

## Cluster-Analyse von LIDAR-Vektordaten zur Erfassung geometrischer Eigenschaften von einzelnen Bäumen

FELIX MORSDORF, ERICH MEIER, BRITTA ALLGÖWER & DANIEL NÜESCH, Zürich

**Keywords:** laser scanning, LIDAR, segmentation, cluster analysis, tree height, reconstruction

**Zusammenfassung:** Basierend auf einem LIDAR Datensatz mit einer Punktdichte von mehr als 20 Punkten pro m<sup>2</sup> versuchen wir die geometrischen Eigenschaften von Einzelbäumen abzuleiten. Die Genauigkeit der LIDAR Messungen wurde durch 6 geometrische Ziele bestimmt, welche 3 m mal 3 m groß und im Mittel ca. 1 m über dem Boden angebracht waren. Die internen Lageoffsets der Laserdaten war kleiner als 0.25 m, die Standardabweichung der Höhe der Punkte auf den Flächen betrug 0.06 m. Wir führen eine Unterteilung der LIDAR Rohdaten mit Hilfe einer überwachten Klassifikation durch und validieren diesen Ansatz mit Feldmessungen von etwa 1200 Bäumen. Dieser Datensatz beinhaltet Baumposition, -höhe sowie den Kronendurchmesser und wurde auf einer Langzeitmessfläche der Eidgenössischen Forschungsanstalt für Wald, Schnee und Landschaft (WSL) erhoben. Eine robuste lineare Regression der aus den LIDAR Daten ermittelten Baumhöhen mit den Feldmessungen ergibt annähernd ein 1:1 Verhältnis mit einer Steigung von 0.92 und einem Achsenabschnitt von 0.47 m, bei einem RMS von 0.6 m. Dieses ergibt eine geringe Unterschätzung der Baumhöhen durch die LIDAR Messungen.

**Summary:** *Cluster Analysis of LIDAR Raw Data for Geometric Property Derivation of Single Trees.* Using a LIDAR dataset having a high point density of more than 20 points per square meter, we test the potential of deriving geometric properties of single trees. The accuracy of the LIDAR data is evaluated using 6 geometric reference targets, being 3 m by 3 m in size and a height of about 1 m above ground. The internal translational offsets of the LIDAR data were found to be less than 0.25 m, and the standard deviation of the height values on the targets were 0.06 m. We carried out a segmentation of the LIDAR raw data using cluster analysis and validated this approach using field measurements for about 1200 trees. The measurements were taken by the Swiss Federal Institute for Forest, Snow and Landscape Research and contained properties as tree-position, -height and crown diameter. A linear robust regression of field measured tree heights and LIDAR derived ones resulted in a relationship close to 1:1, with a slope of 0.92 and an offset of 0.47 m, with the RMS being at 0.6 m. This manifests a slight underestimation of tree height by the LIDAR measurements.

### 1 Einleitung

Das Potenzial des flugzeuggetragenen ‚Laserscannings‘ zur Erfassung von Forstparametern ist in den letzten Jahren intensiv evaluiert worden. Algorithmen zur Bestimmung von strukturellen Forstparametern auf Bestandes- oder Plotebene wurden bereits von einigen Wissenschaftlern erfolgreich implementiert (MEANS et al. 2000,

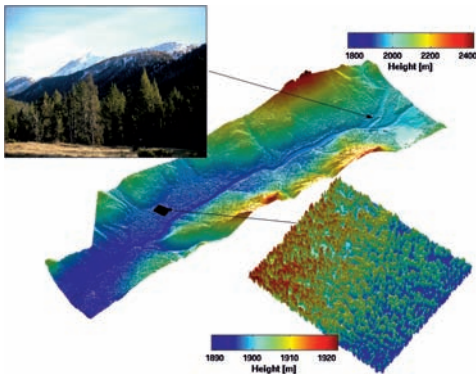
DRAKE et al. 2002, NAESSET & OEKLAND 2002). Da mittlerweile LIDAR Systeme mit sehr hoher Punktdichte (> 10 Punkte/m<sup>2</sup>) verfügbar sind (BALTSAVIAS 1999), wurde die Ableitung von diesen geometrischen Parametern auf Basis von einzelnen Bäumen vermehrt Ziel wissenschaftlicher Arbeiten. Die bisherigen Ansätze zur Detektion von Einzelbäumen basierten zumeist auf einer Unterteilung des Oberflächenmodells

(DOM, *Digitales Oberflächen Modell*) oder Vegetationsmodells (Differenz DOM-DGM *Digitales Gelände Modell*) wie zum Beispiel bei HYYPPÄE et al. 2001 oder PERSSON et al. 2002. Da bei der Interpolation der Rohdaten in die Rastermodelle Information verloren geht, ist man vermehrt dazu übergegangen direkt mit den Rohdaten zu arbeiten (PYSSALO & HYYPPÄE 2002, BRANDENBERG et al. 2003). So haben zum Beispiel ANDERSEN et al. (2002) ellipsoide Kronenmodelle unter Verwendung eines zusätzlichen, stochastischen Terms an die Rohdaten angepasst, und in diesem Prozess die Interaktion zwischen Kronen und Laserpuls statistisch modelliert. Wir werden eine neue Methode vorstellen, um die LIDAR Rohdaten in Echos von einzelnen Bäumen zu unterteilen. Diese ermöglichen dann das direkte Ableiten ihrer geometrischen Eigenschaften aus der dreidimensionalen Punktwolke.

## 2 Testgebiet und Daten

### 2.1 Testgebiet und Feldmessungen

Das Testgebiet befindet sich im Schweizer Nationalpark in Höhen von 1800 bis 2400 m ü. NN. Die Vegetation ist boreal und wird von Beständen aus Bergföhren und Lärchen dominiert, welche von ein paar Alpweiden unterbrochen werden. Auf einer kleinen Teilfläche (ca. 1.5 ha) des Testgebietes unter-



**Abb. 1:** Digitales Oberflächen Modell (DOM) des Ofenpassgebietes im Schweizer National-Park. Die Langzeitmessfläche der WSL ist vergrößert. Das Foto wurde am Tag des LIDAR-Fluges aufgenommen.

hält die Eidgenössische Forschungsanstalt für Wald, Schnee und Landschaft (WSL) eine Langzeitmessfläche, die ungefähr 1200 Bäume mit einem Brusthöhendurchmesser (BHD) von größer als 0.15 m enthält. Zu dieser Fläche existiert eine Datenbank mit Position der Bäume, Baumhöhe, Kronendurchmesser sowie BHD, welche mit GPS und Standard-Forstwerkzeugen wie einem Hypsometer ermittelt wurden. In Abb. 1 ist ein Überblick über das Testgebiet gegeben.

### 2.2 Laser Scanning Daten

Im Oktober 2002 fand eine Befliegung des Testgebietes mit einem Helikopter statt, welche ein Gebiet von ungefähr 14 km<sup>2</sup> abdeckte. Das zur Anwendung kommende LIDAR System war der Falcon II Sensor der deutschen Firma TopoSys ([www.toposys.com](http://www.toposys.com)). Es handelt sich dabei um einen *push broom* Laser Altimeter, welcher sowohl das erste als auch das letzte Echo vom Boden aufzeichnen kann (*first/last pulse*). Der Flug wurde mit einer nominalen Flughöhe von ca. 850 m über Grund durchgeführt, was zu einer mittleren Punktdichte am Boden von mehr als 10 Punkten pro Quadratmeter führte. Eine kleinere Fläche (0.6 km<sup>2</sup>) wurde zusätzlich in einer Höhe von 500 m über Grund überflogen, was eine Punktdichte von mehr als 20 Punkten pro Quadratmeter erbrachte. Die Kombination von beiden Datensätzen ergibt eine sehr hohe Punktdichte von mehr als 30 Punkten pro Quadratmeter für diese Teilfläche, in welcher sich auch die Messfläche der WSL befindet. Der Durchmesser des Laserstrahls am Boden (*foot-print*) liegt bei ungefähr 0.6 m beziehungsweise 0.9 m für die jeweiligen Flughöhen von 500 m und 850 m über Grund. Die Rohdaten, die vom Sensor erfasst wurden (x, y, z-Tripel) wurden von TopoSys mit ihrer eigenen Software in Rastermodelle interpoliert. Das digitale Oberflächenmodell (DOM, beinhaltet Vegetation und Gebäude) wurde aus den *first pulse* Daten ermittelt, das digitale Geländemodell (DGM) wurde mit Hilfe der *last pulse* Echos und Filteralgorithmen erstellt. Die entstehenden Lächer wurden von uns durch nachbarschaftstreu

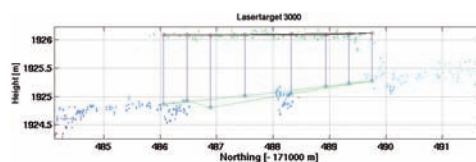
Interpolation geschlossen. Die Gitterweite der Modelle lag bei einem Meter für das Gesamtgebiet und einem halben Meter für das Teilgebiet, wobei die Höhenauflösung in beiden Fällen bei 0.1 m liegt.

### 2.3 Qualitätsanalysen

Eine einfache Qualitätsanalyse der LIDAR Daten wurde mit Hilfe von 6 geometrischen Referenzflächen durchgeführt, welche  $3\text{ m} \times 3\text{ m}$  maßen und ungefähr 1–1.5 m über dem Boden angebracht waren (siehe Abb. 2). Die Flächen wurden mit Hilfe einer digitalen Winkelwaage auf weniger als 0.5 Grad genau horizontal ausgerichtet. Die Positionen der 4 Ecken wurden mit einem Theodoliten bestimmt, der auf GPS Fixpunkte abgestützt wurde, welches die interne Genauigkeit der Vermessung auf ungefähr 2 cm steigen ließ. Mit den Positionen dieser 6 Ziele wurden die fertigen Modelle von TopoSys in Lage und Höhe um einige Meter (1/3.5 m für Hochwert/Rechtswert) rein translatorisch verschoben, so dass die Positionen der Flächen nur noch für Aussagen über die interne Genauigkeit der Laserdaten verwendet werden können. Nach dieser Verschiebung ergab sich für alle Flächen eine Positionsgenauigkeit des Modells von weniger als 0.5 m und eine Höhengenauigkeit von 0.15 m. Mit dieser Methode erfassen wir allerdings nur die Genauigkeit des DOM bzw. des DGM auf freien Flächen, wir können keine Genauigkeit für das DGM an bewaldeten Stellen ableiten, wo die Güte des DGM sehr stark von verwendeten Filteralgorithmen, der Vegetations-



**Abb. 2:** Ansicht des Ofenpassgebietes mit zwei Referenzflächen (die zweite befindet sich links im Hintergrund).



**Abb. 3:** Seitenansicht einer der 6 Referenzflächen, überlagert mit den LIDAR Rohdaten, grün entspricht *first pulse*, rot *last pulse* Daten. Die Punkte, die hier unterhalb der Fläche erscheinen, befinden sich im 3d-Raum vor und hinter der Fläche.

dichte und Beschaffenheit der Topographie abhängt. Andere Studien beschäftigen sich wesentlich eingehender mit diesem Aspekt (KRAUS & PFEIFER 1998).

Weiterhin benutzen wir die Flächen um eine Information über das Rauschen des Sensors auf einer glatten, homogen reflektierenden Fläche zu erhalten. Dieses wird sicherlich nur die untere Schranke für das Sensorrauschen sein, welches auf komplizierteren Oberflächen größer sein wird. Um eine Abschätzung des Rauschens zu bekommen haben wir die Standardabweichung aller Echos, welche von den Flächen reflektiert werden, berechnet (siehe Abb. 3). Die Ungenauigkeit bezüglich der Position der Flächen wurde aus den Schwerpunkten der Flächen und den Schwerpunkten der Punktwolke der Echos von diesen Flächen berechnet. Diese Abweichungen stellen jedoch nur die interne Ungenauigkeit der Laserstreifen dar, da die Modelle (und auch die Rohdaten) schon absolut translatorisch korrigiert

**Tab. 1:** Höhen- und Lageabweichungen der Referenztargets in den LIDAR-Rohdaten.

ID	Anzahl Punkte	$\Delta$ Höhe (cm)	$\sigma$ Höhe (cm)	$\Delta x$ (cm)	$\Delta y$ (cm)
1000	107	3	6.8	9	7
2000	133	-2	5.9	24	-11
3000	75	-2	6.6	6	6
4000	190	1	5.6	15	-3
5000	151	-2	5.8	4	15
6000	138	2	5.2	25	-18

wurden. Die Werte für die Abweichungen und das Rauschen können Tab.1 entnommen werden. Die geringen mittleren Höhenabweichungen von 1 bis 3 cm und ihre nicht systematische Verteilung über das Testgebiet (die Flächen sind von West nach Ost nummeriert) lassen den Schluss zu, dass es keine Verkippung der Laserstreifen in dieser Richtung gibt. Weiterhin ist die Standardabweichung der Echos auf den Flächen mit ca. 5–6 cm erstaunlich klein, und auch hier lassen sich keine signifikanten Differenzen zwischen den einzelnen Flächen feststellen, so dass der Sensor in dieser Hinsicht als sehr stabil bezeichnet werden kann.

### 3 Klassifikation mit *k-means* Cluster-Analyse

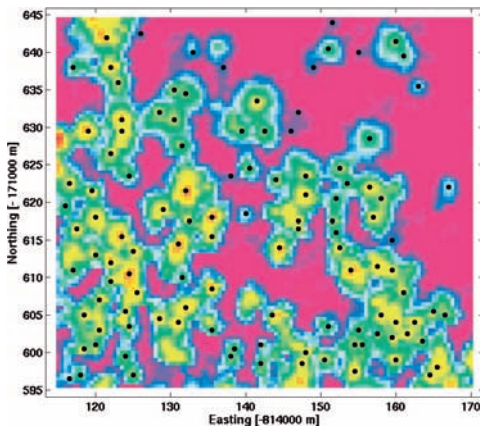
Die *Cluster* Analyse ist ein bekanntes und verbreitetes Werkzeug zur statistischen Analyse von Messdaten. Hierbei wird ein Merkmalsraum in  $k$  Bereiche aufgeteilt, die jeweils ähnliche Werte aufweisen. Die Ähnlichkeit wird meistens über eine spezifische Metrik definiert. In unserem Fall bilden die Koordinaten  $(x, y, z)$  den Merkmalsraum und wir verwenden eine einfache euklidische Distanz als Metrik. Der *k-means* Algorithmus versucht dann iterativ die globale Summe der Distanzen der jeweiligen *Cluster*-Mitglieder zu ihren Mittelpunkten (*centroids* oder *bouys*) zu minimieren. Dieser Prozess ist sehr rechenaufwändig, da bei jeder Iteration alle diese Distanzen neu berechnet werden müssen. Zumeist werden die Mittelpunkte im ersten Schritt der Iteration zufällig verteilt, und das Ergebnis der Analyse hängt dann stark von dieser Verteilung ab, da man das globale Minimum der Distanzen (beste Lösung) dann oft nicht erreicht. Also muss man diese Startpunkte entweder auf eine sinnvolle Art bestimmen oder aufwändig ausprobieren. Da die von uns untersuchten Bergföhren über genau eine definierte Spitze, welche meistens horizontal mittig platziert ist, verfügen, bietet sich die Position der Spitzen zur Verwendung als Startpunkte an. Folglich wird der erste Teil unserer Methode aus der Ermittlung von lokalen Maxima im DOM bestehen, der zweite

umfasst dann die eigentliche *Cluster* Analyse in den Rohdaten, mit den Koordinaten der lokalen Maxima als Startpositionen.

#### 3.1 Finden der lokalen Maxima

Die Extraktion von lokalen Maxima aus einem Tiefenbild ist keine leichte Aufgabe; komplizierte *watershed* Algorithmen, bei denen simulierte Wassertropfen eingesetzt werden, liefern generell die besten Ergebnisse. Ein großes Problem ist die Fehlklassifikation, bei welcher zu viele Maxima oder zu wenige gefunden werden. Wenn jedoch die Maxima in Form und Ausdehnung möglichst homogen sind, so kann man auch mit einfacheren Methoden gute Ergebnisse erzielen.

Unsere Methode basiert auf der Tatsache, dass das Maximum einer Pixelzeile des DOM durch einen Nulldurchgang in der ersten Ableitung (Steigung) und ein Extrem in der zweiten Ableitung (Krümmung) gekennzeichnet ist. Bei einem nicht geglätteten Modell wird man nun zu viele Maxima finden, da Rauigkeit auf den Skalen unterhalb der Kronendurchmesser auch lokale Maxima produziert. Wenn man jedoch diese Rauigkeit aus dem DOM herausfiltert, werden am Ende nur die Maxima der Baumkronen gefunden. Natürlich müssen die Parameter eines solchen Filters an die jeweiligen Kronendurchmesser und Gitterauflösungen angepasst werden. Wir verwenden in unserem Fall, bei einer Auflösung des Modells von 0.5 m und mittleren Kronendurchmessern von 1.5 bis 3 m (aus den Feldmessungen) einen rotationssymmetrischen, Gaußschen Tiefpassfilter mit einem Durchmesser von 12 Gitterpunkten und einer Standardabweichung von 0.5. Die lokalen Maxima werden anschließend über ihre analytischen Kriterien (notwendig und hinreichend) für jede Gitterzeile einzeln gesucht. Dieses geschieht für beide Koordinatenrichtungen, so dass man am Ende zwei Datensätze erhält, einen mit den lokalen Maxima in  $x$ - und einen mit den lokalen Maxima in  $y$ -Richtung. Diese werden dann mit einem logischen UND kombiniert, so dass nur noch die lokalen Maxima aus  $x$  und  $y$  Richtung übrig



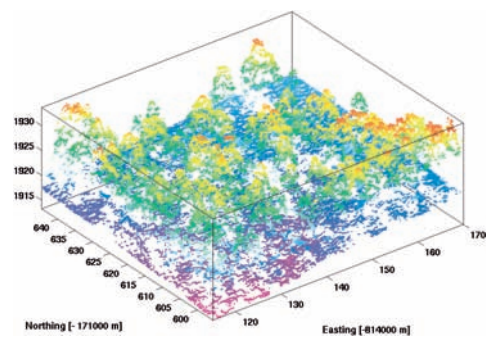
**Abb. 4:** Die Startpunkte der Cluster Analyse, wie von unserem lokalen Maxima Filter im DOM gefunden.

bleiben, die in beiden Bildern an der gleichen Stelle vorhanden sind. Das Ergebnis dieses Filters ist in Abb. 4 dargestellt, wobei die schwarzen Punkte die gefundenen lokalen Maxima repräsentieren. Man kann auch erkennen, dass einige Maxima einen relativ großen Bereich einnehmen, bei dem wahrscheinlich mehrere Bäume zusammen nur ein lokales Maximum bilden. Dieses bringt natürlich Probleme für die Unterteilung mit sich, welche dann in Abschnitt 4.1 genauer besprochen werden.

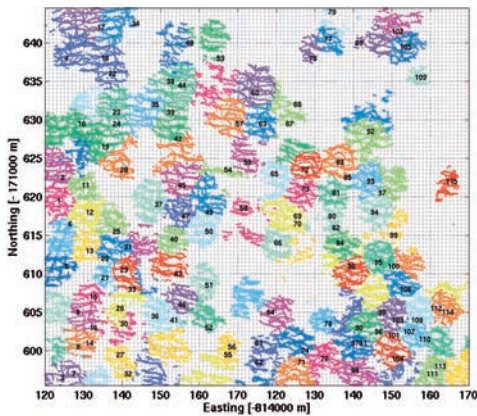
### 3.2 Cluster-Analyse in der dreidimensionalen Punktwolke

Die Verwendung einer euklidischen Metrik favorisiert das Unterteilen des Merkmalsraums in kugelförmigen Objekten, die Kronen der Bergföhren sind aber Ellipsoide mit einem Höhen/Breiten-Verhältnis von 2–3 zu 1. Aus diesem Grund führen wir ausgehend von den Feldmessungen eine Skalierung der Höhe  $z \rightarrow z'$  von  $1/2$  ein. Dieses lässt die Baumkronen im neu geschaffenen Merkmalsraum  $(x, y, z')$  kugelförmig erscheinen. Es werden sowohl *first-* als auch *last pulse-* Daten für die Klassifikation genutzt, ohne zwischen den beiden Echoarten zu unterscheiden. Wir benutzen für die Analyse den *k-means* Algorithmus, implementiert nach SPATH (1985). Der Algorithmus klassifiziert

die Rohdaten in einem iterativen Prozess, welcher in zwei Teile geteilt ist. Im ersten Teil werden so genannte *batch updates* durchgeführt, bei denen jeder Schritt daraus besteht, alle Echos ihrem nächsten *Cluster-Mittelpunkt* zuzuweisen, gefolgt von einer Neubestimmung der Mittelpunkte. Der zweite Teil besteht aus *online updates*, wo Echos individuell neuen Gruppen zugeordnet werden, wenn das die Summe der Distanzen reduziert, wiederum gefolgt von einer Neuberechnung der *Cluster-Mittelpunkte*. Um nicht auch Bodenechos zu klassifizieren, wurde ein Schwellwert von 1 m über der Höhe des DGM angewandt. Abb. 5 zeigt denselben Ausschnitt Abb. 4. Wir benutzen in diesem Beispiel nur die Daten des tiefen Überfluges mit mehr als 20 Echos pro Quadratmeter. Dieses ist im Vergleich zu normalen kommerziellen Datensätzen eine hohe Punktdichte, allerdings sollte erwähnt werden, dass die Baumkronen in der Messfläche sehr kleine Durchmesser von 1.5 bis 3 m aufweisen, so dass die Methode bei größeren Kronendurchmessern auch mit der normalen Punktdichte von ca. 10 Echos pro Quadratmeter gute Resultate liefern sollte. Da *small footprint* LIDAR-Rohdaten schnell bis zu 400 MB pro km<sup>2</sup> umfassen können, ist die Prozessierung dementsprechend rechenaufwändig. Allerdings benötigt keiner der oben aufgeführten Schritte eine spezielle Interaktion des Benutzers. Da



**Abb. 5:** Die LIDAR Rohdaten in einer perspektivischen Ansicht; nur Daten des tiefen Überfluges wurden hier verwendet. Gelbe und rote Werte markieren hohe Elevationen, wohingegen blaue und violette für tiefere z-Werte stehen.



**Abb. 6:** Die klassifizierten Rohdaten in der  $x, y$  Ebene, verschiedene Farben und Nummern zeigen die Cluster Zugehörigkeit an. Gleicher Ausschnitt wie Abb. 4 und 5

eine Klassifikation eines größeren Gebietes sehr lange dauert und auch nicht notwendig ist (man berechnet Abstände von Punkten, von denen man weiß, dass sie nicht im gleichen Cluster liegen können), haben wir den Algorithmus in überlappenden Fenstern von  $50\text{ m} \times 50\text{ m}$  Größe laufen lassen. In einem ebenfalls automatisierten Schritt werden dann doppelte (im Überlappungsbereich von 50%) oder nur partiell vorhandene Cluster (an den Kanten) aussortiert. Für den kleineren Teil des Datensatzes von  $0,6\text{ km}^2$  dauerte die Prozessierung ungefähr einen Tag auf einem state-of-the-art PC. Da die Geometrie des Waldbestandes aber kein sehr schnell ändernder Parameter ist, sollte dieses für die Anwendbarkeit der Methode nicht hinderlich sein.

Das Ergebnis der Klassifikation ist Abb. 6 zu entnehmen. Die Rohdaten sind in die  $x, y$ -Ebene projiziert, um die horizontalen Grenzen besser erkennen zu können. Die Zahlen geben den Cluster Index und die Farben die Cluster Zugehörigkeit an.

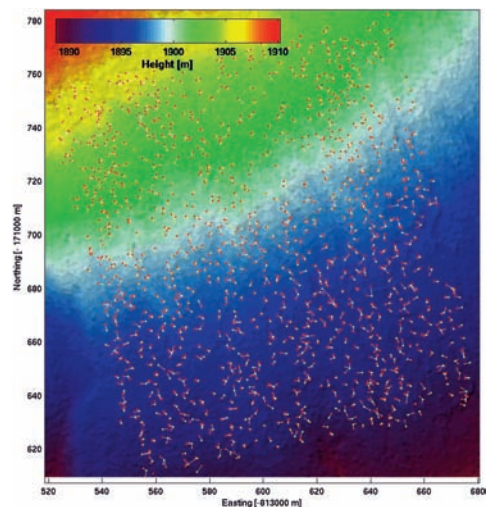
## 4 Ergebnisse

### 4.1 Validierung mit Feldmessungen

Da wir mehr als 1200 Feldmessungen den aus den LIDAR Daten ermittelten Größen

zuordnen mussten, haben wir dafür ebenfalls einen automatischen Prozess implementiert. Wir haben nur ca.  $3/4$  aller Bäume mit  $\text{BHD} > 0,15\text{ m}$  in der Testfläche korrekt detektiert, da es dort häufig zu einem so genannten *clumping* der Bäume kommt, d. h. es stehen mehrere Bäume in einem sehr engen Radius (typischerweise  $< 1,5\text{ m}$ ) zusammen, die im Feldinventar alle einzeln aufgeführt werden, im DOM aber nur ein lokales Maximum ergeben.

Mehrere Bäume in einem sehr kleinen Radius vorzufinden, ist ein typisches Merkmal der Vegetation auf unserer Testfläche im Schweizer Nationalparks (DOBBERTIN et al. 2001). Um dieses Problem in die Zuordnung eingehen zu lassen, gehen wir von den Feldmessungen aus und suchen den von der Position her nächsten Baum in den LIDAR-Daten. Auf diese Weise können mehrere Feldmessungen einem LIDAR-Cluster zugeordnet werden, was genau den Effekt des *clumping* berücksichtigt. Da die gezeigte Messfläche vollständig, d. h. jeder Baum mit



**Abb. 7:** Das Zuordnen der Feldmessungen (rote Punkte) zu denen aus den LIDAR Daten ermittelten (gelbe Punkte) geschieht automatisch über die Distanz. Die kurzen, weißen Linien verbinden zugeordnete Werte; dabei kann eine LIDAR Messung mehreren Feldmessungen zugeordnet werden, um den Effekt des *clumping* um dominante Bäume herum zu berücksichtigen. Im Hintergrund ist eine schattierte Darstellung des DGM zu erkennen.

BHD > 0.12 m, erfasst wurde, ist ein solches Vorgehen möglich. Das Resultat dieser Zuordnung ist Abb. 7 zu entnehmen.

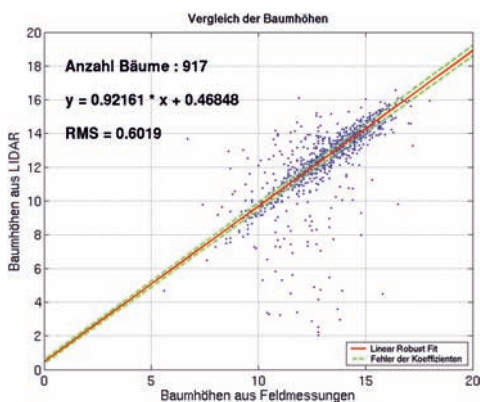
Wir gehen davon aus, dass es für jeden LIDAR-Cluster eine Feldmessung geben muss. Für den Fall, dass mehrere Feldmessungen einem Cluster zugeordnet wurden, ist lediglich der höchste (dominante) Baum der Feldmessungen in die Regression übernommen worden, da der höchste Wert des LIDAR-Clusters diesem Baum entsprechen sollte.

#### 4.2 Regression von Baumhöhen und Kronendurchmessern

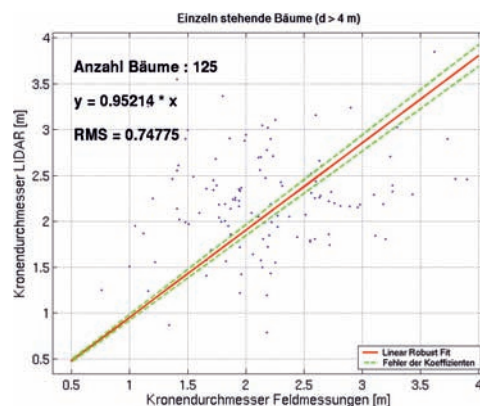
Nach erfolgreicher Zuordnung der Feldmessungen zu unseren LIDAR-Messungen können wir nun eine Regression der Baumhöhen durchführen. Durch den automatischen Zuordnungsprozess sind einige Ausreißer entstanden (Abb. 8), welche wir nicht in die Regression einfließen lassen möchten. Deshalb wählen wir eine robuste lineare Regression (HUBER 1981), welche Ausreißer weniger stark gewichtet. Wir können dieses aus zwei Gründen tun: zum Einen wissen wir um die Ursache der Ausreißer (fehlerhafte Zuordnung) und zum Anderen wird nur ein kleiner Teil der Daten als kompletter Ausreißer klassifiziert (ca. 9%, entnommen

aus dem Histogramm der Gewichte). Die berechneten Fehlergrenzen für die Koeffizienten des linearen Modells ( $y = m * x + b$ ) sind ebenfalls in Abb. 8 zu finden. Die so ermittelte Steigung liegt fast bei 1 (0.92) und besitzt einen Achsenabschnitt von ungefähr einem halben Meter (0.47 m). Dieses manifestiert eine Unterschätzung der Baumhöhen mit dem LIDAR, welche auch von anderen Autoren gefunden wurde (GAVEAU & HILL 2003, PERSSON et al. 2002). Da man nicht alle Baumspitzen notwendigerweise direkt mit dem Laserstrahl trifft (beprob), ist diese systematische Unterschätzung systembedingt. Bei höherer Punktdichte sollte dieser Effekt kleiner werden. Eine zusätzliche Ursache für die Unterschätzung ist die Tatsache, dass der *first pulse* nicht notwendigerweise, selbst wenn er die Spitze trifft, dort detektiert wird, denn es ist eine gewisse Dichte der Vegetation vonnöten, um wirklich ein Echo zu generieren. Dieses wird unter Umständen etwas tiefer als die tatsächliche Baumspitze liegen, abhängig vom Vegetationstyp und Durchmesser des Laserstrahls.

Für eine Regression der Kronendurchmesser ergibt sich ein größeres Problem durch das *clumping* der Bäume als für die Baumhöhen, da ein dominanter Durchmesser für eine Gruppe von Bäumen nicht de-



**Abb. 8:** Robuste Regression der Feldmessungen der Baumhöhen gegen unsere LIDAR Messungen, bei der Ausreißer weniger stark gewichtet werden. Fehlergrenzen für die Koeffizienten des linearen Modells wurden bestimmt und als grüne, gestrichelte Linie eingezeichnet.



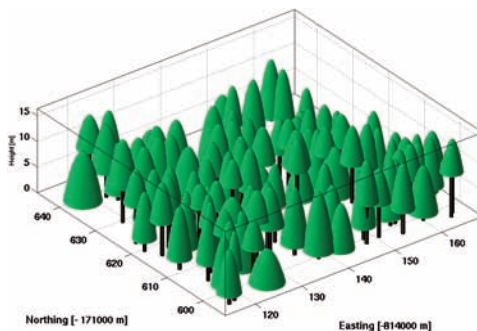
**Abb. 9:** Regression der Kronendurchmesser der Bäume, die mindestens 4 Meter vom nächsten entfernt sind. Es wurde ebenfalls eine robuste Regression durchgeführt, allerdings ohne einen Achsenabschnitt im linearen Modell zuzulassen.

finiert ist und man den Durchmesser der Baumgruppe nur unter Verwendung von allen einzelnen Positionen und Durchmessern berechnen könnte. Um dieses Problem zumindest teilweise zu umgehen, benutzen wir als ersten Schritt nur Bäume für die Regression, die mindestens 4 Meter vom ihrem nächsten Nachbarn entfernt sind (Abb. 9). Dieses reduziert die Zahl der Werte für die Regression auf 125. Wir verwenden ein lineares Modell ohne Achsenabschnitt ( $y = m * x$ ), und nehmen die Regressionsgerade nur als Anhaltspunkt für den Zusammenhang  $y = x$ , um die Messwerte besser einordnen zu können. Die große Streuung der Messwerte um die Ausgleichsgerade kann natürlich auch durch Fehler in den Feldmessungen herrühren, da der Kronendurchmesser wesentlich schwieriger zu bestimmen ist als die Baumhöhe.

#### 4.3 Geometrische Rekonstruktion

Ausgehend von der unterteilten Punktwolke leiten wir die folgenden Größen ab:

- Baumposition: Schwerpunkt der Punktwolke eines *Cluster*,
- Baumhöhe: Maximum der  $z$ -Werte in einem *Cluster*,
- Baumvolumen: Berechnung über die konvexe Hülle der Punkte,
- Kronendurchmesser: Projizierte Fläche in der  $x, y$ -Ebene wird mit Hilfe des Zusammen-



**Abb. 10:** Geometrische Rekonstruktion des zuvor präsentierten Ausschnittes. Es wurden die Werte für Höhe, Position, Durchmesser und Höhe der Kronenunterkante in einem einfachen Modell unter Verwendung eines halben Rotationsparaboloids benutzt.

menhangs für die Kreisfläche in einen Durchmesser überführt und

- Kronenunterkante: 98% Percentil der tiefsten  $z$ -Werte der *Cluster*.

Unter Verwendung der oben aufgeführten Parameter und eines simplen geometrischen Modells können wir nun den Baumbestand rekonstruieren und visualisieren (Abb. 10). Die Baumkrone wird durch einen halben Rotationsparaboloid (eine Parabel als Einhüllende) dargestellt, der Baumstamm durch einen Zylinder, der allerdings für alle Bäume den gleichen Durchmesser hat. Diese Rekonstruktion kann jetzt zum Beispiel für Strahlungstransfermodelle benutzt werden, um die Invertierung von weiteren Forstparametern aus Hyperspektraldaten zu verbessern (KOETZ et al. 2004).

## 5 Zusammenfassung und Diskussion

Wir haben gezeigt, dass es mit einer Kombination aus Bildverarbeitung und Statistischen Analysen möglich ist, geometrische Informationen einzelner Bäumen automatisiert aus LIDAR-Rohdaten zu ermitteln, ohne dabei den Verlust von Information durch Interpolation in ein DOM in Kauf nehmen zu müssen. Sollten vom Anwender bestimmte Parameter auf größeren Skalen verlangt sein, so kann man die auf Einzelbäumen basierenden Werte einfach in einem größeren Raster zusammenfassen. Wir haben diesen Ansatz für Baumhöhen und Kronendurchmesser mit Feldmessungen validiert und zumindest für die Baumhöhen eine gute Übereinstimmung gefunden. Probleme bereitete die Zuordnung der Kronendurchmesser, so dass hier die Qualität der Regression bei weitem nicht so hoch ist wie für die Baumhöhen. Dieses liegt in erster Linie an der speziellen Vegetation im Schweizer Nationalpark, mit vielen eng zusammenstehenden Baumgruppen sowie einem großen Anteil – ca. 20% (DOBBERTIN et al. 2001) – von stehenden toten oder kranken Bäumen, welche nur noch über partielle Kronen verfügen. Allerdings muss an dieser Stelle noch einmal darauf hingewiesen werden, wie feh-



leranfällig Messungen von Kronendurchmessern im Feld sein können. Wir scheinen in einem Bereich der Genauigkeit der Laserscanner angelangt zu sein, wo traditionelle Methoden der Forstinventur zur Bereitstellung von *ground truth* Information anscheinend nicht mehr ausreichend sind. Es stellt sich die Frage, inwiefern terrestrische Laserscanner hier weiterhelfen könnten. Erste Versuche mit Nahbereichsphotogrammetrie unter Verwendung einer Software des Fraunhofer Instituts für Computer Graphik (IGD) in Darmstadt erwiesen sich als nicht hilfreich, da Bäume zu komplexe Strukturen für die Rekonstruktions-Algorithmen darstellten. Als Verbesserungen der Algorithmik ließe sich eine Suche der Startpunkte direkt in den Rohdaten vorstellen, um zumindest einige der Baumgruppen als einzelne Bäume aufzulösen. Weiterhin werden wir den Algorithmus an einem gesunden Bestand mit größeren Kronendurchmessern testen, wo er noch bessere Ergebnisse liefern sollte.

## 6 Danksagung

Dieses Projekt wird durch das EU Projekt *Forest Fire Spread and Mitigation* (SPREAD), EU-Kontrakt Nr. EVG1-CT-2001-00027 und dem Schweizer Bundesamt für Bildung und Wissenschaft (BBW) finanziert; BBW-Contract Nr. 01.0138.

Wir danken STEPHAN HEINER für die Vermessung der Referenzflächen, sowie BENJAMIN KÖTZ und RUEDI HALLER für die Unterstützung bei der Feldarbeit. Ein spezieller Dank geht an TopoSys für die anhaltende Unterstützung und die technischen Informationen. Der Felddatensatz wurde vom Lang-Zeit Waldökosystem Forschungsprogramm LWF bereitgestellt; dies ist ein Partnerschaftsprojekt zwischen der Eidg. Forschungsanstalt für Wald, Schnee und Landschaft (WSL) und dem Bundesamt für Wald, Schnee und Landschaft (BVWAL).

## Literatur

- ANDERSEN, H.-E., REUTEBUCH, S.E. & SCHREUDER, G.F., 2002: Bayesian object recognition for the analysis of complex forest scenes in airborne laser scanner data. – ISPRS Commission III, Symposium 2002, September 9–13, Graz, Austria, pp. A–035 ff (7 pages).
- BALTSAVIAS, E.P., 1999: Airborne laser scanner: existing systems and firms and other resources. – ISPRS Journal of Photogrammetry & Remote Sensing **54** (2–3): 164–198.
- BRANDTBERG, T., WARNER, T.A., LANDENBERGER, R.E. & MCGRAW, J.B., 2003: Detection and analysis of individual leaf-off tree crowns in small footprint, high sampling density LIDAR data from the eastern deciduous forest in North America. – Remote Sensing of Environ. **85** (3): 290–303.
- DOBBERTIN, M., BALTENSWELER, A. & RIGLING, D., 2001: Tree mortality in an unmanaged mountain pine (*Pinus mugo* var. *uncinata*) stand in the Swiss national park impacted by root rot fungi. – Forest Ecology and Management **145**: 79–89.
- DRAKE, J.B., DUBAYAH, R., KNOX, R.G., CLARK, D.B. & CONDIT, R., 2001: Relationship between LIDAR metrics and aboveground biomass in closed-canopy neotropical forests. – International Archives of Photogrammetry and Remote Sensing **XXXIV**-3/W4: 147–154.
- DRAKE, J.B., DUBAYAH, R.O., CLARK, D.B., KNOX, R.G., BLAIR, J.B., HOFTON, M.A., CHAZDON, R.L., WEISHAMPEL, J.F. & PRINCE, S.D., 2002: Estimation of tropical forest structural characteristics using large-footprint LIDAR. – Remote Sensing of Environ. **79**: 305–319.
- FILIN, S., 2002: Surface clustering from airborne laser scanning data. – ISPRS Commission III, Symp. 2002, September 9–13, Graz, Austria, pp. A–119 ff (6 pages).
- GAVEAU, D. & HILL, R., 2003: Quantifying canopy height underestimation by laser pulse penetration in small-footprint airborne laser scanning data. – Canadian Journal of Remote Sensing **29**: 650–657.
- HUBER, P., 1981: Robust Statistics. – Wiley, New York.
- HYYPPE, J., HYYPPE, H., INKINEN, M., ENGBAHL, M., LINKO, S. & ZHU, Y.-H., 1999: Accuracy comparison of various remote sensing data sources in the retrieval of forest stand attributes. – Forest Ecology and Management **128**: 109–120.
- HYYPPE, J., KELLE, O., LEHIKONEN, M. & INKINEN, M., 2001: A segmentation-based method to retrieve stem volume estimates from 3-d tree height models produced by laser scanners. – IEEE Transactions on Geoscience and Remote Sensing **39**: 969–975.

- KOETZ, B., SCHAEPMAN, M., BOWYER, P., MORS-DORF, F., ITTEN, K. & ALLGÖWER, B., accepted: Radiative transfer modeling within a heterogeneous canopy for estimation of forest fire fuel properties. – *Remote Sensing of Environment*.
- KRAUS, K. & PFEIFER, N., 1998: Determination of terrain models in wooded areas with airborne laser scanner data. – *ISPRS Journal of Photogrammetry & Remote Sensing* **53**: 193–203.
- MAAS, H.-G., 2000: Least-squares matching with airborne laserscanning data in a tin structure. – *International Archives of Photogrammetry and Remote Sensing* **33** (3a): 548–555.
- MEANS, J.E., ACKER, S.A., FITT, B.J., RENSLOW, M., EMERSON, L. & HENDRIX, C., 2000: Predicting forest stand characteristics with airborne scanning LIDAR. – *Photogrammetric Engineering & Remote Sensing* **66** (11): 1367–1371.
- MEANS, J.E., ACKER, S.A., HARDING, D.J., BLAIR, J.B., LEFSKY, M.A., COHEN, W.B., HARMON, M.E. & MCKEE, W.A., 1999: Use of large-footprint scanning airborne LIDAR to estimate forest stand characteristics in the western cascades of Oregon. – *Remote Sensing of Environ.* **67**: 298–308.
- NAESSET, E. & BJERKNES, K.-O., 2001: Estimating tree heights and number of stems in young forest stands using airborne laser scanner data. – *Remote Sensing of Environ.* **78**: 328–340.
- NAESSET, E. & OEKLAND, T., 2002: Estimating tree height and tree crown properties using airborne scanning laser in a boreal nature reserve. – *Remote Sensing of Environ.* **79**: 105–115.
- NILSSON, M., 1996: Estimation of tree heights and stand volume using an airborne LIDAR system. – *Remote Sensing of Environ.* **56**: 1–7.
- PERSSON, A., HOLMGREN, J. & SÖDERMAN, U., 2002: Detecting and measuring individual trees using an airborne laser scanner. – *Photogrammetric Engineering & Remote Sensing* **68** (9): 925–932.
- POULIOT, D.A., KING, D.J., BELL, F.W. & PITT, D.G., 2002: Automated tree crown detection and delineation in high-resolution digital camera imagery of coniferous forest regeneration. – *Remote Sensing of Environ.* **82** (2–3): 322–334.
- PYYSALO, U. & HYYPPAE, H., 2002: Reconstructing tree crowns from laser scanner data for feature extraction. – *ISPRS Commission III, Symposium 2002 September 9–13, Graz, Austria*, pp. B–218 ff (4 pp.).
- ROGGERO, M., 2001: Airborne laser scanning: Clustering in raw data. – *International Archives of Photogrammetry and Remote Sensing* **XXXIV-3/W4**: 227–232.
- SCHARDT, M., ZIEGLER, M., WIMMER, A., WACK, R. & HYYPPAE, J., 2002: Assessment of forest parameters by means of laser scanning. – *ISPRS Commission III, Symposium 2002 September 9–13, Graz, Austria*, pp. A–302 ff (8 pp.).
- SPATH, H., 1985: *Cluster Dissection and Analysis: Theory, FORTRAN Programs, Examples*. – Halsted Press, New York, 226 pp.
- ST-ONGE, B.A. & ACHAICHA, N., 2001: Measuring forest canopy height using a combination of LIDAR and aerial photography data. – *International Archives of Photogrammetry and Remote Sensing* **XXXIV-3/W4**: 131–137.
- VOSSELMAN, G., 2002: On the estimation of planimetric offsets in laser altimetry data. – *ISPRS Commission III, Symposium 2002, September 9–13, Graz, Austria*, pp. A–375 ff (6 pp.).

#### Anschriften der Autoren:

Dipl.-Oz. FELIX MORS-DORF  
e-mail: morsdorf@geo.unizh.ch

Dr. ERICH MEIER,

Prof. Dr. DANIEL NÜESCH,

Remote Sensing Laboratories, Department of Geography, University of Zurich

Dr. BRITTA ALLGÖWER, Geographic Information Systems, Department of Geography, University of Zurich

Manuskript eingereicht: Dezember 2003

Angenommen: Januar 2004