

Untersuchungen im 1D- und 2D-Parameterraum von Punktwolken aus Flugzeuglaserscannerdaten

ALEXANDRA D. HOFMANN¹

Zusammenfassung: Über ein Segmentierungsverfahren wurden einzelne, mit hoher Wahrscheinlichkeit gebäudebeinhaltende Punktwolken aus einem Laserscannerdatensatz extrahiert. Die Auswertung ergab, dass zahlreiche Nicht-Gebäudeobjekte mit erfasst wurden. Ziel der Studie war es, diese Objekte zu reduzieren und zusätzlich weitere Informationen von Gebäuden abzuleiten. Die Untersuchungen wendeten statistische Verfahren im 1D und 2D-Parameterraum der Punktwolken an. Die Ergebnisse zeigten, dass die Parameterräume nur bedingt Aufschluss über den Inhalt einer Punktwolke geben und weitere Unterscheidungskriterien für eine eindeutige Klassifikation notwendig sind.

1 Hintergrund

Flugzeuglaserscannerdaten sind zur Zeit ein viel genutztes Medium um Informationen von Gebäuden abzuleiten. Die Erkennung von Gebäuden im Datensatz und deren Extrahierung sind erste Schritte dazu. In früheren Arbeiten (HOFMANN, MAAS & STREILEIN, 2002) wurde festgestellt, dass die auf segmentierten first-pulse Laserscannerdaten basierende Detektierung von Gebäuden ohne Hilfsmittel wie Grundrissinformationen nicht trivial ist. Es zeigte sich, dass nur unter Inkaufnahme eines hohen Anteils an Fehlern 2. Art (als Gebäude erkannte Vegetation) ein Großteil der Gebäude identifiziert werden konnte. Ziel weiterführender Arbeiten war es, diesen Anteil weitestgehend zu reduzieren.

Als Gebäude klassifizierte Segmente wurden mit den rohen Laserscannerdaten verschnitten und die jeweiligen Punktwolken extrahiert. Die Punktwolken wurden zuerst mit einfachen statistischen Parametern, wie der Standardabweichung der Höhenwerte analysiert. Weiterhin wurde ein Ansatz entwickelt, der eine in jede Punktwolke gerechnete TIN-Struktur nutzt. Für jedes der Dreiecke der TIN-Struktur wurden Eigenschaften wie Ausrichtung und Neigungsgrad ermittelt. Die Eigenschaften wurden einzeln mit statistischen Maßen wie Verteilungsformen und auch kombiniert mit einem Ansatz aus der Clusteranalyse untersucht. Mit Hilfe der Clusteranalyse wurden spezifische Unterscheidungen zu den Inhalten einer Punktwolke im 2D Parameterraum erkundet.

Vorteile und Probleme der Ansätze werden im Folgenden ausführlich diskutiert. Es werden außerdem weiterführende Analysen besprochen, die sich mit der Interpretation von Objekteigenschaften (wie die Dachart) in den analysierten Parametern befassen. In einem kurzen Ausblick soll gezeigt werden, welche Möglichkeiten die Weiterführung der Clusteranalyse im 3D-Parameterraum erwarten lässt.

2 Daten

Die verwendeten first-pulse Flugzeuglaserscannerdaten haben eine mittlere Punktdichte von ca. 1,2 m. Die Standardabweichung beträgt ca. 30 cm in x und y Richtung und ca. 20 cm in z. Da planimetrische Fehler der Laserscannerdaten hauptsächlich durch das GPS und INS-System des Flugzeuges hervorgerufen werden, können diese innerhalb kleiner Objekte fehler-

¹ Dipl.-Ing. Alexandra D. Hofmann, Institut für Photogrammetrie und Fernerkundung, Technische Universität Dresden, Mommsenstr. 13, 01062 Dresden, e-mail: Alexandra.Hofmann@mailbox.tu-dresden.de

technisch als korreliert verstanden werden. Das heißt, nur der Fehler in der z-Koordinate ist zu beachten. Diese Annahme kann bei einer mittleren Objektgröße von ca. 300 m², die in dieser Studie angewendet wird, vorgenommen werden.

Die analysierten Studiengebiete sind ländliche Gebiete mit alpinem Charakter - das erste mit vereinzelter Bebauung und das zweite mit dichteren Ortskernen. Der in beiden Studiengebieten vorherrschende Gebäudetyp ist ein Einzelhaus mit Satteldach. Walmdächer und Pultdächer oder Kombinationen treten nur selten auf. Abb. 1 gibt Beispiele. Der Charakter der Region wird durch Felder mit Weidenbäumen und kleineren Waldflächen bestimmt.

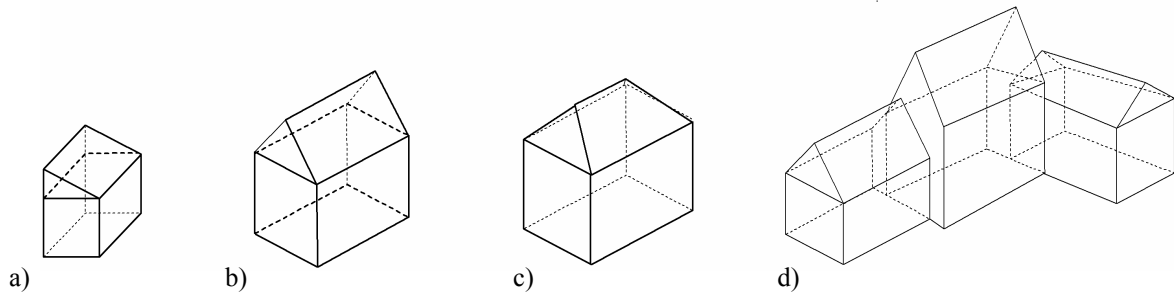


Abb. 1 a) Pultdach b) Satteldach c) Walmdach d) Kombination

Beide Areale, jedes ca. 7 km² groß, sind mit eCognition, wie in HOFMANN, MAAS & STREILEIN (2002) beschrieben, segmentiert und klassifiziert worden. All jene Segmente, die eine größere Wahrscheinlichkeit als 80% für das Vorhandensein eines Gebäudes haben, wurden mit einer Bufferzone von 3 m versehen und mit den rohen Laserscannerpunktdaten verschnitten. Die Laserscannerpunkte der Segmente wurden extrahiert und, wie im Folgenden beschrieben, analysiert. Es sei bemerkt, dass die Segmentgröße die Objektgröße um ein Vielfaches übersteigen kann. Die Untersuchungen wurden an insgesamt ca. 5500 als Gebäude klassifizierten Punktwolken durchgeführt, wovon tatsächlich nur 1310 Punktwolken Gebäude beinhalten.

3 Statistische Verteilungsanalyse im eindimensionalen Parameterraum

Erste Untersuchungen der Punktwolken wurden durch statistische Parameter, wie der Standardabweichung der Höhenwerte oder den mittleren Neigungsgrad realisiert. Es wurde angenommen, dass vor allem Punktwolken, die Waldflächen darstellen, über eine erhöhte Standardabweichung der z-Werte bestimmt und eliminiert werden könnten. Es bestand auch die Annahme, dass über den mittleren Neigungsgrad Wiesenflächen detektierbar sind, da jene i.d.R. eine flächenhafte Struktur besitzen.

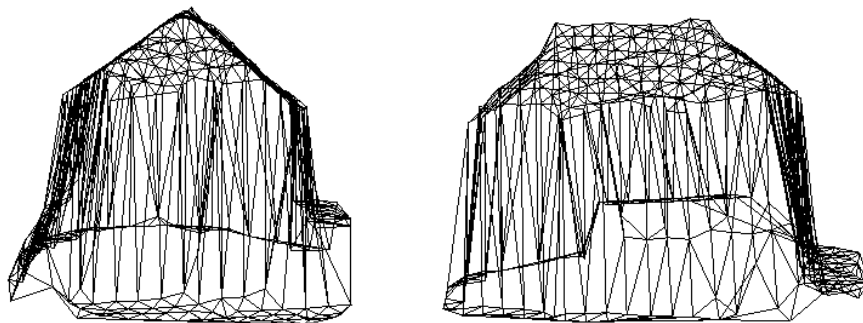


Abb. 2 Beispieldarstellung eines Gebäudes mit Walmdach durch eine TIN-Struktur

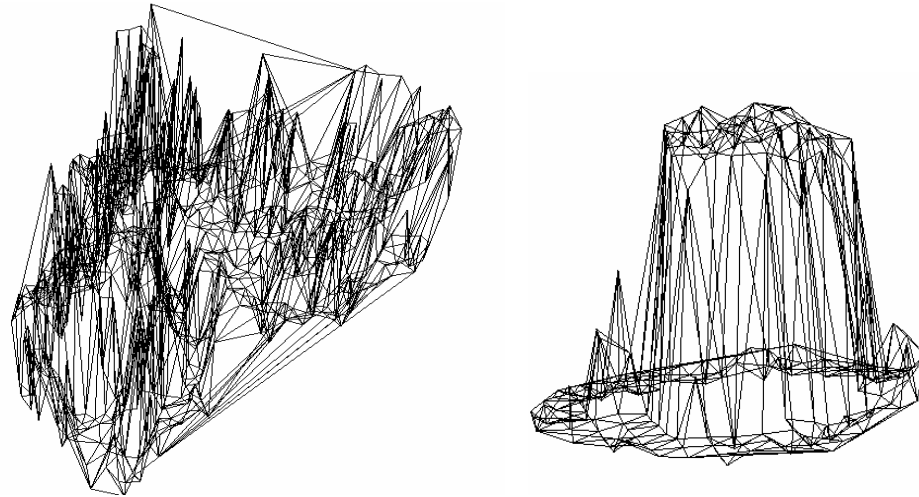


Abb. 3 Beispieldarstellung eines Waldstückes und eines Baumes durch eine TIN-Struktur

Eine Analyse dieser allgemeinen Parameter der Punktwolken brachte keine nutzbringenden Erkenntnisse. Eine detailliertere Untersuchung ist somit nötig. Hierzu wurde in jede Punktwolke mittels einer Delaunay-Triangulation mit dem Modul Triangle von SHEWCHUK (1996) eine TIN-Struktur gerechnet. Abb. 2 zeigt eine Frontal- und Seitenansicht eines Gebäudes mit Walmdach und Abb. 3 ein Waldstück und einen Baum, welche mit einer solchen TIN-Struktur dargestellt sind. Für jedes Dreieck sind Ausrichtung und Neigungsgrad bekannt.

3.1 1D-Parameteranalyse

Aus dem ersten Studiengebiet wurden ca. ein Drittel (800) der Objekte willkürlich ausgewählt, um daraus Parameter abzuleiten, mit denen beide Studiengebiete analysiert werden sollen. Dieses Drittel enthielt eine repräsentative Auswahl an Objekten.

Erste Versuche, Gebäudepunktwolken von Nicht-Gebäudepunktwolken zu trennen, bezogen sich auf die Verteilung der Ausrichtung der Dreiecke aus der TIN-Struktur. Im Histogramm des Gebäudes im Abb. 4 sind gut vier Peaks zu erkennen, wobei die beiden höheren durch Dreiecke die in Dachrichtung liegen hervorgerufen worden.

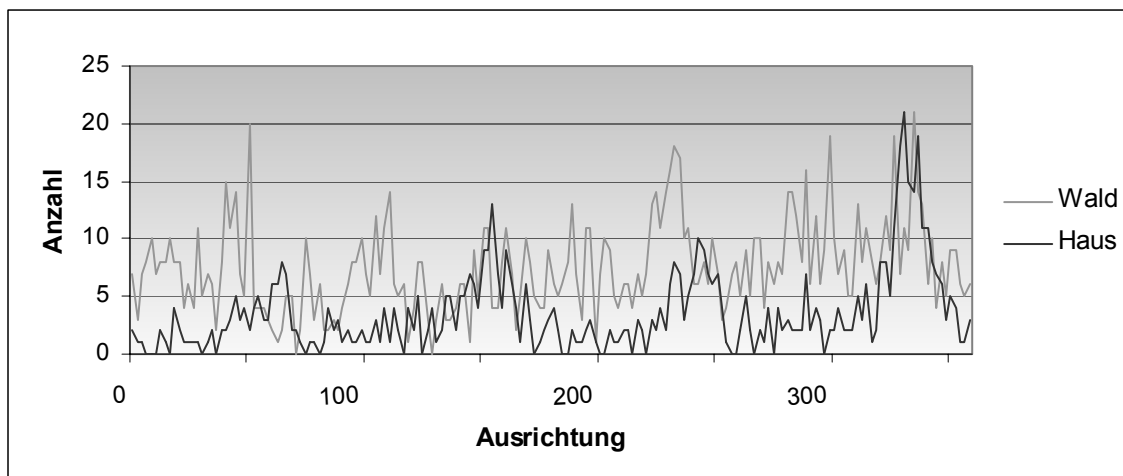


Abb. 4 Histogramm der Ausrichtungsverteilung

Die Besonderheit dieser Peaks ist, dass sie jeweils 90 Grad zueinander versetzt sind. Eine derartige Struktur lässt sich im Histogramm der Waldpunktwolke nur unter Umständen wiederfinden. Für eine statistische Auswertung dieser Kurven treten zu viele Mehrdeutigkeiten

auf. Nachteilig ist vor allem die Schwankung der Intensität der Peaks von Gebäudepunkt-
wolken in Abhängigkeit von Dachgröße, Dachart und Ausdehnung der Punktwolke.

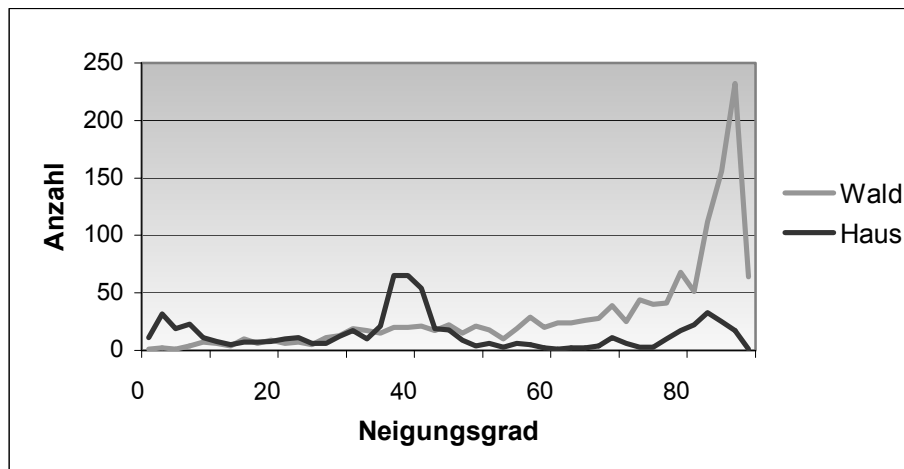


Abb. 5 Histogramm der Verteilung der Neigungsgrade

Die Verteilung der Neigungsgradwerte der Dreiecke der selben Objekte ist in Abb. 5 illustriert. Für das Waldobjekt ist der hohe Anteil an Dreiecken größeren Neigungsgrades auffällig. Dieser Verlauf ist typisch für Waldflächen. Da für Gebäude von einer maximal zulässigen Dachneigung von 75 Grad ausgegangen werden kann, wurde für jede Punkt-
wolke der Anteil an Dreiecken mit einem größeren Neigungsgrad als 76 Grad bestimmt. Aus den gesammelten Informationen konnte empirisch die Regel abgelitten werden, dass Punkt-
wolken mit mehr als 30% solcher Dreiecke Waldobjekte darstellen. Eine Regel für Wiesenflächen konnte nicht gefunden werden, da jene in diesem Parameterraum sehr häufig flachen Gebäuden ähneln.

3.2 Ergebnisse

Beide Studiengebiete wurden mit der oben beschriebenen Regel in Gebäude und Nicht-Gebäudeobjekte unterteilt. Tab. 1 fasst die Resultate zusammen. Die Analyse stellt sich als recht effektiv dar. Das Ziel, keine Gebäudeobjekte (Fehler 1. Art) zu verwerfen, wurde insofern eingehalten, als das die vier eliminierten Gebäude (0,35 %) sich als Objekte mit sehr schlechter Laserscannerdatenqualität herausstellten und somit vegetationsähnliche Eigenschaften aufwiesen.

Tab. 1: Anteil an eliminierten Objekten

Eliminierte Objekte	Studiengebiet 1	Studiengebiet 2
Nicht-Gebäude:	54 %	31 %
Gebäude:	0,5 %	0,2 %
Anteil an eliminierten Fehlern 2. Art		42,5 %
eliminierten Fehlern 1.Art		0,35 %

Das Gebäude-Nichtgebäude-Verhältnis ist für das erste Studiengebiet von 1:5 auf 1:3 und für das zweite Studiengebiet von 1:3 auf 1:2 reduziert worden. Der im ersten Studiengebiet an sich höhere Anteil an Fehlern 2. Art kann nur durch eine ungünstig verteilte Objektvielfalt erklärt werden. In diesem Falle mit einem größeren Anteil an Wiesenflächen.

4 Clusteranalyse im zweidimensionalen Parameterraum

Die Betrachtung der Abb. 4 und 5 impliziert, dass die Darstellung der einzelnen Dreiecke in einem Ausrichtungs-Neigungsgrad-Diagramm, wie in Abb. 6 zu sehen ist, Aufschluss über das Objekt geben kann. Im Gebäudediagramm der Abb. 6 sind die Clusters der beiden Längsseiten des Walmdaches eingekreist. Die Mittelpunkte der Clusters geben die ungefähre Ausrichtung und Neigung der jeweiligen Dachseite wieder. Eine solche Struktur ist nicht im Diagramm der Waldfläche zu finden. Schlussfolgernd könnte aufgrund von logischen Clusterkombinationen auf Gebäude geschlossen werden. Somit könnten in diesem 2D-Parameterraum auch weitere Fehler 2. Art eliminiert werden.

4.1 Clusteralgorithmus

Die Clusterung wurde durch einen agglomerativen Ansatz mit single-linkage Zuweisung realisiert. Diese begrenzt nicht die Größe der Cluster, trennt jedoch dichtere Punktgruppen von entfernter liegenden Punkten. Clusteralgorithmen die die Kenntnis über der Anzahl der Clusters voraussetzen, konnten nicht angewendet werden, da die Clusteranzahl vom Objekt abhängt und somit nicht bekannt ist. Eine automatische Bestimmung der Anzahl der Cluster erwies sich als nicht erfolgreich.

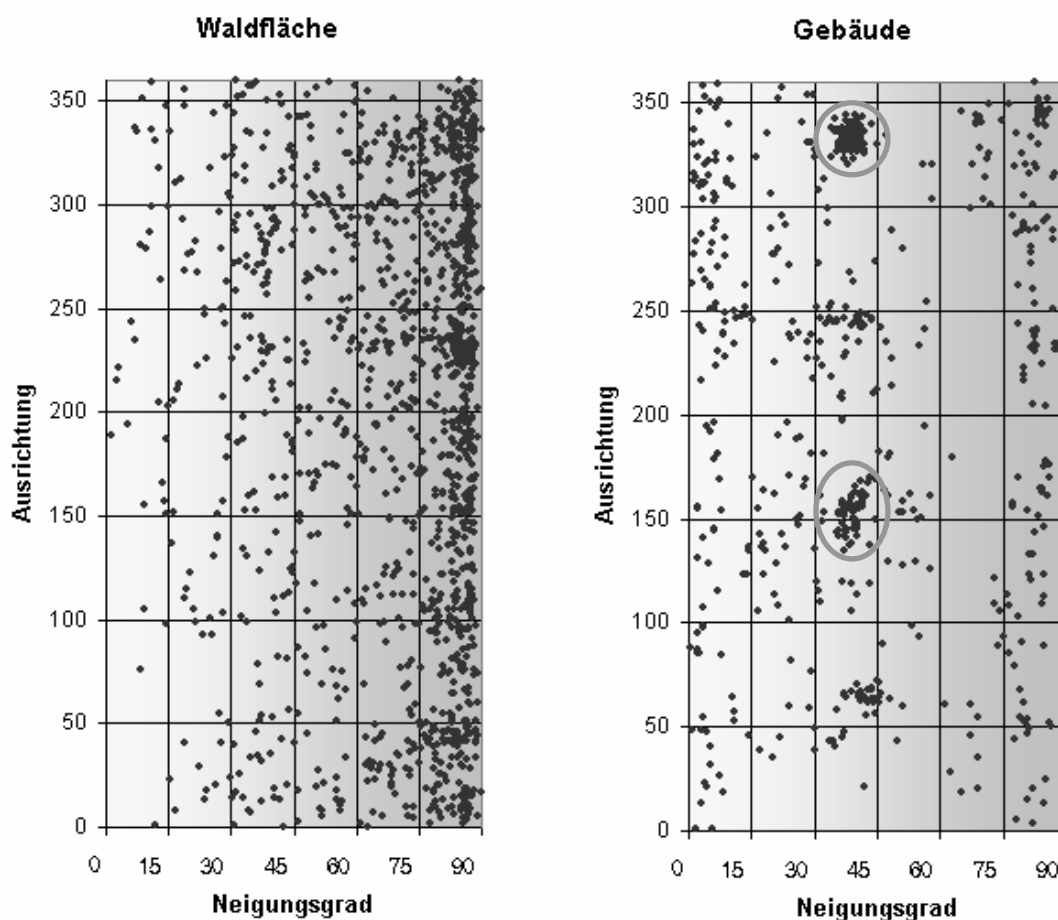


Abb. 6 2D Verteilung der Dreiecke einer Waldfläche und eines Gebäudes mit Walmdach

Damit zu jedem Punkt die direkten Nachbarn und deren Abstand zueinander bekannt sind, wurde ein Netzwerk in die Parameterpunkte gerechnet. Ein willkürlicher Startpunkt wurde unter den Parameterpunkten gewählt und sämtliche Distanzen zu den direkten Nachbarn verglichen. Lag mindestens eine der Distanzen unter einer festgestellten Grenze (maximale Dis-

tanz), wurde ein Cluster gebildet und die entsprechenden nahen Nachbarn zum Startpunkt gruppiert. Von diesem nahen Nachbarn wiederholt sich der Prozess bis kein neuer naher Nachbar mehr gefunden wurde. Der Cluster ist somit abgeschlossen und ein neuer Startpunkt kann gewählt werden. Cluster mit weniger als 13 Parameterpunkten wurden nicht als Dachcluster zugelassen, um eventuelle zufällige Clusterbildungen auszuschließen. Folglich wird eine minimale Dachfläche von 13 Dreiecken pro Dachseite festgelegt.

Als problematisch stellt sich die Ermittlung der maximalen Distanz zwischen nahen Parameterpunkten heraus. Diese wird indirekt durch den Fehler der z-Koordinate der Laserscannerdaten bestimmt. Es ist zu beachten, dass je kleiner der Neigungsgrad, desto größer der Einfluss der Fehler der Laserscannerdaten auf die Ausrichtung und Neigung der Dreiecke. Flächen mit einer Neigung kleiner als 7 Grad bilden keine nachvollziehbaren Cluster. Es gilt auch, dass je größer die Punktwolke, desto mehr Dreiecke und umso kleiner das Signal-Rausch-Verhältnis. Je nach Inhalt der Punktwolke konzentrieren sich die Parameterpunkte mehr oder weniger in Clusters. Die maximale Distanz bestimmt sich nun aus der Streuung der Dreiecke (aufgrund der Fehler der z-Koordinate) in Abhängigkeit vom Neigungsgrad und der Konzentration der Parameterpunkte. Um Clusters auch in Punktwolken mit mehr Rauscheffekten zu detektieren, wurde die Größe der Clusters bestimmt. Wenn mindestens 80% aller Parameterpunkte in einem Cluster gruppiert waren, wurde die Distanz schrittweise reduziert und die Clusterung erneut gestartet.

4.2 2D Parameteranalyse

Jede Punktwolke wurde der Clusteranalyse unterzogen. Gefundene Clusters wurden auf ihre Eigenschaften geprüft. Der logische Neigungsgrad einer Dachseite liegt zwischen 7 Grad und 75 Grad. Die Dachform des Gebäudes sollte aus der Anzahl und Position der Clusters abgeleitet werden. Bei z.B. einem Satteldach sollten sich zwei Clusters mit 180 Grad unterschiedlicher Ausrichtung auffinden. Wenn wenigstens ein Cluster mit einer logischen Position und Größe vorhanden war, wurde der Punktwolke ein Gebäude unterstellt. Folgender Entscheidungsbaum wurde genutzt:

- 0 Cluster: kein Gebäude
- 1 Cluster mit schlüssigem Neigungsgrad: Gebäude mit Pultdach
- 2 Cluster mit sinnvollen Neigungsgraden und ca. 180° versetzter Ausrichtung: Satteldach
- Mehr als 2 Cluster mit sinnvollen Neigungsgraden und logisch versetzten Ausrichtungen (d.h. 180° bzw. 90°): kombiniertes Dach

Aussagen, ob es sich um ein Walmdach oder zwei Satteldächer handelt, können nicht vorgenommen werden. Hierzu ist eine dritte Dimension nötig.

4.3 Ergebnisse

Die Ergebnisse der 1D-Parameteranalyse werden als Vorinformation behalten. Somit werden nur als Gebäude klassifizierte Punktwolken mit der Clusterung analysiert. Tab. 2 zeigt, dass die Clusteranalyse den Anteil erkannter Fehler 2. Art auf 62 % erhöht. Identifizierte Nicht-Gebäudeobjekte sind jetzt neben Waldflächen auch zahlreiche Bäume und Waldrandgebiete. Ein Hauptgrund für den abnehmenden Wirkungsgrad ist, dass teilweise Bäume als Satteldächer und Wiesen- und Waldgebiete als Pultdächer interpretiert wurden. Das heißt, es bildeten sich ein oder mehrere zufällige Clusters im 2D-Parameterraum. Dieses Problem könnte mit einem 3D-Parameterraum, in dem die Position der Dreiecke einbezogen wird, gelöst werden. Die Clusteranalyse verwarf ca. 21 % der Gebäude. Dieses Ergebnis ist so suboptimal und wird näher untersucht. Statistische Proben ergaben, dass 34 von 100 Gebäuden mangelhaft in ihrer Datenqualität sind. Mangelhaft heißt, dass mindestens eine Dachseite im Laserscanner-

datensatz zweimal mit unterschiedlichen Höhen und zum Teil unterschiedlicher Lage aufgenommen wurde. Ein Beispiel ist in MAAS (2003) gegeben. Die Parametereigenschaften der Dachseiten werden dabei verfälscht und können durch die Clusteranalyse nicht mehr erkannt werden. Ca. 50 % der verworfenen Gebäude wiesen diese Mängel in den Laserscannerdaten auf.

Tab. 2: Anteil an eliminierten Objekten

Eliminierte Objekte	Studiengebiet 1	Studiengebiet 2
Nicht-Gebäude:	70 %	55 %
Gebäude:	23 %	20 %
Anteil an		
eliminierten Fehlern 2. Art		62,5 %
eliminierten Fehlern 1.Art		21,5 %

Weitere Gründe für das Verwerfen von Gebäuden sind zu kleine Dachflächen oder weitere in der Punktwolke vorhandene Objekte wie Bäume, die sich wie Rauschen auswirken. Dachclusters können somit nicht eindeutig oder gar nicht identifiziert werden.

Die Richtigkeit der Dachformbestimmung ist nur teilweise mit der Dachform selbst korreliert, sondern eher mit der Größe des Daches und der Umgebung des Gebäudes. Tab. 3 bestätigt diesen Eindruck. Die extrem geringe Trefferzahl für kombinierte Dächer ist mit der erhöhten Punktdichte bei teilweise kleinerer Dachfläche und somit einer schlechteren Trennbarkeit der einzelnen Clusters zu erklären. Die fehlende 3. Dimension ist wieder offensichtlich.

Tab. 3: Ergebnisse der Spezifikation der Dachform durch eine 2D-Clusteranalyse

Dachform	Richtig klassifiziert	Gar nicht klassifiziert
Pultdach	56 %	25 %
Satteldach	29 %	23 %
Kombination	12 %	22 %

Für die Klassifizierung der Dachart konnte nur nach Pult-, Satteldach oder Kombination unterschieden werden. In Tab. 3 bedeutet „Richtig klassifiziert“, wenn die Zuweisung der Dachform korrekt war und „Gar nicht“, dass kein Haus entdeckt wurde; die restlichen Prozente sind mangelhafte oder falsche Zuweisungen. Tab. 3 bestätigt auch, dass das Verwerfen eines Gebäudes nicht mit der Dachform korreliert ist.

5 Zusammenfassung

Die Ergebnisse zeigen, dass allein mit der Analyse des Neigungsgrades einer Punktwolke eine recht gute Entscheidung getroffen werden kann, ob es sich um ein Gebäudeobjekt handelt oder nicht.

Die 2D-Clusteranalyse erweist sich prinzipiell als Möglichkeit, Punktwolken differenzierter zu betrachten, ist allerdings mit der Variabilität von Objekten überfordert. Besonders die ungleichmäßige Größe des Objektfeldes wirkt sich störend auf die Analyse aus. Die Clusteranalyse ist schwer zu automatisieren, wenn starke Schwankungen in der gesuchten Objektgröße und im Rauschverhalten vorliegen. Zudem kann der wichtigste Parameter, die maximale single-linkage Entfernung, nicht auf die Anzahl der Dreiecke angepasst werden, da der

Inhalt der Punktwolke nicht bekannt ist. Dieser Parameteransatz ist außerdem sehr sensibel bezüglich Schwankungen in der Laserscannerdatengenauigkeit. Je fehlerbehafteter die Daten sind, desto größer ist die Clusteraufweitung und die maximale single-linkage Entfernung wird leichter überschritten.

Generell wurde eine höhere Erfolgsquote für die Erkennung der Fehler 2. Art erwartet. Weitere Prüfungen sollen klären, inwiefern Qualitätsmängel der Laserscannerdaten Einfluss auf das Ergebnis hatten. Es ist möglich, dass die Datenqualität ebenfalls Ursache für die hohe Anzahl an Fehlern 2. Art war.

Als Fortführung zur Clusteranalyse soll die Verwendung eines 3. Parameters zur Dachformenklassifizierung diskutiert werden. Es soll gezeigt werden, dass aufgrund der Vielfalt der Objekte und Bautypen auch im 3D-Parameterraum eine klare Klassifizierung nur bedingt möglich ist (HOFMANN, MAAS & STREILEIN, 2003). Es ist angedacht 4. und evtl. 5. Parameter hinzuzunehmen. Weitere Untersuchungen werden auch zeigen, wie exakt die aus der Clusteranalyse gewonnenen Dacheigenschaften sind.

7 Danksagung

Diese Studie wurde vom Schweizer Bundesamt für Landestopographie unterstützt. Die Autorin dankt dem Schweizer Bundesamt für Landestopographie für die Bereitstellung der Flugzeuglaserscannerdaten.

8 Literaturverzeichnis

- HOFMANN, A.D., MAAS, H.-G., STREILEIN, A. 2002: Knowledge-based building detection based on laser scanner data and topographic map information - IAPRS International Archives of Photogrammetry and Remote Sensing and Spatial Information Sciences Vol.34, Part 3A, pp.169-174
- HOFMANN, A.D., MAAS, H.-G., STREILEIN, A. 2003: Derivation of roof types by cluster analysis in parameter spaces of airborne laserscanner point clouds – ISPRS Commission III WG3, Workshop, 3-D reconstruction from airborne laserscanner and InSAR data, 8. - 10. October, Dresden, Germany (eingereicht)
- MAAS, H.-G. 2003: Planimetric and height accuracy of airborne laserscannerdata: User requirements and system performance - Photogrammetric Week 2003, Wichmann Verlag
- SHEWCHUK, J.R. 1996: Triangle: Engineering a 2D Quality Mesh Generator and Delaunay Triangulator - First Workshop on Applied Computational Geometry (Philadelphia, Pennsylvania), pp. 124-133, ACM