

Von Photogrammetrie und Fernerkundung zu Visual Computing

Zusammenfassung

Dieser Beitrag stellt Entwicklungen im Laufe der Jahre dar, die der Autor zusammen mit dem Kollegen KURT BRUNNER an der Universität der Bundeswehr (UniBw) München zum großen Teil am selben Institut verbracht hat. Am Anfang stand Photogrammetrie und Fernerkundung mit automatischer Straßenextraktion aus Satellitenbildern und Generalisierung von 3D-Gebäudemodellen. Auf Anregung von Förstner [2000] begann der Autor, sich mit projektiver Geometrie zu beschäftigen, was zu neuen, vollautomatischen Verfahren zur Bildorientierung und Arbeiten zur Fassadeninterpretation und Baumextraktion führte. Mit der Schließung des Studienganges Geodäsie und Geoinformation und dem Wechsel in die Fakultät für Informatik wurde eine neue Entwicklung in Richtung Visual Computing eingeleitet.

Summary

This paper presents developments over the years, the author has spent together with colleague KURT BRUNNER at Bundeswehr University (UniBw) Munich, most of the time at the same institute. At the begin there was photogrammetry and remote sensing with automatic road extraction from satellite images and generalization of 3D building models. Motivated by Förstner [2000], the author began to study projective geometry, leading to new, fully automatic procedures for image orientation and work on facade interpretation and tree extraction. With closing the studies of Geodesy and Geoinformation and the transition into Computer Science, a new development towards Visual Computing has been started.

1. Einleitung

Ein bedeutendes Ereignis in all den großen Veränderungen im Bereich der topographischen Datenerfassung und Darstellung der letzten Jahre stellt die Veröffentlichung von *Google Maps* im Jahr 2005 dar. Seitdem hat jeder Internetnutzer kostenlosen Zugang zu Geoinformation und zu riesigen Mengen an georeferenzierten, oft hoch aufgelösten Bildern. Andere Firmen, vor allem

Microsoft mit *Virtual Earth* und später *bing Maps* stellen konkurrierende Dienste zur Verfügung.

Die Bereitstellung großer Mengen georeferenzierter Bilder sowie die Ableitung von 3D Information ist eine klassische Domäne von Photogrammetrie und Fernerkundung. So wie *Google* und *Microsoft* keine geodätischen Firmen sind, erfolgt auch die Erfassung und Prozessierung der Daten für *Google Maps* und *bing Maps* in starkem Maße durch Gruppen außerhalb von Photogrammetrie und Fernerkundung. Die große Bedeutung dieser Arbeiten ist u. a. an der „*Special Issue*“ der *IEEE Computer Society* Flaggschiff Zeitschrift „*Computer*“ zu „*Capturing the World*“ im Jahr 2010 zu sehen. Dort wurden Arbeiten zu Community Photo Collections (CPC) [Goesele et al., 2010], d. h., die Erstellung von 3D Modellen aus im Internet frei verfügbaren Bildern, aber auch die Prozessierung von Luftbildern [Leberl et al., 2010] dargestellt, wobei bei Letzterem der Begriff „*Photogrammetry*“ nicht einmal erwähnt wurde.

Der Weg des Autors von der Geodäsie in die Informatik war aber dennoch mehr ein Unfall durch Sparbemühungen, als der Versuch, auf den Lauf der Zeit zu reagieren. Er führte im Endeffekt aber zu einem ähnlichen Ergebnis, nämlich Photogrammetrie mit vielen neuen Inhalten und daher im neuen Umfeld auch zu dem neuen Namen „*Visual Computing*“. Visual Computing ist hierbei ein Überbegriff, der Computer Vision und Graphik umfasst, zwei als für die Zukunft bedeutend angesehene Bereiche, die bisher in der Informatik der UniBw München brach lagen und deren Besetzung daher willkommen war.

Der Beitrag beschreibt die Entwicklung der Arbeiten des Autors von der Straßenextraktion und der 3D Generalisierung über die projektive Geometrie hin zur Fassadeninterpretation und Baumextraktion. Er wird abgeschlossen durch die Darstellung der neuen Ausbildungsinhalte in Visual Computing.

2. Strukturelle Objektextraktion und Maßstabsräume

Der Autor war maßgeblich an den Arbeiten zur Straßenextraktion an der Technischen Universität München (TUM) im Bündel „*Semantische Modellierung*“ der Deutschen Forschungsgemeinschaft (DFG) in den 1990er

Jahren beteiligt. Diese sind grundlegend der strukturellen Objektextraktion zuzuordnen. Objekte wurden manuell modelliert, wobei die Geometrie, z. B. in Form paralleler Kanten, im Vordergrund stand.

Die Arbeiten zur Straßenextraktion führten nach der Berufung des Autors an die UniBw 1999 zur Forschung in zwei Richtungen: Straßenextraktion aus hoch aufgelösten Satellitenbildern und 3D Generalisierung von Gebäuden basierend auf Maßstabsräumen. Letzteres war von der Erkenntnis aus der Straßenextraktion motiviert, dass Ereignisse in Maßstabsräumen für Bilder zu Abstraktion im Bereich der zugeordneten Objekte führen können.

Kerngedanken der Straßenextraktion in hoch aufgelösten Satellitenbildern waren die automatische Generierung von Trainingsgebieten für die multispektrale Klassifikation auf Grundlage Straßen-spezifischer Eigenschaften, wie vor allem die Parallelität der Straßenränder sowie die globale Netzbildung [Bacher/Mayer, 2005]. Im Rahmen des vom Geoinformationsdienst der Bundeswehr geförderten Projektes wurde auch eine Evaluierung verschiedener Ansätze im Rahmen eines Euro SDR (European Spatial Data Research) Projektes durchgeführt [Mayer et al., 2006], welche die Leistungsfähigkeit des entwickelten Ansatzes aufzeigte (s. Abb. 1).

Im Rahmen der Straßenextraktion waren, basierend auf der Maßstabsraumtheorie von Lindeberg [1994], Ideen zur Verknüpfung von Ereignissen im Maßstabsraum des Bildes mit der Abstraktion von Objekten entstanden. Diese wurden im Rahmen eines von der DFG geförderten Projektes auf die Generalisierung von Ge-

bäuden übertragen, insbesondere auch in 3D. Es zeigte sich, dass diese Übertragung grundlegend sinnvoll ist (s. Abb. 2). Die Implementierung von konsistenten Modifikationen von 3D Geometrie ist aber schwierig, und neben den geometrischen Problemen treten zusätzlich auch noch semantische Probleme auf, die für eine allgemeingültige Lösung beachtet werden müssen [Forberg/Mayer, 2006].

3. Direkte Orientierung auf Grundlage von projektiver Geometrie

Die Beschäftigung mit Orientierungsverfahren basierend auf projektiver Geometrie [Mayer, 2002], einem Schwerpunkt von Computer Vision in den 1990ern, begann motiviert durch Förstner [2000]. Es stellte sich heraus, dass sich auch schon der Vorgänger des Autors an der UniBw vor mehr als dreißig Jahren mit mathematischen Konstrukten beschäftigt hat [Dorrer, 1980], die heute als essentielle bzw. Fundamentalmatrix bezeichnet werden. Allerdings fehlten damals wichtige Erkenntnisse der folgenden Jahrzehnte für eine praktische Nutzung.

Auf Grundlage von Experimenten und insbesondere des Buches von Hartley / Zisserman [2000] wurde dem Autor in den ersten Jahren nur langsam klar, dass projektive Geometrie zwar zu direkten Lösungen führt, d. h. es sind keine Näherungen notwendig, diese aber nicht unbedingt euklidisch, d. h. einfach interpretierbar, sein müssen. Daher kam der 5-Punkt Algorithmus von Nistér [2004] als „Retter an kalibrierte, euklidische Ufer“

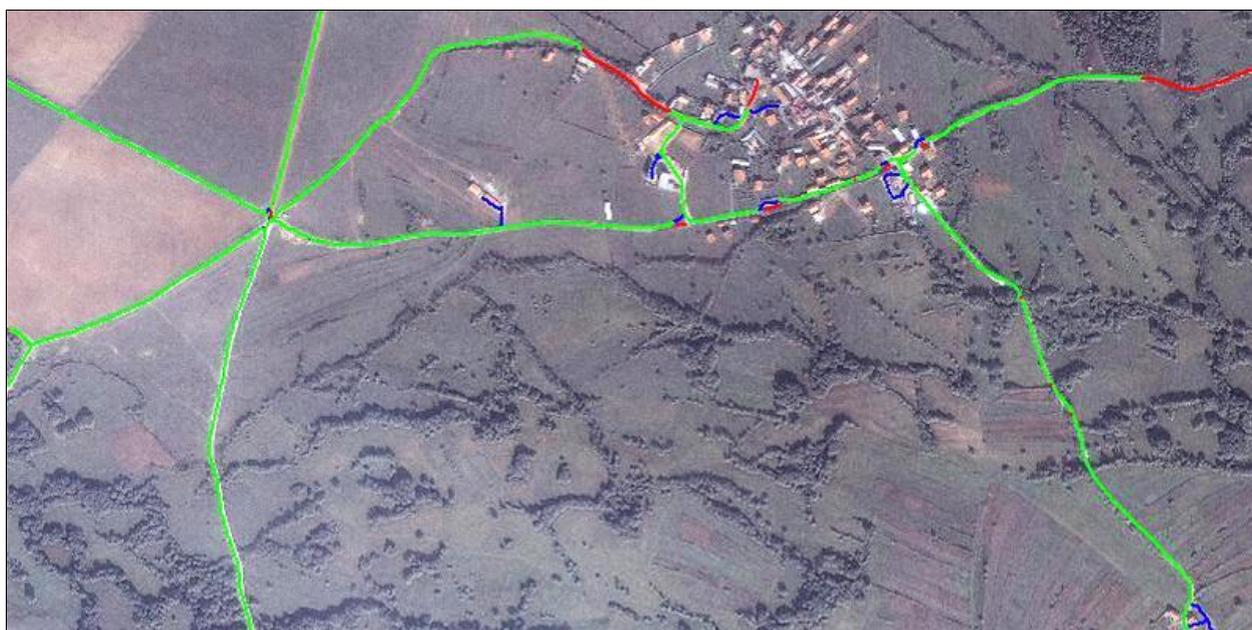


Abb. 1: Ergebnis automatische Straßenextraktion im hoch aufgelösten Satellitenbild und Vergleich mit Referenzdaten [Mayer et al., 2006] (Ausschnitt). Korrekt extrahierte Straßen sind in Grün, fehlende Extraktion in Rot, und zusätzliche extrahierte Straßen in Blau dargestellt. Die Extraktion ist zu 86% vollständig und zu 89% korrekt.

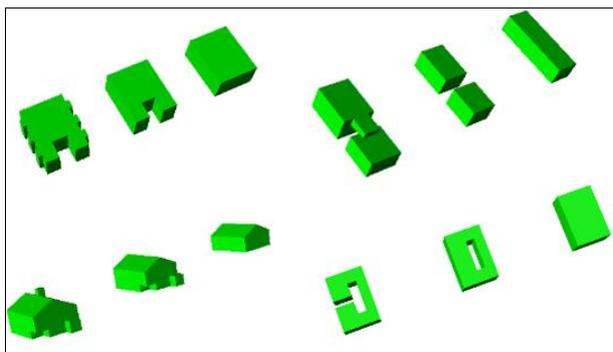


Abb. 2: 3D Generalisierung von Gebäuden mit Originaldaten jeweils links und Ergebnissen einer immer stärkeren Generalisierung (kleiner Maßstab) rechts. Im rechten Teil der Abbildung ist die Aufspaltung und Verschmelzung von Objekt(teil)en bzw. die Bildung und Schließung eines Loches dargestellt.

gerade rechtzeitig. Zusammen mit Random Sample Consensus – RANSAC [Fischler / Bolles, 1981] wird es möglich, bei ungefähr bekannter Kalibrierung, Bilder auch bei Fehlerquoten von 80% und mehr bei den Zuordnungen korrekt relativ zu orientieren.

Fortschritte in der Bildzuordnung besonders in Form des SIFT (Scale Invariant Feature Transform) Algorithmus von Lowe [2004] haben zusammen mit RANSAC und Fünf-Punkt Algorithmus sowie Entwicklungen im Bereich von Parallelrechnung vor allem auf Graphics Processing Units (GPUs) zu beeindruckenden Lösungen geführt. So wird in Frahm et al. [2010] gezeigt, wie Community Photo Collections (CPCs) [Goesele et al., 2010] mit mehreren Millionen Bildern, die in ihrer textlichen Beschreibung auf eine Stadt, wie z. B. Berlin oder Rom, hinweisen, innerhalb eines Tages auf einem einzigen Rechner in visuell ähnliche Cluster unterteilt, diese orientiert und 3D Modelle daraus abgeleitet werden können.

Die vom Autor – seit dem Jahre 2007 auch mit Unterstützung durch den Geoinformationsdienst der Bundeswehr – entwickelte Software erreicht zwar keine solche Größenordnung der Zahl der orientierten Bilder. Allerdings können Bilder mit einer größeren Basis als bei fast allen anderen Ansätzen effizient zugeordnet werden. Solche große Basen treten z. B. bei der Zuordnung von Bildern vom Boden und aus der Luft, z. B. von kleinen Drohnen/Unmanned Aerial Systems (UASs), auf.

Der in Bartelsen/Mayer [2010] vorgestellte Ansatz basiert auf SIFT Punkten, verwendet aber zur Zuordnung Korrelation und affine kleinste Quadrate. Da die Kalibrierung der Kamera als zumindest ungefähr bekannt vorausgesetzt wird, kann der 5-Punkt Algorithmus zusammen mit RANSAC zur Generierung und Verifizierung von Hypothesen für die euklidische relative Orientierung von Bildpaaren und Triplets verwendet werden. Im Gegensatz zu konkurrierenden Ansätzen

wird robuste kleinste Quadrate Bündelausgleichung angewandt, die sich einer Schätzung der Genauigkeit der Beobachtungen bedient. Die Bündelausgleichung wird nicht nur für größere Teile des Blockes, sondern auch bei der Selektion der Hypothesen für Paare und Triplets genutzt. Bei Letzterer wird ein dem Expectation-Maximization-(EM-)Algorithmus ähnelndes Vorgehen in Anwendung gebracht, um Schritt für Schritt, basierend auf einer korrekten, aber ungenauen Lösung, weitere Punkte zu finden und aus diesen eine umfassendere und genauere Lösung zu bestimmen. Abbildung 3 zeigt ein Beispiel für die kombinierte Orientierung von Aufnahmen vom Boden und von einem UAS aus.

Seit 2011 arbeitet die Gruppe an der UniBw im Rahmen des erweiterten Projektes des Geoinformationsdienstes der Bundeswehr mit dem Institut für Robotik und Mechatronik am Deutschen Zentrum für Luft- und Raumfahrt (DLR) zusammen. Insbesondere wurden hierbei die Ergebnisse für die Orientierung so aufbereitet, dass sie als Eingabe für das Semi Global Matching (SGM) von Hirschmüller [2008] genutzt werden können. Durch die hohe Genauigkeit der Bestimmung der Orientierung wird es möglich, hoch qualitative dichte 3D Modelle vollautomatisch zu generieren (s. Abb. 4).

4. Statistisch generative und aussehensbasierte Objektextraktion

Mumford [2000] fordert mit seinem „Dawning of the Age of Stochasticity“ eine stärkere statistische Modellierung. Dies gilt auch für die Interpretation von Bildern und die Objektextraktion. So wird in Dick et al. [2004] die 3D Interpretation von Gebäuden auf Grundlage von

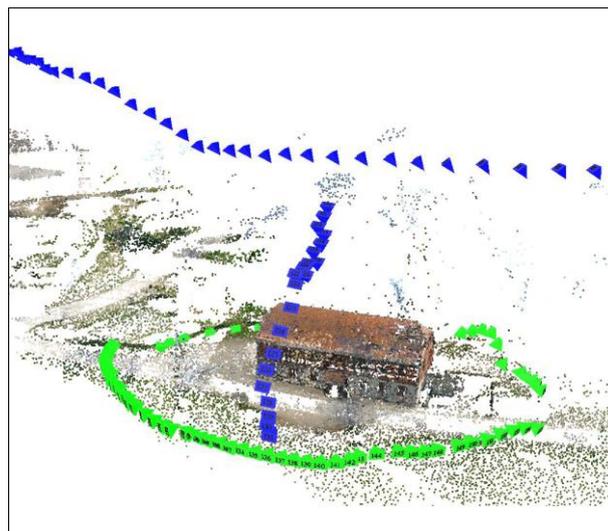


Abb. 3: Automatisch rekonstruierte 3D Punkte und Kamerapositionen – grüne Pyramiden für Aufnahmen vom Boden aus und blaue Pyramiden für Bilder vom UAS (aufsteigender Pfad und Überflug).



Abb. 4: Ergebnis für vollautomatische dichte Tiefenschätzung mittels Semi Global Matching (SGM) [Hirschmüller, 2008] auf Grundlage der hoch genauen Orientierungsbestimmung aus Bartelsen / Mayer [2010].

Bildsequenzen und statistischer Modellierung mittels Markoff Ketten Monte Carlo (MCMC) vorgestellt. Die Modellierung ist generativ, d. h., auf Grundlage des statistischen Priors, der die Annahmen des Modells beschreibt, können mittels zufälliger Stichprobengenerierung typische Instanzen des Modells, d. h. Gebäude, generiert werden. Damit wird es möglich, die gewählte Modellierung durch Simulation zu verifizieren: Wenn die erzeugten Gebäude unnatürlich aussehen, muss der Prior verändert werden.

Für die statistisch generative Objektextraktion werden, basierend auf dem näherungsweise vorliegenden 3D Modell und dessen Reflexionseigenschaften, simulierte Bilder erzeugt, die auf Grundlage einer geeigneten Metrik mit den real vorliegenden Bildern verglichen werden. Durch (statistische) Variation des Modells entstehen verschieden gut bewertete Simulationsergebnisse, die zur Bestimmung der Parameter des Modells genutzt werden. Ein grundlegendes Problem vieler statistischer generativer Ansätze besteht darin, dass das Ergebnis und vor allem die Effizienz der Generierung von gegebenen sinnvollen Näherungen abhängen. Zusätzlich zum Start einer statistisch generativen Objektextraktion auf Grundlage einer diskriminativen Näherung schlagen *Tu et al.* [2005] die Kombination von diskriminativem und generativem Vorgehen vor, indem in die Vorschlagswahrscheinlichkeiten für Änderungen des generativen Modells Ergebnisse diskriminativer Tests eingehen.

Die in Abschnitt 2 behandelte strukturelle Objektextraktion verwendet als Grundlage meist geometrische Merkmale wie Punkte, Kanten und Linien. Während sich diese z. B. für Straßenränder gut eignen, sind sie für die Modellierung anderer Objekte, wie z. B. Gesichter

oder Tiere, deutlich weniger sinnvoll. Hierfür haben sich in den letzten zehn Jahren aussehensbasierte Ansätze etabliert, die Objekte auf Grundlage von (kleinen) Bildausschnitten modellieren. Im Fall der „Bag of Words“ Ansätze spielt die räumliche Anordnung der Ausschnitte keine Rolle. Objekte werden rein an Hand der Häufigkeit und somit der Statistik des Auftretens von dem Modell ähnelnden Bildausschnitten gefunden. In vielen Fällen ist aber neben dem Aussehen auch noch die Anordnung der Bildausschnitte/Teile und sogar die Zuordnung einzelner Pixel der Bildausschnitte zum Objekt (Segmentierung) von Bedeutung [Leibe/Schiele, 2004]. Bei genügend Trainingsdaten können die Modelle oder Klassifikatoren überwacht gelernt werden. Modelle können aber auch unüberwacht gelernt werden, oder durch Übertragung von Priors können leistungsfähige Modelle auf Grundlage von sehr wenigen Trainingsdaten abgeleitet werden [Fei-Fei et al., 2006].

Der Ansatz der UniBw zur automatischen 3D Fassadeninterpretation [Reznik/Mayer, 2008] entstand auf der Grundlage des Teile-basierten Ansatzes zur aussehensbasierten Objektextraktion „Implicit Shape Model“ von Leibe/Schiele [2004]. Grundlegende Motivation war, dass das Aussehen von Fensterecken deutlich anders ist, als das von vielen anderen punktförmigen Strukturen. Weiterhin haben Fensterecken eine charakteristische Anordnung. Unter der Annahme, dass die zu analysierenden Fassaden, z. B. auf Grundlage des vertikalen Fluchtpunktes, horizontal ausgerichtet sind und zudem eine ähnliche Auflösung besitzen, wird für Trainingsdaten, d. h. gegebene Bildausschnitte mit Fenstern, das typische Aussehen von aus dem Bild extrahierten punktförmigen Strukturen als Repräsentanten für Fensterecken sowie deren Anordnung gelernt. Für ein Testbild werden dann wiederum die punktförmigen Strukturen extrahiert und diese mit den gelernten verglichen. Sind diese ähnlich, was über den Kreuzkorrelationskoeffizienten bewertet wird, so wird mit Hilfe des Wissens über die Anordnung auf die Lage des Fensters auf der Fassade geschlossen. Durch Konsens vieler Hypothesen erhält man eine statistische Verteilung für die Lage der Fensterränder. Diese wird mit dem Wissen kombiniert, dass Fenster oft in Form von Zeilen oder Spalten angeordnet sind, und für die entstehenden Einzelfenster, Fensterreihen oder Fensterspalten wird die 3D Lage in Form einer Parallelebene zur Fassade bestimmt (s. Abb. 5).

Die statistisch generative Extraktion von unbelaubten Laubbäumen [Huang/Mayer, 2010] verwendet ein LINDENMAYER(L-)System aus der Computergraphik, um die generative Verzweigungsstruktur von Bäumen zu modellieren. Der statistische Anteil besteht zuerst in den Prior-Verteilungen für die Parameter, wie z. B. Neigung, Dicke oder Länge eines Astes, die aber jeweils wiederum von den Vorgänger-Ästen abhängig sind. Beginnend bei dem deterministisch extrahierten Stamm werden auf Grundlage des L-Systems und der Prior-



Abb. 5: Ergebnis der automatischen 3D Interpretation von drei Fassaden mit Kameras als grüne Pyramiden und Fenstern als rote Rechtecke.

Verteilungen synthetische 3D Modelle für den Baum generiert und damit auf Grundlage von automatisch generierten Hintergrundbildern virtuelle Ansichten erzeugt. Diese werden mit den realen Bildern verglichen, was zur Likelihood des Modells und in Verbindung mit der Prior-Verteilung für die Parameter zu einer Maximum-a-Posteriori Schätzung führt. Zur statistischen Variation des Modells wird eine Kombination von Monte Carlo und MCMC Stichprobengenerierung verwendet. Abb. 6 zeigt ein Resultat des Ansatzes in rot im Vergleich mit dem gegebenen 3D Modell in grün. Die links dargestellte Seitensicht entspricht hierbei einer der verwendeten Kameraperspektiven. Für diese sind die Unterschiede zwischen den beiden Modellen gering. Das automatisch generierte Modell gibt die grobe Struktur wieder, hat aber weniger Details.

Durch die bei statistisch generativen Modellen notwendige Erzeugung von zufälligen Stichproben variieren die Ergebnisse. Allerdings variiert auch die menschliche Erfassung und daher sollte es ausreichend sein, dass die automatisch erzeugten Ergebnisse in der gleichen Grundgesamtheit wie die menschlichen Ergebnisse liegen.

5. Visual Computing-Ausbildung in der Informatik

Nach dem Wechsel in die Informatik der UniBw stand die Integration in das Curriculum von Bachelor und Master an. Da die Begriffe Photogrammetrie und Fernerkundung in der Informatik fast unbekannt sind, war eine Namensänderung naheliegend. An der Informatik der UniBw war neben Computer Vision auch Computer Graphik nicht vertreten. Daher erschien es sinnvoll, beides in Form des Begriffes „Visual Computing“ zu vereinen. Visual Computing wird in dieser Konnotation an verschiedenen Universitäten, u. a. auch an der ETH Zürich verwendet.

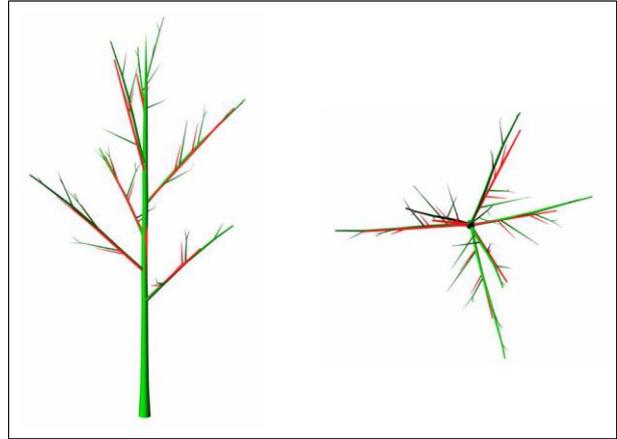


Abb. 6: Vergleich des Ergebnisses der automatischen Baumextraktion (rot) mit gegebenem 3D Modell (grün) für Seitensicht und Draufsicht.

Die Kombination von Computer Vision und Graphik bietet sich aus verschiedenen Gründen an. So stößt man in Computer Vision z. B. in der Bildzuordnung mit einfachen Reflexionsmodellen immer mehr an Grenzen und es werden mehr und mehr Modelle aus der Computer Graphik übernommen. Dort sind detaillierte Reflexionsmodelle schon länger von großem Interesse, um die komplizierten Reflexionseigenschaften natürlicher Objekte darzustellen. Eine weitere enge Kopplung zwischen Computer Vision und Computer Graphik ergibt sich bei der generativen Objektextraktion (siehe Abschnitt 4). Ein wichtiger Teilschritt ist dort die möglichst realitätsnahe Darstellung, um eine gute Übereinstimmung von simulierten und realen Bildern zu erreichen. Die Computer Graphik wird somit als Teil eines Computer Vision Ansatzes verwendet.

Im Studienangebot der Informatik wird Visual Computing in Form von Wahlpflichtfächern angeboten. Diese wurden im ersten Durchgang im Sommer 2011 gut angenommen. Im Bachelor nahm an „Grundzüge von Visual Computing“ im Modul „Geoinformatik und Visual Computing“ ein großer Teil des Jahrganges teil. Im Master werden zwei Module angeboten: Das Modul „Visual Computing“ besteht aus „Bildverarbeitung für Computer Vision“, „Computer Vision“ und „Computer Vision und Graphik“ und das Modul „Fernerkundung“ aus „Optische Fernerkundung“ und „Radar- und Lasermethoden“. An den ersten zwei Vorlesungen nahm immerhin ein Fünftel der Master-Studenten teil.

Im Weiteren werden insbesondere die Schwerpunkte der Vorlesungen angesprochen, die die Entwicklungen im letzten Jahrzehnt widerspiegeln.

5.1 Computer Graphik in Grundzüge von Visual Computing

Die Vorlesung „Grundzüge von Visual Computing“ stellt Grundlagen von Bildverarbeitung, (geometrischer) Computer Vision und Bildinterpretation bzw. Objekt-

extraktion dar. Der Schwerpunkt liegt auf Computer Graphik und hierbei wiederum auf der radiometrischen Abbildung. Die Modellierung der Reflektanz erfolgt für matte und spiegelnde, aber auch halb-spiegelnde Oberflächen. Die Schattierung kann kontinuierlich geändert werden (glatte Übergänge), Spiegelungen können durch Strahlverfolgung (Ray Tracing) realisiert werden und eine gegenseitige Beleuchtung von Oberflächen wird durch Radiosity ermöglicht. Letzteres bedeutet, dass z. B. eine im Schatten liegende Straße neben einer direkt beleuchteten roten Wand einen mehr oder weniger starken Rot-Ton erhält.

Neben der radiometrischen Abbildung ist die Modellierung von Verdeckungen durch „z-Buffering“, bei dem für jedes Pixel ein Objekt nur dargestellt wird, wenn es näher zur Kamera liegt, als das bisherige Objekt, sowie das Nachbilden von Effekten der Atmosphäre, wie z. B. Nebel, wichtig sowohl für Visualisierung als auch generative Modellierung.

5.2 Direkte Orientierung in Computer Vision

Neben der „klassischen“ Bündelausgleichung ist die vollautomatische Orientierung von Bilddaten praktisch ohne Vorwissen auch bei größeren Basen zwischen den Bildern der Schwerpunkt der Vorlesung „Computer Vision“. Zusammen mit dem direkten 5-Punkt Algorithmus spielt RANSAC (siehe Abschnitt 3) und damit die Nutzung von Daten mit vielen groben Fehlern eine zentrale Rolle. Die Zuordnung von Punkten wird mit dem Optischen Fluss eingeleitet. Dieser motiviert nicht nur Punktoperatoren, wie z. B. den FÖRSTNER-Operator [Förstner/Gülch, 1987], sondern erweitert auch die Möglichkeiten hin zu starren (oder auch nicht-starren) Bewegungen für Anwendungen, wie z. B. die Verfolgung von Personen oder Fahrzeugen. SIFT ermöglicht die Bestimmung von Größe und Ausrichtung von Punkten und affine kleinste Quadrate die Zuordnung bei großen Basen. Zuletzt wird mit dem Semi Global Matching von Hirschmüller [2008] eines der momentan leistungsfähigsten Verfahren für die dichte Tiefenschätzung dargestellt.

5.3 Aussehensbasierte Objektkategorisierung in Computer Vision und Graphik

Während die strukturelle Objektextraktion (siehe Abschnitt 2) mit ihrer Modellierung von Objekten und Relationen in der bisherigen Vorlesung „Bildanalyse“ den Schwerpunkt bildete, basiert die Vorlesung „Computer Vision und Graphik“ auf aussehensbasierter Objektkategorisierung. So wird dargestellt, wie beim „Bag of Words“ Ansatz (siehe Abschnitt 4) allein auf Grundlage der Häufigkeit (Statistik) von kleinen Bildausschnitten Objektkategorien in Bildern unterschieden werden können. Wird die Anordnung der Ausschnitte hinzugenommen, wird die Modellierung zwar unflexib-

ler, schränkt aber dafür auch die Lösungsmöglichkeiten ein. Sollen nicht nur Bilder kategorisiert werden, sondern die Objekte auch segmentiert, d. h. extrahiert werden, dann ermöglicht die statistische generative Objektextraktion die umfassendste Modellierung. Allerdings ist diese noch nicht in der Lage, allgemeine Objekte der realen Welt abzubilden, und so bleiben die Objekte und Relationen der strukturellen Objektextraktion doch zumindest eine Quelle der Inspiration.

5.4 Moderne Klassifikationsverfahren in Optischer Fernerkundung

Die Vorlesung „Optische Fernerkundung“ hat als erste Aufgabe, die Verbindung zur Geoinformatik herzustellen, die mit „Visual Computing“ am Institut für Angewandte Informatik vereint ist. Hierzu werden Sensoren vorgestellt sowie Vorgehensweisen für die Datenerfassung. Ein weiterer Schwerpunkt liegt auf der Klassifikation. Neben der klassischen Maximum Likelihood und der auf dem BAYES-Theorem basierenden Maximum-a-Posteriori (MAP) Schätzung werden neue Techniken wie z. B. Support Vector Machines oder Random Forests vorgestellt. Hinter allem steht das Ziel, die Statistik der Daten auszunutzen und wenn möglich die komplexen Modelle unserer Umwelt zu lernen. Eine Modellierung von Hand ist zumeist zu aufwändig.

6. Ausblick

Tempora mutantur, nos et mutamur in illis – mehr als ein Jahrzehnt an der UniBw ist vorbei und es hat sich vieles geändert. Computer Vision und Computer Graphik, d. h. Visual Computing, sind in dieser Zeit stark gewachsen und werden mit Anwendungen wie Surveillance oder auch Fahrerassistenzsystemen [Barth et al., 2010] immer relevanter für die Praxis. Weiterhin gehen wichtige Arbeiten für die Erfassung topographischer Objekte für *Google Maps* oder *bing Maps* von Computer Vision Gruppen, wie z. B. an der TU Graz aus [Leberl et al., 2010]. Zuletzt ist sogar der „heilige Gral“ der Photogrammetrie, nämlich die Bündelausgleichung, in diesem Fall für Tausende von Bildern in allgemeiner Lage in Computer Vision so wichtig geworden, dass ein Vortrag hierzu [Jian et al., 2011] bei der International Conference on Computer Vision (ICCV) 2011 in Barcelona bei einer Annahmequote von unter 4% als Vortrag vorgestellt wurde. Auch ohne den „Unfall“ durch die Sparbemühungen an der UniBw wäre somit eine stärkere Annäherung von Photogrammetrie an Computer Vision (und Computer Graphik) folgerichtig gewesen.

Anzumerken ist, dass der Trend in Richtung Computer Vision an der UniBw schon sehr viel älter ist. Schon der Vorgänger des Autors, EGON DORRER, beschäftigte sich mit Inverser Schummerung – Shape from Shading –, einem in Computer Vision bedeutendem

Thema. Eine gute Darstellung der erreichten Leistungsfähigkeit gibt *Dorrer et al. [2005]*.

Die momentanen Entwicklungen in Theorie und Praxis geben Hoffnung, dass die Generierung und Nutzung von Bildinformation mit Visual Computing eine große Zukunft hat, deren Gestaltung spannend ist und sich lohnt – ob als Photogrammeter oder „Visual Computer“ scheint hierbei weniger wichtig.

7. Literatur

- Bacher, Uwe / Mayer, Helmut [2005]*: Automatic Road Extraction from Multispectral High Resolution Satellite Images. In: *Stilla, Uwe / Rottensteiner, Franz / Hinz, Stefan (Hrsg.)*: Proceedings of the ISPRS Workshop CMRT 2005, Object Extraction for 3D City Models, Road Databases and Traffic Monitoring – Concepts, Algorithms and Evaluation, Vienna – Austria, August 29-30 2005. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Band XXXVI, Teil 3/W24. Technische Universität München, Institut für Photogrammetrie und Kartographie, München, S. 29-34.
ISSN 1682-1777
http://www.isprs.org/proceedings/XXXVI/3-W24/papers/CMRT05_Bacher_Mayer.pdf
letzter Aufruf: 07.11.2011.
- Bartelsen, Jan / Mayer, Helmut [2010]*: Orientation of Image Sequences Acquired from UAVs and with GPS Cameras. In: *Surveying and Land Information Science*, 70. Jhrg., Heft 3, Gaithersburg, S. 151-159.
ISSN 1538-1242
ISSN 1052-2905
http://www.isprs.org/proceedings/XXXVIII/Eurocow2010/euroCOW2010_files/papers/04.pdf
letzter Aufruf: 07.11.2011.
- Barth, Alexander / Siegemund, Jan / Meißner, Annemarie / Franke, Uwe / Förstner, Wolfgang [2010]*: Probabilistic Multi-Class Scene Flow Segmentation for Traffic Scenes. In: *Goesele, Michael / Roth, Stefan / Kuijper, Arjan / Schiele, Bernt / Schindler, Konrad (Hrsg.)*: Pattern Recognition. 32nd DAGM Symposium, Darmstadt, Germany, September 22-24, 2010. Proceedings. Lecture Notes in Computer Science (LNCS), Band 6376. Springer, Berlin/Heidelberg, S. 503-512.
ISBN 978-3-642-15986-2
ISBN 3-642-15985-0
ISBN 978-3-642-15985-5
http://www.ipb.uni-bonn.de/uploads/tx_ikgpublication/DAGM2010_086.pdf
letzter Aufruf: 08.11.2011.
- Dick, Anthony R. / Torr, Philip H. S. / Cipolla, Roberto [2004]*: Modelling and Interpretation of Architecture from Several Images. In: *International Journal of Computer Vision*, 60. Jhrg., Heft 2, Dordrecht, S. 111-134.
ISSN 0920-5691
<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.93.5886&rep=rep1&type=pdf>
letzter Aufruf: 07.11.2011.
- Dorrer, Egon [1980]*: Suffix and Tensor Notation for Analytical Photogrammetry. In: *Manuscripta Geodaetica*, 5. Jhrg., Heft 2, Berlin/Heidelberg, S. 91-143.
ISSN 0340-8825
- Dorrer, Egon / Mayer, Helmut / Haase, Yvonne / Ostrovskiy, Alexej / Renter, Jörg / Rentsch, Matthias / Reznik, Sergej / Neukum, Gerhard / HRSC Co-Investigator Team [2005]*: Verbesserung räumlicher Daten durch „Shape from Shading“. In: *Photogrammetrie – Fernerkundung – Geoinformation (PFG)*, 9. Jhrg., Heft 5, Stuttgart, S. 403-408.
ISSN 1432-8364
- Fei-Fei, Li / Fergus, Rob / Perona, Pietro [2006]*: One-Shot Learning of Object Categories. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28. Jhrg., Heft 4, Washington, S. 594-611.
ISSN 0162-8828
<http://authors.library.caltech.edu/5407/1/LIFeetpam06.pdf>
letzter Aufruf: 07.11.2011.
- Fischler, Martin A. / Bolles, Robert C. [1981]*: Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. In: *Communications of the ACM*, 24. Jhrg., Heft 6, New York, S. 381-395.
ISSN 0001-0782
<http://www.ai.sri.com/pubs/files/836.pdf>
letzter Aufruf: 07.11.2011.
- Förstner, Wolfgang [2000]*: Moderne Orientierungsverfahren. In: *Photogrammetrie – Fernerkundung – Geoinformation (PFG)*, 4. Jhrg., Heft 3, Stuttgart, S. 163-176.
ISSN 1432-8364
- Förstner, Wolfgang / Gülch, Eberhard [1987]*: A Fast Operator for Detection and Precise Location of Distinct Points, Corners and Centres of Circular Features. In: *Grün, Armin / Beyer, Horst A. (Hrsg.)*: Intercommission Conference on Fast Processing of Photogrammetric Data, Interlaken, Switzerland. Proceedings. ETH Zürich, Institut für Geodäsie und Photogrammetrie, Zürich, S. 281-305.
<http://www.ipb.uni-bonn.de/proceedings>
letzter Aufruf: 07.11.2011.
- Forberg, Andrea / Mayer, Helmut [2006]*: Simplification of 3D Building Data. In: *Zeitschrift für Geodäsie, Geoinformation und Landmanagement (ZfV)*, 131. Jhrg., Heft 3, Augsburg, S. 148-152.
ISSN 1618-8950
http://ica.ign.fr/Leicester/oral-presentation/forberg_ICA_WS04-oral.pdf
letzter Aufruf: 08.11.2011.

- Frahm, Jan-Michael / Fite-Georgel, Pierre / Gallup, David / Johnson, Tim / Raguram, Rahul / Wu, Changchang / Jen, Yi-Hung / Dunn, Enrique / Clipp, Brian / Lazebnik, Svetlana / Pollefeys, Marc [2010]: Building Rome on a Cloudless Day. In: Daniilidis, Kostas / Maragos, Petros / Paragios, Nikos (Hrsg.): Computer Vision – ECCV 2010. 11th European Conference on Computer Vision, Heraklion, Crete, Greece, September 5-11, 2010, Proceedings, Part IV. Springer, Berlin/Heidelberg, S. 368-381. ISBN 3-642-15548-0 ISBN 978-3-642-15548-2 (print) ISBN 978-3-642-15549-9 (e-book) <http://www.springerlink.com/content/978-3-642-155604/#section=776899&page=3&locus=21> letzter Aufruf: 07.11.2011.*
- Goesele, Michael / Ackermann, Jens / Fuhrmann, Simon / Klowsky, Ronny / Langguth, Fabian / Muecke, Patrick / Ritz, Martin [2010]: Scene Reconstruction from Community Photo Collections. In: IEEE Computer, 43. Jhrg., Heft 6, Los Alamitos, S. 48-53. ISSN 0018-9162 http://www.cs.cornell.edu/~snave/publications/papers/proc_ieee_scene_reconstruction.pdf letzter Aufruf: 08.11.2011.*
- Hartley, Richard / Zisserman, Andrew [2000]: Multiple View Geometry in Computer Vision. Cambridge University Press, Cambridge, 607 S. ISBN 0-521-62304-9*
- Hirschmüller, Heiko [2008]: Stereo Processing by Semi-Global Matching and Mutual Information. In: IEEE Transactions on Pattern Analysis and Machine Intelligence, 30. Jhrg., Heft 2, Washington, S. 328-341. ISSN 0162-8828 http://elib.dlr.de/55367/1/Stereo_ProcessingHirschm%C3%BCller.pdf letzter Aufruf: 07.11.2011.*
- Huang, Hai / Mayer, Helmut [2010]: Generative Statistical 3D Reconstruction of Unfoliated Trees from Terrestrial Images. In: Annals of GIS, 15. Jhrg., Heft 2, Berkeley, S. 97-105. ISSN 1947-5683 http://www.ikg.uni-hannover.de/fileadmin/ikg/staff/publications/Begutachtete_Zeitschriftenartikel_und_Buchkapitel/huang09.pdf letzter Aufruf: 08.11.2011.*
- Jian, Yong-Dian / Balcan, Doru Christian / Dellaert, F. [2011]: Generalized Subgraph Preconditioners for Large-Scale Bundle Adjustment. In: Proceedings of IEEE 13th International Conference on Computer Vision (ICCV 2011), 6-13 November 2011, Barcelona, Spain, S. 295-302, DVD. ISBN 978-1-4577-1102-2 http://www.cc.gatech.edu/~yjian6/publication/jian2011_iccv.pdf letzter Aufruf: 08.11.2011.*
- Leberl, Franz / Bischof, Horst / Pock, Thomas / Irschara, Arnold / Kluckner, Stefan [2010]: Aerial Computer Vision for a 3D Virtual Habitat. In: IEEE Computer, 43. Jhrg., Heft 6, Los Alamitos, S. 24-31. ISSN 0018-9162 <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&number=5481931> letzter Aufruf: 08.11.2011.*
- Leibe, Bastian / Schiele, Bernt [2004]: Scale Invariant Object Categorization Using a Scale-Adaptive Mean-Shift Search. In: Rasmussen, Carl Edward / Bühlhoff, Heinrich H. / Giese, Martin A. / Schölkopf, Bernhard (Hrsg.): Pattern Recognition. 26th DAGM Symposium, Tübingen, Germany, August/September 2004. Proceedings. Lecture Notes in Computer Science (LNCS), Band 3175. Springer, Berlin/Heidelberg, S. 145-153. ISBN 3-540-22945-0 ISBN 978-3-540-22945-2 <http://www.vision.ee.ethz.ch/~bleibe/papers/leibe-scaleinvariant-dagm04.pdf> letzter Aufruf: 15.11.2011.*
- Lindeberg, Tony [1994]: Scale-Space Theory in Computer Vision. The Kluwer international series in engineering and computer science, Band 256. Kluwer, Boston, 423 S. ISBN 0-7923-9418-6*
- Lowe, David G. [2004]: Distinctive Image Features from Scale-Invariant Keypoints. In: International Journal of Computer Vision, 60. Jhrg., Heft 2, Dordrecht, S. 91-110. ISSN 0920-5691 <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.157.3843&rep=rep1&type=pdf> letzter Aufruf: 07.11.2011.*
- Mayer, Helmut [2002]: Fundamental Matrix and Trifocal Tensor for View Synthesis. In: Bildtechnik / Image Science, 9., Jhrg., Heft 1, Stockholm, S. 129-141. ISSN 1104-5825*
- Mayer, Helmut, / Baltsavias, Emmanuel / Bacher, Uwe [2006]: Automated Extraction, Refinement, and Update of Road Databases from Imagery and Other Data. Final Report. In: EuroSDR Projects. EuroSDR, European Spatial Data Research, Official Publication No. 50. Bundesamt für Kartographie und Geodäsie, Frankfurt/Main, S. 217-280. ISBN 978-90-5179474-8 ISBN 90-5179474-6 <http://www.eurocdr.net/publications/50.pdf> letzter Aufruf: 07.11.2011.*
- Mumford, David [2000]: The Dawning of the Age of Stochasticity. In: Arnold, Vladimir / Atiyah, Michael / Lax, Peter D. / Mazur, Barry (Hrsg.): Mathematics: Frontiers and Perspectives. American Mathematical Society, Providence, S. 197-218. ISBN 0-8218-2070-2*

ISBN 0-8218-2697-2

<http://www.dam.brown.edu/people/mumford/Papers/OverviewPapers/DawningAgeStoch.pdf>

letzter Aufruf: 07.11.2011.

Nistér, David [2004]: An Efficient Solution to the Five-Point Relative Pose Problem. In: IEEE Transactions on Pattern Analysis and Machine Intelligence, 26. Jhrg., Heft 6, Washington, S. 756-770.

ISSN 0162-8828

Reznik, Sergej / Mayer, Helmut [2008]: Implicit Shape Models, Self-Diagnosis, and Model Selection for 3D Facade Interpretation. In: Photogrammetrie – Fernerkundung – Geoinformation (PFG), 12. Jhrg., Heft 3, Stuttgart, S. 187-196.

ISSN 1432-8364

http://www.unibw.de/inf4/professuren/photo/schriften/pubreznik/PFG-Reznik-final.pdf/at_download/down1

letzter Aufruf: 08.11.2011.

Tu, Zhuowen / Chen, Xiangrong / Yuille, Alan L. / Zhu, Song-Chun [2005]: Image Parsing: Unifying Segmentation, Detection, and Recognition. In: International Journal of Computer Vision, 63. Jhrg., Heft 2, Dordrecht, S. 113-140.

ISSN 0920-5691

http://lear.inrialpes.fr/people/triggs/events/iccv03/cdrom/iccv03/0018_chen.pdf

letzter Aufruf: 07.11.2011.

Anschrift des Verfassers:

Prof. Dr.-Ing. Helmut Mayer

Universität der Bundeswehr München, Fakultät für Informatik, Institut für Angewandte Informatik, Professur für Photogrammetrie und Fernerkundung, D-85577 Neubiberg

E-Mail: Helmut.Mayer@unibw.de

