



Punktwolken, und was dann?

Dieter Fritsch, Stuttgart

Kurzfassung

Mit der zunehmenden Automatisierung der Datenerfassung mittels Laserscanning und Photogrammetrie werden hochaufgelöste und hochqualitative Punktwolken zur Verfügung gestellt. Deren Interpretation wird im folgenden Beitrag diskutiert.

Mittels Grundrissinformation können Gebäudewände extrudiert und passende Dachformen vollautomatisch angepasst werden (LoD2). Die Interpretation von Punktwolken für Gebäudefassaden ist mittels formalen Grammatiken durchzuführen (LoD3). Punktwolken in Gebäudeinnenräumen sind mittels der Hypothese „Manhattan-Geometrie“ ebenso vollautomatisch zu rekonstruieren (LoD4). Erweiterungen der Fassadengrammatik hinsichtlich der Besonderheiten von Innenräumen belegen ebenso das Potenzial für automatische Ansätze, die derzeit weiter erforscht werden.

Schlüsselwörter: 3D Computermodelle, Level-of-Details (LoD), Bilddatenerfassung, Laserscanning, 3D Punktwolken, Formale Grammatiken, LoD3 Modellierung, LoD4 Modellierung

Abstract

Laser scanning and photogrammetry are delivering high density and high quality point clouds, in more or less automated processing pipelines. Their interpretation is dealt with in the following.

Using building footprints corresponding walls are extruded and tied together with best-fit roof landscapes estimated from a sparse set of points (LoD2). The interpretation of point clouds for building façades can be utilized using formal grammars (LoD3). Indoor point clouds – fulfilling the Manhattan geometry hypothesis – are reconstructed fully automatically using sweep algorithms for the walls, floors and ceilings. Extensions of formal grammars, so far used for façade reconstructions, allow also for automated prediction of 3D indoors elements, but this will be explored in near future.

Keywords: 3D Computer Models, Level-of-Details (LoD), Image Data Collection, Laser Scanning, 3D Point Clouds, Formal Grammars, LoD3 Modelling, LoD4 Modelling

1. Einleitung

Mit der zunehmenden Automatisierung der Datenerfassung mittels Laserscanning (LS) und Photogrammetrie (airborne, terrestrisch, mobil) hat sich ein Paradigmenwechsel eingestellt, der mit einer einfachen Überschrift beschrieben werden kann: „Vom Punkt zur Punktwolke“. Mit der Renaissance der photogrammetrischen Datenerfassung ergibt sich eine zum Laserscanning ebenbürtige Erfassungsmethode, die mittels dichter Bildzuordnung (Dense Image Matching) sehr dichte und homogene Punktwolken im Objektraum erzeugen kann. Durch eine Integration beider Punktwolken können ebenso Nachteile beider Methoden überwunden werden, wie z. B. fehlende oder schwache Reflektion des LS einerseits und radiometrische Defekte der Photos andererseits.

Doch Punktwolken haben den Nachteil, dass sie – ähnlich wie Rasterdaten – große Datenvolumina belegen und interpretiert (klassifiziert) werden müssen. Dies ist eine große Heraus-

forderung. Zum einen können Filtermethoden eingesetzt werden, um die Datenmengen zu reduzieren und ggfs. die Punktwolken so ausdünnen, dass lediglich objektrelevante Information erhalten bleibt. Zum anderen wird nach Methoden gesucht, die diese Punktwolken in vektorielle Objekte überführen.

Diese Methoden können manuell, halbautomatisch und vollautomatisch die Punktwolken weiterverarbeiten, ganz im Sinne der Überschrift: „Punktwolken, und was dann?“. Am Institut für Photogrammetrie der Universität Stuttgart wird seit zwei Jahrzehnten sehr intensiv an der Entwicklung von Methoden zur Punktwolkengenerierung, -reduzierung und -interpretation von 3D-Stadtmodellen gearbeitet. Der Ursprung der Punktwolke spielt dabei keine Rolle, obwohl jüngst ebenso Methoden zur Integration von Punktwolken erarbeitet worden sind.

Die anfänglichen Arbeiten bezogen sich auf die Interpretation von dünnen Punktwolken des flugzeuggetragenen Laserscannings. Hier wur-



Abb. 1: (a) Laserscanning Punktwolke und Orthophoto (b) Extrudierte Gebäude mit approx. Dachflächen (Schlossplatz Stuttgart)

den lediglich 4 Pkt/m^2 vorausgesetzt, Stand 1997/98. Durch erste Dissertationen [1,2,3] wurden Verfahren erforscht, um vollautomatisch komplexe Dachformen aus Laserscanning-Punktwolken zu extrahieren (siehe Abbildung 1). Ein erster Durchbruch in der vollautomatischen Generierung von 3D-Stadtmodellen wurde mit [3] erzielt. Diese Methode wurde in [8] erneut auf den Prüfstand gestellt und durch weitere Ansätze ersetzt. Eine daraus resultierende Software wird derzeit von einem privaten Anbieter zur vollautomatischen Rekonstruktion von 3D-Stadtmodellen erfolgreich eingesetzt.

Im terrestrischen Laserscanning überwiegt derzeit noch die manuelle Interpretation durch CAD-Modellierung, die anschließend texturiert wird (siehe Abbildung 2).

Die Computergraphik nutzt neben Punktwolken auch Photos und deren Fluchtpunktgeometrie zur 3D-Rekonstruktion von LoD2. Dabei kann die Kameraposition rein grafisch als auch rechnerisch durch Structure-from-Motion Verfah-

ren [22] ermittelt werden (siehe Abbildung 3). Heutzutage stehen Millionen von Photos zum Download im Internet bereit, durch volontär aufgefüllte Bilddatenbanken wie Flickr, Picasa, Panoramio u.v.m. Eine einfache Google-Suche findet dann die Bilder, die für die Erzeugung von LoD2-Modellen eingesetzt werden können.

Es wird der Versuch unternommen, Ansätze zur vollautomatischen Interpretation von Punktwolken aufzuzeigen. Dachlandschaften können durch vorgegebene Bauelemente zu komplexen Einheiten zusammengefasst und approximativ aus zugehörigen Punkten des Oberflächenmodells berechnet werden. Eine vollautomatische Interpretation mittels Fassadengrammatiken ist in der Lage, anspruchsvolle LoD3-Modelle zu entwickeln. LoD4, also Innenraummodelle, sind ebenso automatisch zu generieren, wenn man sich zum einen auf Manhattan-Geometrien beschränkt und andererseits Innenraum-Grammatiken einsetzt. Abschließend wird versucht, die neuen Entwicklungen von Virtual und Augmented Reality in den Ausblick zu integrieren.

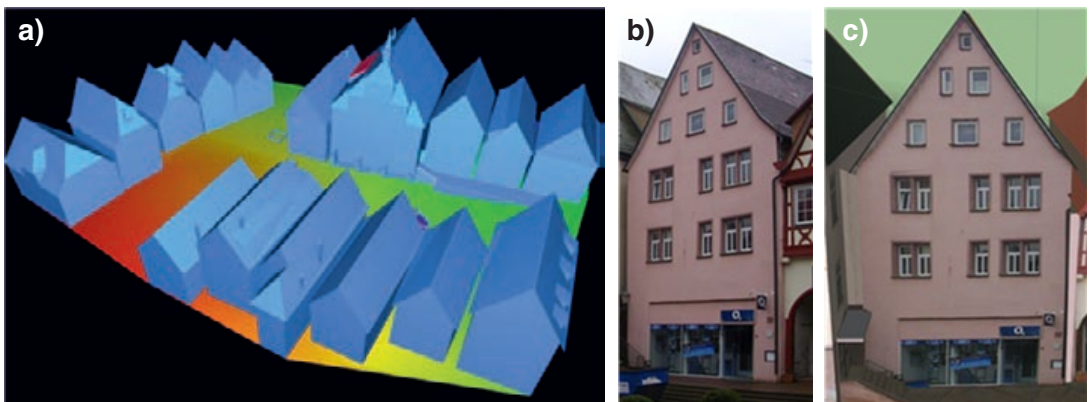


Abb. 2: (a) CAD-Modell aus LS-Punktwolke (b) Photo (c) Texturiertes 3D-Modell

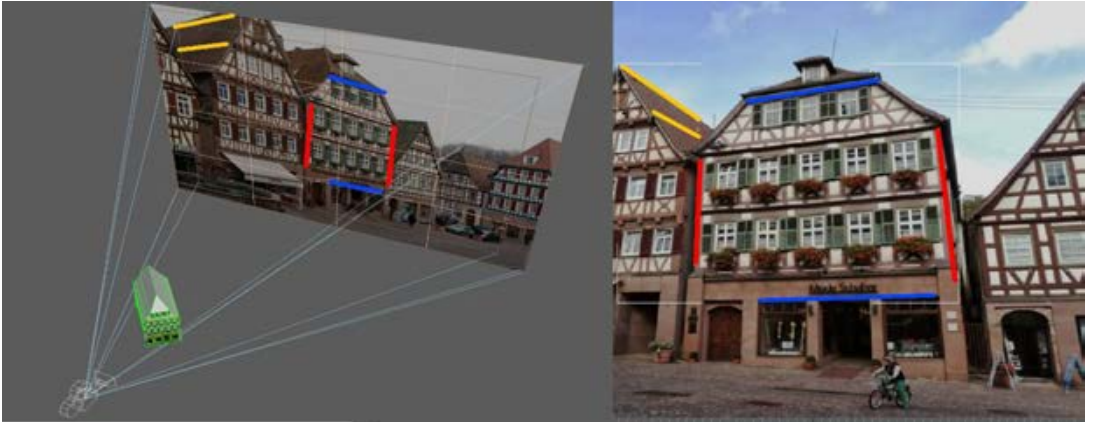


Abb. 3: Manuelle Bestimmung der Kameraposition mittels Fluchtpunktgeometrie

2. Laserscanning

Mit der Einführung des Lasers wurden unter dem Akronym LiDAR (Light Detection and Ranging) in den 1960er Jahren die Voraussetzungen zur Entwicklung von Laser-Entfernungsmessern geschaffen [14]. Darunter versteht man die automatische Entfernungsmessung mit gepulsten oder kontinuierlich strahlenden Laserlichtquellen. Im ersten Fall wird die Entfernung durch Laufzeitmessung von gepulsten Laserphotonen ermittelt,

die aufgrund ihrer hohen Energie reflektiert werden und auf diese Weise das Zeitintervall zwischen ausgesandtem und empfangenem Photon(s) die Messgröße darstellt. Im zweiten Fall stellt die Phasendifferenz zwischen ausgesandtem und reflektiertem Signal die Messgröße dar, die mit mm-Genauigkeit die Reststrecke bestimmt, und zu der ein Vielfaches der Wellenlängen hinzuaddiert werden muss. Aufgrund von Sicherheitsbestimmungen, speziell im Hinblick



Abb 4: Typische Punktwolke des Mobile Mapping „auf der Straße“

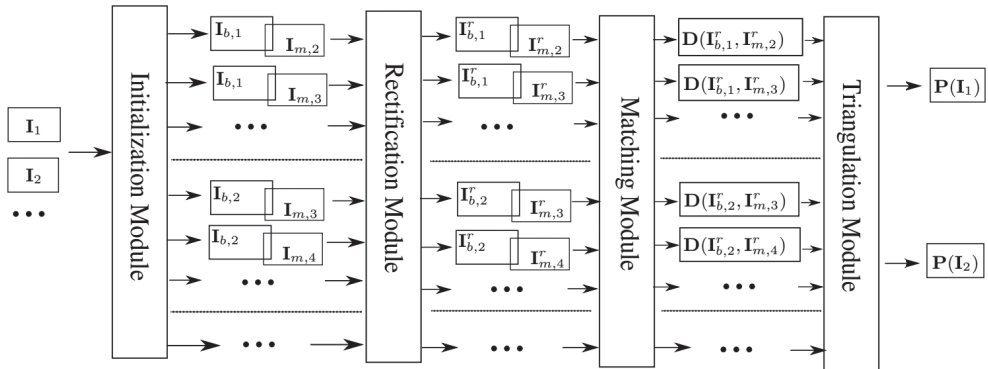


Abb. 5: Prozessierungsschritte bei dichter Bildzuordnung, am Beispiel von SURE

auf Augensicherheit, sind die gepulsten Systeme überwiegend im Einsatz. Neuerdings kombiniert man gepulste und kontinuierlich strahlende Laserlichtquellen und kommt auf Datenraten von 1 Million Messpunkte pro Sekunde. Diese hohe Punktdichte hat auch in der Vermessungstechnik zum Paradigmenwechsel geführt: „Vom Punkt zur Punktwolke“.

Obwohl LiDAR bereits in den 1980er Jahren im Flugzeug und/oder Helikopter zur Erfassung von Höhenprofilen, seit den 1990er Jahren zur flächendeckenden Erfassung (Scanning) von Oberflächenmodellen eingesetzt wird, dauerte es weitere 10 Jahre, bis ab 2005 erste kommerzielle System für das Mobile Mapping zur Verfügung standen [20].

Die Motivation für die Anwendung von Mobile-Mapping liefert die Punktwolke selbst (siehe Abbildung 4): Mittels der sehr dichten 3D-Informationen können Häuserfassaden, Straßen- und Schienenoberflächen, Tunnel und weitere Objekte mm- bzw. cm-genau erfasst werden. Diese erleichtern ebenso die Inventur von Stadtmobiliar wie z.B. Straßenschilder, Verkehrsampeln, Hydranten und vieles mehr. Es lassen sich ebenso Aussagen zur Qualität der Straßenoberflächen wie auch -markierungen ableiten wie auch eine Detektion und Analyse von Oberflächenschäden durchführen. Beim Schienenmonitoring werden Lichtraumprofile sowie Längs- und Querneigungen der Gleisanlagen erfasst, eine Überwachung von Kabelschächten und Tunnelröhren ist leicht möglich wie auch die exakte Positionierung der Oberleitung.

3. Dichte photogrammetrische Bildzuordnung

Mit der Entwicklung des Semi-Global Matching (SGM) [5,9] wurden die Grundlagen für eine

Renaissance der Photogrammetrie gelegt. Die Photogrammetrische Woche 2011 stand ganz im Kontext dieser Entwicklung und hat die Frage aufgeworfen, welches Verfahren letztlich die besseren Punktwolken liefert. Dabei wurde festgestellt, dass beide Verfahren Stärken und Schwächen haben, die vielfach komplementär sind und daher eine Stützung durch Integration gewünscht ist [13].

Beim SGM erfolgt die Verarbeitung von Stereo und Mehrfach-Stereo-Aufnahmen Pixel für Pixel in einem ausgeklügelten Algorithmus (siehe Abbildung 5). Das am Institut für Photogrammetrie der Universität Stuttgart entwickelte Softwarepaket SURE (Surface REconstruction from Imagery) nutzt eine etwas abgeänderte Version des Semi-Global Matching [17]. Nach der Orientierung der Bilder, z.B. durch einen Kalman-Filter (beim Mobile Mapping) oder aber durch Verfahren des Computer Vision wie Structure-from-Motion (SfM) erfolgt ein Initialisierungsschritt, um die einander zuzuordnenden Bildpaare aufzubereiten. Die Bilder werden von Verzerrungen und Abbildungsfehlern befreit, entzerrt und in eine Epipolarbildgeometrie umgerechnet. Danach erfolgt der Zuordnungsschritt, bei dem die Parallaxen/Disparitäten entlang der Stereobasis berechnet und mit diesen im sogenannten Triangulationsmodul durch Vorwärtsschnitt die Punkte der Punktwolke berechnet werden. Neuere Entwicklungen filtern die resultierenden Punktwolken, um bereits zu diesem Zeitpunkt das Datenvolumen deutlich zu reduzieren.

Zur bildbasierten vollautomatischen Fassadenrekonstruktion sollen hier beispielhaft drei wesentliche Arbeitsschritte demonstriert werden. In einem ersten Schritt sind Bilder der Fassade(n) zu erfassen. Dabei ist darauf zu ach-

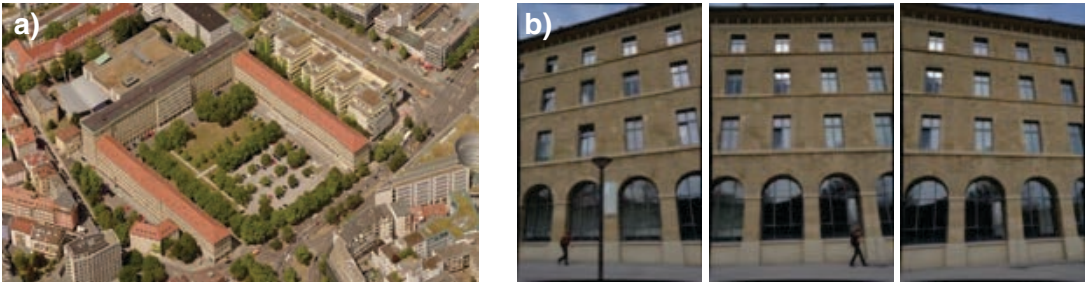


Abb. 6: (a) Rotenbühlbau Stuttgart (b) Auf jeder Station wurden 3 Bilder erfasst

ten, dass das Objekt mit hoher Redundanz photogrammetrisch erfasst wird. In einem zweiten Schritt sind die Linsenverzeichnungen sowie die Position und Orientierung der Bilder zu berechnen – heutzutage durch Structure-from-Motion. Die resultierenden relativen äußeren Positionen und Rotationen der Einzelbilder werden in einem gemeinsamen Euklidischen Koordinatensystem definiert. Die (minimale) innere Orientierung repräsentiert die kameraspezifischen Parameter wie die Brennweite (Kammerkonstante) und die Lage des Bildhauptpunkts. Mittels dieser Informationen können bereits dünne Punktwolken mit 3D-Fassadenpunkte aus den 2D-Bildern extrahiert werden. Der letzte Schritt berechnet pro Bildpixel einen Fassadenpunkt und ergibt sehr dichte Punktwolken für die Fassade(n). Jedoch ist diese hohe Punktdichte nicht immer zielführend und sollte daher für eine Interpretation mittels Fassadengrammatik entsprechend ausgedünnt werden.

3.1 Photogrammetrische Datenerfassung

Zur bildbasierten Fassadenrekonstruktion wurden mit einer Nikon D7000, Brennweite 20mm, insgesamt 96 Photos aufgenommen, von 33 Stationen aus. Um eine hinreichende Bildüberdeckung zu gewährleisten, sind von jedem Standpunkt aus jeweils 3 Bilder mit leicht variierenden Blickrichtungen aufgenommen worden. Jedes Bild hat eine Auflösung von 16,2 MPixel. Die Bilddatenerfassung eines Fassadenausschnitts von ~50 m Breite, ~20 m Höhe hat nicht mehr als 15 Minuten in Anspruch genommen – ein Stativ wurde dabei nicht eingesetzt. Abbildung 6 gibt einerseits einen Luftbildausschnitt des erfassten Gebäudes wieder und zeigt andererseits drei Bilder von einem Standpunkt. Falls möglich, wurden die Bilder etwa im tangentialen Abstand von ~1,5 m und mit einer radialen Distanz von ~14,5 m aufgenommen. Diese Aufnahmeanordnung garantiert ähnlichen Bildinhalt von benachbarten Standpunkten und letztlich eine

ordentliche Berechnung der Punktwolken aus den Bildern. Unter der Annahme einer in etwa parallelen Ausrichtung quer zur Fassade erhält man eine Bildüberdeckung von bis zu 87 % bei einer Ground Sampling Distance (GSD) von weniger als 3,5mm.

3.2 Photogrammetrische Bildorientierung

Die Abbildungsverzeichnungen und Orientierungen werden mittels dem frei verfügbaren Softwarepaket Visual Structure-from-Motion (VisualSfM) [19] berechnet. Bei dieser Software ist eine inkrementelle Bündelausgleichung implementiert, die jeweilige Bildverbände analysiert und deren Überdeckungen feststellt. Dabei werden homologe Bildpunkte für den gesamten Bildverband identifiziert und in der lokalen wie auch globalen Bündelausgleichung in den Objektraum vorwärts eingeschnitten. Innerhalb dieser Auswertungen werden radiale Verzeichnungsfehler korrigiert und die Parameter der äußeren Orientierung berechnet. Als Nebenprodukt ergibt sich eine dünn besetzte Punktwolke im Objektraum – die vorwärtseingeschnittenen homologen Punkte des Bildraums – siehe Abbildung 7. Diese Punktwolke ist noch nicht skaliert, was auch nicht unbedingt für eine grammatikalische Anwendung notwendig ist.

3.3 Dichte Bildzuordnung und Multi-View-Triangulation

Mit den zuvor erhaltenen Bildorientierungen können dichte 3D-Punktwolken berechnet werden. Hierzu wird die institutseigene Software SURE eingesetzt, die mittlerweile von nFrames für professionelle Anwendungen vertrieben wird. Diese basiert auf *libtsgm*, einer Bibliothek, die eine Speicher- und zeiteffiziente Implementierung des SGM-Algorithmus bereitstellt.

SGM ist eine dichte Bildzuordnungsstrategie, die zugehörige Pixel (welche ein und dasselbe Objekt in zwei verschiedenen Bildern darstellen)

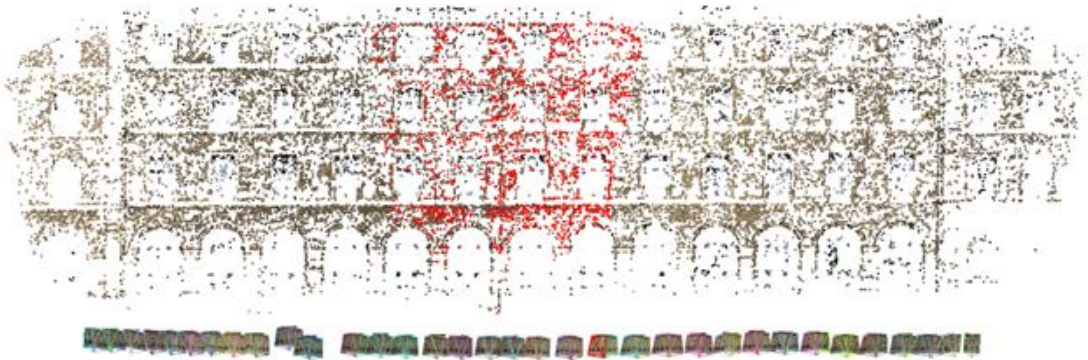


Abb. 7: Dünn besetzte Punktwolke als Ergebnis von VisualSfM

auffindet. Idealerweise wird für jedes Pixel eines Referenzbildes ein korrespondierendes Pixel in einem zweiten Bild – dem Stereopartner – gesucht. Die Basisidee sucht dabei auf intelligente Weise eine Approximation einer globalen Kostenfunktion, die die Güte der Zuordnung für alle Pixel in der Nachbarschaft berechnet und das Minimum als optimales Ergebnis darstellt. Die globale Kostenfunktion ist zusammengesetzt aus dem pixelweisen Vergleich oder intensitätsbasierten Ähnlichkeitsmassen wie z.B. die Census-Korrelation, die Mutual Information oder den Daisy-Deskriptor [9]. Pixelweise abgeleitete Kostenfunktionen präsentieren die Ähnlichkeit von zwei korrespondierenden Pixel und ihrer Nachbarschaft. Darüber hinaus wird ein Strafterm eingeführt, um Mehrdeutigkeiten zu minimieren wie auch eine gewisse Glattheit der zu rekonstruierenden Oberflächen zu garantieren. Die überragenden Eigenschaften dieses Algorithmus bestehen in dem Erhalt von Tiefeninformation speziell bei Kanten und moderaten Berechnungszeiten. Bei der klassischen Anwendung der globalen Kostenfunktion werden die pixelweisen Kosten für jedes Pixel im Referenzbild und einer konstanten Anzahl von möglichen Partnern im zuzuordnenden Bild berechnet und minimiert. SURE/libtsgm implementiert abweichend davon eine hierarchische Vorgehensweise. Basierend auf der Oberflächenberechnung von grob aufgelösten Gittern/Bildern, deren Berechnung sehr schnell ist, kann die Suche nach potenziellen Matchingpartnern

entsprechend eingegrenzt werden, infolge der vorher durchlaufenen Berechnung innerhalb einer Bildpyramide. Dieses impliziert eine nochmals reduzierte Berechnungszeit und auch Speicheranforderung, und erlaubt auf diese Weise die Berechnung auch von großen Disparitäten. Für die Fassadenanwendung in diesem Beitrag wurden insgesamt 405 Stereomodelle berechnet. Jedes Bild wurde im Durchschnitt 4 potenziellen Partnern zugeordnet. Die Prozessierungszeit betrug 2.75h (i7, 4 x 3.4 GHz), Input und Output mit eingerechnet.

Da sich die Bilder hochgradig überlappen, wird ein Teil der Fassade mehrfach in den einbezogenen Bildern abgebildet. Innerhalb der dichten Bildzuordnung werden jetzt die Bilder paarweise ausgewertet und die 3D-Punktwolken extrahiert. Als Ergebnis ergeben sich mehrfache Punktwolken für ein und denselben Fassadenausschnitt. Diese Redundanz wird nun in der Multiview-Implementierung ausgenutzt, was zu einer Erhöhung der Genauigkeit im Objektraum beiträgt. Dabei ist bedeutend, dass fehlerhafte Tiefeninformationen zuverlässig aufgedeckt und eliminiert werden können. Letztlich ergeben sich hochqualitative Punktwolken, die nur noch wenige Ausreißer enthalten. Für die hier prozessierten 405 Stereomodelle konnte letztlich eine Punktwolke für alle 95 Bilder berechnet werden. Die Rechenzeit betrug 1.33h (i7, 4 x 3.4 GHz), Input und Output eingeschlossen.

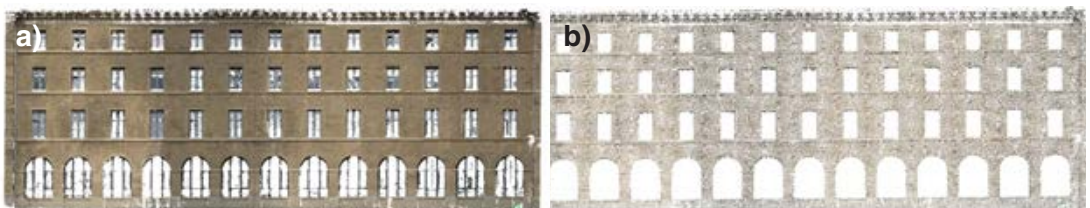


Abb. 8: SURE Punktwolken (a) unverdünn't (b) gefiltert auf GSD 5 cm

4. Von der Punktwolke zum 3D-Modell

Ganz generell können formale Grammatiken zur 3D-Objektrekonstruktion aus Punktwolken eingesetzt werden, speziell um die Plausibilität und topologische Korrektheit der rekonstruierten Elemente sicherzustellen. Eine ausgezeichnete Grammatik für diese Anwendungen stellen die Lindenmayer Systeme (L-Systeme) dar [12]. Auch wenn diese originär für die Beschreibung des Wachstums von Pflanzen eingesetzt worden sind, stellen sie eine gute Ausgangsbasis für die Weiterentwicklung von Grammatiken zur Beschreibung von architektonischen Elementen zur Verfügung. Zum Beispiel leiten [6,7] eine detaillierte Gebäudehülle ab, ohne jedwede Sensordaten, sondern lediglich durch eine Figuringrammatik.

4.1 Formale Grammatiken zur 3D-Rekonstruktion

In [12] wird eine formale Grammatik zur Generierung von Fassadenstrukturen eingesetzt, wobei nur teilweise Sensordaten (Punktwolken) zur Verfügung stehen. Prinzipiell bestehen formale Grammatiken aus einem Vokabular und einer Menge von Produktionsregeln. Das Vokabular stellt verschiedene Symbole (Typen) zur Verfügung, die auch als *non-terminals* bezeichnet werden – diese können durch andere Symbole ersetzt werden, oder durch komplementäre Elemente wie die *terminals*. Das non-terminal Symbol, welches den Ausgang für alle weiteren Beschreibungen darstellt, wird auch als *Axiom* bezeichnet. Die Eigenschaften einer Grammatik hängen von der Definition ihrer Produktionsregeln ab. Diese könne z.B. deterministisch oder stochastisch sein, jedoch auch parametrisch und kontext-sensitiv. Eine allgemeine Formulierung zum Einsatz einer Fassadengrammatik ist gegeben durch:

$$id: lc < pred > rc : cond \rightarrow succ : prob$$

Die Produktionsregel, gekennzeichnet durch label *id*, gibt den Ersatz des Vorgängers *pred* durch den Nachfolger *succ* an. Da ein Vorgänger jeweils linken und rechten Kontext einbezieht, hier ausgedrückt durch *lc* und *rc*, ist die Regel kontext-sensitiv. Falls die Bedingung *cond* sich als wahr (true) herausstellt, wird eine Substitution durchgeführt mit der Wahrscheinlichkeit *prob*. Auf diesen Definitionen aufbauend kann eine Fassadengrammatik abgeleitet werden, die uns erlaubt, neue synthetische Fassaden zu generieren, unabhängig von Figur und Größe. Das Axiom bezieht sich dann auf die neue, zu model-

lierende Fassade und beinhaltet Informationen zum Fassadenpolygon. Die Menge der *terminals* und *non-terminals* sowie die Produktionsregeln werden dabei automatisch durch die Analyse von Punktwolken gelernt. Zu Beginn der Anwendung von Grammatiken wurden unsere Algorithmen auf Laserscanning-Punktwolken angewandt, doch wie nachstehend aufgezeigt wird, funktioniert der Ansatz auch bei photogrammetrisch erfassten Punktwolken.

Bestehende Systeme zur grammatikbasierten Rekonstruktion von 3D-Gebäudemodellen, welche zur Interpretation von Bildern oder Modellen eingesetzt werden, sind i.d.R. als halbautomatisch zu klassifizieren. Unser Ansatz ist von Beginn an auf die vollautomatische Gebäude-rekonstruktion für verschieden Architekturstile ausgelegt.

4.2 Post-Processing von Punktwolken

Der im Folgenden beschriebene Algorithmus wurde erstmals für die Analyse von Laserscanning-Punktwolken eingesetzt. Das Ergebnis sind Strukturinformationen der Gebäudefassaden.

Er basiert auf der Annahme, dass Fensterflächen als Löcher in den LiDAR-Punktwolken auftreten. Diese Annahme ist dadurch gerechtfertigt, dass typischerweise LiDAR-Impulse Glasflächen durchdringen und daher kein richtiges Echo zur Verfügung steht. Im Vergleich zur Photogrammetrie ergeben sich hier jedoch viele Punkte auch im Fensterbereich. Jedoch sind diese i.d.R. beträchtlich hinter Wandflächen gelegen und können daher leicht identifiziert und entfernt werden. Für komplexere Fassadengeometrie sind u.U. alternative Verfahren anzuwenden. Neben Filtern basierend auf der Bildradiometrie, können auch solche Anwendung finden, die das lokale Rauschen eliminieren sollen. Die Qualität der dichten Bildzuordnung hängt überwiegend von der Bildgüte und Bildtextur ab. Normalerweise weisen Fensterflächen eine geringe Textur auf und führen daher auf ein hohes Rauschniveau der Punktwolke. Eliminiert man die Punkte von Fensterflächen, wird die Annahme von Löchern als Fenster erfüllt.

Wie bereits in 3.2 angedeutet, liefert die dichte Bildzuordnung Punktwolken von beliebigem Maßstab. Da der in der Rekonstruktion eingesetzte Algorithmus metrische Informationen einbezieht, kann der korrekte Maßstab jederzeit nachträglich abgeleitet werden. Daher wurden die Dimensionen der Fensteröffnungen am Testgebäude vor Ort ermittelt. Da große Teile von Fensterflächen zur Verfügung stehen, konnten

dessen Dimensionen auch in der Punktwolke ermittelt werden. Der resultierende Maßstabsfaktor wurde daher auf alle Koordinaten angewendet.

Ein weiterer Vorverarbeitungsschritt betrifft die Reduktion der Anzahl von Punkten. Der Vergleich zwischen LiDAR- und Photogrammetrie-Punktwolke kommt vielfach zu dem Schluss, dass letztere vielfach dichtere Punktwolken liefert. Für die Grammatikentwicklung wird jedoch diese hohe Punktdichte nicht benötigt, da diese unnötige Rechenzeit und Speicherplatz in Anspruch nehmen würde. Daher wurden die Ergebnisse nochmals umgerechnet (resampled), auf eine minimale Punktdistanz von 5 cm.

4.3 Grammatikbasierte Fassadenrekonstruktion

Der Algorithmus startet mit einer datengetriebenen Extraktion von Fassadenstrukturen aus 3D-Punktwolken (Abschnitt A). Mittels der photogrammetrisch berechneten Punktwolken werden die resultierenden Fassadenelemente als Wissensbasis zur weiteren Verarbeitung eingesetzt (Abschnitt B). Dominante und repetitive Merkmale als auch ihre hierarchische Beziehungen zueinander werden aufgedeckt. Zur gleichen Zeit können Produktionsregeln gefunden werden, die, zusammen mit den gefundenen 3D-Elementen (Sprach)Umfang einer Formalen Grammatik bilden. Aus diesem Grund nennen wir diese Grammatik auch Fassadengrammatik. Diese Elemente enthalten alle notwendigen Informationen, um den Fassadenstil bzw. alle Strukturelemente eines Gebäudes zu modellieren und zu präzisieren (Abschnitt C).

A. Punktwolkeninterpretation

Die Vorgehensweise einer datengetriebenen Fassadenrekonstruktion zielt ab auf die Verfeinerung eines bestehenden, groben Gebäudemodells (gegeben als LoD2). Fenster, Türen und

Vorsprünge/Erker werden aus der Punktwolke extrahiert, unter der Annahme dass diese sich nicht auf den planaren Fassadenabschnitten befinden. Dieser Modellierungsprozess verwendet eine 3D-Objektrepräsentation als Zellenzerlegung. Die grundlegende Idee ist dabei die Zerlegung einer existierenden Gebäudehülle mit planarer Fassadenfläche in 3D-Zellen. Jede 3D-Zelle repräsentiert entweder eine Fassadenkachel oder ein Fensterelement. Nach einer Klassifikation der 3D-Zellen in Fenster- und Wandkacheln werden die Fensterkacheln eliminiert und die verbleibenden Wandkacheln zusammengefasst. Eine Schwierigkeit dabei ist dabei das Auffinden von Schwellenwerten, um gute Arbeitshypothesen zu gewährleisten. Da unser Fokus auch auf der Rekonstruktion von Fenstern liegt, müssen diese Schwellenwerte aus den Punktwolken in der Nähe der Fenster gesucht werden. Für den beispielhaften Datensatz "Rotebühlbau, Stuttgart", ist dies in der Abbildung 9a wiedergegeben. Aus Übersichtsgründen wird an dieser Stelle bereits das Endergebnis (Abbildung 9b) dargestellt.

Punktwolken-Segmentierung

Gewöhnlicherweise sind die Glasflächen von Fenstern von der Fassadenebene zurückgesetzt bzw. eingelassen. Von dieser Hypothese ausgehend können die Punktwolken hinsichtlich der Abstände von den Fassadenebenen klassifiziert werden. Falls nur die Punkte in Betracht gezogen werden, die innerhalb oder vor einer Fassade zu liegen kommen, werden Fenster als Löcher dargestellt. Diese Nichtdaten-Flächen sind dann in der Klassifikation zum Auffinden von Fensterecken einzusetzen. Zum Beispiel, falls Punkte einer linken Fensterecke detektiert werden sollen, wird vorausgesetzt, dass keine Nachbarn der rechten Fensterecke in einem vorgegebenen Suchradius vorhanden sind. In einem weiteren Schritt werden horizontale und vertikale Linien geschätzt, von nichtisolierten Punkten auf Ecken.

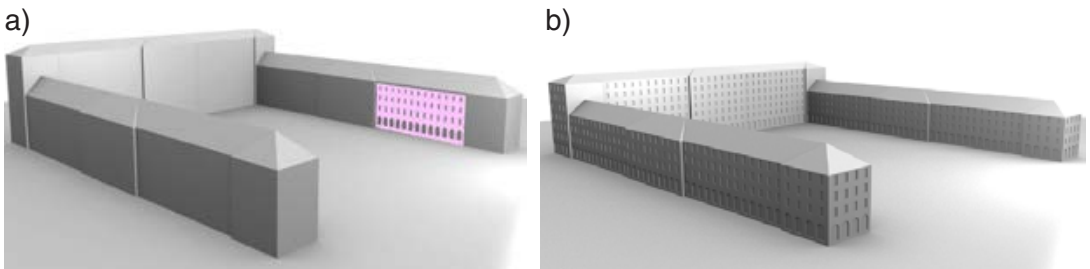


Abb. 9: Rotebühlbau Stuttgart: (a) Gebäudehülle und 3D Punktwolken durch dichte Bildzuordnung (b) Endergebnis der grammatikbasierten Rekonstruktion

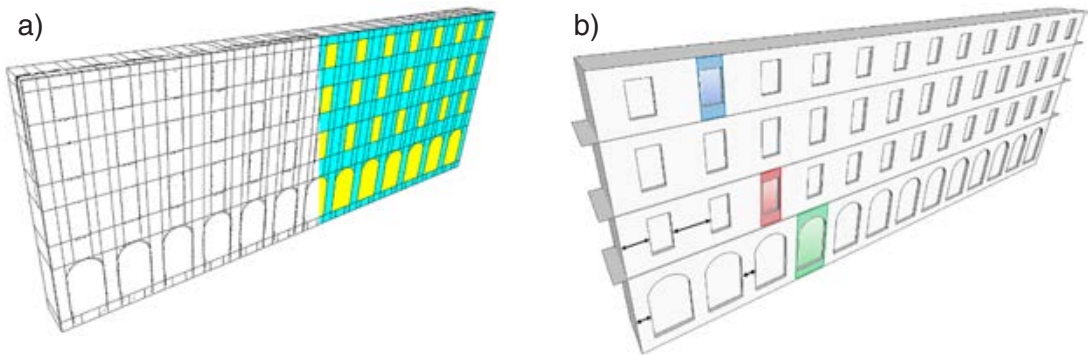


Abb. 10: Datengetriebene Fassadenrekonstruktion: (a) Erkannte Liniensegmente und Fenster (b) Fassadenmodell

Abbildung 10a zeigt die extrahierten Punkte von Fensterecken als auch die abgeleiteten horizontalen und vertikalen Linien. Aufbauend auf diesen Fensterlinien, planare Delimeter/Schwellwerte können dann für eine nachfolgende räumliche Zerlegung generiert werden. Jede Begrenzungslinie definiert eine Unterteilungsebene, die rechtwinklig zur Fassade steht. Um die Fenstertiefe zu berechnen, sind zusätzliche Unterteilungsebenen zu nutzen, die aus den Punkten der Punktwolke approximiert werden, die sich gerade auf Fensterflächen befinden. Diese Punkte werden dadurch gefunden, indem in parallelen Flächen zur Fassadengeometrie gesucht wird. Die Menge der Partitionen liefert die Strukturinformation für den Zellenzerlegungsprozess.

Klassifikation

Um die 3D-Zellen entsprechend als Fassaden- oder Fensterzellen zu klassifizieren, ist eine Punkteverfügbarkeitskarte einzusetzen. Das ist ein Binärbild mit grober Auflösung, wobei jedes Pixel ein Gitterelement der Fassade wiedergibt. Die optimale Gittergröße sollte etwas größer sein als die GSD. Gitterelemente, für die 3D-Punkte zur Verfügung stehen, produzieren Fassadenpixel; die Gitterelemente, die zu Nichtdaten-Flächen gehören, präsentieren Nichtfassadenpixel. Die Klassifikation ist durchzuführen, indem das Verhältnis zwischen Fassadenpixel und Nichtfassadenpixel für jede 3D-Zelle berechnet wird. Zellen, die mehr als 70% Fassadenpixel enthalten, werden als Fassadensolitäre bezeichnet, wobei Zellen mit weniger als 10% Fassadenpixel Fenstersolitäre darstellen. Während zumeist die 3D-Zellen zuverlässig klassifiziert werden können, gibt es lediglich ungewisse Ergebnisse in der Nachbarschaft von Fenstern oder in Gebieten mit unzulänglicher Datenerfassung.

Jedoch sorgt die Integration von Nachbarschaftsrelationen und Bedingungen dafür, dass auch solche Zellen klassifiziert werden können. Die rechte Fassade in Abbildung 10a zeigt die klassifizierten 3D-Zellen: Fassadenzellen (Cyan) und Fensterzellen (gelb). Während die Fenster – als Ergebnis der Ebenenzerlegung – natürlicherweise als polyhedrale Zellen dargestellt werden, sind ebenso gekrümmte Primitive in den Rekonstruktionsprozess einzubeziehen. Dies ist durch die gebogenen Fenster im Erdgeschoss demonstriert.

Modellierung

Innerhalb eines Modellierungsschritts werden die Fensterkacheln aus der existierenden groben Gebäudehülle herausgeschnitten bzw. tiefergelegt. Daher erscheinen diese im Fassadenmodell als „Einlässe“ (Abbildung 10). Jedoch ist der Modellierungsprozess nicht nur auf das Unterscheiden bzw. Auffinden von Fensterflächen beschränkt – es können viel komplexere Elemente beschrieben werden.

B. Automatische Inferenz der Fassadengrammatik

Aufbauend auf dem datengetriebenen Ergebnis der Rekonstruktion, kann die Fassadengrammatik automatisch abgeleitet werden, indem nach terminals, ihren interrelationships (Beziehungen), und Produktionsregeln gesucht wird.

Suche nach Terminalen

Um eine aussagekräftige Menge von terminals für die Fassadengrammatik abzuleiten, wird die Gebäudefassade in eine Menge von Basiselementen unterteilt, die nicht weiter substituiert werden können. Für diesen Zweck wird eine räumliche Zerlegung angewandt, die die Fassade unterteilt in Stockwerke und jedes Stockwerk in Kacheln.

mente komplettiert, die bisher nur unvollständig vorliegen. Drittens, gänzlich abgeschattete Fassaden oder Fassadenteile – d.h. nicht beobachtete Fassaden – können durch das gelernte Wissen in einem Produktionsprozess komplett rekonstruiert werden.

Dieser Produktionsprozess startet iterativ mit einer beliebigen Fassade, auch Axiom genannt, und läuft folgendermaßen ab: (1) Wähle ein non-terminal Element im gegenwärtigen String, (2) suche eine Produktionsregel mit diesem non-terminal als Vorgänger, (3) ersetze das non-terminal

Element mit den Regeln seines Nachfolgers, (4) beende den Produktionsprozess, falls alle non-terminal Elemente ersetzt worden sind, andernfalls gehe zurück zu (1).

Während der Produktion werden die non-terminal Elemente sukzessiv mit den zugehörigen Produktionsregeln überschrieben. Falls mehr als eine Produktionsregel möglich ist, wird diejenige mit der höchsten Wahrscheinlichkeit ausgewählt. Sobald der Fassadenstring nur noch Terminale enthält, wird die Produktion beendet und der String in eine 3D-Darstellung überführt. Das

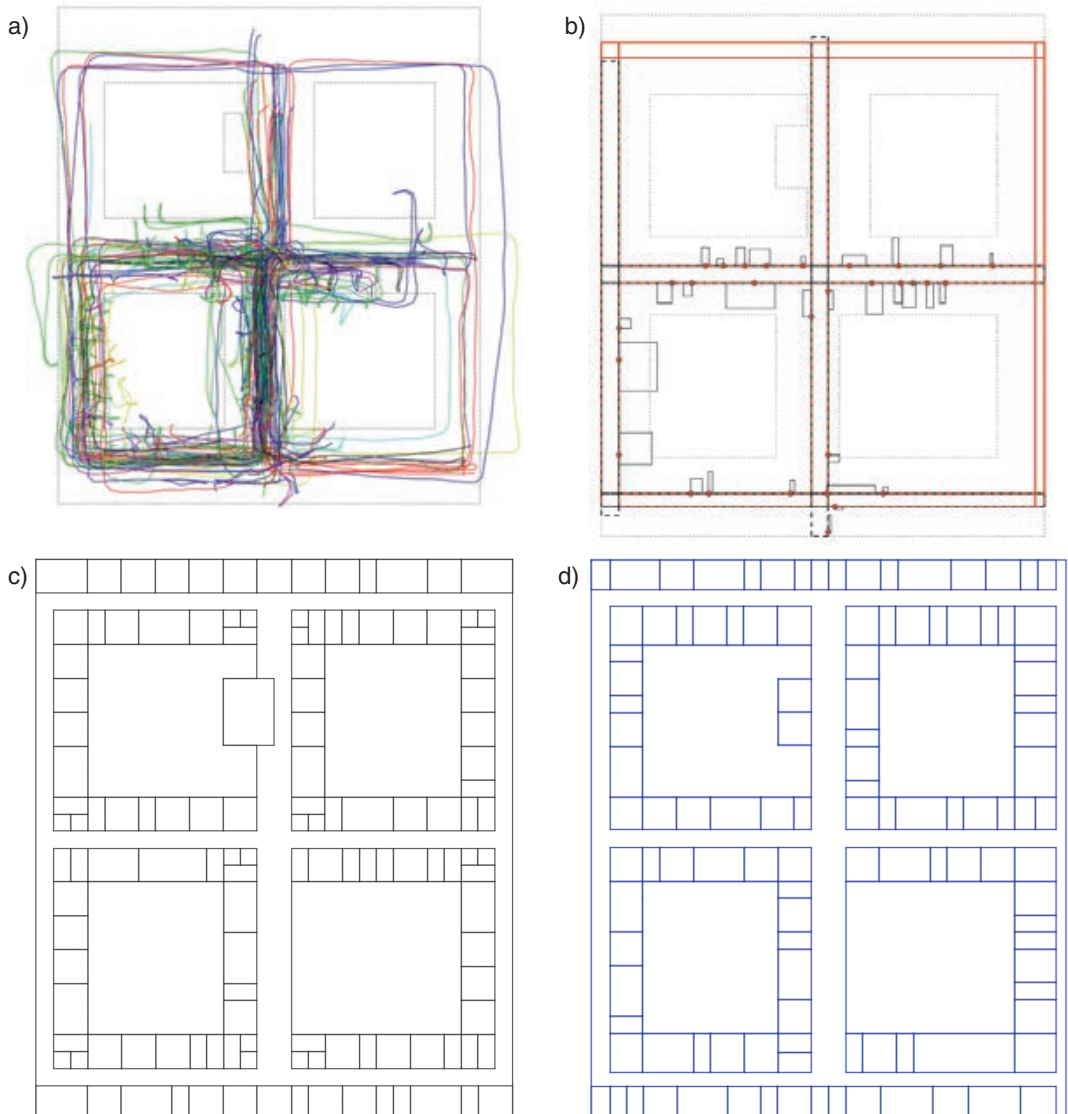


Abb. 11: Grammatik-Anwendung im Innenraum: (a) Bewegungsspuren (b) Ableitung genäherter Flurelemente (c) Splitgrammatik und datengetriebene Flure (d) Splitgrammatik und komplettierte Flure

Endergebnis ist bereits mit Abbildung 9b wiedergegeben.

5. Zusammenfassung und Ausblick

Der Beitrag hat mit generellen Betrachtungen zur 3D-Rekonstruktion von Stadt- und Gebäude-Modellen begonnen, für die mittels LiDAR und Photogrammetrie dichte Punktwolken erfasst werden. Eine echte Revolution in der Photogrammetrie ist durch die Einführung des Semi-Global Matching eingeleitet worden, welches hochqualitative und sehr dichte Punktwolken zur Verfügung stellt. Die Verarbeitung von Punktwolken mittels formeller Grammatiken ist seit einigen Jahren Gegenstand der Forschung im Bereich Photogrammetrie und Computer Vision. Deren Ergebnisse sind sehr vielversprechend, so dass für diesen Beitrag gerade die Verbindung von photogrammetrisch erfassten Punktwolken und die grammatikbasierte 3D-Rekonstruktion in den Vordergrund gestellt werden. Die Bilder sind mit einer klassischen DSLR-Kamera aufgenommen worden, die nicht teuer ist und eine hervorragende Bildqualität liefert. Nach der Kalibration werden die Bilder einem Orientierungsprozess zugeführt – heutzutage sind dies die Algorithmen des sogenannten Structure-from-Motion (SfM). Hierzu wurde das OpenSource Paket VisualSfM eingesetzt. Die dichte Bildzuordnung liefert eine immense Punktwolke, etwa 145 Mio. Punkte für den ausgewählten Fassadenausschnitt eines Gebäudes (Rotebühlbau) in Stuttgart.

Die anschließende Verarbeitung mit dem Grammatik-Ansatz interpretiert die zur Verfügung gestellte Punktwolke und berechnet schließlich alle Fassadenelemente, auch an den Stellen, für die keine Daten vorliegen. Die Ergebnisse sind sehr vielversprechend – es scheint dass zur Komplettierung von LoD3-Modellen auch die Nahbereichsphotogrammetrie erfolgreich eingesetzt werden kann. Zudem liefern die neuen Schrägbild-Kamerasysteme der Luft-

bildphotogrammetrie neben dem klassischen Nadirbild mindesten 4 Schrägluftbilder, die zur Punktwolkengenerierung für Fassadenelemente eingesetzt werden können. Dies ist eine weitere Bestätigung dafür zu sagen, dass „Fassadenrekonstruktionen mittels Punktwolken und Grammatiken ein sehr erfolgversprechendes Duo darstellen, welches in einem daten- und wissensgetriebenen Prozess die gewünschten Informationen bereitstellen kann.“

Eine Erweiterung der Fassadengrammatik in den Innenraum ist in [21] untersucht und erfolgreich angewendet worden. Erste Ergebnisse sind in Abbildung 11 wiedergegeben. Dabei wurden Bewegungspuren im Informatikgebäude der Universität Stuttgart, Pfaffenwaldring 39 analysiert. Als erstes konnten genäherte Flursegmente als Axiome definiert werden, indem ein grammatikalisches L-System generiert wurde. Die Anwendung einer Splitgrammatik auf Nichtflurelemente liefert folgendes Ergebnis: Rein datengetrieben konnten 29 Zimmer rekonstruiert werden, eine Anwendung der Splitgrammatik auf datengetriebene Flure liefert 92 Zimmer (Abbildung 11c). Wendet man die Splitgrammatik auf komplettierte Flure an, erhält man immerhin schon 116 Zimmer (Abbildung 11d).

Mit den neuen Entwicklungen im Bereich Augmented und Virtual Reality ergeben sich ganz neue Möglichkeiten zur Visualisierung von 3D-Modellen (LoD2, LoD3 und LoD4). Einen kleinen Vorgeschmack gibt Abbildung 12, indem zu einem Stadtplan (Stadt Calw, Deutschland) die Nikolausbrücke (Abbildung 12a) – rekonstruiert aus Photos und Punktwolken – auf einem beliebigen Smartphone oder Tablet in 3D dargestellt werden kann (Abbildung 12b).

Referenzen

- [1] Haala, N. (1996): Gebäuderekonstruktion durch Kombination von Bild- und Höhendaten. Deutsche Geodätische Kommission, Reihe C, Nr. 460, München, 96S.

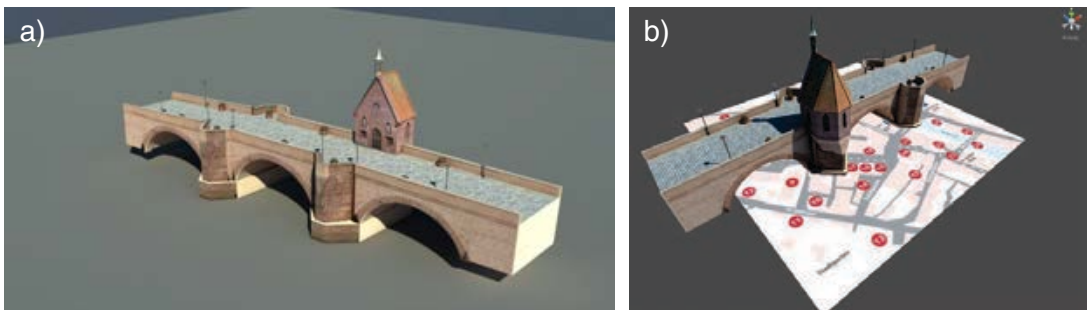


Abb. 12: (a) 3D-Rekonstruktion Nikolausbrücke Calw (b) Zuordnung 3D Modell-Stadtplan in Unity Game Engine

- [2] *Ameri, Shrabai, B. (2000): Automatic Recognition and 3D Reconstruction of Buildings from Computer Vision and Digital Photogrammetry. Deutsche Geodätische Kommission, Reihe C, No. 526, München, 110p.*
- [3] *Brenner, C. (2000): Dreidimensionale Gebäuderekonstruktion aus digitalen Oberflächenmodellen und Grundrissen. Deutsche Geodätische Kommission, Reihe C, Nr. 530, München, 124S.*
- [4] *Bekins, D., Aliagia, D. (2005): Build-by-Number: Rearranging the Real World to visualize Novel Architectural Spaces. IEEE Visualization, pp. 143-150.*
- [5] *Hirschmueller, H. (2005): Accurate and Efficient Stereo Processing by Semi-Global Matching and Mutual Information. IEEE Conference on Computer Vision and Pattern Recognition, June 2005, San Diego, CA, USA, Vol. 2, pp. 807-814.*
- [6] *Müller, P., Zeng, G., Wonka, P. und Van Gool, L. (2006): Procedural Modelling of Buildings. ACM Transactions on Graphics (TOG), 25(3), pp 331-343.*
- [7] *Van Gool, L., Zeng, G., van den Borre, F., Müller, P. (2007): Towards Mass-produced Building Models. Conf. Computer Vision and Pattern Recognition, pp. 1-8.*
- [8] *Kada, M. (2007): Zur massstabsabhängigen Erzeugung von 3D-Stadtmodellen. Dissertation Universität Stuttgart. 111S.*
- [9] *Hirschmüller, H. (2008): Stereo Processing by Semi-global Matching and Mutual Information. IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 328-341.*
- [10] *Eastman, C., Teicholz, P., Sacks, R., Liston, K. (2008): BIM Handbook: A Guide to Building Information Modeling for Owners, Managers, Designers, Engineers and Contractors, 2008, John Wiley & Sons, Inc., New Jersey.*
- [11] *Furukawa, Y., Curless, B., Seitz, S., Szeliski, R.: Manhattan-World Stereo, IEEE Conference on Computer Vision and Pattern Recognition, 2009*
- [12] *Becker, S. (2011): Automatische Ableitung und Anwendung von Regeln für die Rekonstruktion von Fassaden aus heterogenen Sensordaten. Deutsche Geodätische Kommission, Reihe C, Nr. 658, München, 156S.*
- [13] *Fritsch, D. (2011): Photogrammetric Week '11. Wichmann, VDE Verlag,*
- [14] *Carswell, A. (2011): Lidar Imagery – From Simple Snapshots to Mobile 3D Panoramas. In: Photogrammetric Week '11, Ed. D. Fritsch, Wichmann, VDE Verlag, Berlin und Offenbach, pp. 3-14.*
- [15] *Haala, N., Fritsch, D., Peter, M., Khosravani, A.: Pedestrian navigation and modeling for indoor environments 7th International Symposium on Mobile Mapping Technology, Crakow, Poland, 2011*
- [16] *Budroni, A. (2013): Automatic Model Reconstruction of Indoor Manhattan World Scenes from Dense Laser Range Data. Deutsche Geodätische Kommission, Reihe C, Nr. 715, 104p.*
- [17] *Rothermel, M., Wenzel, K., Fritsch, D., Haala, N. (2012): SURE: Photogrammetric Surface Reconstruction from Imagery. Proceedings LC3D Workshop Berlin.*
- [18] *Fritsch, D., Becker, S., Rothermel, M. (2013): Modeling Facade Structures Using Point Clouds from Dense Image Matching. Proceedings Int. Conf. Advances in Civil, Structural and Mechanical Engineering. Inst. Research Engineers and Doctors, pp 57-64. ISBN: 978-981-07-7227-7.*
- [19] *Wenzel, K., Rothermel, M., Fritsch, D., Haala, N. (2013): Image Acquisition and Model Selection for Multi-View Stereo. Proceedings 3D Arch Conference, Trento, Italy.*
- [20] *Fritsch, D. (2014): Mobile Mapping – Eine Revolution im Vermessungswesen. Stiftung Bauwesen, Stuttgart, Vorträge 22.11.2013.*
- [21] *Philipp, D., Baier, P., Dibak, C. Dürr, F., Rothermel, K., Becker, S., Peter, M. Fritsch, D. (2014): MapGENIE: Grammar-enhanced Indoor Map Construction from Crowd-sourced Data. Proceedings International Conference on Pervasive Computing and Communications (PerCom 2014, Budapest, pp. 1-9.*
- [22] *Fritsch, D., Klein, M. (2015): Augmented Reality 3D Reconstruction of Buildings – Reconstructing the Past. Int. Journal Multimedia Tools and Applications (MTAP), under review.*

Anschrift des Autors

Prof. Dr.-Ing. Dieter Fritsch, Institut für Photogrammetrie, Universität Stuttgart, Geschwister-Scholl-Strasse 24D, 70174 Stuttgart, Deutschland.
E-Mail: dieter.fritsch@ifp.uni-stuttgart.de