

Entwicklung eines Informationssystems für Laserscannerdaten mit OpenSource-Software

Bernhard HÖFLE, Thomas GEIST, Armin HELLER und Johann STÖTTER

Dieser Beitrag wurde nach Begutachtung durch das Programmkomitee als „reviewed paper“ angenommen.

Zusammenfassung

In der Geographie wird das flugzeuggestützte Laserscanning erst in den letzten Jahren verstärkt als Verfahren für die Geländeaufnahme angewandt. Die hochauflösende und genaue flächendeckende Aufnahme der Erdoberfläche bietet neue Möglichkeiten für die Analyse von räumlichen Phänomenen. Die große Datendichte und die damit verbundene Datenmenge der Laserscannerdaten stellen aber auch eine Herausforderung für vektorbasierte Verfahren dar, die keine Reduzierung der Punkte anstreben. Die Zusammenführung eines Geographischen Informationssystems (GIS), einer räumlichen Datenbank und einer Statistiksoftware im OpenSource-Bereich ermöglicht eine breite Palette an Analysemöglichkeiten für die Verwaltung, Verarbeitung und Visualisierung speziell dieser Daten.

1 Einleitung

Die „Entwicklung eines Informationssystems für Laserscannerdaten mit OpenSource-Software“ wurde im Rahmen einer Diplomarbeit am Institut für Geographie der Universität Innsbruck durchgeführt (HÖFLE 2005). Am Institut für Geographie wird in den letzten Jahren verstärkt an der Auswertung der Daten des flugzeuggestützten Laserscannings (Airborne Laserscanning) für glaziologische Fragestellungen gearbeitet (siehe z. B. GEIST et al. 2003). Am *alpS – Zentrum für Naturgefahren Management* in Innsbruck läuft seit Juli 2004 ein dreijähriges Forschungsprojekt, das sich mit der Analyse von Laserscannerdaten im Hinblick auf die Bestimmung von Oberflächeneigenschaften beschäftigt (www.alps-gmbh.com). Die angestrebten Ergebnisse sollen zur Verbesserung von Modellen beitragen, die in der Naturgefahrenforschung eingesetzt werden. Das „Informationssystem“ für Laserscannerdaten kommt in diesem Projekt operationell zum Einsatz und wird ständig weiterentwickelt. In dieser Arbeit wird der Entwicklungsstand zu Beginn des Jahres 2005 beschrieben, dabei wird auf folgende Fragen näher eingegangen: „Warum wird ein Informationssystem benötigt?“, „Wie ist dieses System aufgebaut?“ und „Was kann es leisten?“

2 Ausgangssituation

Das flugzeuggestützte Laserscanning liefert eine nahezu flächendeckende Erfassung des gescannten Gebiets mit einer Genauigkeit, die für horizontale Flächen im Submeterbe-

reich liegt (KRAUS 2004). Das Primärresultat einer Laserscannerbefliegung ist eine sehr große Anzahl (bis zu mehreren Punkten pro m^2) von gemessenen Punktkoordinaten, die eine zusätzliche Information über die Intensität des reflektierten Signals enthalten (x, y, z, I).

Die traditionelle Vorgehensweise besteht darin, diese Punkte als Stützpunkte für die Berechnung von Rasterdaten heranzuziehen (SITHOLE & VOSSELMAN 2003). Die Methoden der rasterbasierten Auswertungen sind weit fortgeschritten und erlauben die Trennung von Gelände und Objekten, die sich auf dem Gelände befinden. Das Resultat sind *Digitale Geländemodelle* (ohne Objekte) und *Digitale Oberflächenmodelle* (mit Objekten). Diese Vorgehensweise liefert brauchbare Ergebnisse, weil es die Genauigkeit und die Möglichkeiten der bisherigen Methoden der großflächigen Oberflächenerfassung übertrifft und weil es die ursprüngliche Datenmenge der Messpunkte um vieles verringert. Bei der Überführung der Laserpunkte in das Rasterdatenmodell treten unweigerlich Verluste in der Genauigkeit auf. Ferner enthalten die einzelnen Laserpunkte sowie ihre Verteilung im Raum zusätzliche Information, die für die Beschreibung der aufgenommenen Erdoberfläche nützlich sein kann (VOSSELMAN et al. 2004).

Die Originaldaten bzw. „Rohdaten“ einer Befliegungskampagne sind das eigentliche Resultat einer Laserscannerbefliegung. Oft werden sie jedoch vom Kunden nicht gewünscht, da die daraus abgeleiteten Digitalen Gelände- und Oberflächenraster ausreichen. Die große Datenmenge (ca. 500 MByte für 10.000.000 Laserpunkte) und die Struktur der Rohdaten erlauben keine sinnvolle Verarbeitung für große Gebiete. Bei einer Punktdichte von 5 Punkten/ m^2 und einer Fläche von $100 km^2$ würden die Rohdaten ca. 24 GByte einnehmen.

Am Markt gibt es bis zum jetzigen Zeitpunkt kein befriedigendes kommerzielles Softwarepaket, das die Verwaltung der Laserpunkte im Vektordatenmodell, die Visualisierung dieser Punktwolken, die statistische Auswertung und die Entwicklung und Einbindung eigener Auswertestrategien in einem einzigen System zusammenführt. Der Begriff „Informationssystem“ steht repräsentativ für ein System, das die Datenverwaltung, die Datenverarbeitung und die Visualisierung der Laserscanner-Rohdaten vereint. Es wird ein flexibles System benötigt, das den schnellen und einfachen Zugriff auf die Laserscannerdaten und deren deskriptive Informationen ermöglicht.

Die ausschließliche Verwendung von freier, nichtkommerzieller Software steht im Vordergrund der technischen Umsetzung. Die offene Struktur dieser Software eröffnet neue Möglichkeiten und erlaubt jedem, die Auswertestrategien ohne Lizenzkosten selbst umzusetzen (CHRISTL & TRAKAS 2004).

3 Datengrundlagen

Im Rahmen des Projekts wird mit Laserscannerdaten gearbeitet, die vom Land Vorarlberg, das in Österreich eine Vorreiterrolle in der flächendeckenden Landeserfassung mit Laserscanning einnimmt, zur Verfügung gestellt wurden. Die Befliegungen wurden von der Firma TopScan GmbH (Rheine, Deutschland) durchgeführt.

Die Struktur, in der die Rohdaten ausgeliefert werden, wird von der Befliegungsfirma festgelegt. Generell können Flugzeugdaten und Laserpunktdaten unterschieden werden. Die Flugzeugdaten bestehen aus der Position (x, y, z) und Orientierung (*roll, pitch, heading, wander angle*) des Flugzeugs. Für jeden Laserpunkt wird eine x-, y- und z-Koordinate sowie die Intensität abgespeichert. Das Aufnahmesystem ALTM 2050 (www.optech.on.ca) zeichnet die erste und letzte Reflexion (*first, last pulse*) für jeden Lichtimpuls auf. Die Flugzeugpositionen und die Laserpunkte besitzen ein Attribut *Zeit*, das in weiterer Folge für die Verknüpfung dieser beiden Datensätze herangezogen wird.

Die Auslieferung der Laserpunkte – die Dateistruktur – kann entweder nach räumlichen Kriterien (z. B. 1×1 km Kacheln) oder nach ihrer Chronologie (z. B. in Flugstreifen) erfolgen. Das Problem der großen Datenmenge (große Dateien!) und die schwierige Navigation in den Rohdaten bleiben bei beiden Auslieferungsvarianten unverändert.

4 Konzeption

4.1 Systemanforderungen

Die wichtigsten Eigenschaften des Informationssystems können folgendermaßen zusammengefasst werden:

- Ausschließliche Verwendung von *freier Software* (Systemkomponenten und Programmiersprachen) mit Schnittstellen zu kommerziellen Softwarepaketen
- 4-dimensional: *zeitliche* und *räumliche Verortung* aller Datensätze (multitemporale Analysen)
- Speicherung und Verarbeitung aller Datensätze in der *maximalen Genauigkeit* (Vektordatenmodell)
- Verarbeitungsstrategien, die eine *ansprechende Performance* mit sich bringen (Software- und Hardwaretuning)
- Client-Server-Architektur (*plattformunabhängig*)

Die Grundfunktionalität des Informationssystems kann mit den Aufgaben beschrieben werden, die damit bewältigt werden sollen:

- *Import* der Rohdaten (Flugzeugpositionen und Laserpunkte)
- *Verwaltung* der Laserscannerdaten (Geometrie- und Sachdaten)
- *Bearbeitung* der Laserscannerdaten bzw. Analysen in der Punktwolke (Steuerung über räumliche Kriterien oder über Attributwerte) mit GIS-Funktionalität und statistischen Methoden
- *Visualisierung* (2D, 3D) im Vektor- und Rasterdatenformat
- *Export* aus dem System in GIS-Standardformate (= Austauschformat), Bildformate, usw.

4.3 Datenverwaltung

Eine zentrale Anforderung an das System ist es, die Rohdaten der Laserscannerbefliegungen so abzulegen, dass ein schneller Zugriff und die Selektion der Daten leicht möglich ist. Der Zugriff auf die Daten steuert nicht nur den Export, sondern ist auch bei Berechnungen entscheidend für die Performance. Der Datenbankserver PostgreSQL mit der Erweiterung PostGIS werden für das Datenmanagement herangezogen. Bei der Speicherung von großen Datenmengen in einer PostgreSQL-Datenbank gibt es wenig Einschränkungen (Tab. 2).

Tab. 2: Einschränkungen des PostgreSQL-Servers
(<http://www.postgresql.org/users-lounge/limitations.html>)

Maximum size for a database:	unlimited (4 TB databases exist)
Maximum size for a table:	16 TB on all operating systems → 23.900 CDs
Maximum size for a row:	1.6 TB
Maximum size for a field:	1 GB
Maximum number of rows in a table:	unlimited
Maximum number of columns in a table:	250–1600 depending on column types
Maximum number of indexes on a table:	unlimited

Die große Anzahl der Datensätze in den Tabellen (z. B. mehrere hundert Millionen Laserpunkte) führt zu der Anwendung eines einfachen Datenbankschemas. Relationstabellen zwischen den Entitäten verlangsamen die Performance bei Abfragen und bringen einen erheblichen Speicherplatzverbrauch mit sich. RAMSEY (2005) unterstreicht die Wichtigkeit der Datenbankoptimierung, die vor allem durch die Indexierung der Geometriespalten erreicht werden kann.

„*Indexes are what make using a spatial database for large databases possible*“ (RAMSEY 2005, 20)

PostGIS verwendet den GiST-Index (Generalized Search Tree) für die Indexierung der Geometriespalten (JURGEIT et al. 2004).

Der aktuelle Stand der Technik des flugzeuggestützten Laserscannings erlaubt die Erstellung eines sehr allgemeinen Datenbankschemas, das für verschiedenste Aufnahmeverfahren Gültigkeit behält. Die Grundeinheit ist immer die Befliegung (vgl. Abb. 2). Eine Befliegung weist immer ein oder mehrere Flugstreifen auf. Um ein Gebiet flächendeckend zu erfassen, werden sich überlappende Streifen geflogen, die schlussendlich das gesamte Gebiet abdecken sollten. Ein Flugstreifen besitzt viele Laserpunkte, die wiederum in *first* und *last pulse* Punkte unterteilt werden können. Die Flugzeugpositionen werden ursprünglich der Befliegung zugeordnet. Mithilfe der Start- und Endzeit eines jeden Streifens können die Flugzeugpositionen mit den Flugstreifen verknüpft werden. Da in den Rohdaten die zeitliche Auflösung der Flugzeugpositionen 200 Hz – im Gegensatz zu der Auflösung der Laserpunkte mit 50.000 Hz – beträgt, wird mit einem linearen Interpolationsansatz für jeden Laserpunkt die Flugzeugposition zum Zeitpunkt der Aufnahme rekonstruiert (HÖFLE 2005).

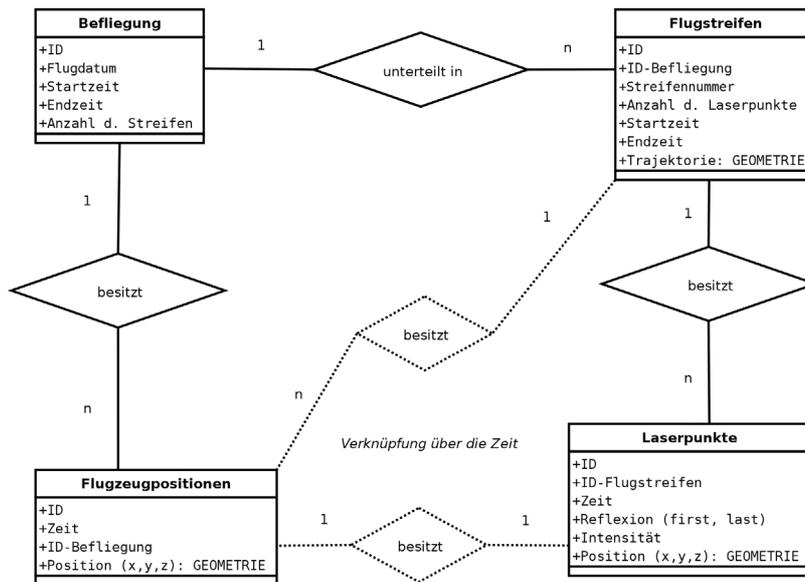


Abb. 2: Entitäten und Relationen der Laserscannerdaten (verändert nach HÖFLE 2005, 56)

4.4 Datenimport

Ein direkter Import der Laserscannerdaten, wie sie derzeit von der Befliegungsfirma geliefert werden, ist nicht möglich. Das vorliegende Rohdatenformat wird mit einem in der objektorientierten Programmiersprache Python geschriebenen Programm gelesen, ausgewertet und in die Datenbank geschrieben.

Die einzelnen Rohdatenfiles werden mit Lesezugriff geöffnet. Alle Zeilen in dieser ASCII-Datei werden systematisch gelesen. Das festgelegte Rohdatenformat erlaubt es, die Zeile mithilfe von so genannten „regulären Ausdrücken“ in ihre Komponenten zu zerlegen (WEIGEND 2004). Bei den Flugzeugpositionen werden zum Beispiel die Zeit und die dreidimensionalen Koordinaten herausgefiltert. Die herausgefilterten Daten werden zu einer Zeichenkette neu zusammengesetzt, die von der Datenbank gelesen werden kann.

Der Datenbankserver PostgreSQL verfügt über den Befehl COPY, der sich besonders für den Import und Export von großen Datenmengen eignet. Das Python-Modul „pg“ erlaubt die Verwendung des COPY-Befehls in einem Python-Programm (GESCHWINDE & SCHÖNIG 2002).

4.5 Datenverarbeitung

Der größte Teil der Datenverarbeitung findet in der räumlichen Datenbank (PostgreSQL/PostGIS) statt. Die prozedurale Sprache PL/pgSQL erlaubt die Entwicklung von komplexen Funktionen mit einfachen Methoden (EISENTRAUT 2003). Die Funktionalität der Statistiksoftware R kann mit der prozeduralen Sprache PL/R angesprochen werden

(www.joeconway.com/plr/). Dieser Funktionalitätstransfer von R in die Datenbank verhindert rechen- und zeitintensive Datentransfers. Die entwickelten Funktionen können in Form von SQL-Statements aufgerufen werden. Die GIS-Funktionalität von PostGIS, die statistischen Methoden und Visualisierungstools von R und die Standardfunktionen der Datenbank stellen ein breites Spektrum an Auswertemöglichkeiten zur Verfügung. SQL-Funktionen eignen sich besonders gut für die Client-seitige Anwendung (z. B. Internet-Client, Python-Applikation, GIS-Programm, Graphical User Interface), da der Multi-User-Betrieb vom Datenbankserver verwaltet wird. Die Datenbank übernimmt nicht nur die Verwaltung und Speicherung der Laserscannerdaten, sondern wird gleichzeitig gezielt für die Analyse der Daten eingesetzt. Das Geographische Informationssystem GRASS kann direkt auf die Laserscannerdaten in der Datenbank zugreifen und damit arbeiten. Es findet wiederum kein Datentransfer statt. Die Laserpunkte können zum Beispiel als Stützpunkte für die Interpolation eines Rasters herangezogen werden.

Die Laserscannerdaten sind räumlich und zeitlich verortet, können im Multi-User-Betrieb Client-seitig abgefragt werden und bleiben bei allen Verarbeitungsschritten an derselben Stelle – in der Datenbank. Jede Änderung der Daten wird sofort wirksam. Wird zum Beispiel ein neues Attribut für die Laserpunkte angelegt, steht die neue Information sofort für alle Clients zur Verfügung.

4.6 Visualisierung und Datenexport

Der Export der Daten des Informationssystems ist in die verschiedensten Datenformate (z. B. GIS-Formate und Grafik-Formate bzw. Vektor- und Rasterformate, etc.) möglich. Die Entwicklergemeinden von OpenSource-Software achten sehr darauf, dass ihre Programme möglichst offen und kompatibel mit anderer Software gestaltet werden. Die meisten Programmiersprachen ermöglichen den Zugriff auf die PostgreSQL-Datenbank und die meisten OpenSource-GIS-Programme können PostGIS-Tabellen lesen und direkt importieren. Es gibt zwei grundlegende Exportvarianten. Einerseits können die Daten von der Datenbank in ein entsprechendes Format geschrieben werden („direkter Datenbankexport“), das vom Zielprogramm gelesen werden kann, und andererseits kann mit geeigneter Software direkt auf die Daten in der Datenbank zugegriffen werden. Das Zielprogramm liest die Daten ein und kann sie dann wiederum in ein anderes Format umwandeln („indirekter Datenbankexport“). Die Daten können direkt mit der Exportfunktionalität von PostgreSQL, PostGIS oder von PL/R (R-Funktionalität) aus der Datenbank geschrieben werden. Die indirekten Möglichkeiten gehen über das Geographische Informationssystem GRASS, über die Statistiksoftware R und über das Visualisierungsprogramm GGobi (= Paket für R).

5 Ergebnisse

5.1 Auswertung des Intensitätssignals

Die Intensität, die für jeden Laserpunkt zur Verfügung steht, gibt die Stärke der Reflexion des Lichts in Integerwerten von 0 bis 255 (8-Bit) an. Im Intensitätswert sind die Reflexi-

onseigenschaften der Oberfläche, die Größe des Streuquerschnitts und die Scangeometrie enthalten (WAGNER et al. 2003).

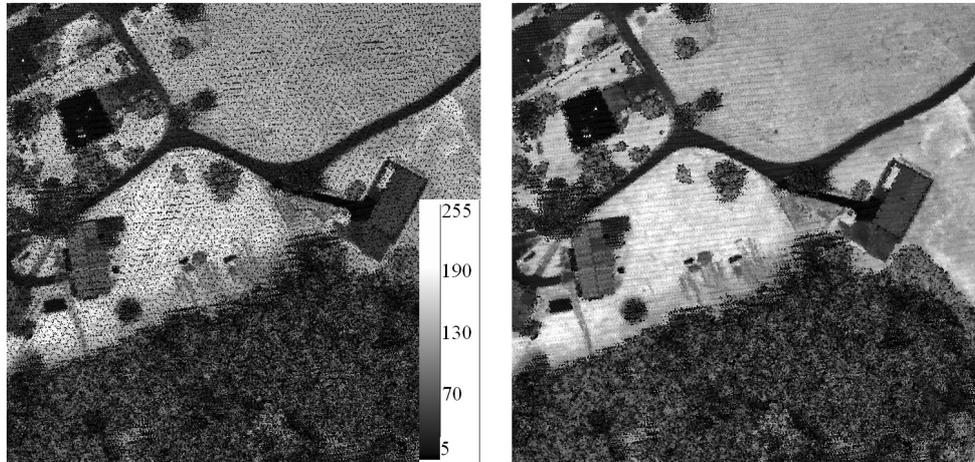


Abb. 3: Intensitätsraster: links: Interpolation mit allen Laserpunkten; rechts: korrigierter, rauschfreier Intensitätsraster (HÖFLE 2005)

Die Raster bzw. Bilder, die aus den Intensitäten berechnet werden, sind vor allem für Gebiete, die nur sehr kontrastarme Luftbilder liefern (z. B. Gletscher), interessant. Mit den statistischen Methoden und den Interpolationsverfahren des Informationssystems können rausch- und streifenfreie Intensitätsraster erstellt werden (Abb. 3).

5.2 Visualisierung der Punktwolke

Die Projektion der Laserpunkte auf eine Fläche, macht ab einem bestimmten Maßstab keinen Sinn mehr, da die Symbole der Punkte die komplette Karte bzw. den ganzen Bildschirm ausfüllen. Erst eine starke Reduktion der Punkte, die mit der Entwicklung einer Funktion für die Erstellung von Profildiagrammen umgesetzt wurde, führt zu einem befriedigenden Ergebnis (Abb. 4). Die Auswahl der Punkte wird mit einer PL/pgSQL-Funktion durchgeführt. Die ausgewählten Punkte werden einer PL/R-Funktion übergeben, die dann die Visualisierung des Diagramms übernimmt. Dabei kommen die graphischen Fähigkeiten von R zum Einsatz. Das R-Zusatzpaket „GGobi“ kann direkt auf die Laserpunkte in der Datenbank zugreifen und sie in Form einer virtuellen „3D-Punktwelt“, in der die Abfrage der Laserpunktattribute möglich bleibt, darstellen.

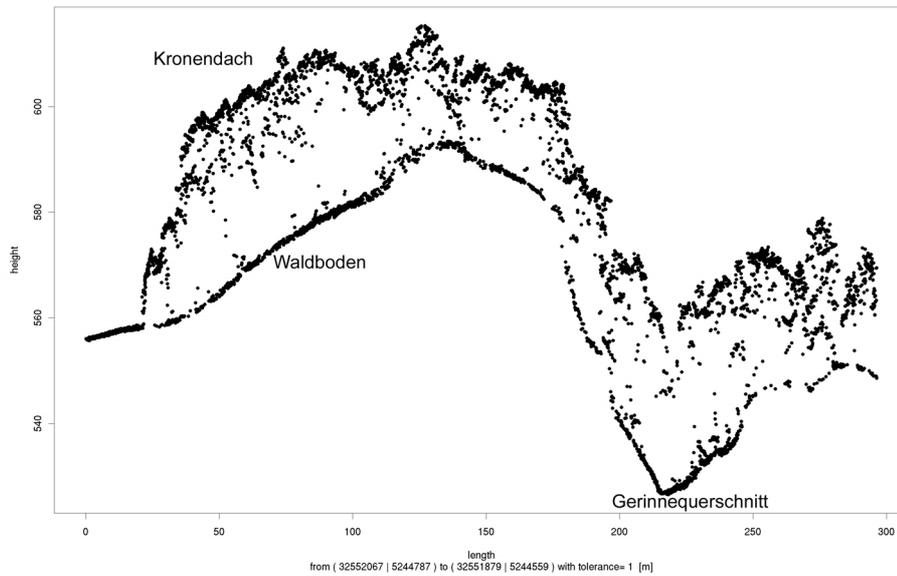


Abb. 4: 2D-Profil durch die Punktwolke mit 1 m Toleranz. Deutlich zu erkennen sind der Waldboden (= Gelände), die Kronenstruktur (= Oberfläche) und der Gerinnequerschnitt eines Wildbachs (HÖFLE 2005)

6 Ausblick

Die reibungslose Zusammenarbeit der drei Hauptkomponenten des Informationssystems (Datenbank, GIS, Statistiksoftware) stellt ein großes Potenzial für wissenschaftliche Anwendungen bereit. Die offene Struktur des Informationssystems erlaubt die Entwicklung eigener Funktionalität sowie die Anbindung zusätzlich erforderlicher Programme, die dem OpenSource-Standard entsprechen (www.opensource.org). Das Informationssystem wurde für Laserscannerdaten entwickelt, ungeachtet von welcher Befliegungsfirma oder von welchem System (z. B. terrestrische Aufnahme oder System mit mehr als zwei Pulsen) die Daten stammen. Das Grundcharakteristikum, dass ein Laserpunkt eine zeitliche und eine räumliche Verortung sowie zusätzliche Attribute besitzt, ist allgemein gültig. Die Komponenten des Informationssystems sind sehr wohl in der Lage, mit den Datenmengen von großen Befliegungskampagnen (über eine Milliarde Punkte) umzugehen. Mit gezieltem Softwaretuning des Betriebssystems und einer verbesserten Einstellung der einzelnen Systemkomponenten (v. a. Datenbanktuning) kann die Performance effektiv gesteigert werden. Die neu geschriebenen Funktionen wurden zahlreichen Tests unterzogen und optimiert, um eine möglichst kurze Berechnungszeit zu erreichen. Die Inwertsetzung der Rohdaten, die mit dem vorgestellten Informationssystem möglich wird, bringt einen Mehrwert für die meisten Kunden von Laserscannerbefliegungen. Das Informationssystem zielt auf die Verwaltung und Verarbeitung von großflächigen Datenbeständen (mehrere hundert km²) ab. Die Umsetzung des Konzepts in eine benutzerfreundliche Software wird angestrebt.

Internetverweise

alpS – Zentrum für Naturgefahren Management: <http://www.alps-gmbh.com>
 GRASS GIS: <http://grass.itc.it>
 Open Source Initiative (OSI): <http://www.opensource.org>
 Optech ALTM: <http://www.optech.ca/prodaltm.htm>
 PL/R: <http://www.joeconway.com/plr>
 PostGIS: <http://postgis.refractive.net>
 PostgreSQL: <http://www.postgresql.org>
 Python: <http://www.python.org>
 Statistiksoftware R: <http://www.r-project.org>
 TopScan GmbH: <http://www.topscan.de> (*Stand der Links: 1.02.2005*)

Literatur

- CHRISTL, A. & A. TRAKAS (2004): Open Source und Freie Software – Argumente jenseits der Kostendiskussion. In: STROBL, J., BLASCHKE T. & G. GRIESEBNER (Hrsg.): *Angewandte Geoinformatik 2004. Beiträge zum 16. AGIT-Symposium Salzburg*, 73-78. Wichmann Verlag, Heidelberg.
- EISENTRAUT, P. (2003): *PostgreSQL – Das offizielle Handbuch*. mitp-Verlag, Bonn.
- GEIST, T., LUTZ, E. & J. STÖTTER (2003). Airborne Laser Scanning Technology and its Potential for Applications in Glaciology. In: *International Archives of Photogrammetry and Remote Sensing*, 34(3/W13), Dresden, 101-106.
- GESCHWINDE, E. & H. J. SCHÖNIG (2002): *PostgreSQL Developer's Handbook*. SAMS Publishing, Indianapolis.
- HÖFLE, B. (2005): *Entwicklung eines Informationssystems für Laserscannerdaten mit open source Software*. Unveröffentlichte Diplomarbeit, Institut für Geographie, Innsbruck.
- JURGEIT, F., FÖRSTER, K. & A. HELLER (2004): *Tirol Atlas – Technisches Framework auf Basis von OpenSource-Technologien*. In: STROBL, J., BLASCHKE T. & G. GRIESEBNER (Hrsg.): *Angewandte Geoinformatik 2004. Beiträge zum 16. AGIT-Symposium Salzburg*, 293-302. Wichmann Verlag, Heidelberg.
- KRAUS, K. (2004): *Photogrammetrie, Band 1: Geometrische Informationen aus Photographien und Laserscanneraufnahmen*. De Gruyter Verlag, Berlin.
- RAMSEY, P. (2005): *PostGIS-Manual*. <http://postgis.refractive.net/docs/postgis.pdf> (1.02.2005)
- SITHOLE G. & G. VOSSELMAN (2003): Report: ISPRS Comparison of Filters. <http://enterprise.lr.tudelft.nl/frs/isprs/filtertest/>.
- VOSSELMAN, G., GORTE, B. G. H., SITHOLE, G. & T. RABBANI (2004): Recognising Structure in Laser Scanner Point Clouds. In: *International Archives of Photogrammetry and Remote Sensing*, 36(8/W2), Freiburg, 33-38.
- WAGNER, W., ULLRICH, A. & C. BRIESE (2003): Der Laserstrahl und seine Interaktion mit der Erdoberfläche. In: *Österreichische Zeitschrift für Vermessung & Geoinformation*, 4, 223-235.
- WEIGEND, M. (2004). *Objektorientierte Programmierung mit Python*. mitp-Verlag, Bonn.