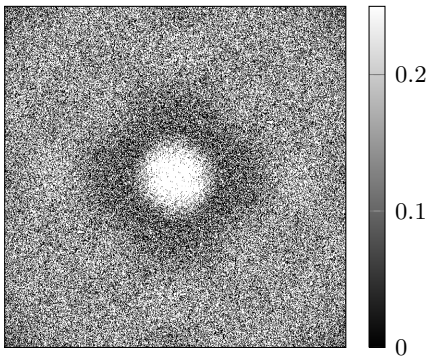


(e) $|\{b_{i,j,rem.}\}|$ for $LMS_{A,7\times7}$

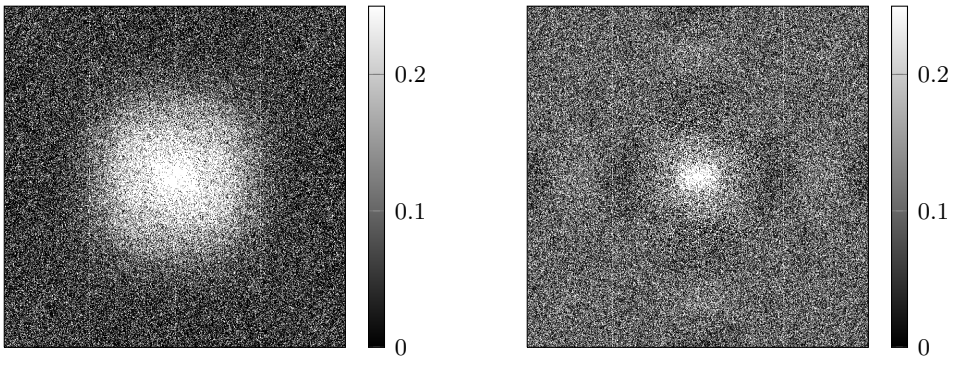
Figure 10.69.: Remaining DSNU patterns for the introduced methods and sequences with low valued nonuniformities (Parameter set according to tab. 10.7).



(a) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $\min(\nabla, BL)$ (b) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $\min(\nabla, MN)$



(c) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $\min(\triangle, MN)$



(d) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $LMS_{A,3\times3}$ (e) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $LMS_{A,7\times7}$

Figure 10.70.: Spectral damping coefficients for the DSNU for the introduced methods and sequences with low valued nonuniformities (Parameter set according to tab. 10.7).

10.2.6. Analysis of the Combined PRNU and DSNU Method

The results of the parameter search for the combined PRNU and DSNU methods for the low valued nonuniformity image material is shown in tab. 10.8. A good performance for other parameter sets than the optimized led to high valued negative correction rates. Such a result could be expected with the already discussed parameter instabilities of this method (see sec. 8). Fig. 10.71-10.73 show the best-case performance analysis in the measures $p_{\mathrm{c,Comb.}}, p_{\mathrm{c,PRNU}}$ and $p_{\mathrm{c,DSNU}}$. Given the $p_{\mathrm{c,Comb.}}$ all variants of the new method result in the same correction rates considering the error bars. For the $p_{\mathrm{c,DSNU}}$-norm the $LMS_{A,7\times7}$ method is however up to $10\,\%$ better, comparing the mean values, than the best of the new methods, the $\min_{\mathrm{comb.}}(\triangle, MN)$ method. The reason for this becomes obvious if one considers the best-case $p_{\mathrm{c,PRNU}}$ performance, where the reference methods average only to around $0\,\%$ while the new methods correct up to $5\,\%$ of the PRNU. Only the $\min_{\mathrm{comb.}}(\triangle, MN)$ method fails to correct and results in a correction performance around $0\,\%$ as well. Consequently the parameter set for the reference methods is adapted better to for DSNU correction, while the new methods do not neglect the PRNU correction.

In the above evaluations a sensitivity of the parameter choices was observed and the difference between the best-case performances and the temporal evaluations with fixed parameters is therefore expected to be large and not friendly towards the new methods. Especially with consideration of the discussion of the parameter instability for the combined estimation methods from sec. 8.4.

Method	Parameter
$\min_{\mathrm{comb.}}(\nabla, BL)$	$0.00781 \pm (0.01422, 182\%)$
$\min_{\mathrm{comb.}}(\nabla, MN)$	$0.005168 \pm (0.01108, 214.4\%)$
$\min_{\mathrm{comb.}}(\triangle, MN)$	$0.04669 \pm (0.1787, 382.7\%)$

Table 10.8.: Optimized Parameter set for the low nonuniformitiy sequences

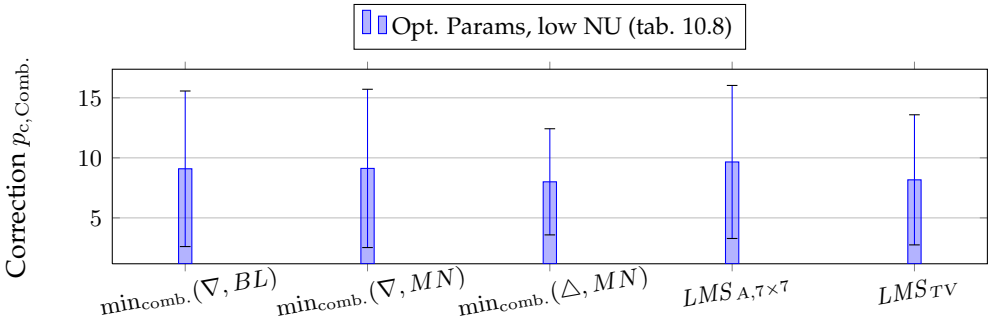


Figure 10.71.: Best-case performance with respect to the $p_{c,Comb.}$ quality for low valued nonuniformity sequences

The figures 10.74 -10.77 show the temporal evaluations for the fixed, optimized parameter set in the $p_{\mathrm{c,Comb.}}$-norm. As supposed above, the combined estima-

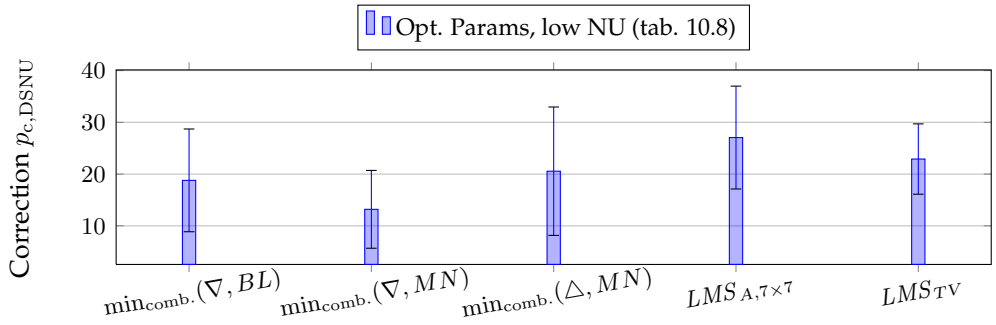


Figure 10.72.: Best-case Performance with respect to the $p_{c,DSNU}$ quality for low valued nonuniformity sequences

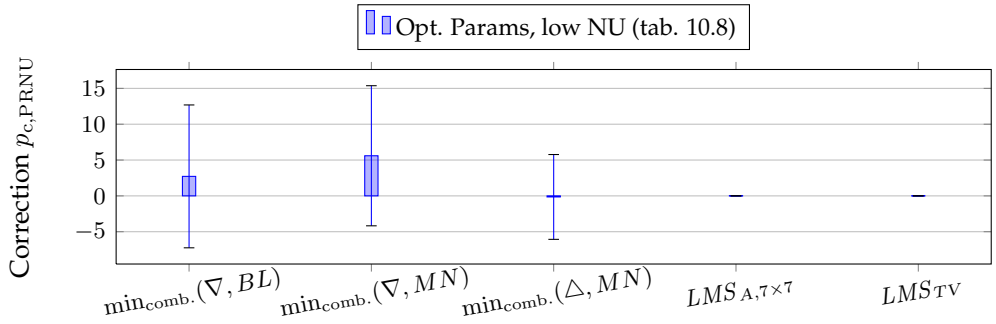


Figure 10.73.: Best-case Performance with respect to the $p_{c,PRNU}$ quality for low valued nonuniformity sequences

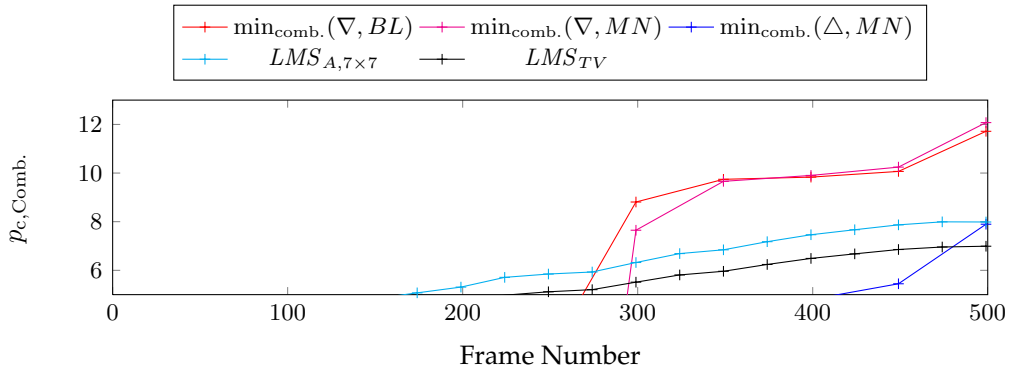


Figure 10.74.: Temporal evaluation for the Basler A602f with $t_{exp.}$=1ms @ 22°C

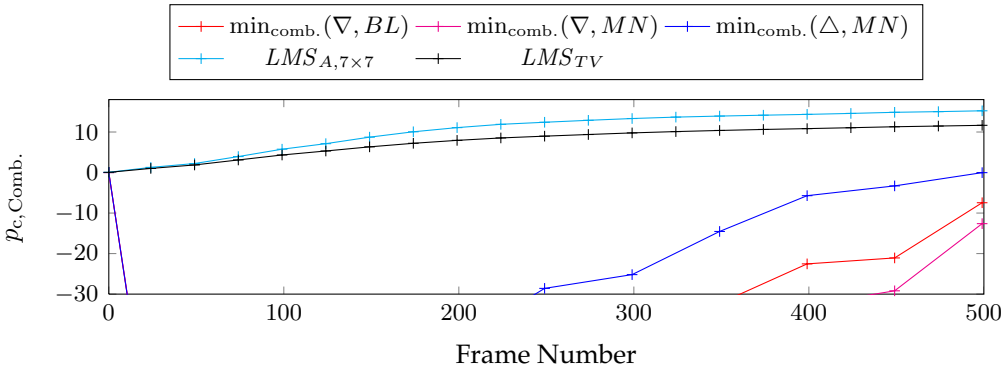


Figure 10.75.: Temporal evaluation for the Basler A602f with $t_{\text{exp.}}$=10ms @ 50°C

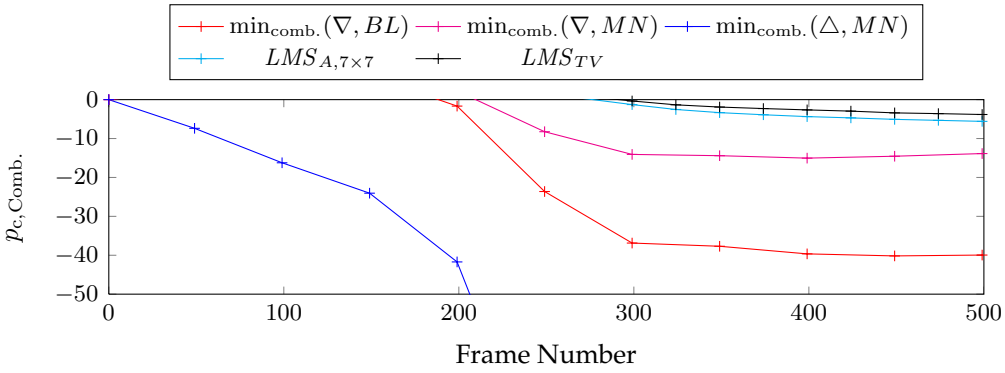


Figure 10.76.: Temporal evaluation for the MV-640-66-CL with $t_{\text{exp.}}$=5ms @ 25°C

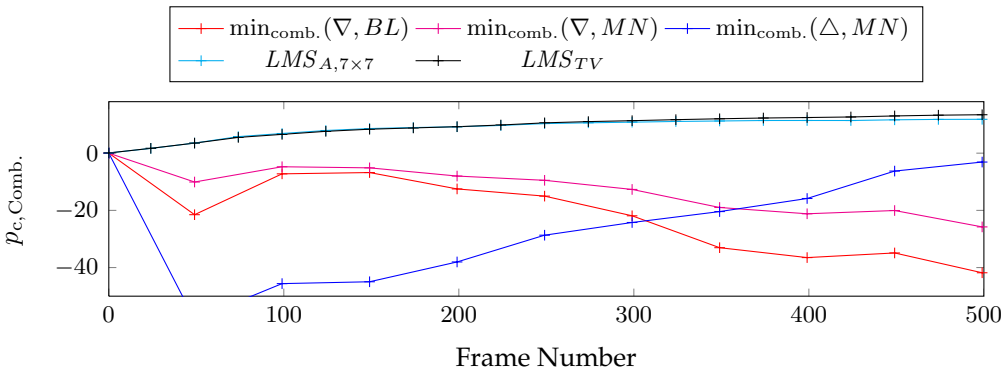


Figure 10.77.: Temporal evaluation for the MV-640-66-CL with $t_{\text{exp.}}$=10ms @ 50°C

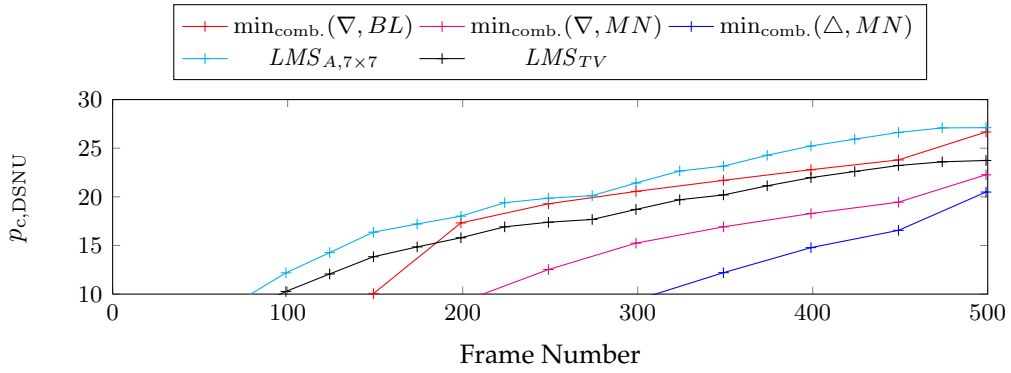


Figure 10.78.: Temporal evaluation for the Basler A602f with $t_{\text{exp.}}$=1ms @ 22°C

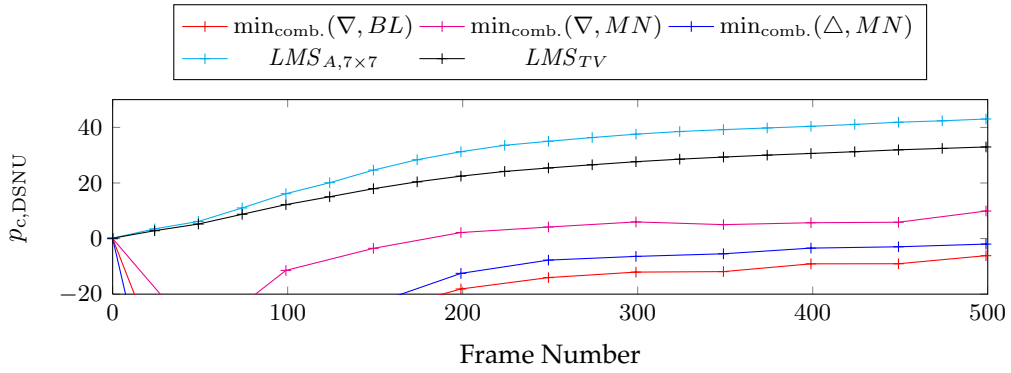


Figure 10.79.: Temporal evaluation for the Basler A602f with $t_{\text{exp.}}$=10ms @ 50°C

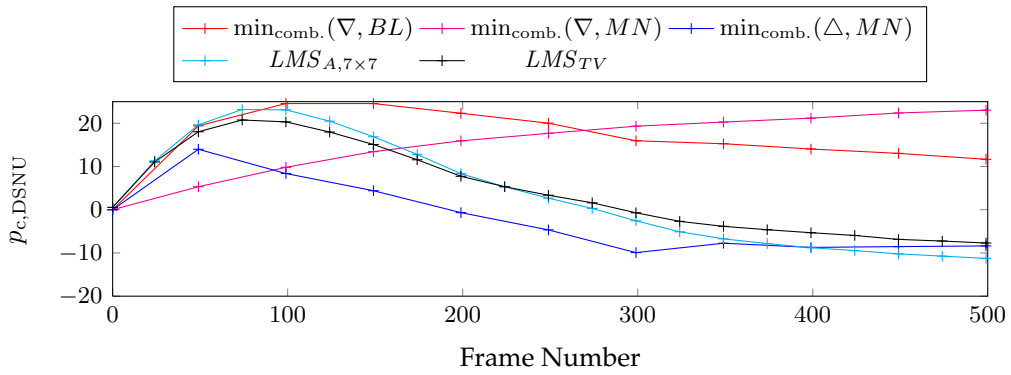


Figure 10.80.: Temporal evaluation for the MV-640-66-CL with $t_{\text{exp.}}$=5ms @ 25°C

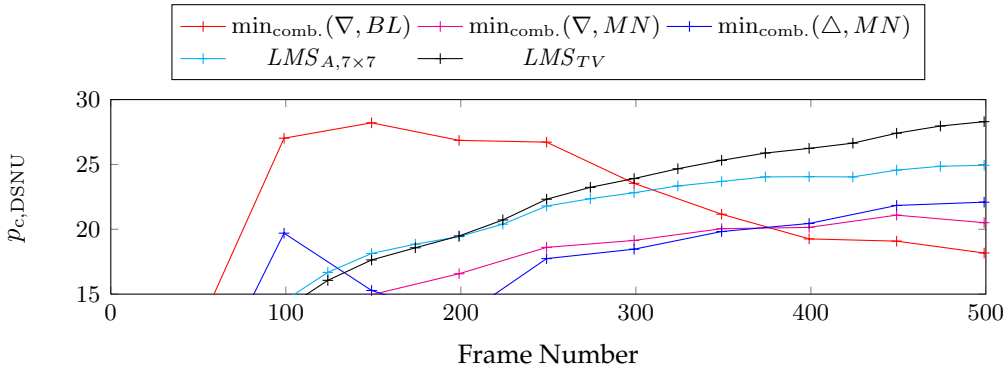


Figure 10.81.: Temporal evaluation for the MV-640-66-CL with $t_{\text{exp.}}$=10ms @ 50°C

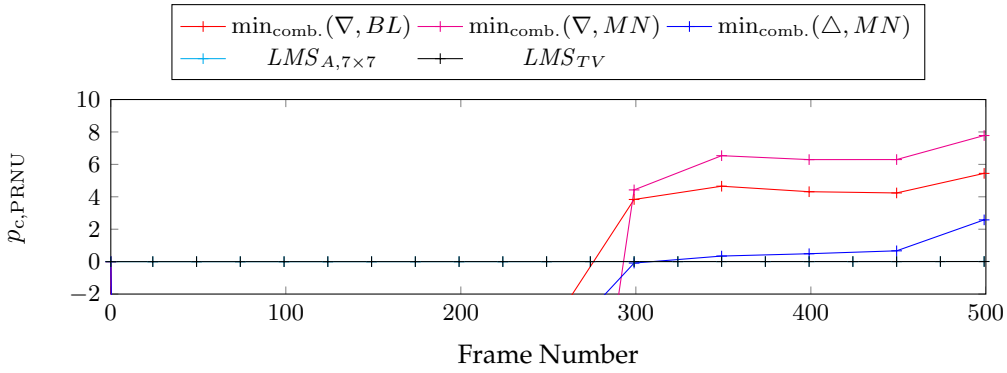


Figure 10.82.: Temporal evaluation for the Basler A602f with $t_{\text{exp.}}$=1ms @ 22°C

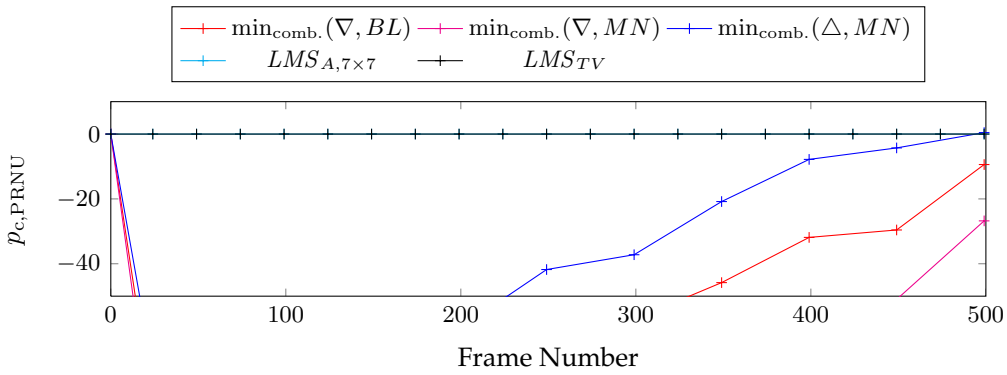


Figure 10.83.: Temporal evaluation for the Basler A602f with $t_{\text{exp.}}$=10ms @ 50°C

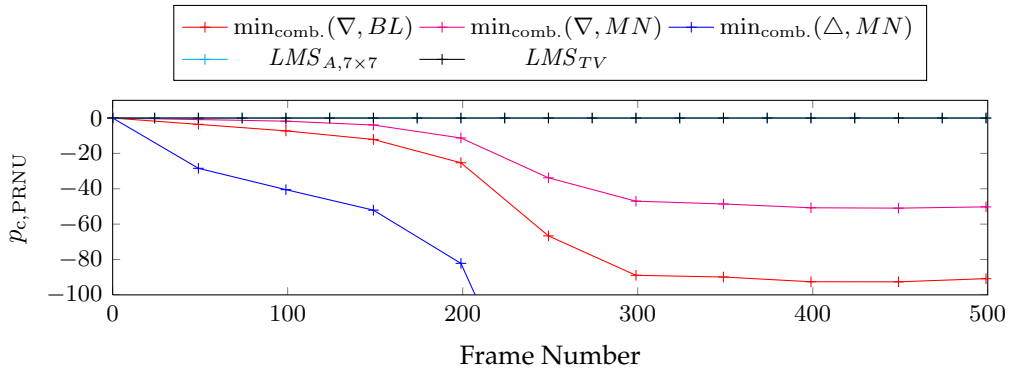


Figure 10.84.: Temporal evaluation for the MV-640-66-CL with $t_{\text{exp.}}$=5ms @ 25°C

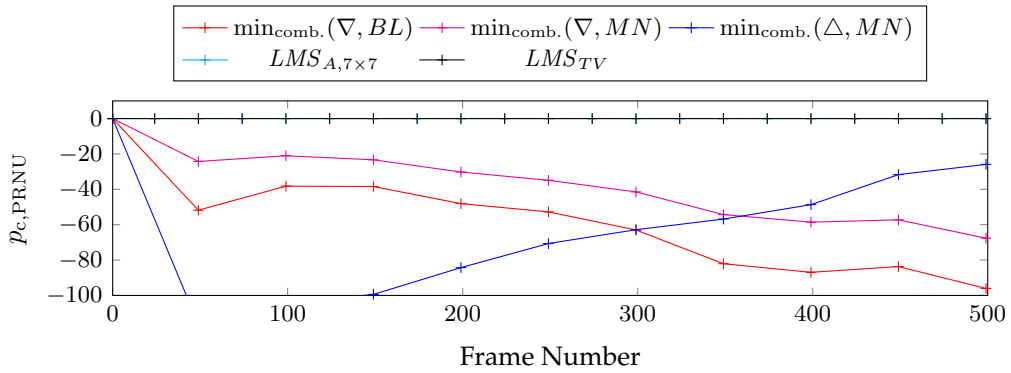


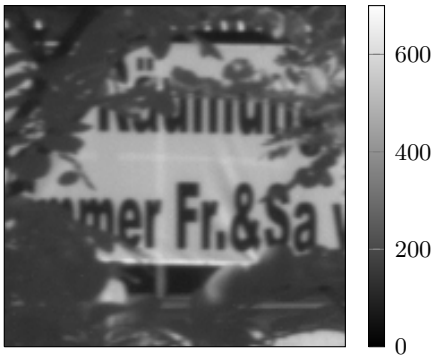
Figure 10.85.: Temporal evaluation for the MV-640-66-CL with $t_{\text{exp.}}$=10ms @ 50°C

tion methods become unstable towards the slight misalignment of the parameters that results from the fixed parameters. For the low temperature experiment of the Basler A602f sequences in fig. 10.74, the $\min_{\text{comb.}}(\nabla, BL)$ and $\min_{\text{comb.}}(\nabla, MN)$ methods outperform the reference methods after 300 frames and the $\min_{\text{comb.}}(\triangle, MN)$ method reaches just after 500 frames their performance. For all the other experiments however the new methods cannot compete with the reference methods and results in negative correction rates. The same is observed for the DSNU correction rates in fig. 10.78 -10.81 and for the PRNU corrections in fig. 10.82 -10.85.

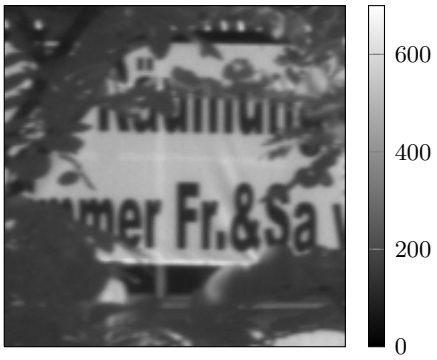
As mentioned before, these low correction performances are most likely the results of the inexactly estimated $\{a_{i,j}\}$ probability distribution as a Gaussian (see sec. 8.4.2). The best-case evaluations however showed that the new methods are able to reach the performance of the reference methods, but not with slightly inexact parameters, as shown in the recent plots. This parameter instability for the combined estimation method needs to be solved in future research. The further plots of the other temperature and exposure times are omitted as they do not provide further information.

Fig. 10.86 shows the visual correction performances and despite the bad correction parameters, no difference in the corrected images can be found. The remaining DSNU patterns are as expected and show a correction of spatial high frequencies and a slightly higher remaining amplitude when compared to the reference methods (see fig. 10.87). In the spectral damping analysis in fig. 10.88 the new methods confirm this statement with a generally good damping coefficient, and an improvement in the areas of spatial low frequencies when compared to the reference methods. However the spatial higher frequencies are corrected better by the reference methods.

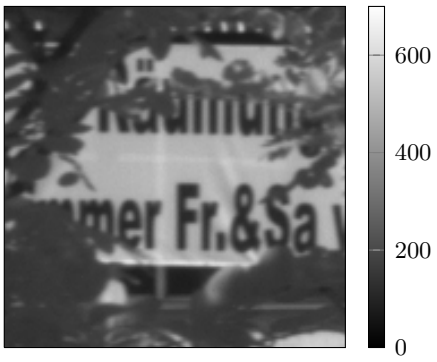
For the PRNU the remaining patterns are shown in fig. 10.89 and a visual difference to the reference methods is hardly visible. Fig. 10.90 finally shows the spectral damping analysis for the PRNU correction and here the new methods show surprisingly a superior correction performance over the reference methods. However, the spatial low frequency components are corrected in favor for spatial high frequency components which is not the usual mode of correction of these methods. Although this result is an isolated case, it shows the possibilities of the new method in cases of correctly selected parameters.



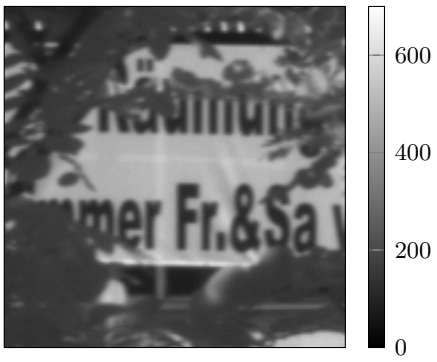
(a) Ground Truth Corrected



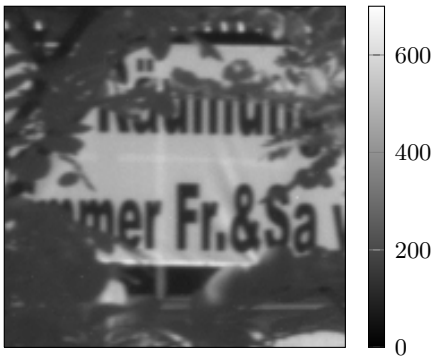
(b) Corrected with $\min(\nabla, BL)$ Method



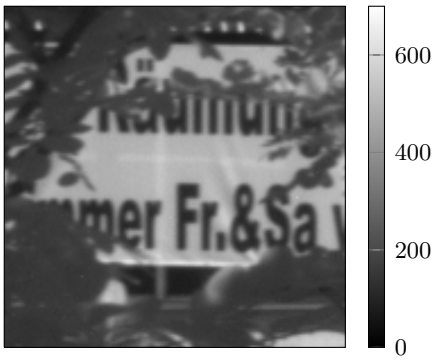
(c) Corrected with $\min(\nabla, MN)$ Method



(d) Corrected with $\min(\triangle, MN)$ Method

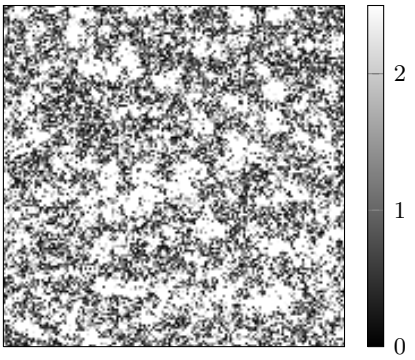


(e) Corrected with $LMS_{A,3\times3}$ Method

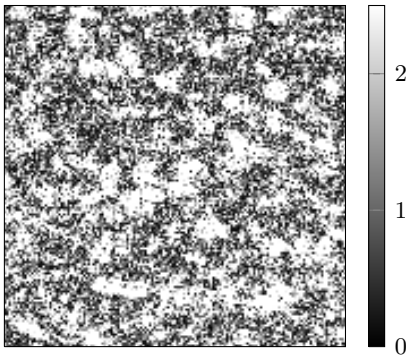


(f) Corrected with $LMS_{A,7\times7}$ Method

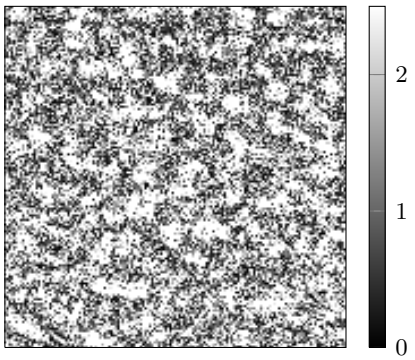
Figure 10.86.: Corrected images by the introduced methods for sequences with low valued nonuniformities (Parameter set according to tab. 10.8).



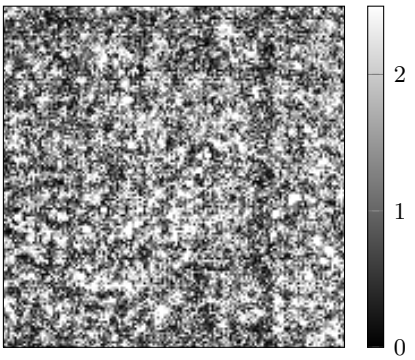
(a) $|\{b_{i,j,rem.}\}|$ for $\min(\nabla, BL)$



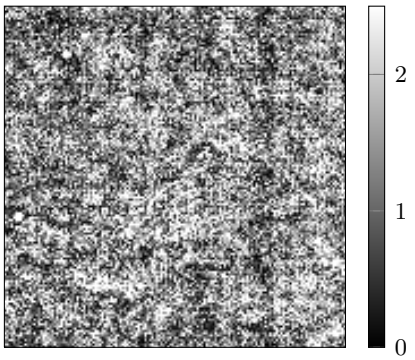
(b) $|\{b_{i,j,rem.}\}|$ for $\min(\nabla, MN)$



(c) $|\{b_{i,j,rem.}\}|$ for $\min(\triangle, MN)$

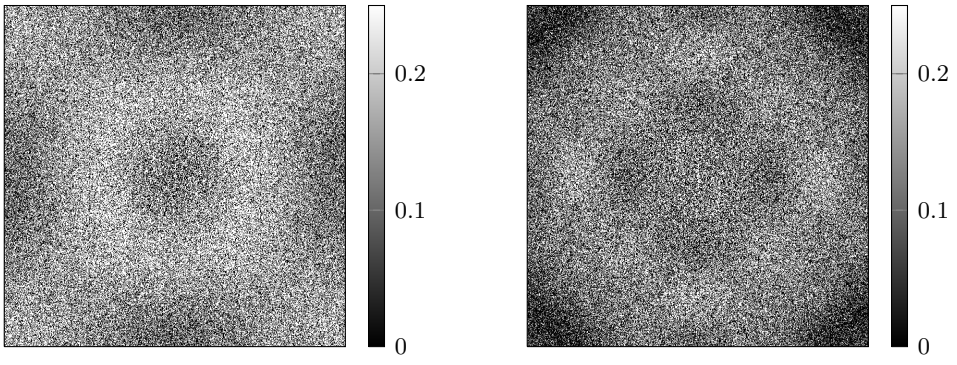


(d) $|\{b_{i,j,rem.}\}|$ for $LMS_{A,3\times3}$

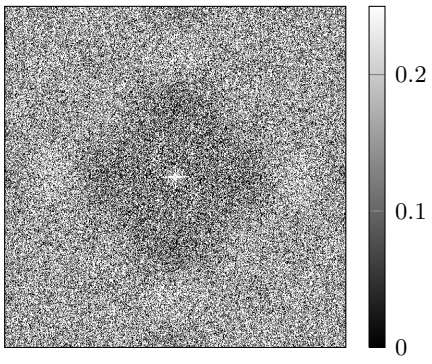


(e) $|\{b_{i,j,rem.}\}|$ for $LMS_{A,7\times7}$

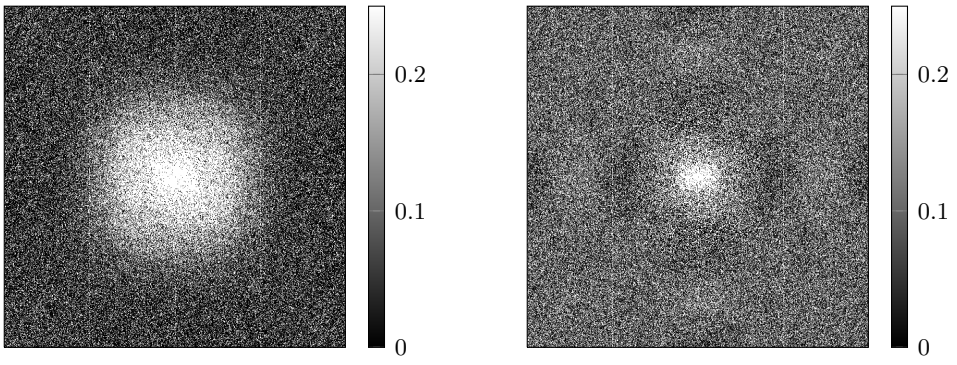
Figure 10.87.: Remaining DSNU patterns for the introduced methods and sequences with low valued nonuniformities(Parameter set according to tab. 10.8).



(a) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $\min(\nabla, BL)$ (b) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $\min(\nabla, MN)$

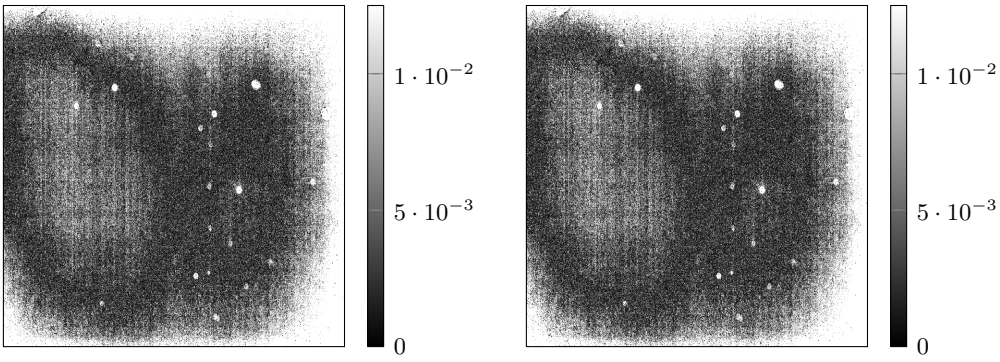


(c) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $\min(\triangle, MN)$

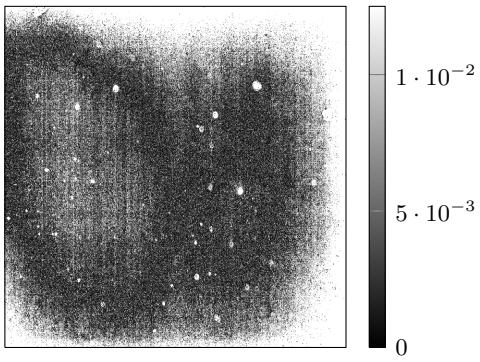


(d) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $LMS_{A,3\times3}$ (e) $\hat{s}_{\text{spectr.,DSNU},\mu,\nu}$ for $LMS_{A,7\times7}$

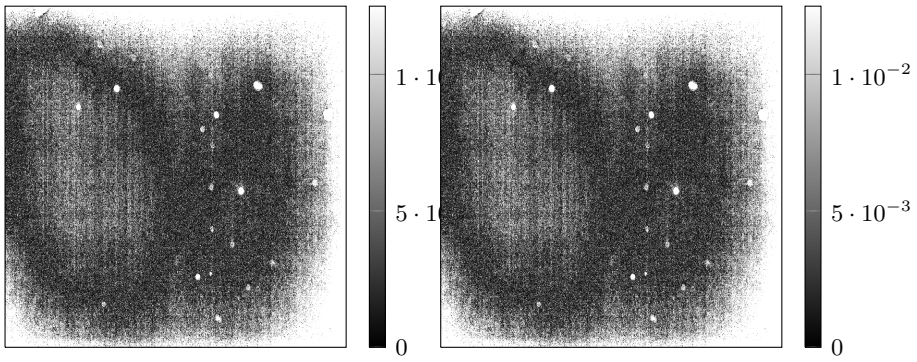
Figure 10.88.: Spectral damping coefficients for the DSNU for the introduced methods and sequences with low valued nonuniformities (Parameter set according to tab. 10.8).



(a) $|1 - \{a_{i,j,rem.}\}|$ for $\min(\nabla, BL)$ (b) $|1 - \{a_{i,j,rem.}\}|$ for $\min(\nabla, MN)$

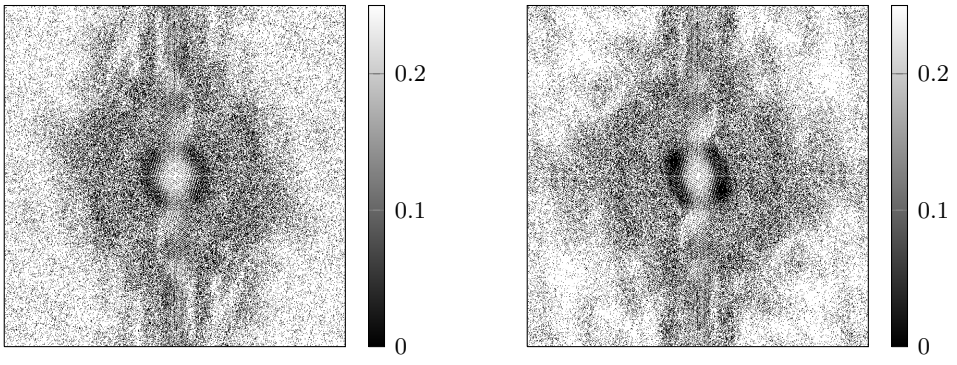


(c) $|1 - \{a_{i,j,rem.}\}|$ for $\min(\triangle, MN)$

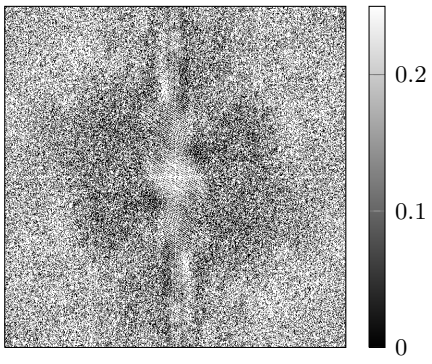


(d) $|1 - \{a_{i,j,rem.}\}|$ for $LMS_{A,3\times3}$ (e) $|1 - \{a_{i,j,rem.}\}|$ for $LMS_{A,7\times7}$

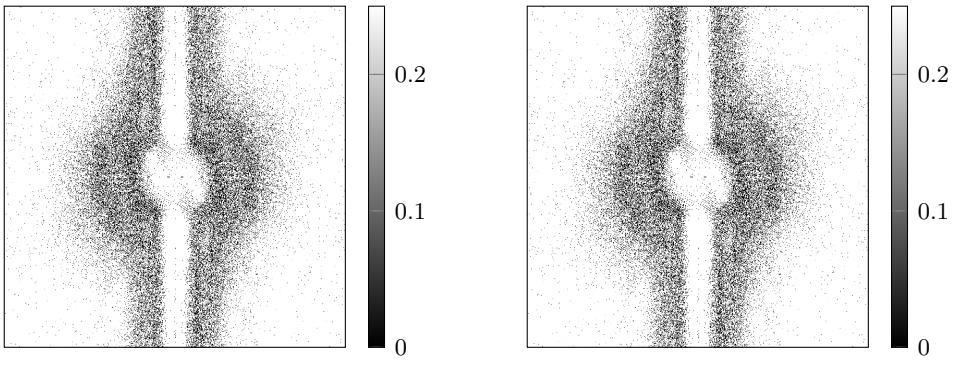
Figure 10.89.: Remaining PRNU patterns for the introduced methods and sequences with low valued nonuniformities (Parameter set according to tab. 10.8).



(a) $\hat{s}_{\text{spectr.,PRNU},\mu,\nu}$ for $\min(\nabla, BL)$ (b) $\hat{s}_{\text{spectr.,PRNU},\mu,\nu}$ for $\min(\nabla, MN)$



(c) $\hat{s}_{\text{spectr.,PRNU},\mu,\nu}$ for $\min(\triangle, MN)$



(d) $\hat{s}_{\text{spectr.,PRNU},\mu,\nu}$ for $LMS_{A,3\times3}$ (e) $\hat{s}_{\text{spectr.,PRNU},\mu,\nu}$ for $LMS_{A,7\times7}$

Figure 10.90.: Spectral damping coefficients for the PRNU for the introduced methods and sequences with low valued nonuniformities (Parameter set according to tab. 10.8).

10.2.7. Intermediate Conclusions

The correction performance for low valued nonuniformity input is very sensitive to slight parameter changes and does not reach the high numbers of up to $70\,\%$ as for the high valued nonuniformities. This is however not necessary, as only $1\,\mathrm{DN}\,\mathrm{to}\,2\,\mathrm{DN}$ of visible noise is present in the images. Because of the discussed sensitivity to parameter changes, a parameter search to find the optimal parameters is recommended and needs to be verified with respect to its stability. As the available parameter set is a mix of different camera types, exposure times and temperatures, the presented results may be improved if more specific material for the intended application is available.

The above section can therefore be seen as a proof that the new methods are able to correct low valued nonuniformities given the different setups. Especially the best-case analysis confirmed this.

11. Outlook and further Work

11.1. Summary

This thesis introduced a new method for a scene based maximum likelihood estimation for the digital image sensors' nonuniformity components PRNU and DSNU. This method was first developed and improved in a DSNU-only correction mode and later extended for a combined PRNU and DSNU correction. A detailed review of the reference methods showed their potential for further improvements in the fields of their theoretical assumptions and their way on choosing the free parameters. Also the correction performance of the methods showed room for further improvements, especially with concern to the achieved PRNU correction. The investigation further showed that a good correction performance for the reference methods could only be achieved with an optimized parameter set for the given type of input sequences.

With concern to the performance measure for the nonuniformity corrections, the published measures in literature depend on the actual image data, which has certain discussed drawbacks. A review of the physical background of the digital imaging process, combined with the EMVA1288 definition for the nonuniformity quantification, allowed to define a new, image data independent performance measure. This new measure has the further advantage to regard the DSNU and PRNU effects separately.

The new maximum likelihood estimation method is then derived by using simple and well-motivated assumptions. The main assumption is that the DSNU and PRNU result in spurious edge and gradient like features that can be detected with a linear high pass operator. In the further development this leads to a demand for temporally independent frames and a method input that, in the DSNU-only variant, consists of the temporally averaged sensor responses. All other parameters of the method could be extracted from the statistics of the recorded scenes and the ground truth EMVA1288 calibration of the image sensor. This makes expensive and partly unreliable parameter searches not necessary any more. The method further includes a convolution of the averaged measurements with a high pass filter, which is directly derived from the used assumptions and the signal reconstruction method. In the first approach, the method did not reach the performance of the best reference methods. However, an improvement that uses a motivated weighted average of the sensor inputs, based on the EMVA1288 measurement standards, yielded to a superior performance compared to all the reference methods.

However, the superior performance of the maximum likelihood estimation method is achieved on costs of a high computational effort. This drawback could be overcome by translating the method's mathematics into a cellular neural network (CNN) of the type introduced by Leon Chua. The non-linearity of the network allowed to keep the resulting correction parameters in useful boundaries, leading to an overall performance increase of the corresponding results. As most practical solutions suffer from additional row- and column-wise nonuniformities, the new approach was improved with an advanced sensor model that includes row- and column-wise nonuniformities. This allowed to choose if the correction should only consider only row-, column-, or pixel-wise nonuniformities, or any combination of those. For row- or column-wise only corrections, the DSNU maximum likelihood approach could be simplified from the two dimensional array of coefficients into a problem that has to consider only one dimension of the nonuniformity parameters. This reduced the computational demands drastically and prepared the method for implementations on embedded systems.

Next, the yet DSNU-only approach is extended into a combined PRNU and DSNU approach, which again showed a superior correction performance over the reference methods, if random frames are used as inputs. At this point the statically derived parameter set did however not yield to a reasonable correction performance and an optimized parameter set had to be used. As additional input for this extended variant, the sensor-wise covariances are needed.

Finally the new approach could be linked mathematically to the reference methods and the reference methods can be interpreted as simplified versions of the introduced maximum likelihood approach. This enabled a deeper understanding of the reference methods' implicitly made assumptions. Last the experimental section is extended to high speed consecutive frame sequences that violate the demands of independent frames, and to image data that shows only low valued nonuniformities. Both cases show the expected limits of the maximum likelihood approach and the best reference methods gain slight advantages or comparable correction rates. Given the discovered link between the methods, the reasons for the observed different behavior of the methods can be understood and explained by their implicit assumptions.

11.2. Conclusions

The newly introduced variants of the maximum likelihood estimation method showed to be an extension to both, the constant statistics and the least mean square (LMS) method. This can be seen as the constant statistic method utilizes the temporal average of the sensor responses, while the LMS methods consider the spatial dependencies of the sensor responses. The new method however uses both types of information.

Further, three variants of the method have been investigated, where the first two depend on gradient-like features while the third variant bases on edge-like fea-

tures. The gradient based variants are in general able to achieve higher performances, even with a smaller 3×3 neighborhood, the 7×7-mask based reference methods are outperformed. The edge-based variant however showed a great stability towards changes in the parameters and in cases of consecutively dependent frames as input. With the presented adaptations for row- and column-wise corrections the typical use cases in industrial applications have been considered as well. The new method gives best results for random frame input.

11.3. Further Work

The further work for this thesis has already been pointed out at the end of the corresponding chapters . To summarize, one possibility to improve the quality of the corrections could be the analysis of other features in extend to the Nabla- and Laplace-based ones. Furthermore, other reconstruction functions than the introduced ones may then become necessary as well, which may result in improved high pass filter masks. Another idea leads towards the installed temporal averaging filter, where further improvements may allow a better selection of input information that is considered *independent*. This may allow to transform the methods' advantages onto the consecutive frame input as well. However, for longer consecutive frame sequences the majority of frames will be independent to each other and a good parameter estimation should be possible again.

A relevant point of the new method is the observed parameter instability for the combined estimation of the DSNU and PRNU. Here the statistically derived parameters did not yield to satisfying results and as possible reasons a misfit between the actual parameter distributions and the assumed ones have already been discussed. A better parametrization of the assumed probability distributions, especially for the PRNU, looks like a promising approach.

Up to now the parameters of the methods are estimated from ground truth corrected images that obey the same statistics as the evaluation sequences. For the DSNU estimations these statistically derived parameters give already a performance that is superior to the reference methods and a further improvement could try to build-up these statistics while the algorithm is running. As the influence of the nonuniformities may be neglectable for the evaluations of the chosen features, such an improvement sounds feasible. Such a physically and statistically well motivated and parameter free method has of course a big advantage for practical applications.

With concern to the practical applications, the row- and column-wise simplifications of the method could be implemented directly due to their low computational demands. For a hardware implementation, a first approach could be the usage of the CNN structure which translates directly into analogue circuity. Therefore new generations of digital sensors could be equipped with a hardware realization of the described algorithm.

Bibliography

[1] EMVA Commitee, "EMVA1288: A standard for Characterization of Image Sensors and Cameras," Dec. 2010, release 3.0, http://www.emva.org.

[2] J. Pezoa and O. Medina, "Spectral model for fixed-pattern-noise in infrared focal-plane arrays," in *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*, ser. Lecture Notes in Computer Science, C. San Martin and S.-W. Kim, Eds. Springer Berlin Heidelberg, 2011, vol. 7042, pp. 55–63. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-25085-9_6

[3] D. Scribner, K. Sarkady, M. Kruer, J. Caulfield, J. Hunt, M. Colbert, and M. Descour, "Adaptive retina-like preprocessing for imaging detector arrays," in *Neural Networks, 1993., IEEE International Conference on*, 1993, pp. 1955 –1960 vol.3.

[4] E. Vera and S. Torres, "Fast adaptive nonuniformity correction for infrared focal-plane array detectors," *EURASIP J. Appl. Signal Process.*, vol. 2005, pp. 1994–2004, Jan. 2005.

[5] R. C. Hardie, F. Baxley, B. Brys, and P. Hytla, "Scene-based nonuniformity correction with reduced ghosting using a gated LMS algorithm," *Opt. Express*, vol. 17, no. 17, pp. 14 918–14 933, Aug 2009.

[6] E. Vera, P. Meza, and S. Torres, "Total variation approach for adaptive nonuniformity correction in focal-plane arrays," *Opt. Lett.*, vol. 36, no. 2, pp. 172–174, Jan 2011.

[7] J. Harris and Y.-M. Chiang, "Nonuniformity correction of infrared image sequences using the constant-statistics constraint," *IEEE Trans. Image Process.*, vol. 8, no. 8, pp. 1148 –1151, Aug 1999.

[8] S. N. Torres and M. M. Hayat, "Kalman filtering for adaptive nonuniformity correction in infrared focal-plane arrays," *J. Opt. Soc. Am. A*, vol. 20, no. 3, pp. 470–480, Mar 2003. [Online]. Available: http://josaa.osa.org/abstract.cfm?URI=josaa-20-3-470

[9] H.-X. Zhou, H. lin Qin, Y. bo Jian, B. jian Wang, and S. qian Liu, "Improved kalman-filter nonuniformity correction algorithm for infrared focal plane arrays," *Infrared Physics and Technology*, vol. 51, no. 6, pp. 528 – 531, 2008. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1350449508000388

[10] F. León and M. H. [Hrsg.], *Forum Bildverarbeitung 2012*. KIT Scientific Publishing, Karlsruhe, 2012.

[11] M. Geese, P. Ruhnau, and B. Jähne, "Cnn based dark signal non-uniformity estimation," *CNNA, Turin*, 2012.

[12] ——, "A Globally Optimal Scene Based Maximum Likelihood DSNU Estimation Method," *Journal of Electronic Imaging*, no. submitted for review, 2012.

[13] ——, "A Scene Based Combined PRNU and DSNU Estimation Method," *Measurement Science and Technology*, no. submitted for review, 2012.

[14] M. Geese and P. Ruhnau, "Verfahren und Vorrichtung zur Abschätzung eines Fliegengittereffekts einer Bilderfassungseinrichtung," Robert Bosch GmbH, 2010, Patent Application at Deutsches Patent und Markenamt; Publication follows: DE 102010061864 A1.

[15] D. Meschede and C. Gerthsen, *Gerthsen Physik: Dieter Meschede*, ser. Springer-Lehrbuch Series. Springer-Verlag, 2003.

[16] B. Jähne, *Digital Image Processing: Concepts, Algorithms and Scientific Applications*, 6th ed. Berlin: Springer, 2005.

[17] R. Müller, *Grundlagen der Halbleiter-Elektronik*, ser. Halbleiter-Elektronik. Springer, 1995.

[18] E. Hering and G. Schönfelder, *Sensoren in Wissenschaft und Technik: Funktionsweise und Einsatzgebiete*. Vieweg+Teubner Verlag, 2012.

[19] F. Li and A. Nathan, *CCD Image Sensors in Deep-Ultraviolet: Degradation Behavior and Damage Mechanisms*, ser. Microtechnology and MEMS. Springer, 2005.

[20] R. Korde and J. Geist, "Quantum efficiency stability of silicon photodiodes," *Appl. Opt.*, vol. 26, no. 24, pp. 5284–5290, Dec 1987.

[21] I. Daniele Passeri, Univeristy of Perugia, "Complementary metal oxide semiconductor active pixel sensor chip layout„" http://www.picoquant.com/application_gallery/.

[22] A. E. Gamal, B. A. Fowler, H. Min, , and X. Liu, "Modeling and estimation of FPN components in CMOS image sensors," *SPIE Solid State Sensor Arrays: Development and Application II*, vol. 3301, pp. 168–177, 1998.

[23] F. Inc., "X3 technology for direct image sensors," http://www.foveon.com/article.php?a=67.

[24] D. Wolf, *Signaltheorie: Modelle Und Strukturen*. Springer, 1999.

[25] J. Goodman, *Introduction To Fourier Optics*, ser. McGraw-Hill physical and quantum electronics series. Roberts & Company, 2005.

[26] D. P. Mitchell and A. N. Netravali, "Reconstruction filters in computer graphics," *Computer Graphics*, vol. 22, no. 4, pp. 221–228, 1988.

[27] G. Hopkinson, T. Goodman, and S. Prince, *A Guide To The Use And Calibration Of Detector Array Equipment*, ser. Press Monographs. SPIE Press, 2004.

[28] M. Schoberl, C. Senel, S. Fossel, H. Bloss, and A. Kaup, "Non-linear dark current fixed pattern noise compensation for variable frame rate moving," *EUSIPCO*, 2009.

[29] H. Rinne, *Tschenbuch der Statistik*, 4th ed. Frankfurt am Main: Verlag Harri Deutsch, 2008.

[30] G. Healey and R. Kondepudy, "Radiometric CCD camera calibration and noise estimation," *IEEE Trans. of PAMI*, vol. 16, no. 3, pp. 267–276, 1994.

[31] R. Gonzalez and R. Woods, *Digital image processing*, 3rd ed. Prentice Hall, 2008.

[32] A. Rossi, M. Diani, and G. Corsini, "Bilateral filter-based adaptive nonuniformity correction for infrared focal-plane array systems," *Optical Engineering*, vol. 49, 2010.

[33] S. N. Torres, E. M. Vera, R. A. Reeves, and S. K. Sobarzo, "Adaptive scene-based nonuniformity correction method for infrared-focal plane arrays," G. C. Holst, Ed., vol. 5076, no. 1. SPIE, 2003, pp. 130–139. [Online]. Available: http://link.aip.org/link/?PSI/5076/130/1

[34] S. Meister, B. Jähne, and D. Kondermann, "Outdoor stereo camera system for the generation of real-world benchmark data sets," *Opt. Eng.*, vol. 51, no. 2, p. 021107, 2012.

[35] Photonfocus AG. (2012, Aug.) Homepage of the photonfocus ag (vistited 5th of august 2012). [Online]. Available: http://photonfocus.com/

[36] Basler AG. (2012, Aug.) Homepage of basler ag (vistited 5th of august 2012). [Online]. Available: http://www.baslerweb.com/

[37] *Basler A60xf*, Basler AG, Dec 2010, 9.

[38] *User Manual MV-D640 Series*, Photonfocus, Mar 2008, 2.1.

[39] *MV1-D1312(I) CameraLinkő Series*, Photonfocus, Apr 2009, 1.2.

[40] *Peltier-Controller TC3224 -Bedienerhandbuch*, CoolTronic, Mar 2010, v.1.02.

[41] M. Erz, "Charakterisierung von Laufzeitkamerasystemen für Lumineszenzlebensdauermessungen," Ph.D. dissertation, Univ. of Heidelberg, IWR, Fakultät für Physik und Astronomie, 2011.

[42] Basler AG. (2012, Aug.) Homepage of basler ag (vistited 5th of august 2012). [Online]. Available: http://www.baslerweb.com/media/documents/BD00037602_Basler_A60xf_EMVA_Standard_1288.pdf

[43] B. W. und Marcian Edward Hoff, "Adaptive switching circuits," *IRE WESCON Convention Record, Los Angeles*, vol. 4, pp. 96–104, 1960.

[44] NVIDIA Corporation. (2012, Oct.) CUDA Homepage of NVIDIA Corporation (vistited 28th of October 2012). [Online]. Available: http://www.nvidia.com/object/cuda_home_new.html

[45] W. Hackbusch, *Iterative Solution of Large Sparse Systems of Equations*. Springer, 1993.

[46] L. O. Chua and T. Roska, *Cellular neural Networks and visual computing*. Cambridge, 2002.

[47] A. Rodriguez-Vazquez, G. Linan-Cembrano, L. Carranza, E. Roca-Moreno, R. Carmona-Galan, F. Jimenez-Garrido, R. Dominguez-Castro, and S. Meana, "Ace16k: the third generation of mixed-signal simd-cnn ace chips toward vsocs," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 51, May 2004.

[48] S. Carey and P. Dudek, "General-purpose 128x128 simd processor array with integrated image sensor," *Electronic Letters*, vol. 42, June 2006.

[49] A. Lopich and P. Dudek, "Architecture of a vlsi cellular processor array for synchronous/asynchronous image processing," in *ISCAS*, 2006.

[50] A. Rodriguez-Vazquez, S. Espejo, R. Dominguez-Castron, J. Huertas, and E. Sanchez-Sinencio, "Current-mode techniques for the implementation of continuous- and discrete-time cellular neural networks," *Circuits and Systems II: Analog and Digital Signal Processing, IEEE Transactions on*, vol. 40, no. 3, pp. 132 –146, mar 1993.

A. Fixed Pattern Noise

A.1. Transformation of the EMVA PRNU Definition

This appendix describes how to transform the EMVA1288 definitions of PRNU from the given measurements $C_{\mathrm{M},i,j,\mathrm{dark}}$ and $C_{\mathrm{M},i,j,50\%}$ onto the sensor model parameters $\{a_{i,j}\}$ and $\{b_{i,j}\}$.

The definition of PRNU by EMVA1288 is given as:

$$PRNU_{EMVA1288} = \frac{\sqrt{s^2(\mu(C_{\mathrm{M},i,j,50\%})) - s^2(\mu(C_{\mathrm{M},i,j,\mathrm{dark}}))}}{e(\mu(C_{\mathrm{M},i,j,50\%})) - e(\mu(C_{\mathrm{M},i,j,\mathrm{dark}}))} \cdot 100\% \,. \tag{A.1}$$

The offset and gain parameters as well as the measurements are considered random variables and transform as given in [29] to:

$$e(\mu(C_{\mathrm{M},i,j,\mathrm{dark}})) = e(b_{i,j}) \tag{A.2}$$

$$e(\mu(C_{\mathrm{M},i,j,50\%})) = \frac{d}{2} \tag{A.3}$$

$$s^2(\mu(C_{\mathrm{M},i,j,\mathrm{dark}})) = s^2(b_{i,j}) \tag{A.4}$$

$$s^2(\mu(C_{\mathrm{M},i,j,50\%})) = s^2(a_{i,j}C_{S,i,j,50\%} + b_{i,j}) \tag{A.5}$$

$$= s^2(a_{i,j}C_{S,i,j,50\%}) + s^2(b_{i,j}) \tag{A.6}$$

$$= s^2(a_{i,j})e^2(C_{S,i,j,50\%}) + \underbrace{s^2(C_{S,i,j,50\%})}_{=0} e^2(a_{i,j}) + s^2(b_{i,j}) \tag{A.7}$$

$$= s^2(a_{i,j})e^2(C_{S,i,j,50\%}) + s^2(b_{i,j}) \tag{A.8}$$

$$e(C_{\mathrm{M},i,j,50\%}) = d/2 \tag{A.9}$$

$$e(C_{\mathrm{M},i,j,50\%}) = e(a_{i,j}C_{S,i,j,50\%} + b_{i,j}) \tag{A.10}$$

$$= e(a_{i,j})e(C_{S,i,j,50\%}) + e(b_{i,j}) \tag{A.11}$$

$$\Leftrightarrow e(C_{S,i,j,50\%}) = \frac{d/2 - e(b_{i,j})}{e(a_{i,j})} \tag{A.12}$$

These basic transformations of expectation values and variances then result in:

$$PRNU_{EMVA1288} = \frac{\sqrt{s^2(a_{i,j})e^2(C_{S,i,j,50\%}) + s^2(b_{i,j}) - s^2(b_{i,j})}}{\frac{d}{2} - e(b_{i,j})} \cdot 100\% \tag{A.13}$$

$$= \frac{s(a_{i,j})e(C_{S,i,j,50\%})}{\frac{d}{2} - e(b_{i,j})} \cdot 100\% \tag{A.14}$$

$$= \frac{s(a_{i,j})(d/2 - e(b_{i,j}))}{(d/2 - e(b_{i,j}))e(a_{i,j})} = \frac{s(a_{i,j})}{e(a_{i,j})} \cdot 100\% \tag{A.15}$$

A.2. A Scaling of the Parameter Sets for Reference Methods

This appendix shows the derivation of the scaling function of the estimated parameter sets $\{a_{i,j,\text{est.}}\}$ and $\{b_{i,j,\text{est.}}\}$ as used in sec. 3.5. The scaling has to depend only on the statistical moments of the sets of the following parameters: Measured image $\{C_{\text{M},i,j}\}$, estimated DSNU $\{b_{i,j,\text{est.}}\}$, estimated PRNU $\{a_{i,j,\text{est.}}\}$, reference DSNU $\{b_{i,j,ref}\}$, reference PRNU $\{a_{i,j,ref.}\}$:

At first, the target is claimed, that the corrected image with the scaled parameters should be identical to the reference image in its first two statistical moments:

$$\begin{aligned} e(C_{\text{S},i,j,\text{est.}}) &= e(C_{\text{S},i,j,\text{ref.}}) \\ s^2(C_{\text{S},i,j,\text{est.}}) &= s^2(C_{\text{S},i,j,\text{ref.}}) \end{aligned} \tag{A.16}$$

Then we use the definitions of those images to analyze their dependencies on the used correction parameters:

$$C_{\text{S},i,j,\text{ref.}} = \frac{1}{a_{i,j,\text{ref.}}}(C_{\text{M},i,j} - b_{i,j,\text{ref.}}) \tag{A.17}$$

$$C_{\text{S},i,j,\text{est.}} = \frac{1}{a_{i,j,\text{est.}}}(C_{\text{M},i,j} - b_{i,j,\text{est.}})\,. \tag{A.18}$$

The sets of the parameters $\{a_{i,j,\text{est.}}\}$, $\{b_{i,j,\text{est.}}\}$, $\{C_{\text{M},i,j}\}$ and $\{C_{S,i,j}\}$ are interpreted as realizations of the corresponding random variables. We achieve the demands of eq. A.16 by standardizing and re-spreading the random variable in the following form:

$$C_{\text{S},i,j,\text{scal.}} = \frac{C_{\text{S},i,j,\text{est.}} - e(C_{\text{S},i,j,\text{est.}})}{s(C_{\text{S},i,j,\text{est.}})} \cdot s(C_{\text{S},i,j,\text{ref.}}) + e(C_{\text{S},i,j,\text{ref.}}) \quad \forall i,j \tag{A.19}$$

The definition of $C_{i,j,\text{est.}}$ helps to transform this equation into the following form:

$$C_{\text{S},i,j,\text{scal.}} = \frac{C_{\text{S},i,j,\text{est.}}s(C_{\text{S},i,j,\text{ref.}})}{s(C_{\text{S},i,j,\text{est.}})} - \frac{e(C_{\text{S},i,j,\text{est.}})s(C_{\text{S},i,j,\text{ref.}})}{s(C_{\text{S},i,j,\text{est.}})} + e(C_{\text{S},i,j,\text{ref.}}) \tag{A.20}$$

$$= \frac{C_{\text{M},i,j}s(C_{\text{S},i,j,\text{ref.}})}{s(C_{\text{S},i,j,\text{est.}})a_{i,j,\text{est.}}} - \frac{b_{i,j,\text{est.}}s(C_{\text{S},i,j,\text{ref.}})}{s(C_{\text{S},i,j,\text{est.}})a_{i,j,\text{est.}}} - \frac{e(C_{\text{S},i,j,\text{est.}})s(C_{\text{S},i,j,\text{ref.}})}{s(C_{\text{S},i,j,\text{est.}})} + e(C_{\text{S},i,j,\text{ref.}}) \quad \forall i,j \tag{A.21}$$

The above equation allows to extract a linear, point wise transformation:

$$C_{\mathrm{S},i,j,\mathrm{scal.}} = \frac{1}{a_{i,j,sc.}}(C_{\mathrm{M},i,j} - b_{i,j,sc.}) \tag{A.22}$$

$$a_{i,j,\mathrm{sc.}} = h_{sc.,PRNU}(V; a_{i,j,\mathrm{est.}}) = \frac{s(C_{\mathrm{S},i,j,\mathrm{est.}})a_{i,j,\mathrm{est.}}}{s(C_{\mathrm{S},i,j,\mathrm{ref.}})} \tag{A.23}$$

$$b_{i,j,\mathrm{sc.}} = h_{sc.,DSNU}(V; b_{i,j,\mathrm{est.}}) \tag{A.24}$$

$$= \frac{s(C_{\mathrm{S},i,j,\mathrm{est.}})a_{i,j,\mathrm{est.}}}{s(C_{\mathrm{S},i,j,\mathrm{ref.}})}\left(\frac{b_{i,j,\mathrm{est.}}s(C_{\mathrm{S},i,j,\mathrm{ref.}})}{s(C_{\mathrm{S},i,j,\mathrm{est.}})a_{i,j,\mathrm{est.}}}\right. \tag{A.25}$$

$$\left. + \frac{e(C_{\mathrm{S},i,j,\mathrm{est.}})s(C_{\mathrm{S},i,j,\mathrm{ref.}})}{s(C_{\mathrm{S},i,j,\mathrm{est.}})} - e(C_{\mathrm{S},i,j,\mathrm{ref.}})\right) \tag{A.26}$$

As mentioned in sec. 3.5, the set V should only contain the moments of the known variables, which is a requirement not yet met as the values of $s(C_{i,j,\mathrm{ref.}})$ and $e(C_{i,j,\mathrm{ref.}})$ are unknown. The same applies for the ground truth parameter sets $\{a_{i,j,\mathrm{ref.}}\}$ and $\{b_{i,j,\mathrm{ref.}}\}$, which are not known to the algorithms. However, their statistic moments can be assumed to be known, as they usually do not vary from image sensor to image sensor if the manufacturing process is stable.

If considering the above unknown realizations of random variables, then the known moments of the missing variables can be used by assuming that the variables are not correlated [29]. Transforming eq. A.17 leads to:

$$e(C_{i,j,\mathrm{ref.}}) = e\left(\frac{1}{a_{i,j,\mathrm{ref.}}}(C_{\mathrm{M},i,j} - b_{i,j,\mathrm{ref.}})\right) \tag{A.27}$$

$$= e\left(\frac{1}{a_{i,j,\mathrm{ref.}}}\right) * e((C_{\mathrm{M},i,j} - b_{i,j,\mathrm{ref.}}))) \tag{A.28}$$

$$= e\left(\frac{1}{a_{i,j,\mathrm{ref.}}}\right)(e(C_{\mathrm{M},i,j}) - e(b_{i,j,\mathrm{ref.}})) \tag{A.29}$$

For the inverted random variable the expectation value estimates as

$$e\left(\frac{1}{a_{i,j,\mathrm{ref.}}}\right) = \frac{1}{e(a_{i,j,\mathrm{ref.}})}\left(1 + \left(\frac{s(a_{i,j,\mathrm{ref.}})}{e(a_{i,j,\mathrm{ref.}})}\right)^2\right), \tag{A.30}$$

and the spatial variance is estimated by the complex expression:

$$s^2(C_{i,j,\mathrm{ref.}}) = s^2\left(\frac{1}{a_{i,j,\mathrm{ref.}}}(C_{\mathrm{M},i,j} - b_{i,j,\mathrm{ref.}})\right) \tag{A.31}$$

$$= s^2\left(\frac{1}{a_{i,j,\mathrm{ref.}}}\right)e(C_{\mathrm{M},i,j} - b_{i,j,\mathrm{ref.}}) + s^2(C_{\mathrm{M},i,j} - b_{i,j,\mathrm{ref.}})e\left(\frac{1}{a_{i,j,\mathrm{ref.}}}\right) \tag{A.32}$$

The only not yet described transformation is the variance $s^2\left(\frac{1}{a_{i,j,\mathrm{ref.}}}\right)$, which can be approximated by [29]:

$$s^2\left(\frac{1}{a_{i,j,\mathrm{ref.}}}\right) = \left(\frac{s(a_{i,j,\mathrm{ref.}})}{e^2(a_{i,j,\mathrm{ref.}})}\right)^2 \tag{A.33}$$

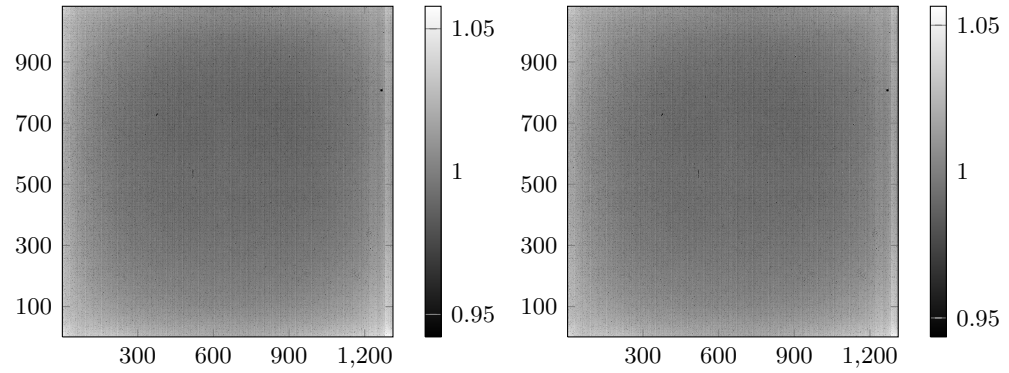
A combination of all these equations allows to fully define the point wise linear transformations $h_{sc.,PRNU}(V; a_{i,j,\mathrm{est.}})$ and $h_{sc.,DSNU}(V; b_{i,j,\mathrm{est.}})$. The actually quite complex expressions above compress to real numbers that should be calculated in advance for a better performance.

B. Image Data

B.1. Further Calibration Data for the Photonfocus MV1-D1312-160-CL

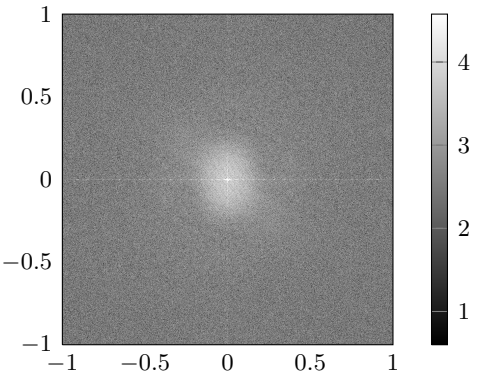
This appendix shows the additional results for the Photonfocus MV1-D1312-160-CL camera that are summarized in the following table:

Sequence Name	$t_{\text{exp.}}$ [ms]	T [°C]	Scene	DSNU_{DN}	PRNU
MV1-D1312-1a,b	5	25	L600a in HD	50.81	1.92
MV1-D1312-1a,b	10	25	L600a in HD	50.83	1.88
MV1-D1312-1a,b	0.05	50	L600a in HD	51.07	7.39
MV1-D1312-1a,b	5	50	L600a in HD	51.19	1.91

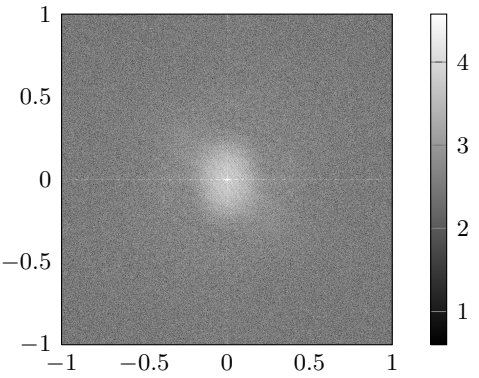


(a) Spatial PRNU for (5 ms @ 25 °C)

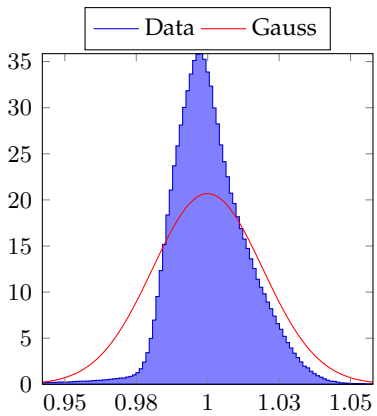
(b) Spatial PRNU for (10 ms @ 25 °C)



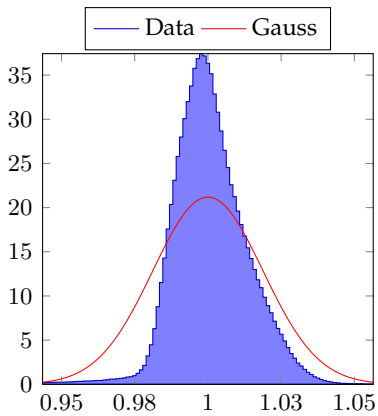
(c) Spectral:log $|\hat{a}_{\mu,\nu}|$ for (5 ms @ 25 °C)



(d) Spectral:log $|\hat{a}_{\mu,\nu}|$ for (10 ms @ 25 °C)

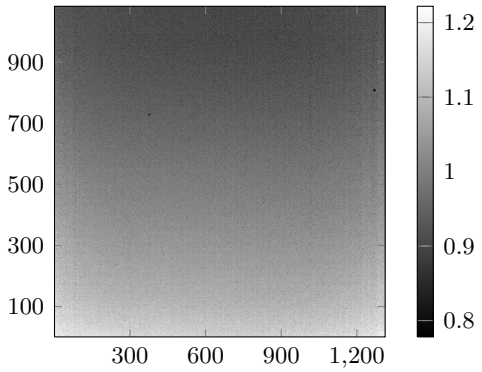


(e) Histogram of $\{a_{i,j}\}$ for (5 ms @ 25 °C)

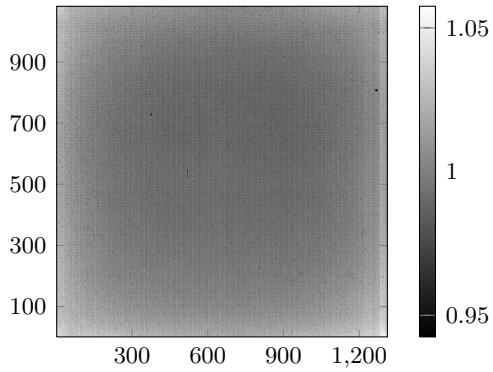


(f) Histogram of $\{a_{i,j}\}$ for (10 ms @ 25 °C)

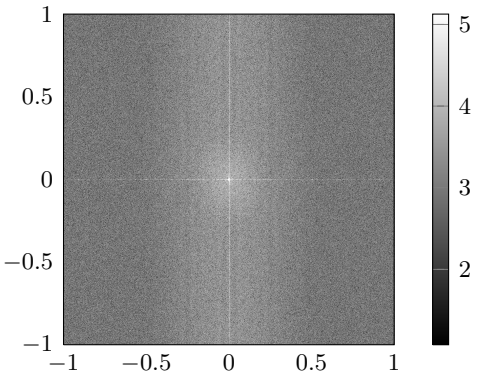
Figure B.1.: Further representations of the $\{a_{i,j}\}$ correction patterns of the EMVA1288 calibration measurements with the Photonfocus MV1-D1312-160-CL Camera. Including a visual spatial and spectral analysis as well as a spatial histogram of the realizations at the given temperatures and exposure times.



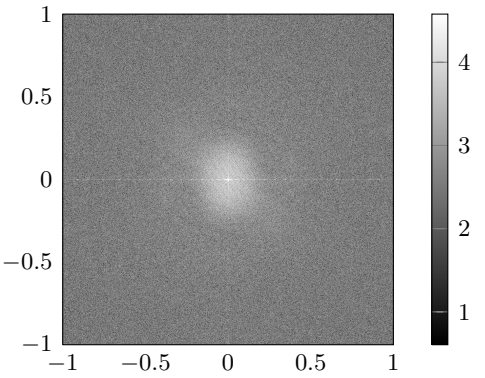
(a) Spatial PRNU for $(0.05\,\mathrm{ms} \,@\, 50\,^{\circ}\mathrm{C})$



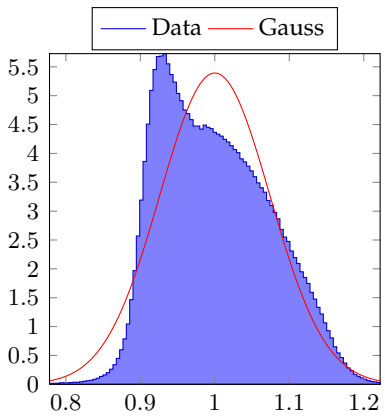
(b) Spatial PRNU for $(5\,\mathrm{ms} \,@\, 50\,^{\circ}\mathrm{C})$



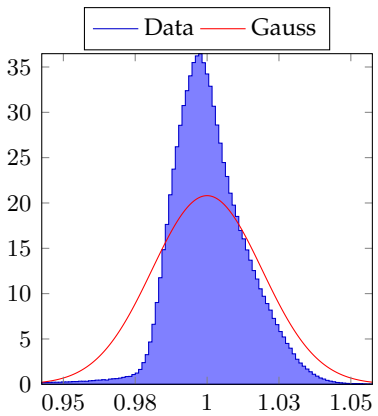
(c) Spectral:$\log|\hat{a}_{\mu,\nu}|$ for $(0.05\,\mathrm{ms} \,@\, 50\,^{\circ}\mathrm{C})$



(d) Spectral:$\log|\hat{a}_{\mu,\nu}|$ for $(5\,\mathrm{ms} \,@\, 50\,^{\circ}\mathrm{C})$

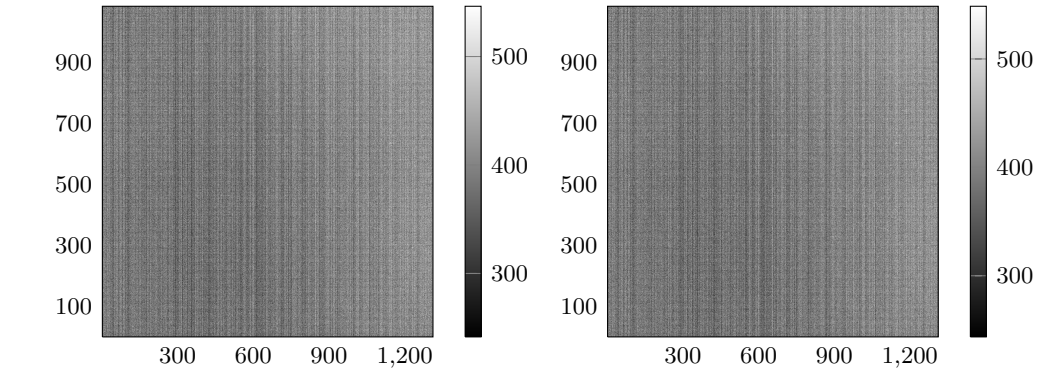


(e) Histogram of $\{a_{i,j}\}$ for $(0.05\,\mathrm{ms} \,@\, 50\,^{\circ}\mathrm{C})$



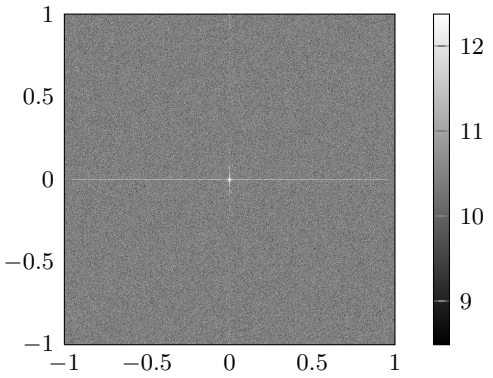
(f) Histogram of $\{a_{i,j}\}$ for $(5\,\mathrm{ms} \,@\, 50\,^{\circ}\mathrm{C})$

Figure B.2.: Further representations of the $\{a_{i,j}\}$ correction patterns of the EMVA1288 calibration measurements with the Photonfocus MV1-D1312-160-CL Camera. Including a visual spatial and spectral analysis as well as a spatial histogram of the realizations at the given temperatures and exposure times.

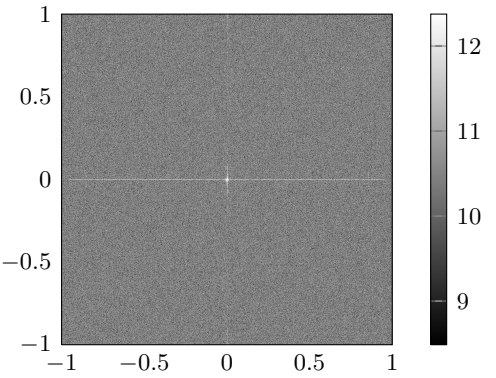


(a) Spatial DSNU for (5 ms @ 25 °C)

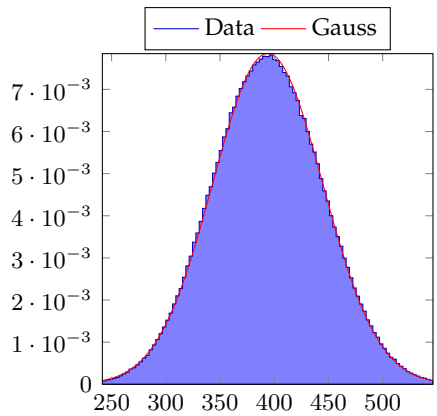
(b) Spatial DSNU for (10 ms @ 25 °C)



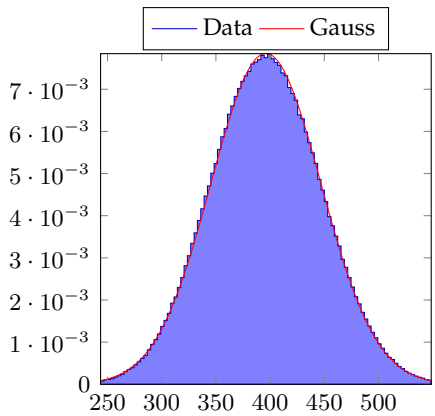
(c) Spectral:log $|\hat{b}_{\mu,\nu}|$ for (5 ms @ 25 °C)



(d) Spectral:log $|\hat{b}_{\mu,\nu}|$ for (10 ms @ 25 °C)

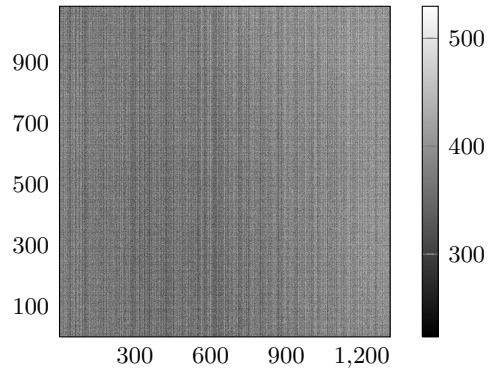


(e) Histogram of $\{b_{i,j}\}$ for (5 ms @ 25 °C)

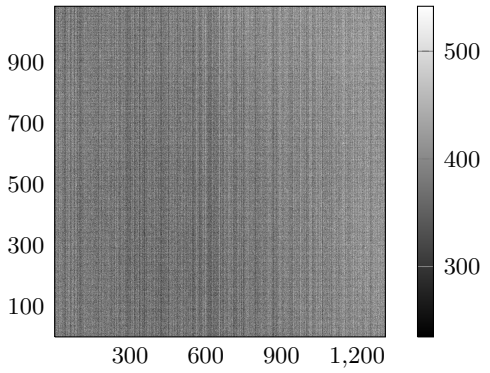


(f) Histogram of $\{b_{i,j}\}$ for (10 ms @ 25 °C)

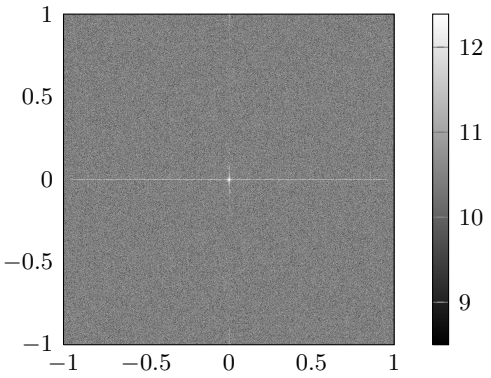
Figure B.3.: Further representations of the $\{b_{i,j}\}$ correction patterns of the EMVA1288 calibration measurements with the Photonfocus MV1-D1312-160-CL Camera. Including a visual spatial and spectral analysis as well as a spatial histogram of the realizations at the given temperatures and exposure times.



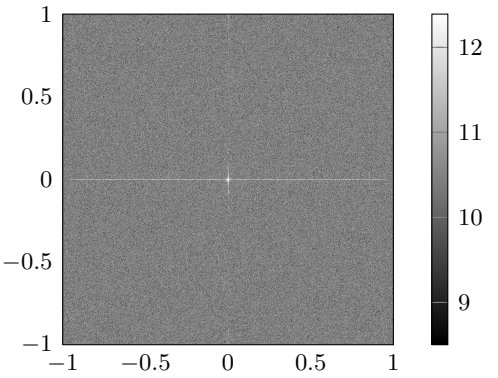
(a) Spatial DSNUfor $(0.05\,\mathrm{ms} @ 50\,^{\circ}\mathrm{C})$



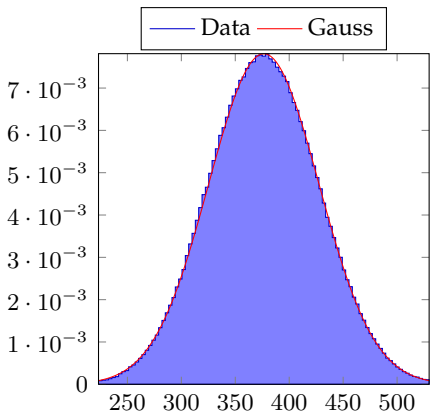
(b) Spatial DSNU for $(5\,\mathrm{ms} @ 50\,^{\circ}\mathrm{C})$



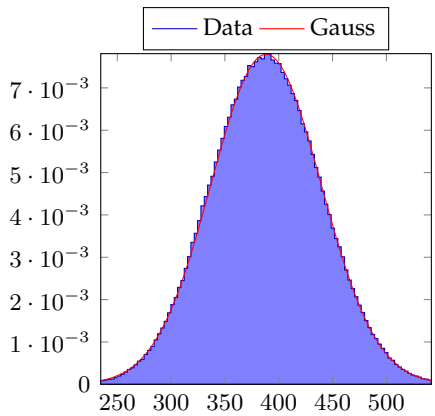
(c) Spectral: $\log|\hat{b}_{\mu,\nu}|$for $(0.05\,\mathrm{ms} @ 50\,^{\circ}\mathrm{C})$



(d) Spectral: $\log|\hat{b}_{\mu,\nu}|$ for $(5\,\mathrm{ms} @ 50\,^{\circ}\mathrm{C})$



(e) Histogram of $\{b_{i,j}\}$for $(0.05\,\mathrm{ms} @ 50\,^{\circ}\mathrm{C})$



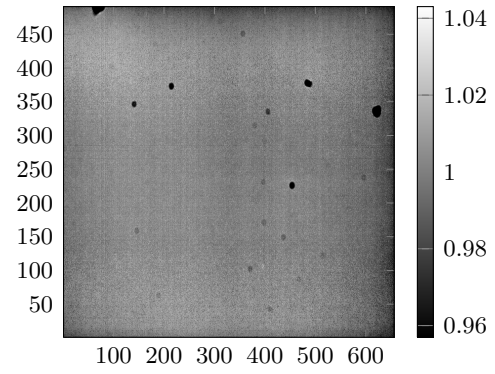
(f) Histogram of $\{b_{i,j}\}$ for $(5\,\mathrm{ms} @ 50\,^{\circ}\mathrm{C})$

Figure B.4.: Further representations of the $\{b_{i,j}\}$ correction patterns of the EMVA1288 calibration measurements with the Photonfocus MV1-D1312-160-CL Camera. Including a visual spatial and spectral analysis as well as a spatial histogram of the realizations at the given temperatures and exposure times.

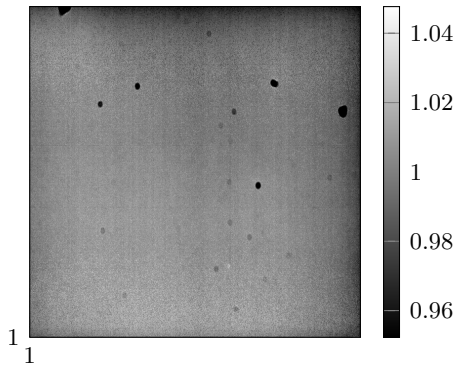
B.2. Further calibration data for the Basler A602f Camera

This appendix shows the additional results for the Basler A602f Camera that are summarized in the following table:

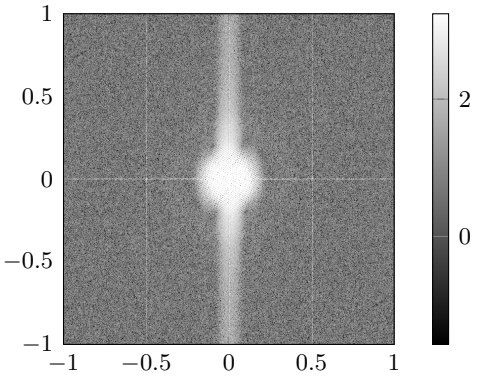
Sequence Name	$t_{\text{exp.}}$ [ms]	T [°C]	Scene	DSNU_{DN}	PRNU
A602f-4a, A602f-4b	10	22	L600a in HD	1.29	1.43
A602f-2a, A602f-2b	1	30	L600a in HD	0.89	1.58
A602f-5a, A602f-5b	10	30	L600a in HD	0.75	1.44
A602f-3a, A602f-3b	1	49	L600a in HD	2.09	1.78



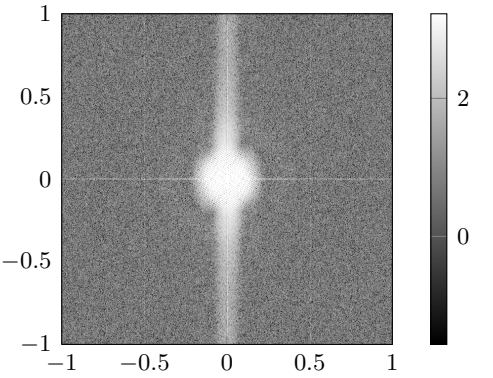
(a) Spatial PRNU for (10 ms @ 22 °C)



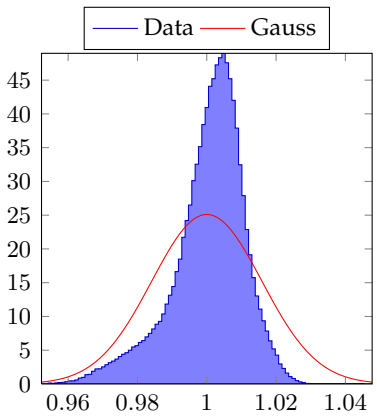
(b) Spatial PRNU for (1 ms @ 30 °C)



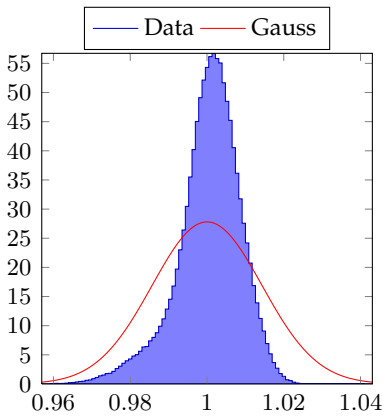
(c) Spectral:log $|\hat{a}_{\mu,\nu}|$ for (10 ms @ 22 °C)



(d) Spectral:log $|\hat{a}_{\mu,\nu}|$ for (1 ms @ 30 °C)

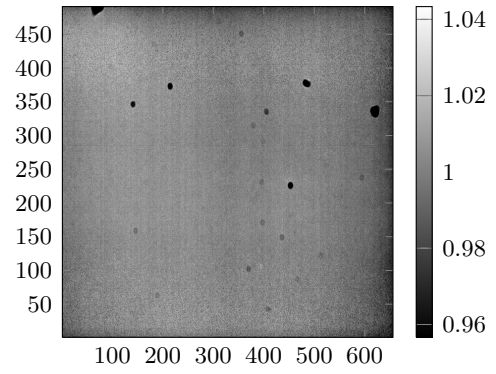


(e) Histogram of $\{a_{i,j}\}$ @ $(20°C, 1ms)$

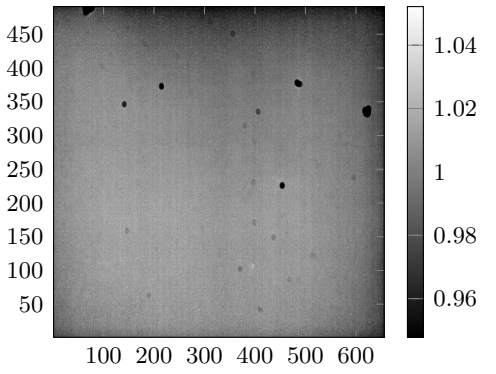


(f) Histogram of $\{a_{i,j}\}$ for (10 ms @ 22 °C)

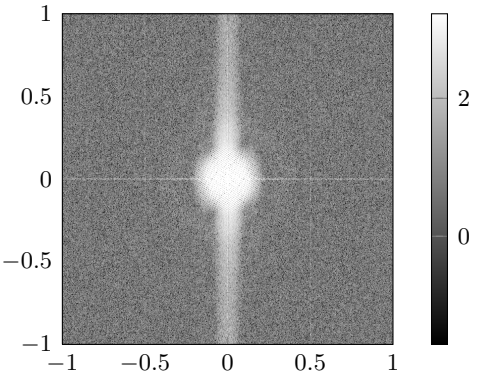
Figure B.5.: Further representations of the $\{a_{i,j}\}$ correction patterns of the EMVA1288 calibration measurements with the Basler A602f Camera. Including a visual spatial and spectral analysis as well as a spatial histogram of the realizations at the given temperatures and exposure times.



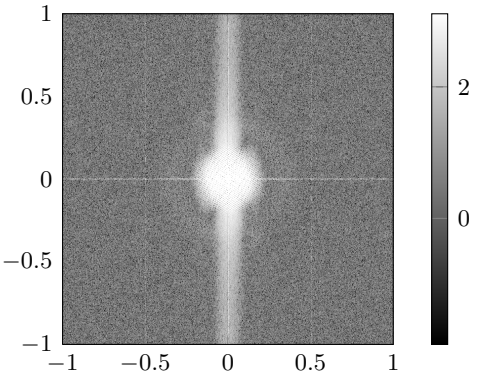
(a) Spatial PRNU for (10 ms @ 30 °C)



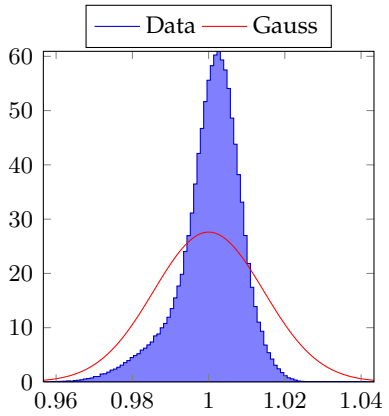
(b) Spatial PRNU for (1 ms @ 50 °C)



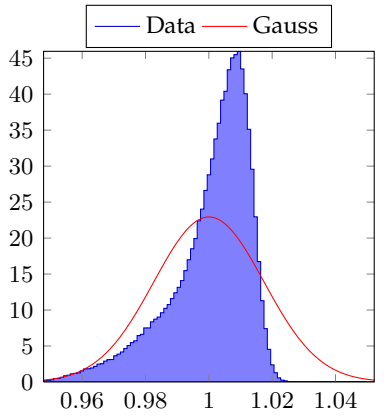
(c) Spectral:log $|\hat{a}_{\mu,\nu}|$ for (10 ms @ 30 °C)



(d) Spectral:log $|\hat{a}_{\mu,\nu}|$ for (1 ms @ 50 °C)

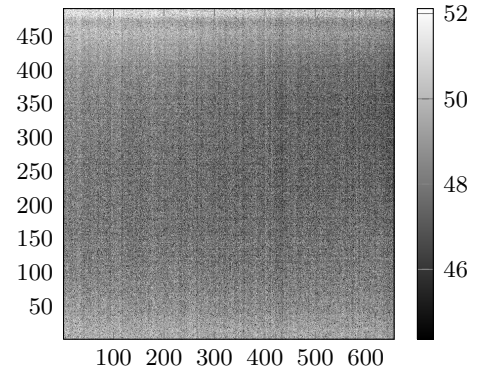


(e) Histogram of $\{a_{i,j}\}$ for (10 ms @ 30 °C)

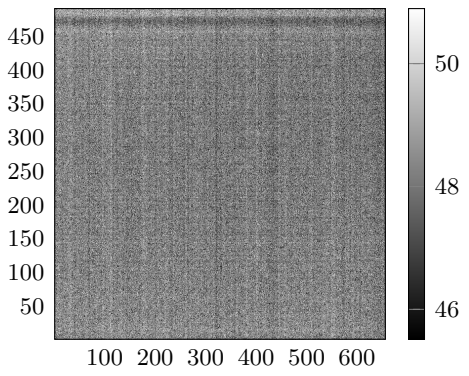


(f) Histogram of $\{a_{i,j}\}$ for (1 ms @ 50 °C)

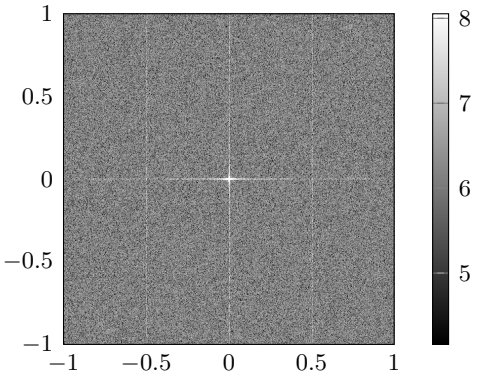
Figure B.6.: Further representations of the $\{a_{i,j}\}$ correction patterns of the EMVA1288 calibration measurements with the Basler A602f Camera. Including a visual spatial and spectral analysis as well as a spatial histogram of the realizations at the given temperatures and exposure times.



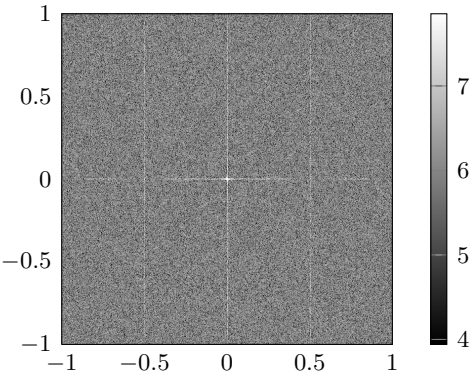
(a) Spatial DSNU for (10 ms @ 22 °C)



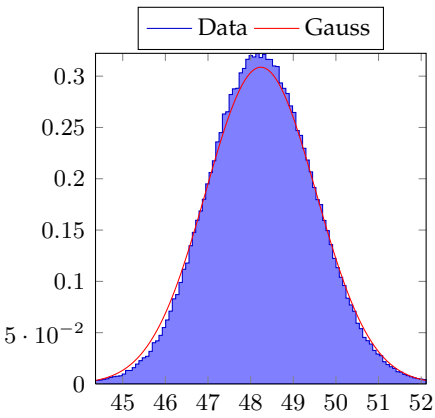
(b) Spatial DSNU for (1 ms @ 30 °C)



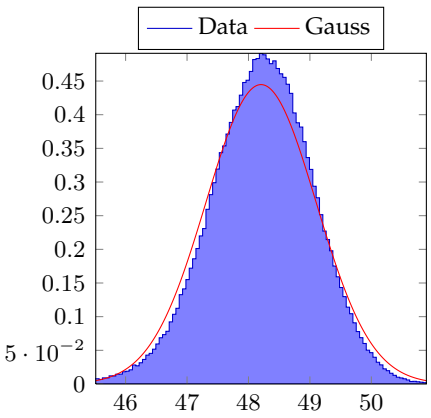
(c) Spectral:log $|\hat{b}_{\mu,\nu}|$ for (10 ms @ 22 °C)



(d) Spectral:log $|\hat{b}_{\mu,\nu}|$ for (1 ms @ 30 °C)

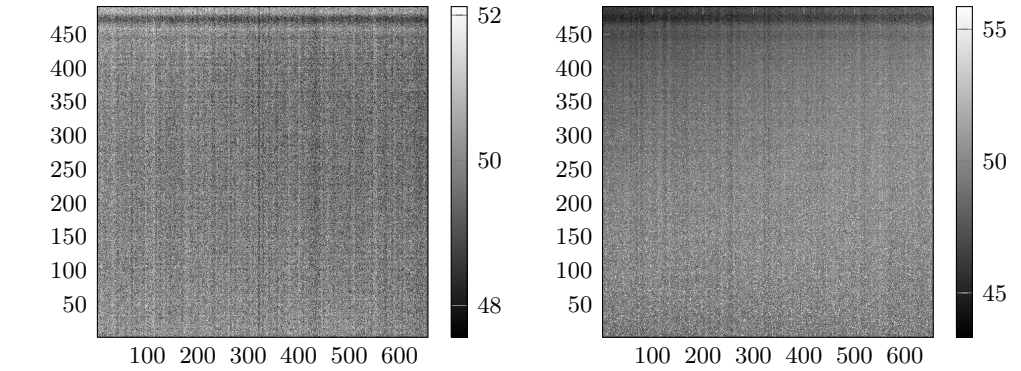


(e) Histogram of $\{b_{i,j}\}$ for (10 ms @ 22 °C)



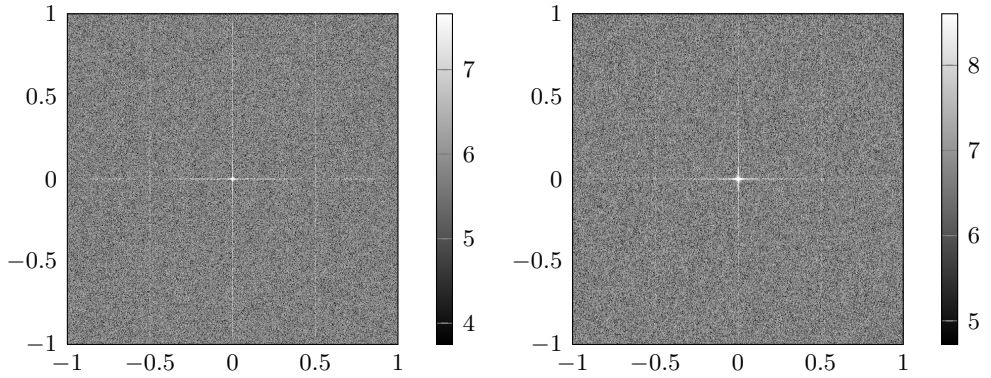
(f) Histogram of $\{b_{i,j}\}$ for (1 ms @ 30 °C)

Figure B.7.: Further representations of the $\{b_{i,j}\}$ correction patterns of the EMVA1288 calibration measurements with the Basler A602f Camera. Including a visual spatial and spectral analysis as well as a spatial histogram of the realizations at the given temperatures and exposure times.



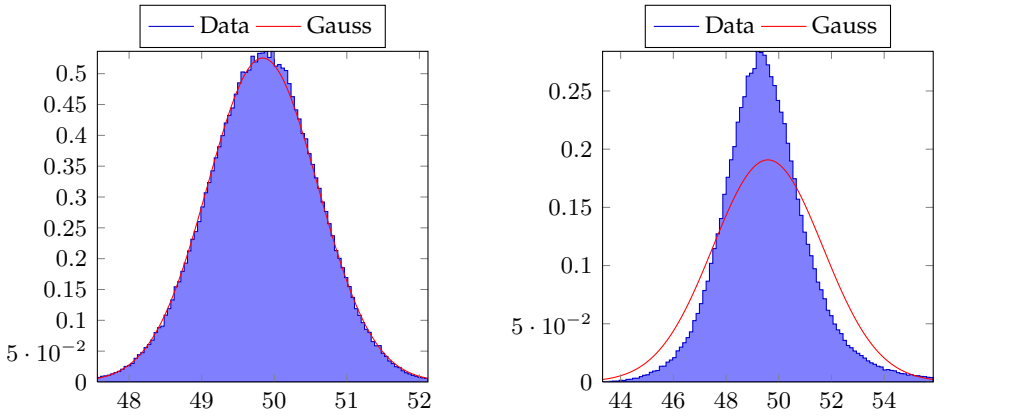
(a) Spatial DSNU for (10 ms @ 30 °C)

(b) Spatial DSNU for (1 ms @ 50 °C)



(c) Spectral:log $|\hat{b}_{\mu,\nu}|$ for (10 ms @ 30 °C)

(d) Spectral:log $|\hat{b}_{\mu,\nu}|$ for (1 ms @ 50 °C)



(e) Histogram of $\{b_{i,j}\}$ for (10 ms @ 30 °C)

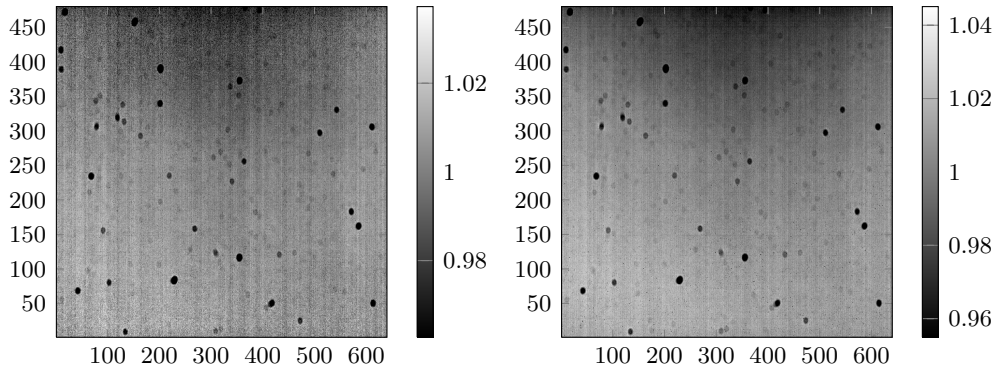
(f) Histogram of $\{b_{i,j}\}$ for (1 ms @ 50 °C)

Figure B.8.: Further representations of the $\{b_{i,j}\}$ correction patterns of the EMVA1288 calibration measurements with the Basler A602f Camera. Including a visual spatial and spectral analysis as well as a spatial histogram of the realizations at the given temperatures and exposure times.

B.3. Further calibration data for the Photonfocus MV-D640-66-CL

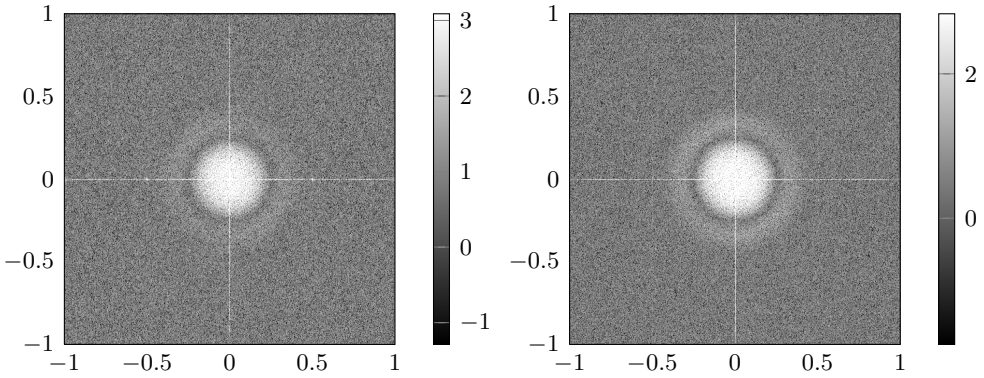
This appendix shows the additional results for the Photonfocus MV-D640-66-CL camera that are summarized in the following table:

Sequence Name	$t_{\text{exp.}}$ [ms]	T [°C]	Scene	DSNU_{DN}	PRNU
MV-D640-66-CL-2	10	25	L600a in HD	0.55	1.24
MV-D640-66-CL-3	5	50	L600a in HD	1.27	1.50



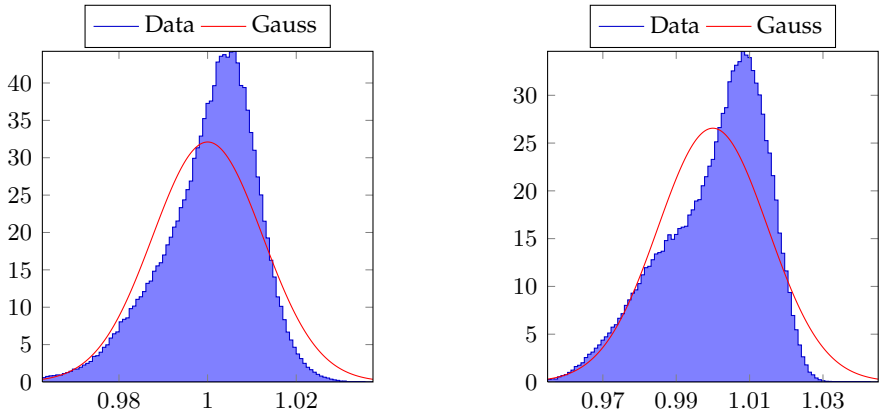
(a) Spatial PRNU for (10 ms @ 25 °C)

(b) Spatial PRNU for (5 ms @ 50 °C)



(c) Spectral: $\log|\hat{a}_{\mu,\nu}|$ for (10 ms @ 25 °C)

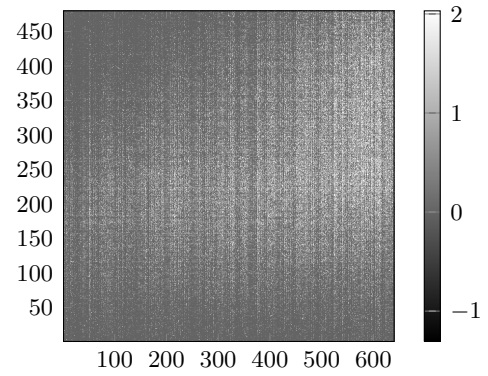
(d) Spectral: $\log|\hat{a}_{\mu,\nu}|$ for (5 ms @ 50 °C)



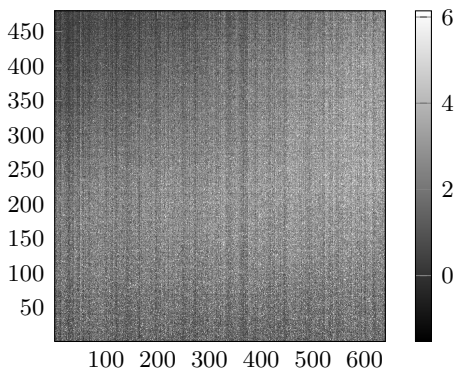
(e) Histogram of $\{a_{i,j}\}$ for (10 ms @ 25 °C)

(f) Histogram of $\{a_{i,j}\}$ for (5 ms @ 50 °C)

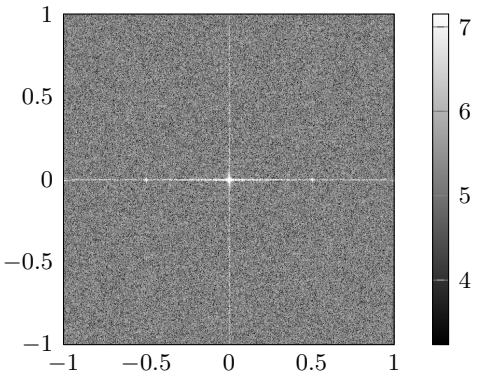
Figure B.9.: Further representations of the $\{b_{i,j}\}$ correction patterns of the EMVA1288 calibration measurements with the MV-D640-66-CL camera. Including a visual spatial and spectral analysis as well as a spatial histogram of the realizations at the given temperatures and exposure times.



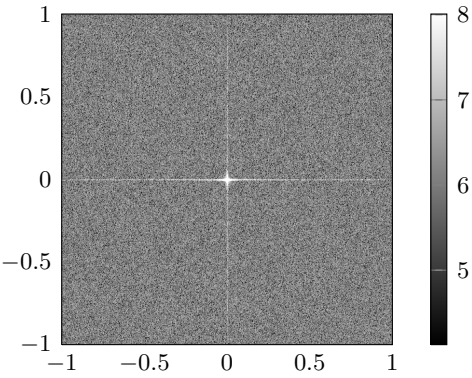
(a) Spatial DSNU for (10 ms @ 25 °C)



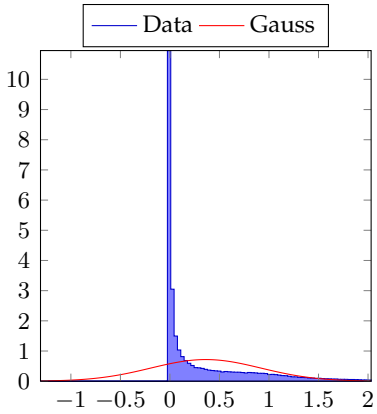
(b) Spatial DSNU for (5 ms @ 50 °C)



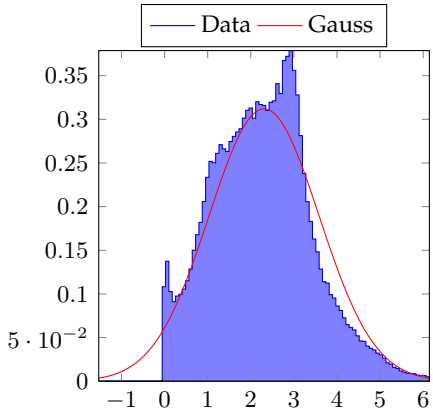
(c) Spectral:log $|\hat{b}_{\mu,\nu}|$ for (10 ms @ 25 °C)



(d) Spectral:log $|\hat{b}_{\mu,\nu}|$ for (5 ms @ 50 °C)



(e) Histogram of $\{b_{i,j}\}$ for (10 ms @ 25 °C)



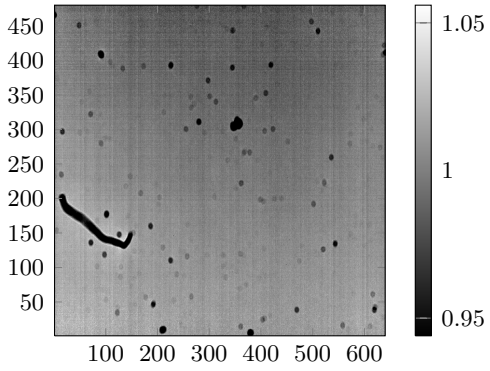
(f) Histogram of $\{b_{i,j}\}$ for (5 ms @ 50 °C)

Figure B.10.: Further representations of the $\{b_{i,j}\}$ correction patterns of the EMVA1288 calibration measurements with the MV-D640-66-CL camera. Including a visual spatial and spectral analysis as well as a spatial histogram of the realizations at the given temperatures and exposure times.

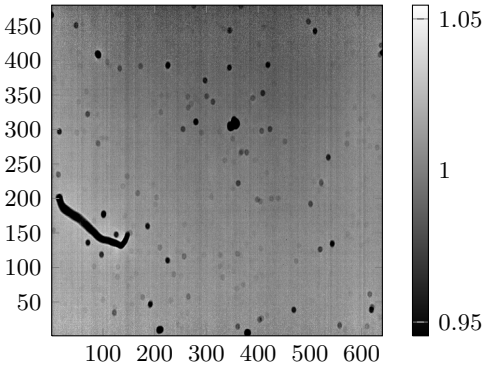
B.4. Further calibration data for the Photonfocus MV-D640-66-CLLinLog

This appendix shows the additional results for the Photonfocus MV-D640-66-CLLinLog camera that are summarized in the following table:

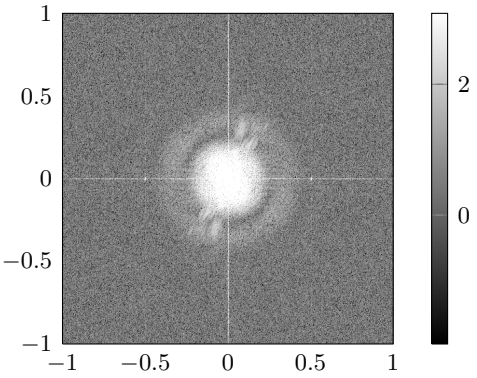
Sequence Name	$t_{\text{exp.}}$ [ms]	T [°C]	Scene	DSNU_{DN}	PRNU
MV-D640-66-CL-LinLog-2	5	25	L600a in HD	1.55	1.86
MV-D640-66-CL-LinLog-3	10	25	L600a in HD	1.55	1.81
MV-D640-66-CL-LinLog-4	0.082	50	L600a in HD	1.97	10.71
MV-D640-66-CL-LinLog-5	5	50	L600a in HD	1.95	1.87



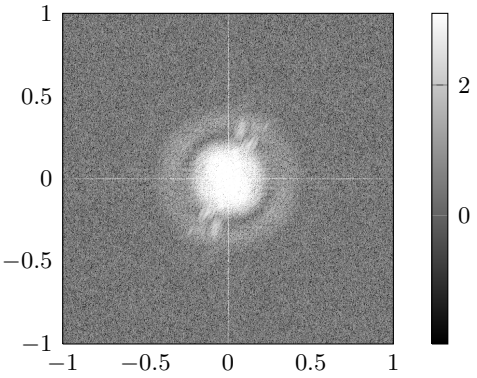
(a) Spatial PRNU for (5 ms @ 25 °C)



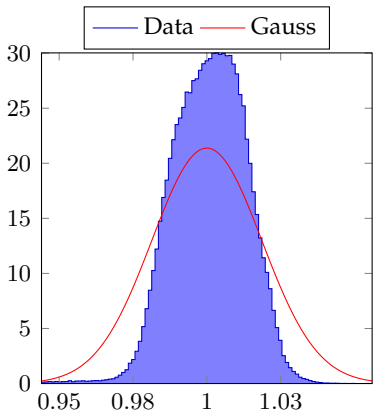
(b) Spatial PRNU for (10 ms @ 25 °C)



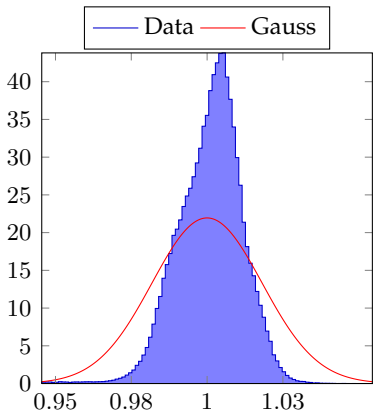
(c) Spectral:log $|\hat{a}_{\mu,\nu}|$ for (5 ms @ 25 °C)



(d) Spectral:log $|\hat{a}_{\mu,\nu}|$ for (10 ms @ 25 °C)

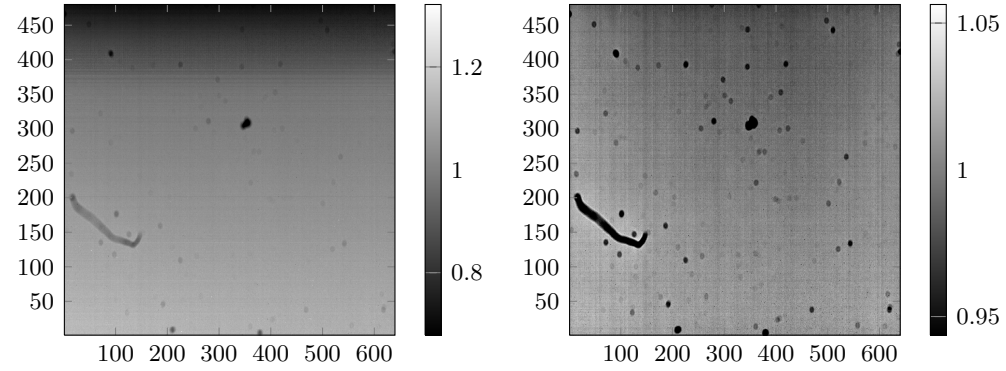


(e) Histogram of $\{a_{i,j}\}$ for (5 ms @ 25 °C)



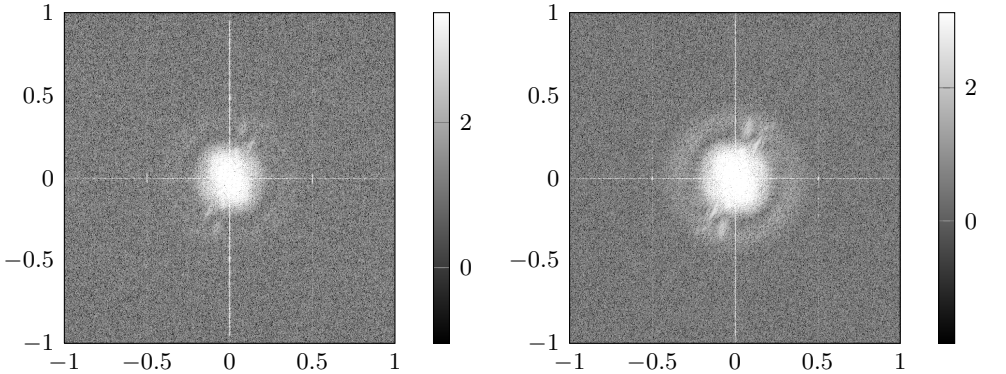
(f) Histogram of $\{a_{i,j}\}$ for (10 ms @ 25 °C)

Figure B.11.: Further representations of the $\{a_{i,j}\}$ correction patterns of the EMVA1288 calibration measurements with the MV-D640-66-CL-LinLog camera. Including a visual spatial and spectral analysis as well as a spatial histogram of the realizations at the given temperatures and exposure times.



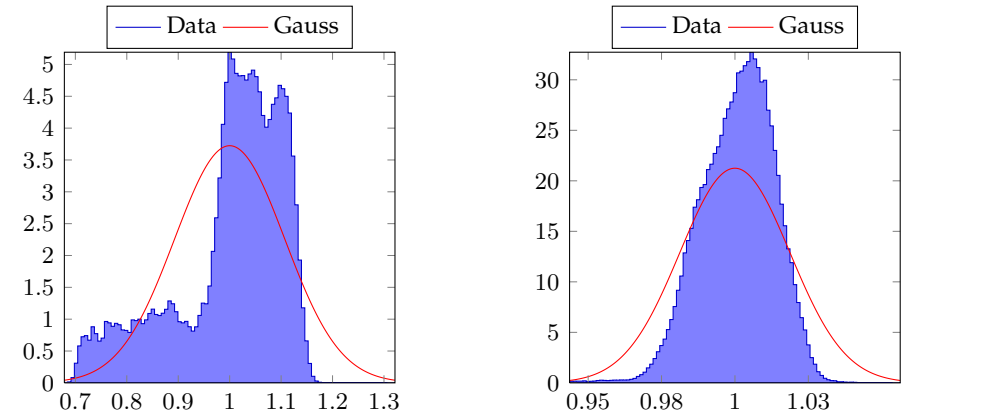
(a) Spatial PRNU for (0.082 ms @ 50 °C)

(b) Spatial PRNU for (5 ms @ 50 °C)



(c) Spectral:log $|\hat{a}_{\mu,\nu}|$ for (0.082 ms @ 50 °C)

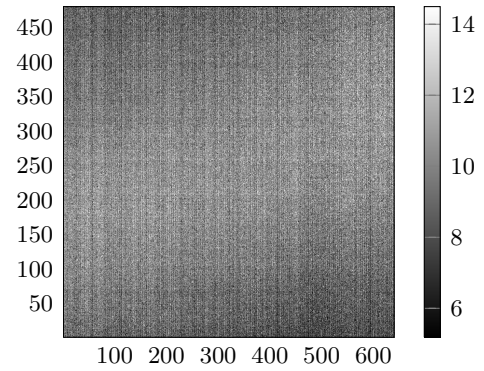
(d) Spectral:log $|\hat{a}_{\mu,\nu}|$ for (5 ms @ 50 °C)



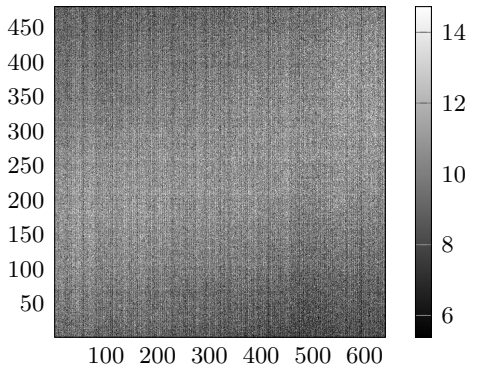
(e) Histogram of $\{a_{i,j}\}$ for (0.082 ms @ 50 °C)

(f) Histogram of $\{a_{i,j}\}$ for (5 ms @ 50 °C)

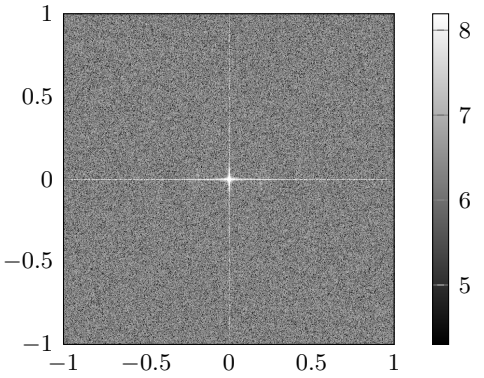
Figure B.12.: Further representations of the $\{a_{i,j}\}$ correction patterns of the EMVA1288 calibration measurements with the MV-D640-66-CL-LinLog camera. Including a visual spatial and spectral analysis as well as a spatial histogram of the realizations at the given temperatures and exposure times.



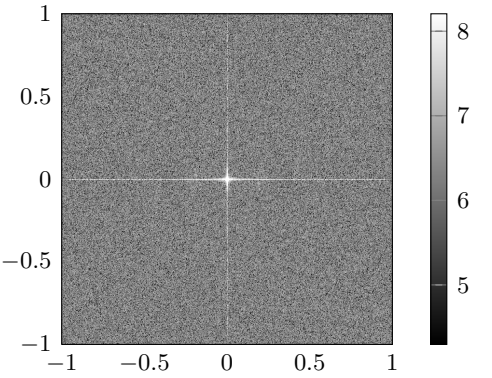
(a) Spatial DSNU for (5 ms @ 25 °C)



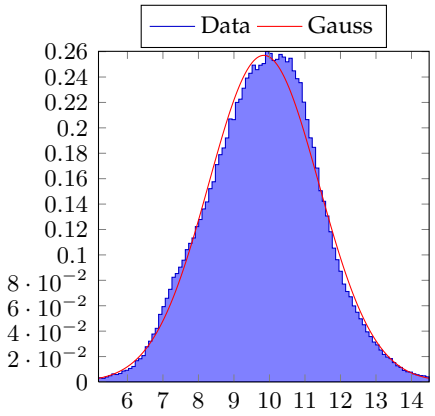
(b) Spatial DSNU for (10 ms @ 25 °C)



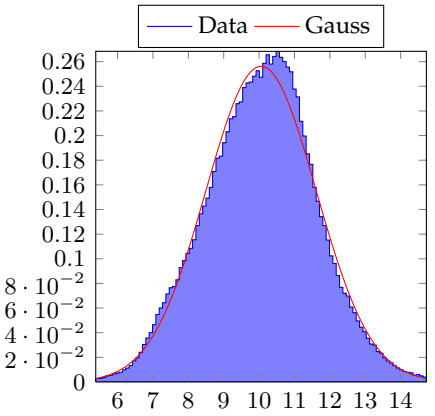
(c) Spectral:log $|\hat{b}_{\mu,\nu}|$ for (5 ms @ 25 °C)



(d) Spectral:log $|\hat{b}_{\mu,\nu}|$ for (10 ms @ 25 °C)

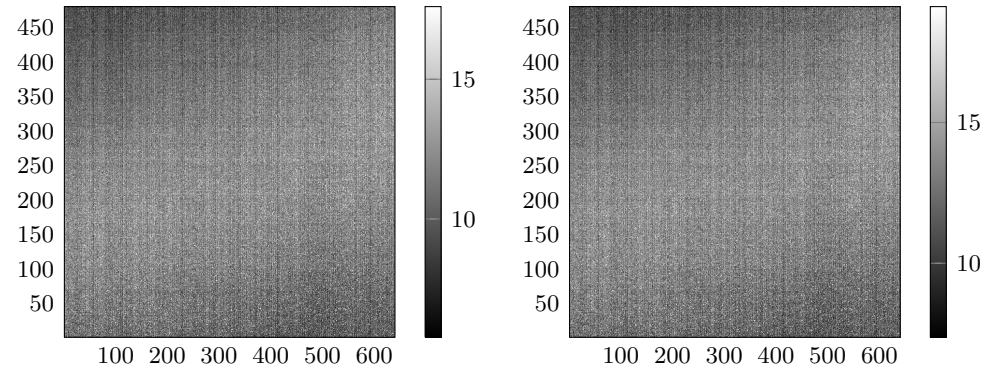


(e) Histogram of $\{b_{i,j}\}$ for (5 ms @ 25 °C)



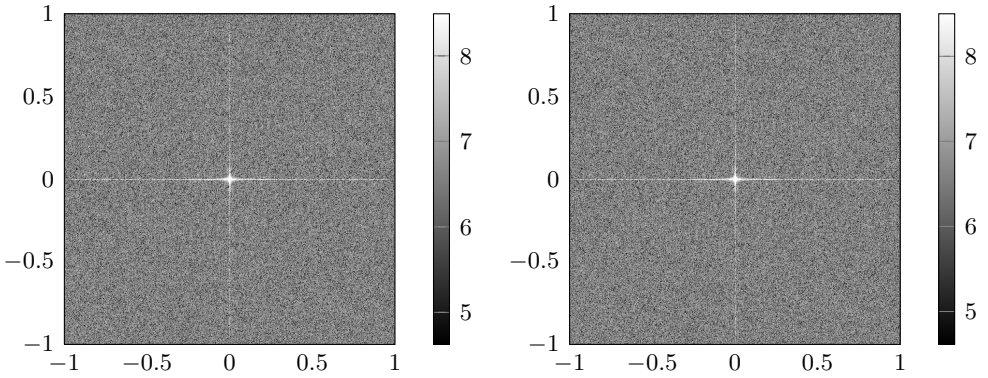
(f) Histogram of $\{b_{i,j}\}$ for (10 ms @ 25 °C)

Figure B.13.: Further representations of the $\{b_{i,j}\}$ correction patterns of the EMVA1288 calibration measurements with the MV-D640-66-CL-LinLog camera. Including a visual spatial and spectral analysis as well as a spatial histogram of the realizations at the given temperatures and exposure times.



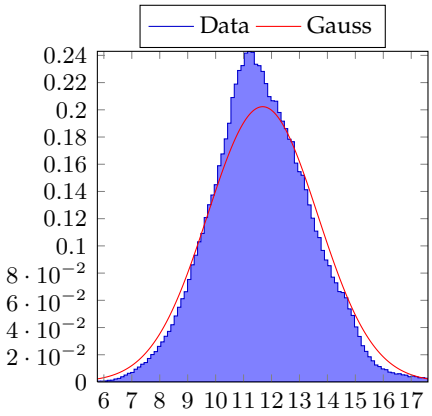
(a) Spatial DSNU for (0.082 ms @ 50 °C)

(b) Spatial DSNU for (5 ms @ 50 °C)

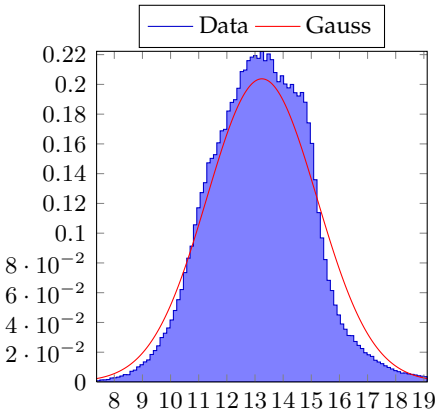


(c) Spectral:log $|\hat{b}_{\mu,\nu}|$ for (0.082 ms @ 50 °C)

(d) Spectral:log $|\hat{b}_{\mu,\nu}|$ for (5 ms @ 50 °C)



(e) Histogram of $\{b_{i,j}\}$ for (0.082 ms @ 50 °C)



(f) Histogram of $\{b_{i,j}\}$ for (5 ms @ 50 °C)

Figure B.14.: Further representations of the $\{b_{i,j}\}$ correction patterns of the EMVA1288 calibration measurements with the MV-D640-66-CL-LinLog camera. Including a visual spatial and spectral analysis as well as a spatial histogram of the realizations at the given temperatures and exposure times.

C. Related Methods

C.1. Further Quality Metrics for the Best-Case Performance Analysis

Fig. C.1 and C.2 show the results of the best-case performance analysis discussed in sec. 5.3.2 in the MAE and RMSE quality norms.

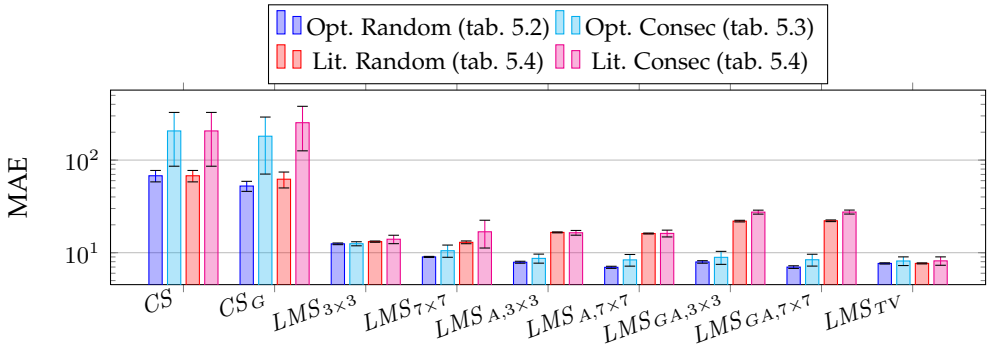


Figure C.1.: Best-case performance analysis with respect to the MAE quality metric

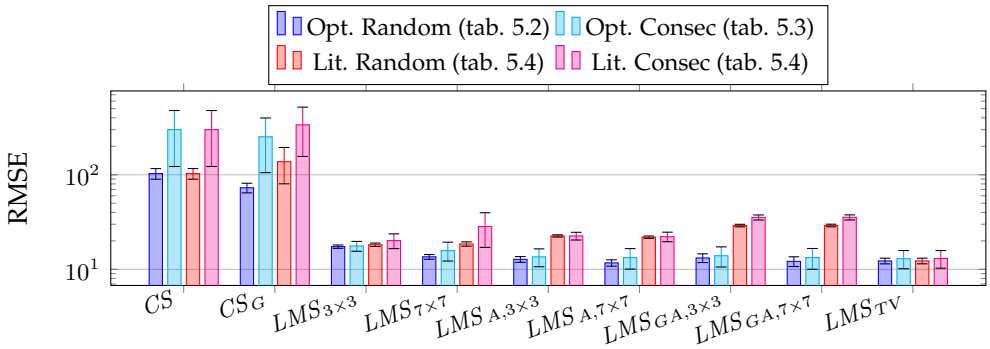


Figure C.2.: Best-case performance analysis with respect to the RMSE quality metric

C.2. Further Evaluations for the LMS-Type Reference Methods

This appendix shows further plots of the temporal evaluations for the reference methods, that are discussed in sec. 5.3.4. Fig. C.3 and C.4 show the results for the RMSE and MAE quality metrics The actual ordering of the methods with respect to the achieved quality does not differ from the already shown plots in sec. 5.3.4. Thus these measures are just presented for for completeness and a further discussion is not necessary.

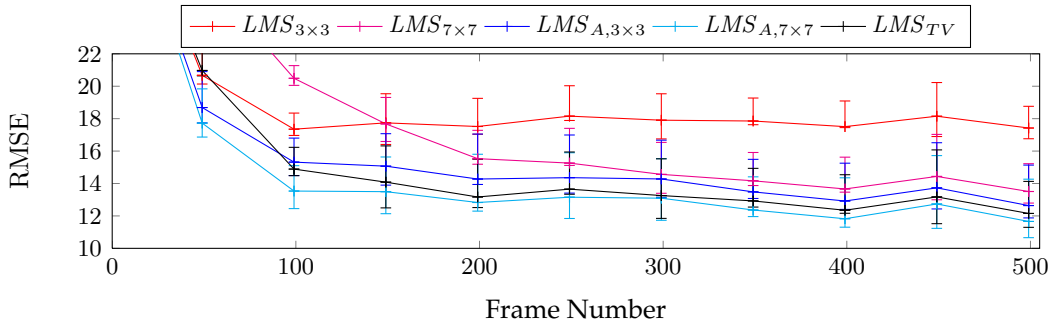


Figure C.3.: The LMS methods' evaluations for the RMSE metric with random frame sequences as input.

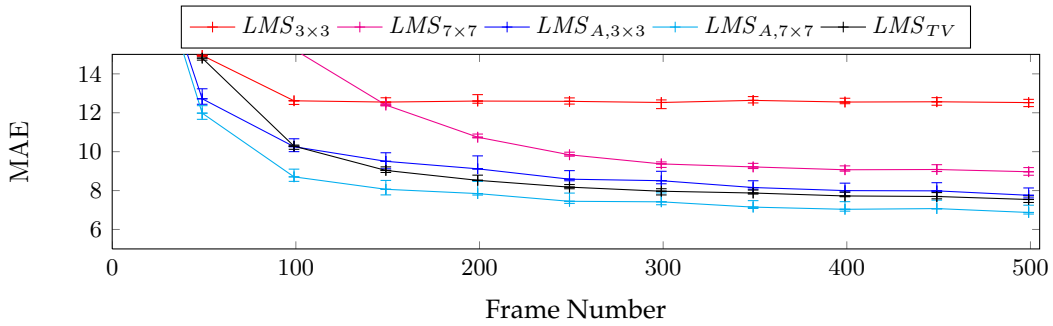


Figure C.4.: The LMS methods' evaluations for the MAE metric with random frame sequences as input.

Fig. C.5, C.6, C.7 and C.8 show the results for the 1000-frame sequences for the quality measures $p_{\mathrm{c,DSNU}}$, SNR, RMSE and MAE. In sec. 5.3.4 the plot for the $p_{\mathrm{c,PRNU}}$ was shown to proof that the PRNU correction performance is not converging. As it can be seen in the other plots, the performances are converging, which is the same result as already stated and proved by the $p_{\mathrm{c,Comb.}}$ evaluation. Again, these plots are just shown for readers which are more familiar to these type of error measurements.

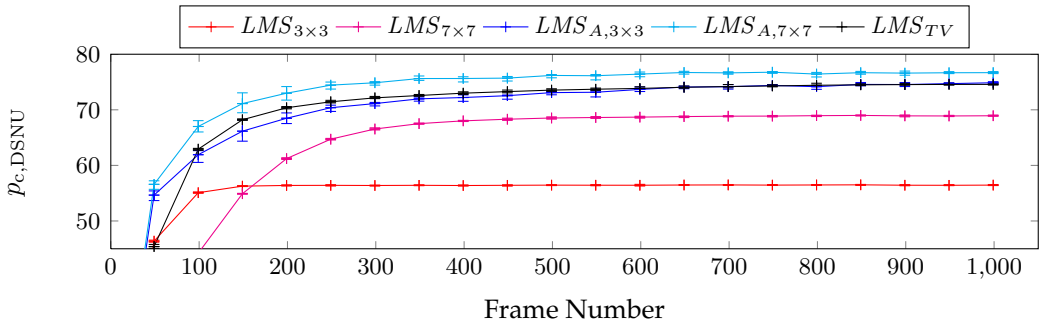


Figure C.5.: The LMS methods' evaluations for the $p_{c,DSNU}$ metric with random frame sequences of 1000 frames length.

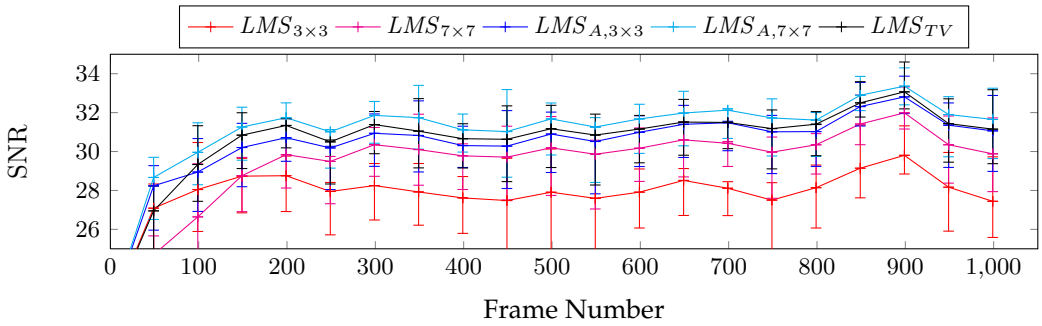


Figure C.6.: The LMS methods' evaluations for the SNR metric with random frame sequences of 1000 frames length.

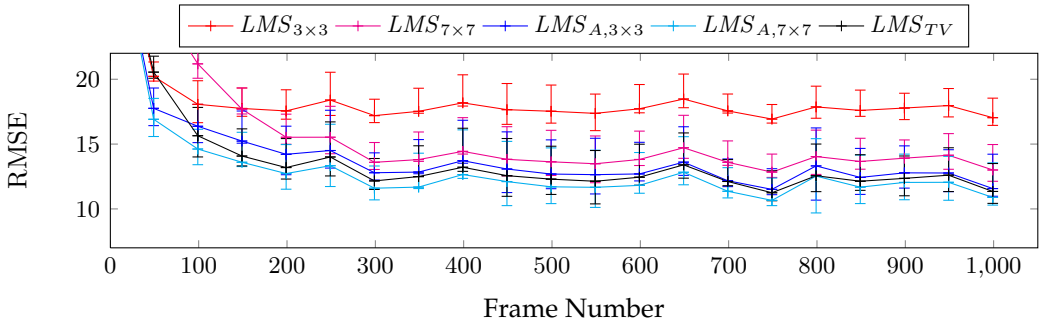


Figure C.7.: The LMS methods' evaluations for the RMSE metric with random frame sequences of 1000 frames length.

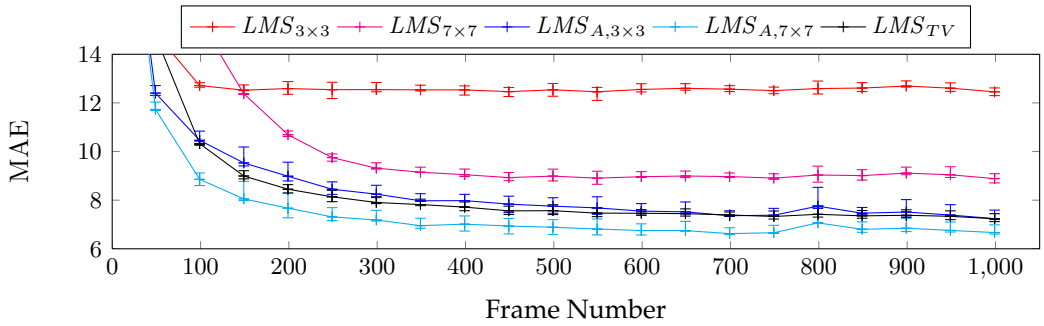


Figure C.8.: The LMS methods' evaluations for the MAE metric with random frame sequences of 1000 frames length.

D. The Maximum Likelihood Appraoch

D.1. Log-Likelihood Energy Transformations

This appendix shows the intermediate steps in the definitions of the energies E_P and E_D according to eq. 6.53 and 6.56 in sec. 6.3 on p. 124. For E_P the transformations go as it follows by using the product to sum conversion of the logarithm as well as the linearity of the sums:

$$
\begin{aligned}
E_P &= -\log f_{\mathbf{DSNU}}(\{b_{i,j}\}) && \text{(D.1)}\\
&= -\log \prod_{i,j} f_{\mathbf{b}_{i,j}}(b_{i,j}) && \text{(D.2)}\\
&= -\log \prod_{i,j} \frac{1}{\sigma_b\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{b_{i,j}-\mu_b}{\sigma_b}\right)^2} && \text{(D.3)}\\
&= \sum_{i,j} -\log \frac{1}{\sigma_b\sqrt{2\pi}} - \log e^{-\frac{1}{2}\left(\frac{b_{i,j}-\mu_b}{\sigma_b}\right)^2} && \text{(D.4)}\\
&= \underbrace{\sum_{i,j} -\log \frac{1}{\sigma_b\sqrt{2\pi}}}_{\text{const.}} + \sum_{i,j} \frac{1}{2}\left(\frac{b_{i,j}-\mu_b}{\sigma_b}\right)^2 && \text{(D.5)}\\
&= \sum_{i,j} \underbrace{\frac{1}{2\sigma_b^2}}_{\beta} \left(b_{i,j}-\mu_b\right)^2 + \text{const.} && \text{(D.6)}\\
&= \sum_{i,j} \beta \left(b_{i,j}-\mu_b\right)^2 + \text{const.} && \text{(D.7)}
\end{aligned}
$$

With the same mathematics as used above, E_D transforms to:

$$
\begin{aligned}
E_D &= -\log f_{\mathbf{I}_{\text{World}}|\mathbf{DSNU}=\{b_{i,j}\}}(I_{\text{World}}) && \text{(D.8)}\\
&= -\log \prod_{i,j}\prod_{t_n} \lambda e^{-\lambda F_{i,j,t_n}} && \text{(D.9)}\\
&= \sum_{i,j}\sum_{t_n} (-\log\lambda - \log e^{-\lambda F_{i,j,t_n}}) && \text{(D.10)}\\
&= \underbrace{\sum_{i,j}\sum_{t_n} -\log\lambda}_{\text{const.}} + \sum_{i,j}\sum_{t_n} \lambda F_{i,j,t_n} && \text{(D.11)}
\end{aligned}
$$

$$= \text{const.} + \sum_{i,j} \sum_{t_n} \lambda F_{i,j,t_n} \tag{D.12}$$

D.2. Gradient Derivations

This appendix shows how to derive the gradient of the energy term E_D, needed in sec. 6.3 on p. 334.

The following transformations basically exploits the linearity of the derivative operator and the product rule, which leads then to trivial derivatives of the sums:

$$\begin{aligned}
\nabla_{\{b_{i,j}\}} E_D &= \nabla_{\{b_{i,j}\}} (\sum_{t_n} \sum_{i,j} \lambda (C_M - b - \chi) \otimes ((C_M - b - \chi) * HP(\hat{L}, \Phi)) + \text{const.}) \\
\partial_{b_{m,n}} E_D &= (\sum_{t_n} \sum_{i,j} \lambda (C_M - b - \chi) \otimes ((C_M - b - \chi) * HP(\hat{L}, \Phi)) + \text{const.})
\end{aligned} \tag{D.13}$$

$$= \sum_{t_n} \lambda \partial_{b_{m,n}} \left(\sum_{(i,j) \in S} (C_{M,i,j} - b_{i,j} - \chi_{i,j}) \right.$$

$$\left. \cdot \sum_{(k,l) \in S} (C_{M,k,l} - b_{k,l} - \chi_{k,l}) \Psi(|k-i|, |l-j|) \right) \tag{D.14}$$

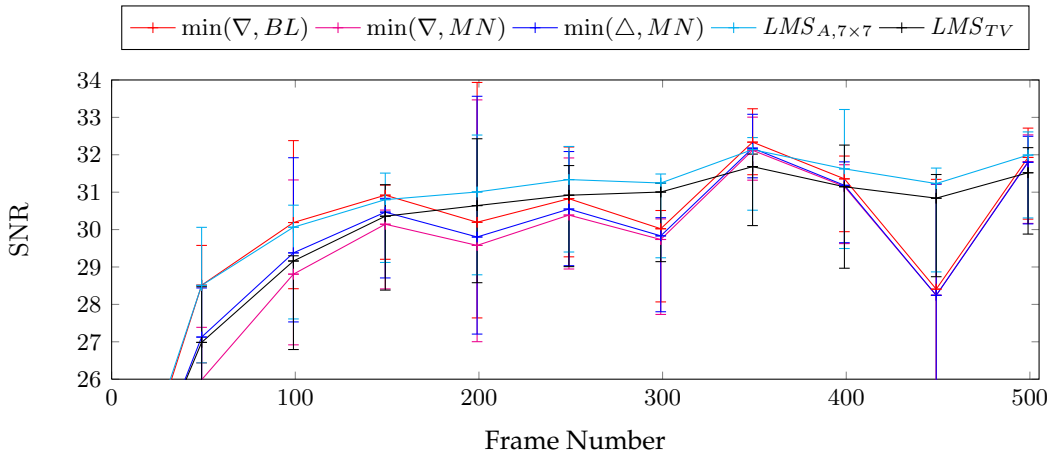
$$= \sum_{t_n} \lambda \left((\partial_{b_{m,n}} \sum_{(i,j) \in S} (C_{M,i,j} - b_{i,j} - \chi_{i,j})) \sum_{(k,l) \in S} (C_{M,k,l} - b_{k,l} - \chi_{k,l}) \Psi(\cdot) \right.$$

$$\left. + \sum_{(i,j) \in S} (C_{M,i,j} - b_{i,j} - \chi_{i,j}) (\partial_{b_{m,n}} \sum_{(k,l) \in S} (C_{M,k,l} - b_{k,l} - \chi_{k,l}) \Psi(\cdot)) \right) \tag{D.15}$$

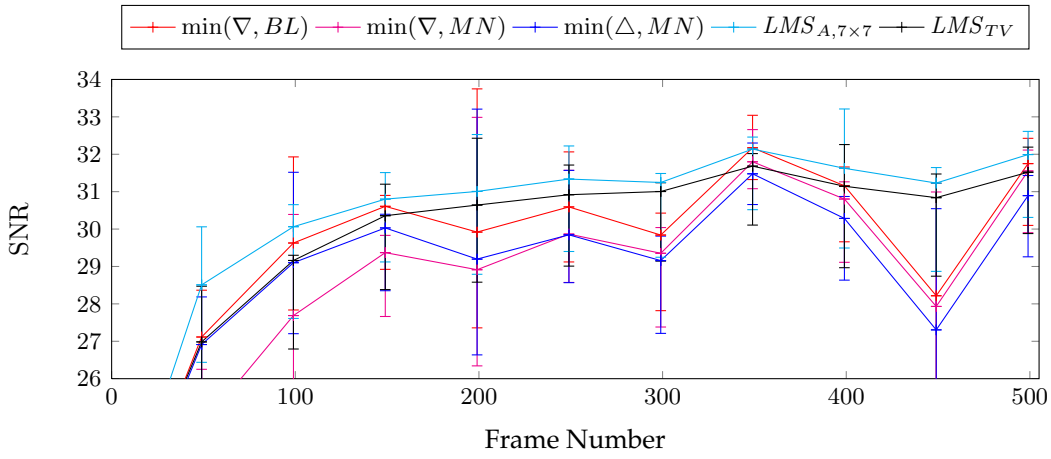
$$= \sum_{t_n} \lambda \left(- \sum_{(k,l) \in S} (C_{M,k,l} - b_{k,l} - \chi_{k,l}) \Psi(|k-m|, |l-n|) \right. \tag{D.16}$$

$$\left. + \sum_{(i,j) \in S} (C_{M,i,j} - b_{i,j} - \chi_{i,j}) (-\Psi(|m-i|, |n-j|)) \right) \tag{D.17}$$

$$= -2\lambda \sum_{t_n} \sum_{(i,j) \in S} (C_{M,i,j} - b_{i,j} - \chi_{i,j}) \Psi(|i-m|, |j-n|)) \tag{D.18}$$



(a) Parameter set: Opt. Random (tab. 6.3)

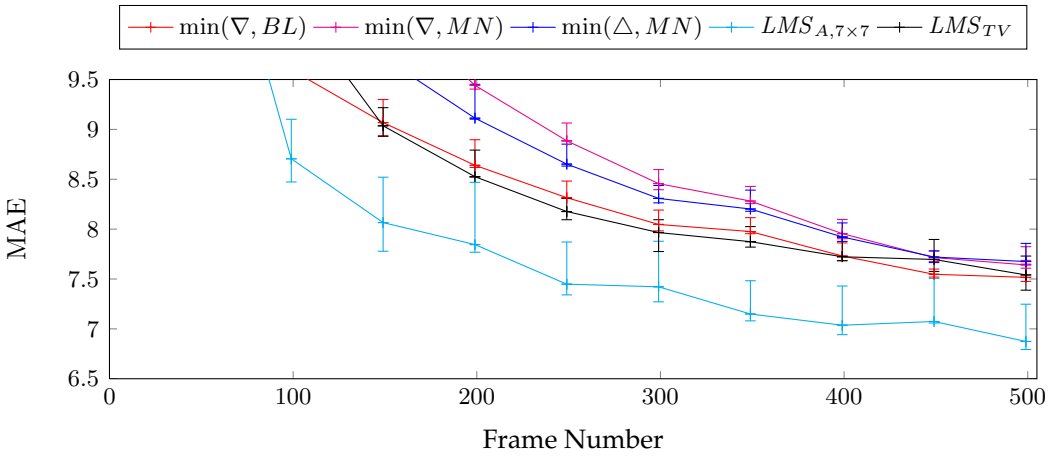


(b) Parameter set: Stat. Params (tab. 6.1)

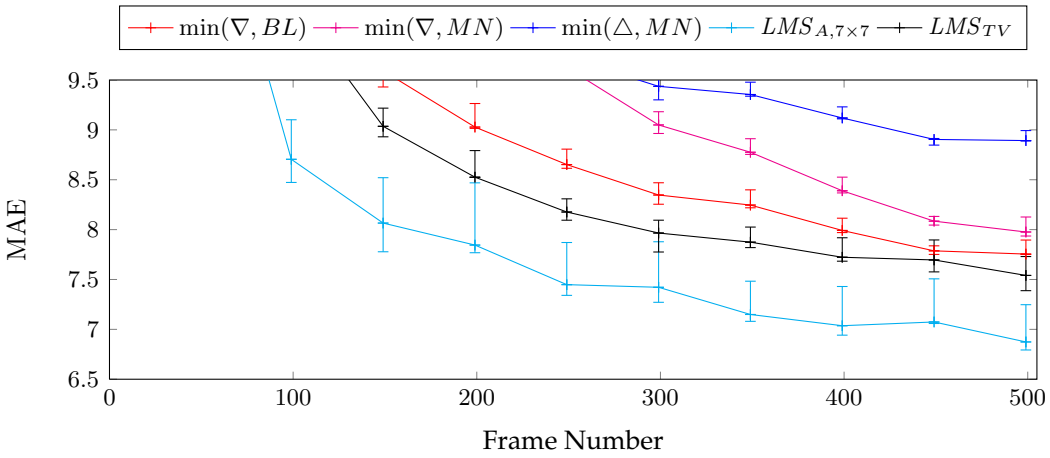
Figure D.1.: Temporal evaluations with the optimized (fig. D.1a) and the statistic (fig. D.1b) parameter sets in the SNR metric for random frame sequences.

D.3. Further Results

This appendix presents additional quality metrics for the temporal performance analysis presented in sec. 6.4.2 on p.129. The discussion can here be directly transferred onto the additionally displayed quality norms.

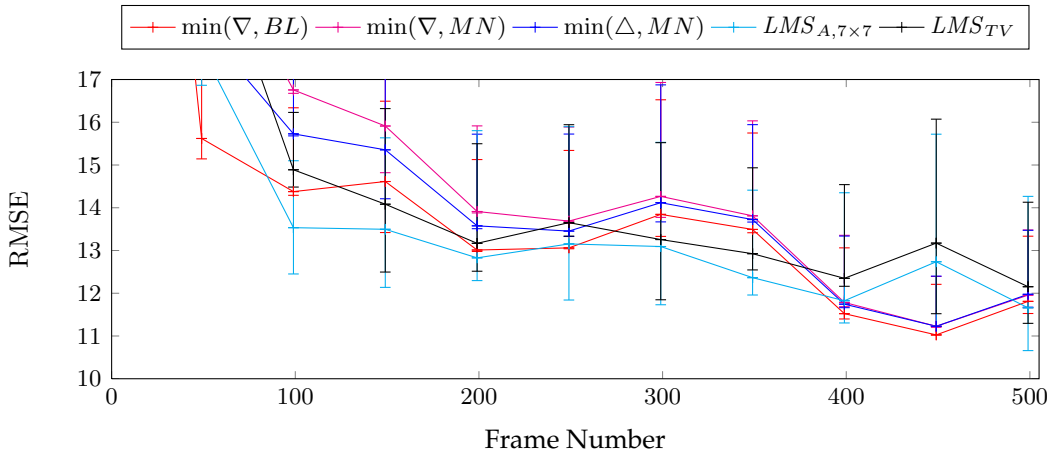


(a) Parameter set: Opt. Random (tab. 6.3)

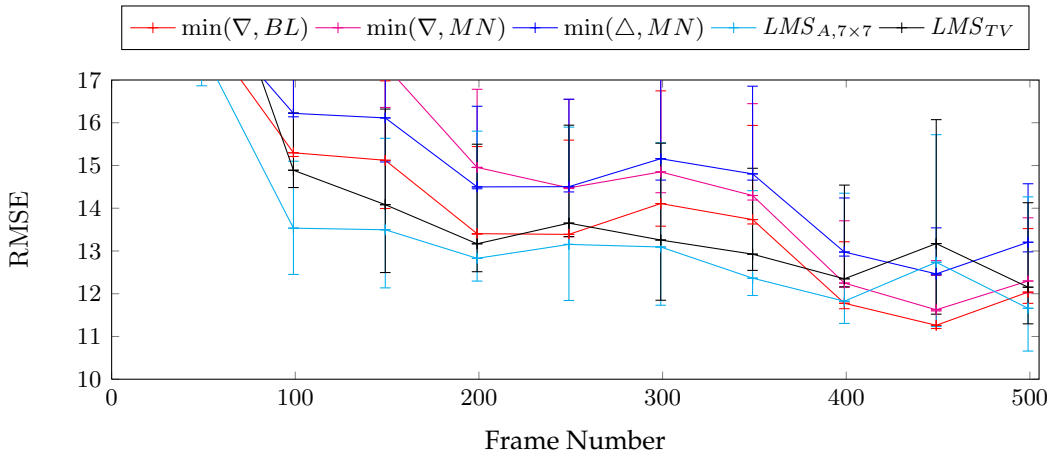


(b) Parameter set: Stat. Params (tab. 6.1)

Figure D.2.: Temporal evaluations with the optimized (fig. D.2a) and the statistic (fig. D.2b) parameter sets in the MAE metric for random frame sequences.



(a) Parameter set: Opt. Random (tab. 6.3)



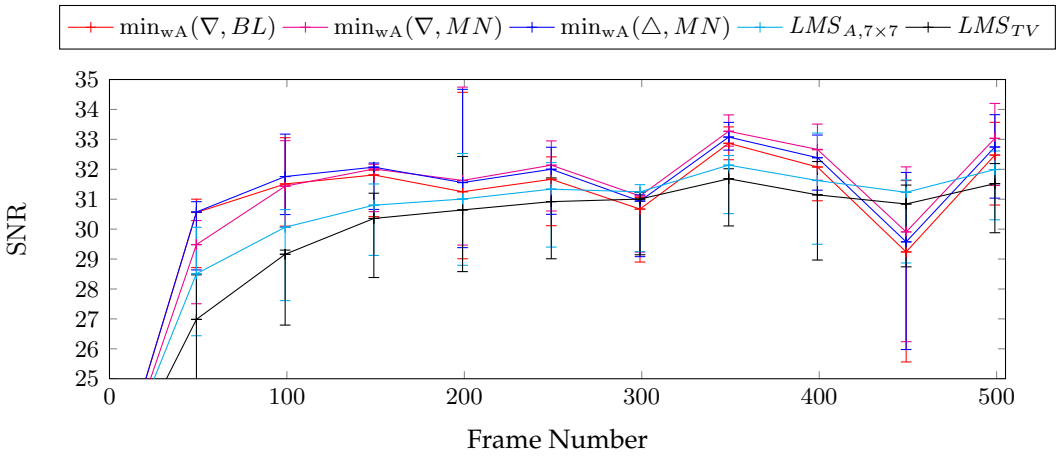
(b) Parameter set: Stat. Params (tab. 6.1)

Figure D.3.: Temporal evaluations with the optimized (fig. D.3a) and the statistic (fig. D.3b) parameter sets in the RMSE metric for random frame sequences.

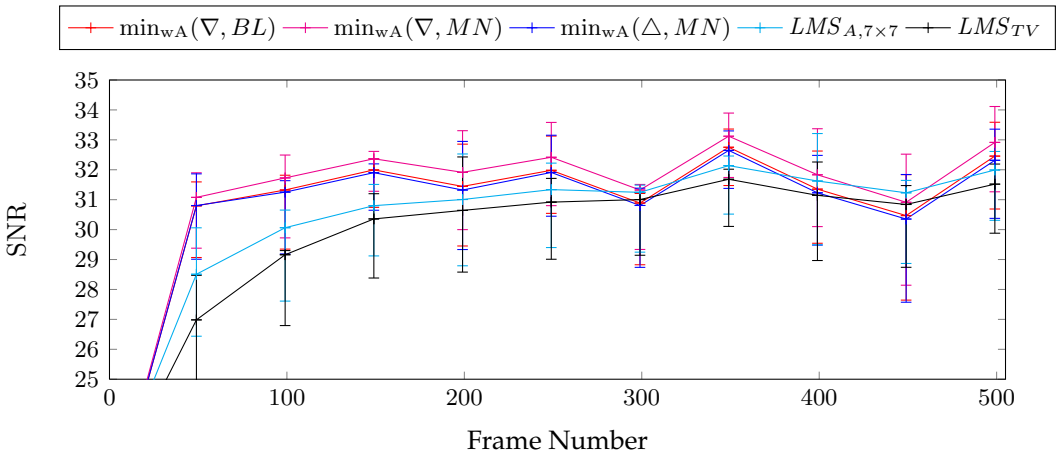
E. The Improved DSNU Estimation Variants

E.1. Further Results for the Weighted Averaging Approach

This appendix presents additional quality metrics for the temporal performance analysis presented in sec. 7.1.3 on p.143. The discussion of the results transfers directly to the here shown quality norms. Furthermore, the differences of the methods show up much better in the discussed quality metric and these results are only shown for readers that are more familiar with these image dependend quality norms.

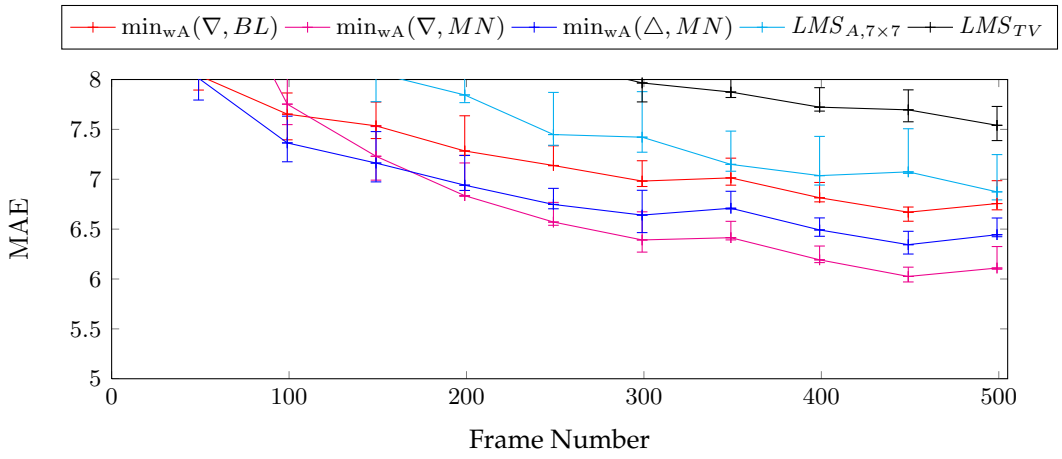


(a) Opt. Random (tab. 7.3)

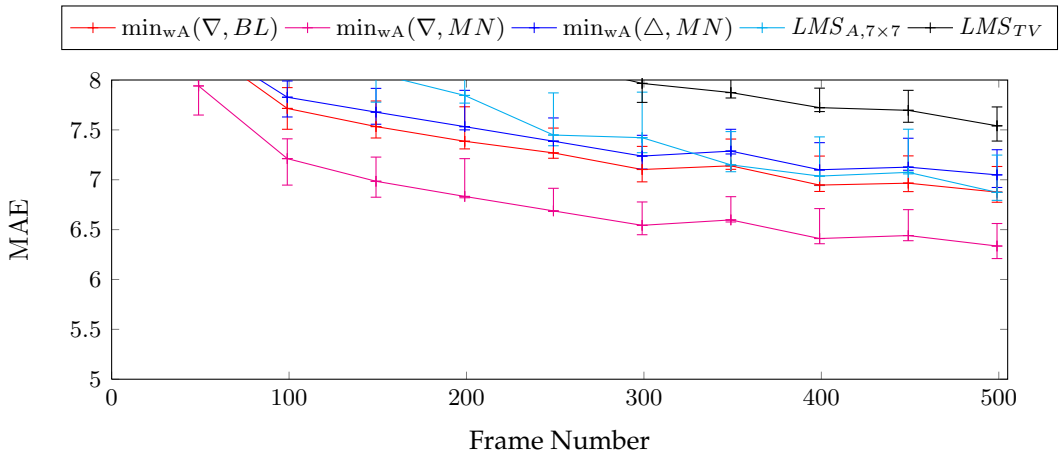


(b) Weighted Stat. (tab. 7.1)

Figure E.1.: Evaluations for random frame sequences with experimentally optimized parameters (fig. E.1a) and the parameter set from the weighted statistic histogram (fig. E.1b) in the SNR metric.

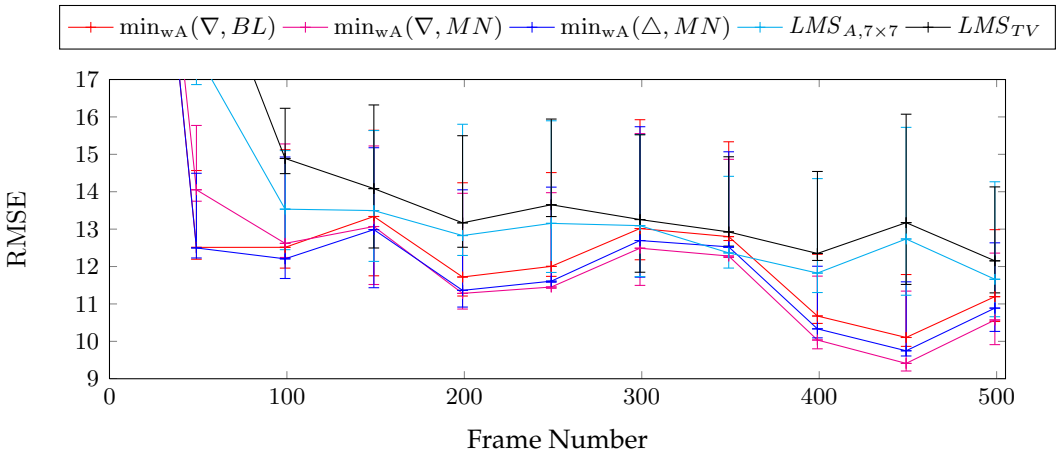


(a) Opt. Random (tab. 7.3)

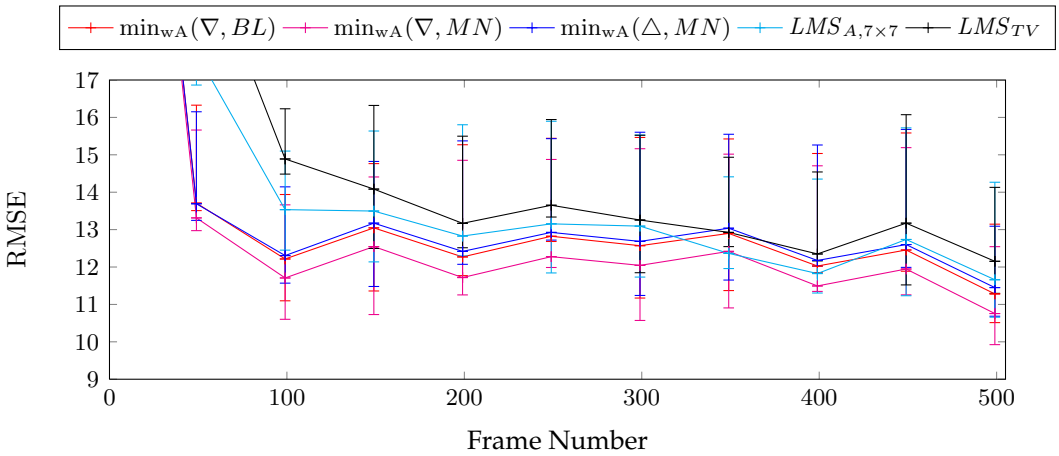


(b) Weighted Stat. (tab. 7.1)

Figure E.2.: Evaluations for random frame sequences with experimentally optimized parameters (fig. E.2a) and the parameter set from the weighted statistic histogram (fig. E.2b) in the MAE metric.



(a) Opt. Random (tab. 7.3)



(b) Weighted Stat. (tab. 7.1)

Figure E.3.: Evaluations for random frame sequences with experimentally optimized parameters (fig. E.3a) and the parameter set from the weighted statistic histogram (fig. E.3b) in the RMSE metric.

E.2. Further Results for the CNN Approach

This appendix presents additional quality metrics for the temporal performance analysis presented in sec. 7.2.3 on p.156. These results are given without further discussion, as the general discussion of these results transfers directly to the here shown quality norms. Furthermore, the differences of the methods show up much better in the discussed quality metric and these results are only shown for readers that are more familiar with these image dependend quality norms.

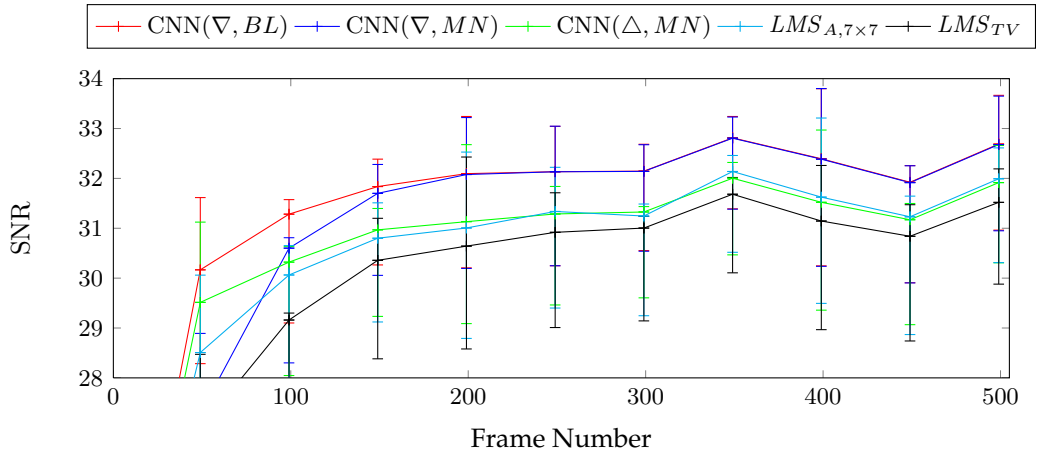


Figure E.4.: Methods evaluations with experimental optimized parameters (tab. 7.6) for the SNR metric and random frame sequences as input

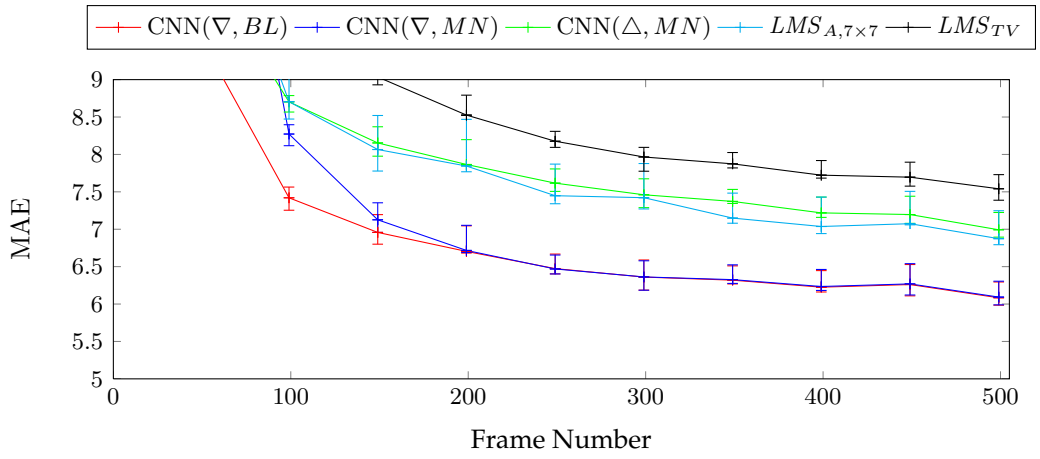


Figure E.5.: Methods evaluations with experimental optimized parameters (tab. 7.6) for the MAE metric and random frame sequences as input

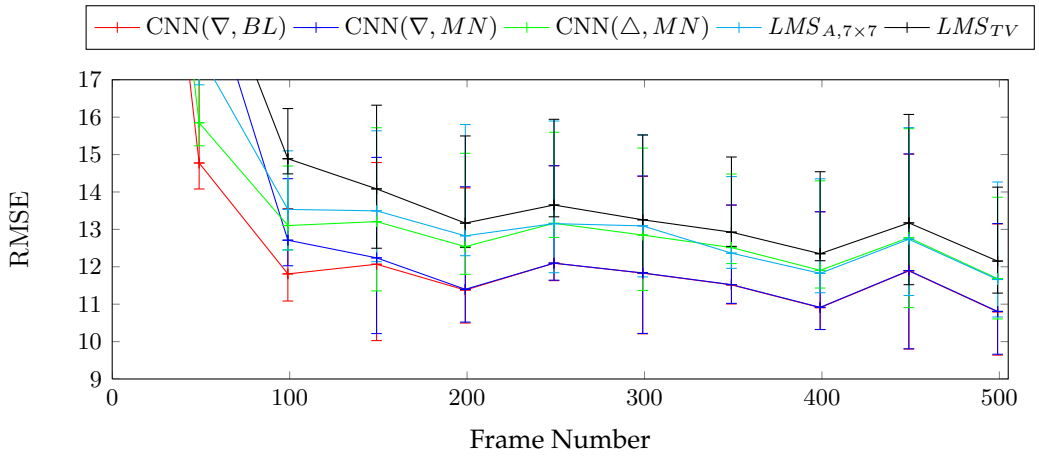
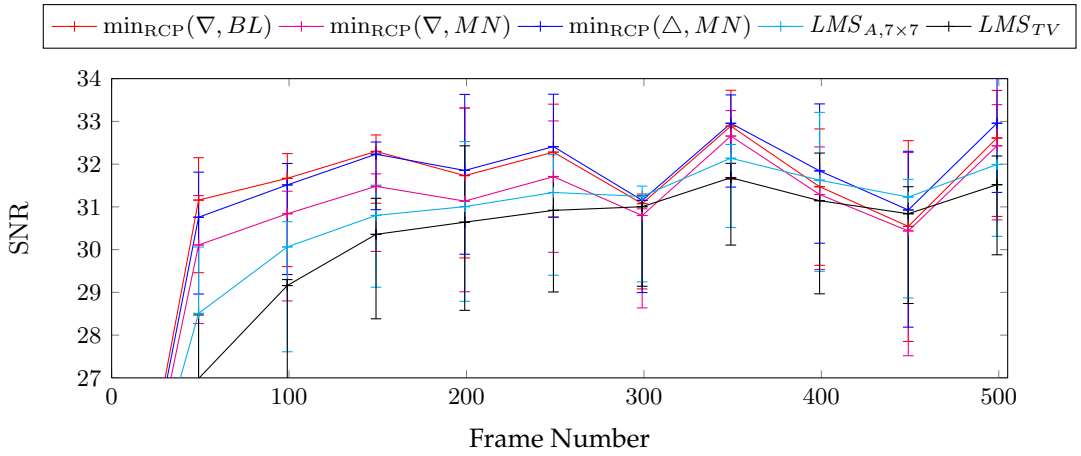


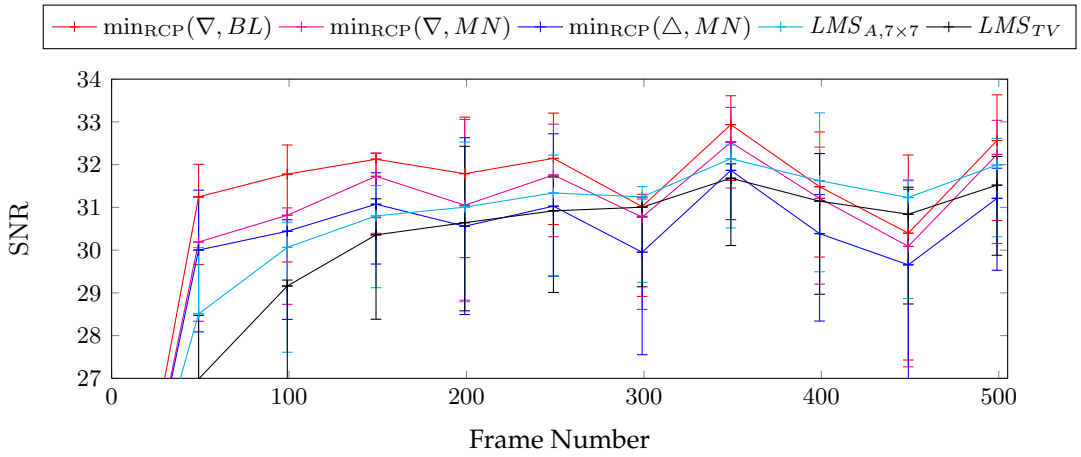
Figure E.6.: Methods evaluations with experimental optimized parameters (tab. 7.6) for the RMSE metric and random frame sequences as input

E.3. Further Results for the RCP Approach

This appendix presents additional quality metrics for the temporal performance analysis presented in sec. 7.3.3 on p.169. These results are given without further discussion, as the general discussion of these results transfers directly to the here shown quality norms. Furthermore, the differences of the methods show up much better in the discussed quality metric and these results are only shown for readers that are more familiar with these image dependend quality norms.

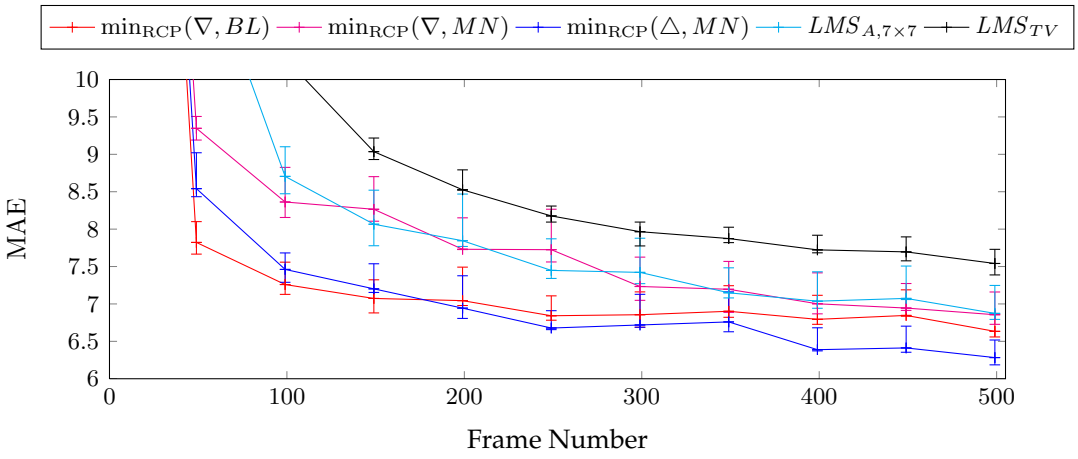


(a) Parameter set: Opt. RCP Random (tab. 7.9)

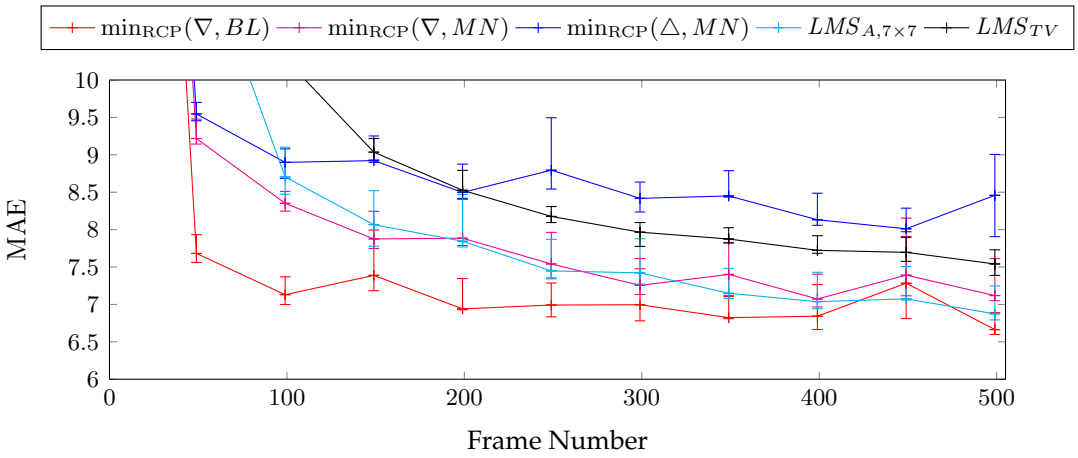


(b) Parameter set: Weighted Stat. (tab. 7.1)

Figure E.7.: Evaluations for random frame sequences with experimentally optimized parameters (fig. E.7a) and the parameter set from the weighted statistic histogram (fig. E.7b) in the SNR metric.

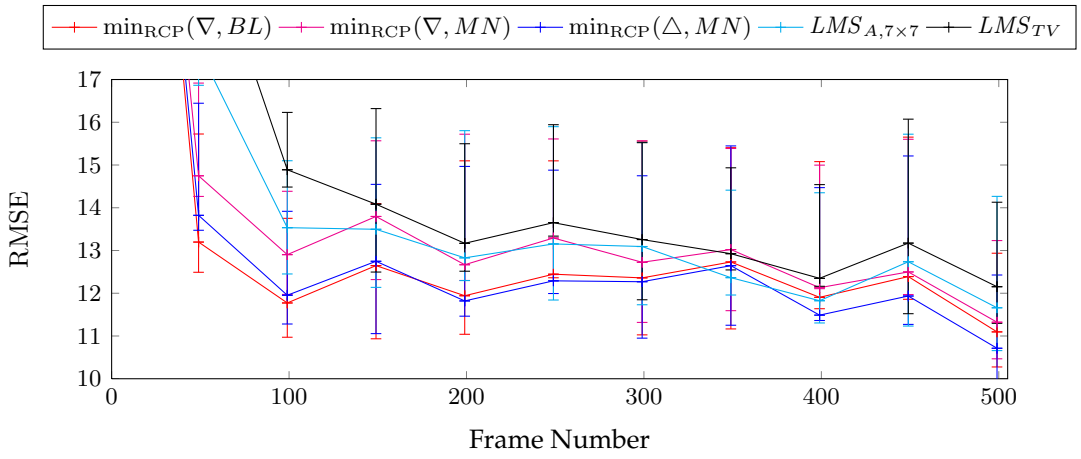


(a) Parameter set: Opt. RCP Random (tab. 7.9)

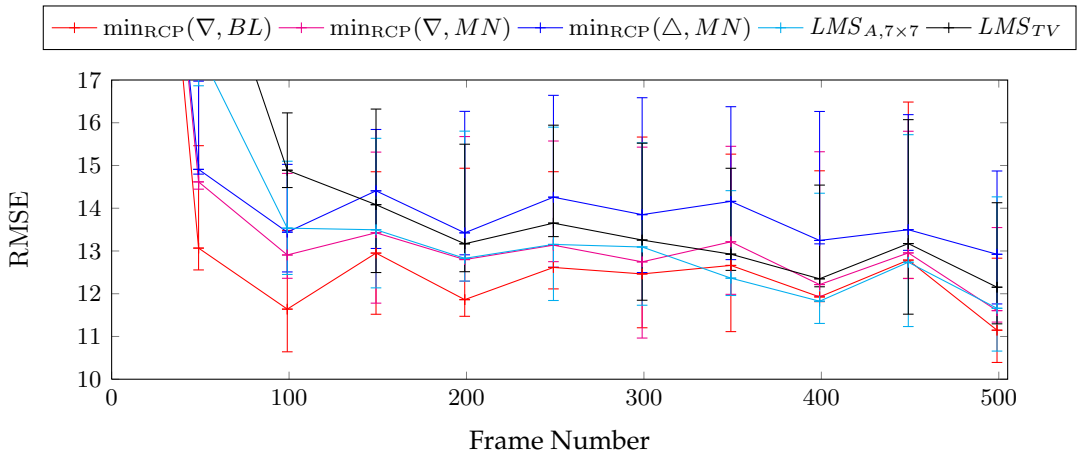


(b) Parameter set: Weighted Stat. (tab. 7.1)

Figure E.8.: Evaluations for random frame sequences with experimentally optimized parameters (fig. E.8a) and the parameter set from the weighted statistic histogram (fig. E.8b) in the MAE metric.



(a) Parameter set: Opt. RCP Random (tab. 7.9)



(b) Parameter set: Weighted Stat. (tab. 7.1)

Figure E.9.: Evaluations for random frame sequences with experimentally optimized parameters (fig. E.9a) and the parameter set from the weighted statistic histogram (fig. E.9b) in the RMSE metric.

F. The Combined Estimation Method

F.1. Log-Likelihood Energy Transformations

This appendix shows the intermediate steps of the transformation of the energy E_P in sec. 8.3 p. 179. Unsing the definitions of the eq. 8.5f. the energy E_P transforms as it follows:

$$
\begin{aligned}
E_P &= -\log f_{\mathbf{NU}}(\{a_{i,j}, b_{i,j}\}) && \text{(F.1)}\\
&= -\log \prod_{i,j} f_{\mathbf{a}_{i,j}}(a_{i,j}) f_{\mathbf{b}_{i,j}}(b_{i,j}) && \text{(F.2)}\\
&= -\log \prod_{i,j} \frac{1}{\sigma_b\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{b_{i,j}-\mu_b}{\sigma_b}\right)^2} \frac{1}{\sigma_a\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{a_{i,j}-\mu_a}{\sigma_a}\right)^2} && \text{(F.3)}\\
&= \sum_{i,j} -\log \frac{1}{\sigma_b\sqrt{2\pi}} - \log e^{-\frac{1}{2}\left(\frac{b_{i,j}-\mu_b}{\sigma_b}\right)^2} - \log \frac{1}{\sigma_a\sqrt{2\pi}} - \log e^{-\frac{1}{2}\left(\frac{a_{i,j}-\mu_a}{\sigma_a}\right)^2} && \text{(F.4)}\\
&= \underbrace{\sum_{i,j} -\log \frac{1}{\sigma_b\sqrt{2\pi}} - \log \frac{1}{\sigma_a\sqrt{2\pi}}}_{\text{const.}} + \sum_{i,j} \frac{1}{2}\left(\frac{b_{i,j}-\mu_b}{\sigma_b}\right)^2 \frac{1}{2}\left(\frac{a_{i,j}-\mu_a}{\sigma_a}\right)^2 && \text{(F.5)}\\
&= \sum_{i,j} \underbrace{\frac{1}{2\sigma_b^2}}_{\beta} (b_{i,j}-\mu_b)^2 + \sum_{i,j} \underbrace{\frac{1}{2\sigma_a^2}}_{\alpha} (a_{i,j}-\mu_a)^2 + \text{const.} && \text{(F.6)}\\
&= \beta \sum_{i,j} (b_{i,j}-\mu_b)^2 + \alpha \sum_{i,j} (a_{i,j}-\mu_a)^2 + \text{const.} && \text{(F.7)}
\end{aligned}
$$

F.2. Gradient calculations

F.2.1. The gradient $\nabla_{\{b_{i,j}\}} E_D$

This appendix shows the intermediate steps of the derivation of the gradient in sec. 8.3 p. 179. The transformations go according to sec. 6.3 and use the chain rule and the linearity of the derivatives:

$$\begin{aligned}
\nabla_{\{b_{i,j}\}} E_D =& \nabla_{\{b_{i,j}\}} \sum_{t_n} \sum_{i,j} \lambda (C_M - b - \chi) \oslash a \\
& \otimes \left(((C_M - b - \chi) \oslash a) * HP(\hat{L}, \Phi) \right) + \text{const.} \\
\partial_{b_{m,n}} E_D =& \partial_{b_{m,n}} (\sum_{t_n} \sum_{i,j} \lambda (C_M - b - \chi) \oslash a \\
& \otimes ((C_M - b - \chi) \oslash a * HP(\hat{L}, \Phi)) + \text{const.}) \qquad \text{(F.8)} \\
=& \sum_{t_n} \lambda \partial_{b_{m,n}} \Bigg(\sum_{(i,j)\in S} \frac{1}{a_{i,j}} (C_{M,i,j} - b_{i,j} - \chi_{i,j}) \\
& \cdot \sum_{(k,l)\in S} \frac{1}{a_{k,l}} (C_{M,k,l} - b_{k,l} - \chi_{k,l}) \Psi(|k-i|, |l-j|) \Bigg) \qquad \text{(F.9)} \\
=& \sum_{t_n} \lambda \Bigg((\partial_{b_{m,n}} \sum_{(i,j)\in S} \frac{1}{a_{i,j}} (C_{M,i,j} - b_{i,j} - \chi_{i,j})) \\
& \cdot \sum_{(k,l)\in S} \frac{1}{a_{k,l}} (C_{M,k,l} - b_{k,l} - \chi_{k,l}) \Psi(\cdot) + \sum_{(i,j)\in S} \frac{1}{a_{i,j}} (C_{M,i,j} - b_{i,j} - \chi_{i,j}) \\
& \cdot (\partial_{b_{m,n}} \sum_{(k,l)\in S} \frac{1}{a_{k,l}} (C_{M,k,l} - b_{k,l} - \chi_{k,l}) \Psi(\cdot)) \Bigg) \qquad \text{(F.10)} \\
=& \sum_{t_n} \lambda \Bigg(\frac{-1}{a_{m,n}} \sum_{(k,l)\in S} \frac{1}{a_{k,l}} (C_{M,k,l} - b_{k,l} - \chi_{k,l}) \Psi(|k-m|, |l-n|) \\
& + \sum_{(i,j)\in S} (C_{M,i,j} - b_{i,j} - \chi_{i,j}) \frac{-1}{a_{m,n}} \Psi(|m-i|, |n-j|)) \Bigg) \qquad \text{(F.11)} \\
=& -2\lambda \sum_{t_n} \frac{1}{a_{m,n}} \sum_{(i,j)\in S} \frac{1}{a_{i,j}} (C_{M,i,j} - b_{i,j} - \chi_{i,j}) \Psi(|i-m|, |j-n|)) \\
& \qquad \text{(F.12)}
\end{aligned}$$

F.2.2. The gradient $\nabla_{\{a_{i,j}\}} E_D$

This appendix shows the intermediate steps of the derivation of the gradient in sec. 8.3.2 p. 182. The transformations go according to sec. 6.3 and use the chain rule and the linearity of the derivatives:

$$\nabla_{\{a_{i,j}\}} E_D = \nabla_{\{a_{i,j}\}} \sum_{t_n} \sum_{i,j} \lambda (C_M - b - \chi) \oslash a \\ \otimes ((C_M - b - \chi) \oslash a * HP(\hat{L}, \Phi)) + \text{const.} \tag{F.13}$$

$$\partial_{a_{m,n}} E_D = \partial_{a_{m,n}} (\sum_{t_n} \sum_{i,j} \lambda (C_M - b - \chi) \oslash a \\ \otimes ((C_M - b - \chi) \oslash a * HP(\hat{L}, \Phi)) + \text{const.}) \tag{F.14}$$

$$= \sum_{t_n} \lambda \partial_{a_{m,n}} \Bigg(\sum_{(i,j)\in S} \frac{1}{a_{i,j}} (C_{M,i,j} - b_{i,j} - \chi_{i,j}) \\ \cdot \sum_{(k,l)\in S} \frac{1}{a_{k,l}} (C_{M,k,l} - b_{k,l} - \chi_{k,l}) \Psi(|k-i|, |l-j|) \Bigg) \tag{F.15}$$

$$= \sum_{t_n} \lambda \Bigg((\partial_{a_{m,n}} \sum_{(i,j)\in S} \frac{1}{a_{i,j}} (C_{M,i,j} - b_{i,j} - \chi_{i,j})) \\ \cdot \sum_{(k,l)\in S} \frac{1}{a_{k,l}} (C_{M,k,l} - b_{k,l} - \chi_{k,l}) \Psi(\cdot) + \sum_{(i,j)\in S} \frac{1}{a_{i,j}} (C_{M,i,j} - b_{i,j} - \chi_{i,j}) \\ \cdot (\partial_{a_{m,n}} \sum_{(k,l)\in S} \frac{1}{a_{k,l}} (C_{M,k,l} - b_{k,l} - \chi_{k,l}) \Psi(\cdot)) \Bigg) \tag{F.16}$$

$$= 2\lambda \sum_{t_n} \Bigg((\partial_{a_{m,n}} \sum_{(i,j)\in S} \frac{1}{a_{i,j}} (C_{M,i,j} - b_{i,j} - \chi_{i,j})) \\ \cdot \sum_{(k,l)\in S} \frac{1}{a_{k,l}} (C_{M,k,l} - b_{k,l} - \chi_{k,l}) \Psi(|i-m|, |j-n|) \Bigg) \tag{F.17}$$

$$= 2\lambda \sum_{t_n} \Bigg(\frac{-1}{a_{m,n}^2} (C_{M,m,n} - b_{m,n} - \chi_{m,n}) \\ \cdot \sum_{(k,l)\in S} \frac{1}{a_{k,l}} (C_{M,k,l} - b_{k,l} - \chi_{k,l}) \Psi(\cdot) \Bigg) \tag{F.18}$$

$$= -2\lambda T_N \frac{1}{T_N} \sum_{t_n} \Bigg(\frac{1}{a_{m,n}^2 a_{k,l}} \sum_{(k,l)\in S} (C_{M,m,n} - b_{m,n} - \chi_{m,n}) \\ \cdot (C_{M,k,l} - b_{k,l} - \chi_{k,l}) \Psi(|i-m|, |j-n|) \Bigg) \tag{F.19}$$

And next the intermediate steps after the temporal averaging follow (used in eq. 8.41 p.183):

$$\begin{aligned}
\partial_{a_{m,n}} E_D &= -2\lambda T_N \Big(\frac{1}{a_{m,n}^2 a_{k,l}} \Big(\sum_{(k,l)\in S} A_{M,m,n,k,l}^2 - A_{M,m,n} b_{k,l} - \\
&\qquad b_{m,n} A_{M,k,l} + b_{m,n} b_{k,l} \Big) \Psi(\cdot) \Big) && \text{(F.20)} \\
&= -2\lambda T_N \Big(\frac{1}{a_{m,n}^2 a_{k,l}} \Big(\sum_{(k,l)\in S} A_{M,m,n,k,l}^2 - A_{M,m,n} A_{M,k,l} \\
&\quad + A_{M,m,n} A_{M,k,l} - A_{M,m,n} b_{k,l} - b_{m,n} A_{M,k,l} + b_{m,n} b_{k,l} \Big) \Psi(\cdot) \Big) && \text{(F.21)} \\
&= -2\lambda T_N \Big(\frac{1}{a_{m,n}^2 a_{k,l}} \Big(\sum_{(k,l)\in S} A_{M,m,n,k,l}^2 - A_{M,m,n} A_{M,k,l} \\
&\quad + \big((A_{M,m,n} - b_{m,n})(A_{M,k,l} - b_{k,l}) \big) \Psi(\cdot) \Big) && \text{(F.22)}
\end{aligned}$$

G. Further Low Valued Nonuniformity Evaluations

This appendix shows further temporal evaluations for the low valued nonuniformity image material. The below sections name the methods and the discussion follows as described in the corresponding subsections in the thesis (see sec. 10.2). A further discussion is therefore omitted. Additionally are the results of the Photonfocus MV-D640-66-CL-LinLog camera (see sec. 4.3.4), as this camera does not obey the assumed sensor model and furthermore does not result in discussable results. As seen below, for some of the evaluations even better correction performances of up to $40\,\%$ have been achieved, compared the ones presented in the chapter.

G.1. For the Reference Methods

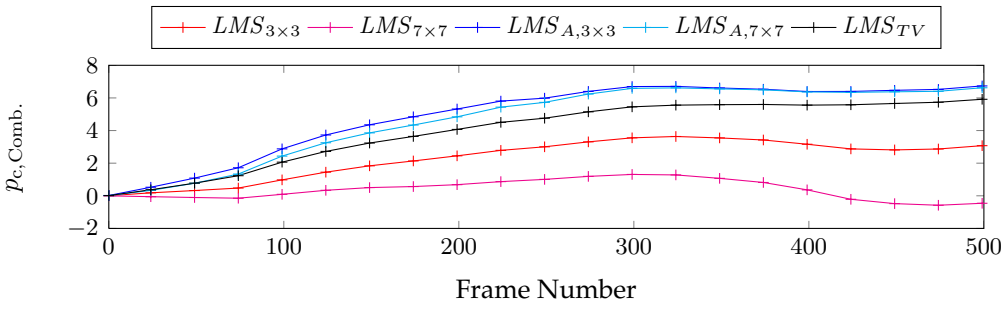


Figure G.1.: Evaluation for the Basler A602f with $t_{exp.}$=10 ms @ 22 °C

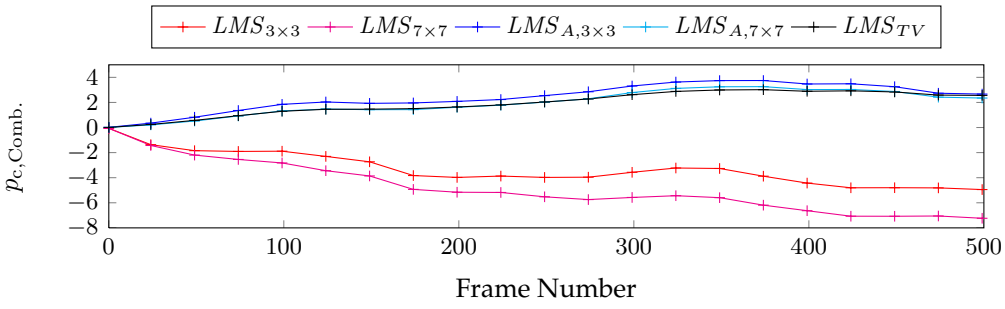


Figure G.2.: Evaluation for the Basler A602f with $t_{exp.}$=1 ms @ 30 °C

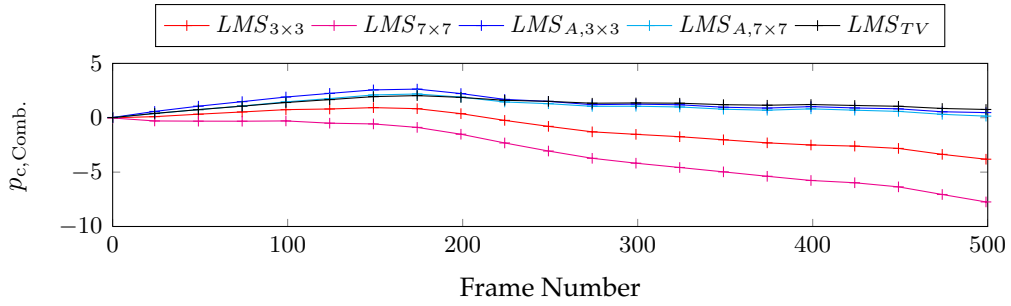


Figure G.3.: Evaluation for the Basler A602f with $t_{exp.}$=10 ms @ 30 °C

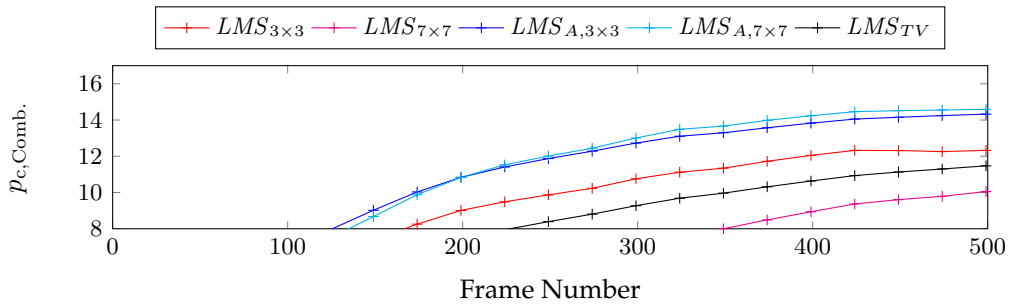


Figure G.4.: Evaluation for the Basler A602f with $t_{exp.}$=1 ms @ 50 °C

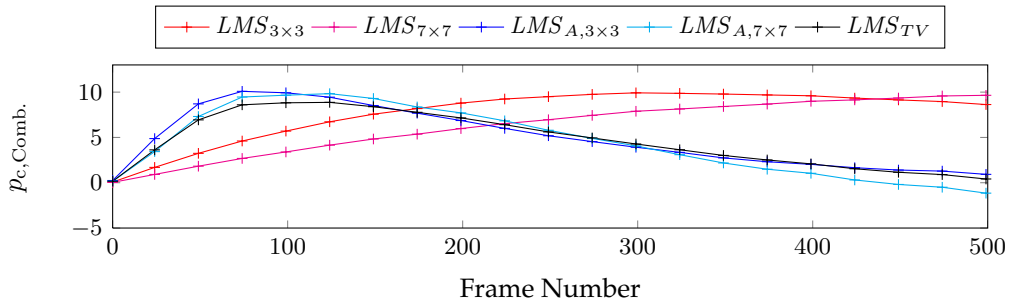


Figure G.5.: Evaluation for the MV-640-66-CL with $t_{exp.}$=10 ms @ 25 °C

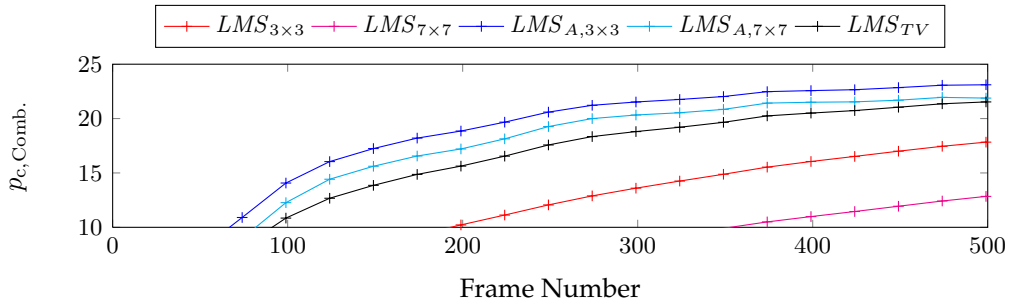


Figure G.6.: Evaluation for the MV-640-66-CL with $t_{exp.}$=5 ms @ 50 °C

G.2. For the Weighted Average Improvement

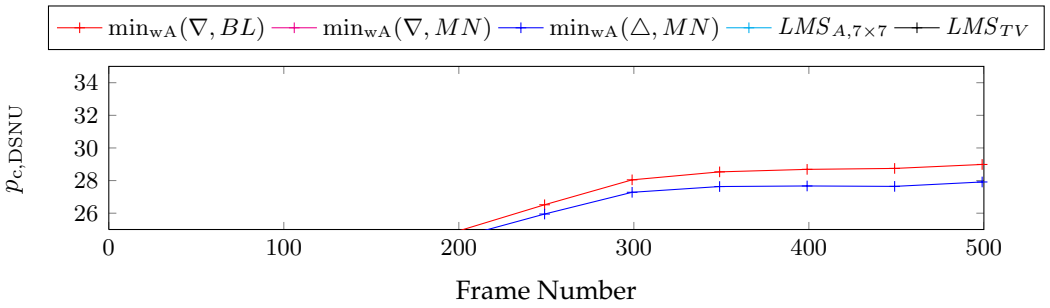


Figure G.7.: Evaluation for the Basler A602f with $t_{exp.}$=10 ms @ 22 °C

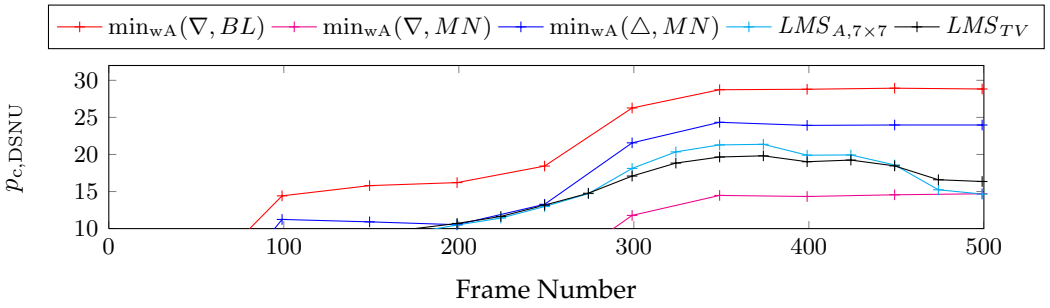


Figure G.8.: Evaluation for the Basler A602f with $t_{exp.}$=1 ms @ 30 °C

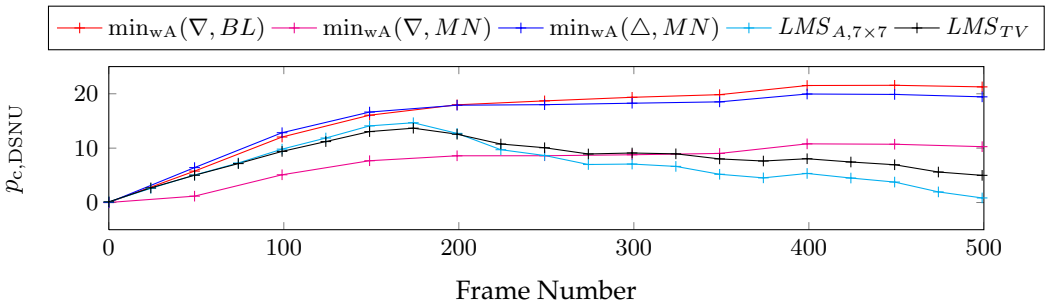


Figure G.9.: Evaluation for the Basler A602f with $t_{exp.}$=10 ms @ 30 °C

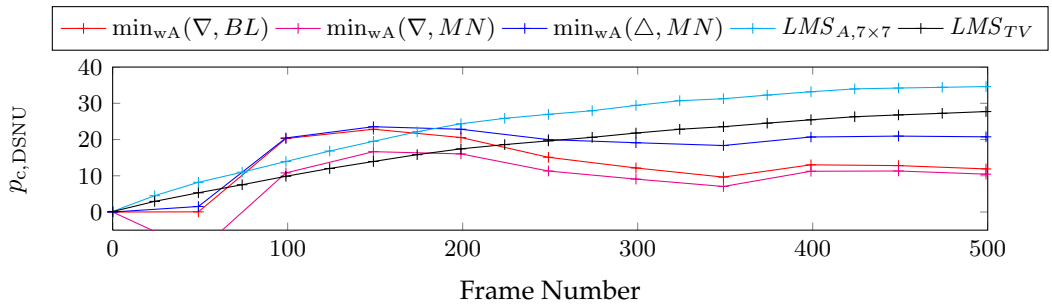


Figure G.10.: Evaluation for the Basler A602f with $t_{exp.}$=1 ms @ 50 °C

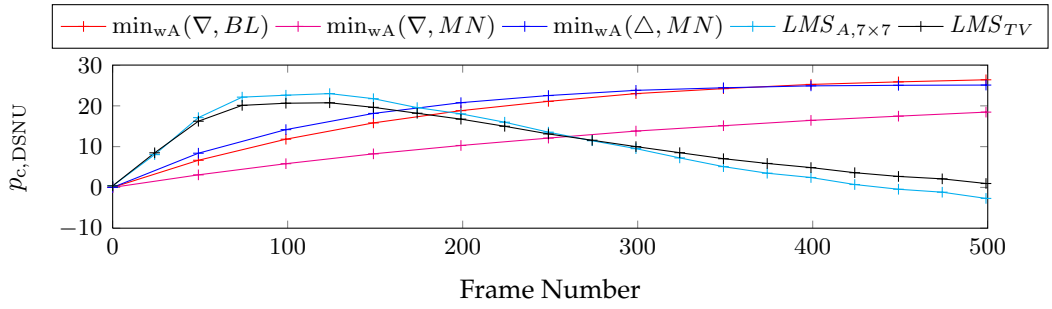


Figure G.11.: Evaluation for the MV-640-66-CL with $t_{exp.}$=10 ms @ 25 °C

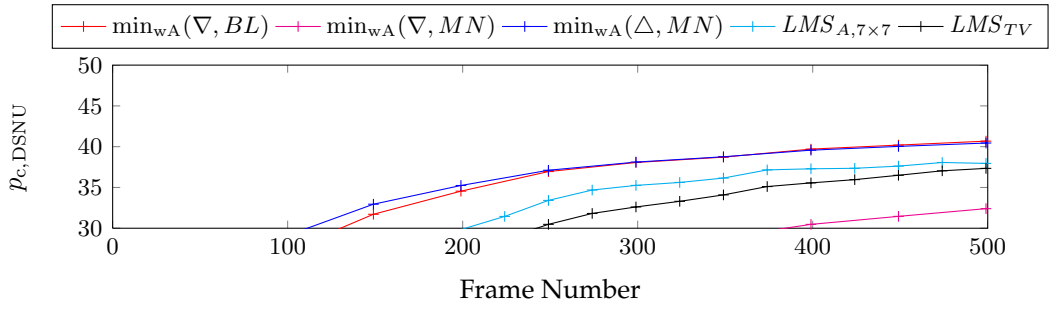


Figure G.12.: Evaluation for the MV-640-66-CL with $t_{exp.}$=5 ms @ 50 °C

G.3. For the CNN-Type Methods

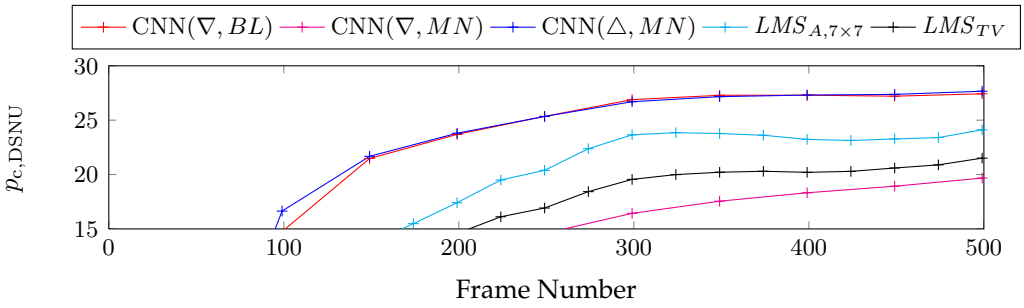


Figure G.13.: Evaluation for the Basler A602f with $t_{exp.}$=10 ms @ 22 °C

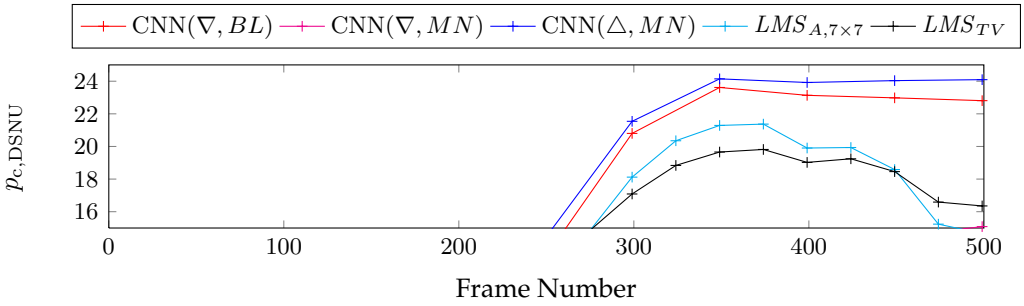


Figure G.14.: Evaluation for the Basler A602f with $t_{exp.}$=1 ms @ 30 °C

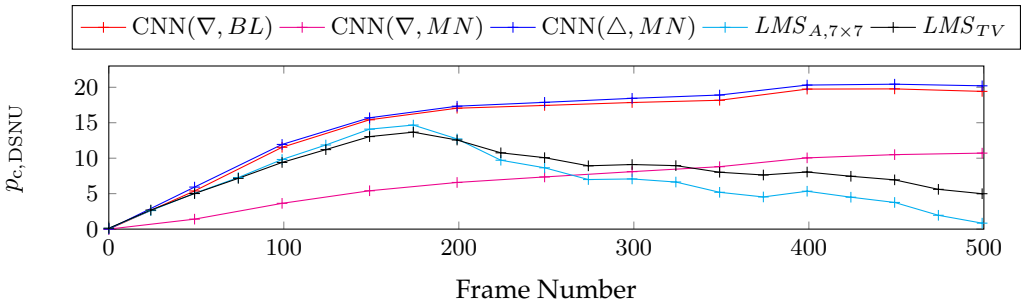


Figure G.15.: Evaluation for the Basler A602f with $t_{exp.}$=10 ms @ 30 °C

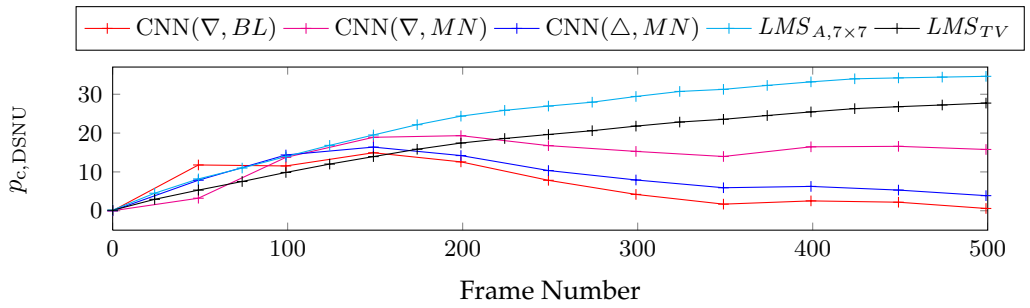


Figure G.16.: Evaluation for the Basler A602f with $t_{exp.}$=1 ms @ 50 °C

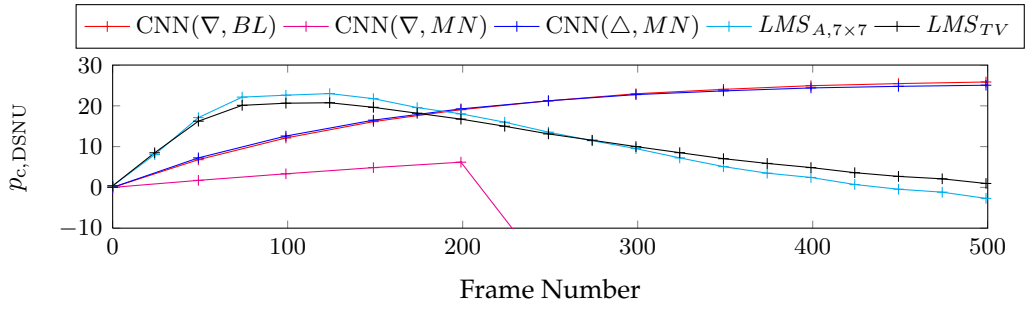


Figure G.17.: Evaluation for the MV-640-66-CL with $t_{exp.}$=10 ms @ 25 °C

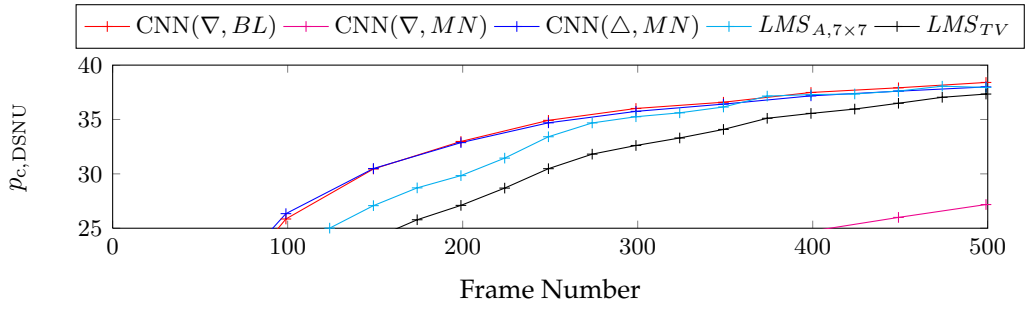


Figure G.18.: Evaluation for the MV-640-66-CL with $t_{exp.}$=5 ms @ 50 °C

G.4. For the RCP Methods

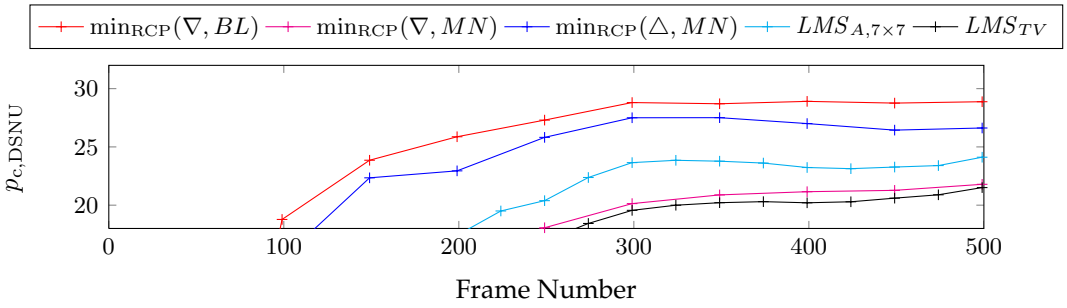


Figure G.19.: Evaluation for the Basler A602f with $t_{exp.}$=10 ms @ 22 °C

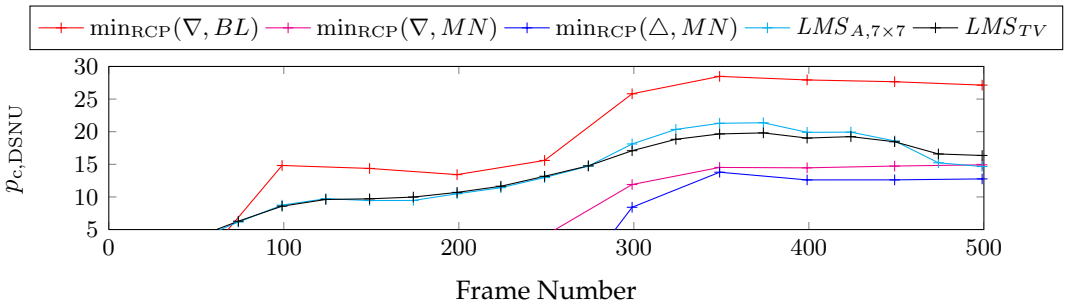


Figure G.20.: Evaluation for the Basler A602f with $t_{exp.}$=1 ms @ 30 °C

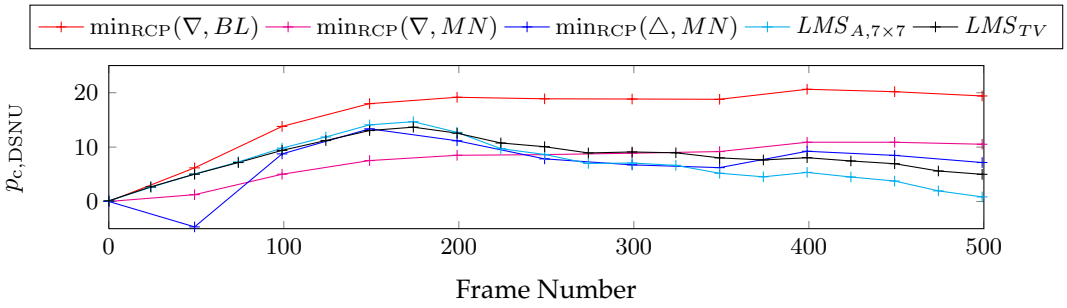


Figure G.21.: Evaluation for the Basler A602f with $t_{exp.}$=10 ms @ 30 °C

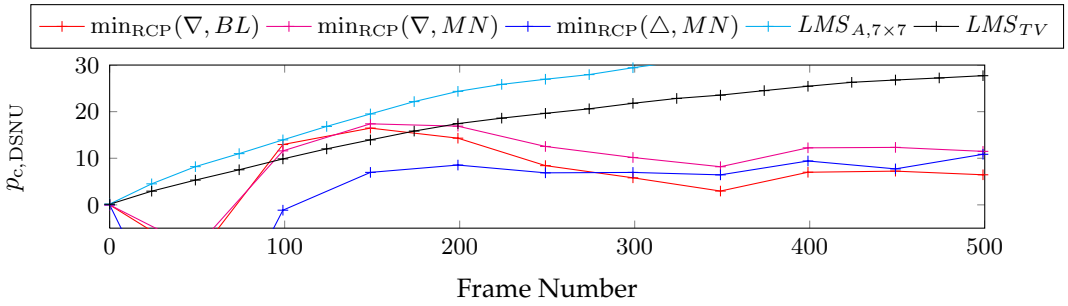


Figure G.22.: Evaluation for the Basler A602f with $t_{exp.}$=1 ms @ 50 °C

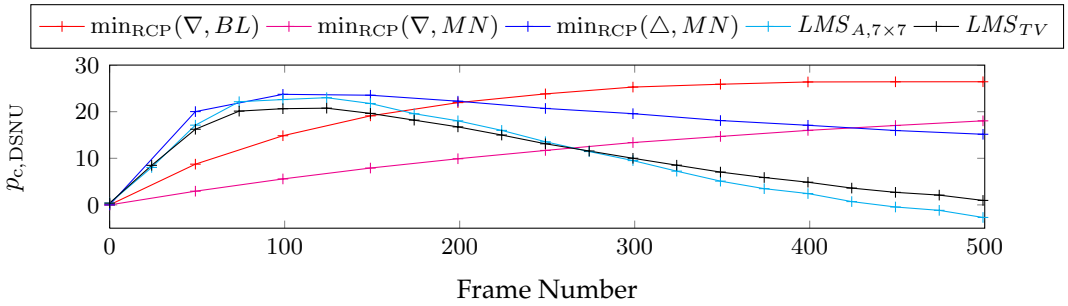


Figure G.23.: Evaluation for the MV-640-66-CL with $t_{exp.}$=10 ms @ 25 °C

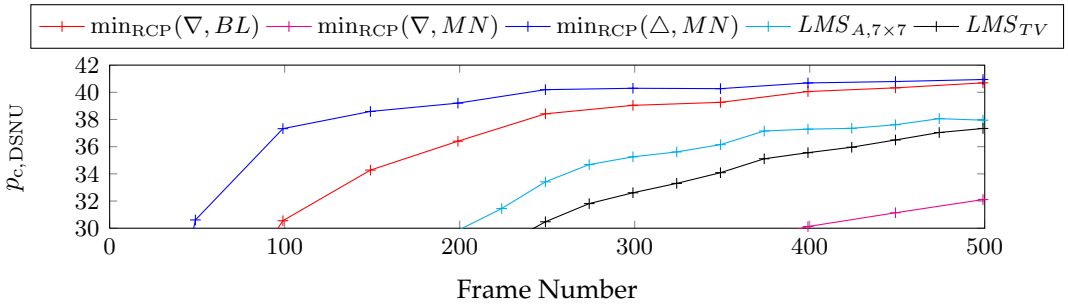


Figure G.24.: Evaluation for the MV-640-66-CL with $t_{exp.}$=5 ms @ 50 °C