



Illustration: * siehe Artikelende

„Du sollst Dir ein Bild machen“

Millionen Schnappschüsse geistern durchs Internet – Computergraphiker nutzen sie nun erstmals, um dreidimensionale Ansichten von Bauwerken und Plätzen zu berechnen.

Als Urlauber entkommt man den Motiven nicht: etwa der Dresdner Frauenkirche unter frühling-blauem Himmel, der Golden Gate Bridge im wabernden Morgennebel über der San Francisco Bay oder der New Yorker Freiheitsstatue im tagesmüden Abendlicht. Welcher Tourist griffe nicht schnell zu Digitalkamera

oder Fotohandy, um sie ins rechte Bild (und Licht) zu setzen?

Heute werden Urlaubsfotos kaum mehr auf dem Wohnzimmertisch ausgebreitet, sondern auf kostenfreien Online-Fotobörsen wie Panoramio, Fotocommunity und Flickr eingestellt, wo sie betrachtet, kommentiert und nicht zuletzt herunter-

geladen werden können. Mehr als 6,5 Milliarden Schnappschüsse aus aller Welt bietet heute das Fotoshare-Portal Flickr (www.flickr.com); allein unter dem Stichwort „Statue of Liberty New York“ sind es mehr als 100 000 Ansichten. Ein gigantisches Daten-Eldorado, das mit Sicherheit ein Manko hat: die Schnappschüsse

– mal mit winkenden Touristen im Vordergrund, mal mit herabstürzenden Taubengeschwadern – sind von höchst unterschiedlicher Motivqualität.

„Für uns ist das ein ganz großer und produktiver Schatz“, sagt Michael Goesele. Der 39-Jährige ist seit Juni 2011 Professor (zuvor Juniorprofessor) für Graphisch Interaktive Systeme an der TU Darmstadt. Er bedient sich der Amateurfotos („eigentlich der Worst Case für einen Computergraphiker“), um aus ihnen dreidimensionale Ansichten von Bauwerken, Plätzen und Stadtlandschaf-

ten zu berechnen. Informatiker nennen das „Szenen-Rekonstruktion“.

Seit drei Jahren arbeitet Goesele, Experte für Visual Computing, mit seiner fünfköpfigen Emmy Noether-Nachwuchsgruppe der DFG zur „exakten Szenerekonstruktion aus extrem großen Bildmengen“.

Für Michael Goesele, selbst passionierter Schwarz-Weiß-Fotograf, sind die Internetfotos das Datenmaterial und gewissermaßen auch sein Spielmaterial für die zentrale Frage seiner Forschungsarbeit: Wie kann aus zweidimensionalen Digitalbildern die Tiefeninformation herausgelesen

werden, um Szenen räumlich darzustellen? Dieses Problem treibt Informatiker seit langem um.

Bislang ist mit kontrollierten Daten gearbeitet worden. Häufig sind es mit einer kalibrierten Kamera aufgenommene Innenaufnahmen – mit definiertem Objektstand, festgeschriebenem Fokus, optimal gesetztem Licht. Aus solchen Laboransichten, die zeit- und kostenaufwendig sind, lassen sich Modelle generieren. „Doch wie kann das auch mit unkontrollierten Daten gelingen?“, fragt Goesele. Der Hintergedanke: Erst wenn der Schritt von vordefinierten

Labor- zu komplexen Realitätsdaten gelänge, wäre ein alltagstaugliches, weil robustes Verfahren gefunden.

An eine solche Software knüpfen Fachleute hohe Erwartungen. Sie könnte zum Beispiel bei der Dokumentation archäologischer Ausgrabungsstätten oder musealer Sammlungen helfen, Entwurfsplanungen im Automobilbau unterstützen oder Visualisierungen in der Raum- und Stadtplanung vereinfachen. Möglicherweise wäre auch eine Alternative zum kostspieligen Laserscannen dreidimensionaler Objekte sowie zur Stereofotografie gefunden – den bislang gängigen Möglichkeiten, „dreidimensionalen Content direkt zu generieren“.

Die Erde und ihre Hotspots im Blick: Computergraphiker Michael Goesele und sein Mitarbeiter Simon Fuhrmann.



Ausgangspunkt aller Studien ist das Stereoprinzip – Grundlage der räumlichen Wahrnehmung. Mit seinen Augen betrachtet der Mensch seine Umgebung aus zwei Blickwinkeln. So vermag das Gehirn allen Objekten eine Entfernung zuzuordnen, und es entsteht ein räumliches Bild. „Umgangssprachlich wird Stereoskopie fälschlicherweise mit 3-D gleichgesetzt“, unterstreicht Goesele, „obwohl es sich genau genommen nur um zweidimensionale Abbildungen handelt, die einen räumlichen Eindruck vermitteln.“

Das Prinzip des stereoskopischen Sehens macht sich die Computergraphik zunutze. Goeseles Internetfotos sind unter allen möglichen und unmöglichen Bedingungen aufgenommen worden: der Abstand der Kamera zum Objekt, der Blickwinkel, die Lichtverhältnisse, der Bildausschnitt, die Schärfenstrecke – sie alle unterscheiden sich. Um die Tiefeninformationen aus diesen Digitalbildern herauszuziehen, durchlaufen sie zunächst ein „Registrierungsverfahren“: Mit dem „Structure from Motion“-Verfahren wird für jedes Bild der Standort des Fotografen, sein Blickwinkel und weitere wichtige Kameraparameter (Zoom, Verzerrungen etc.) ermittelt. Dies geschieht mit einer Software namens „Bundler“, die vor einigen Jahren an der Universität Washington in Seattle entwickelt wurde und die auch ein Grobmodell der Szene erzeugt.

Um zu einem detailgenauen Modell zu gelangen, nutzt Goesele dann einen „Multi View Stereo“-Algorithmus, eine Software, die er zusammen mit Kollegen von der University of Washington und Microsoft Research entwickeln konnte. Diese Innovation trug ihm 2008 den „Eurographics Young Researcher Award“ ein. Ausgehend von einem Bildpunkt eines Fotos, zum Beispiel einem Fensterkreuz,



macht sich die Software Pixel für Pixel auf die Suche nach korrespondierenden Punkten, die stereotauglich sind. Auch der Laie versteht, dass das ein gigantischer Puzzelauf in einem automatisierten Verfahren sein muss.

Mithilfe der ermittelten Kameraposition kann aus der Geometrie der Anordnung die Tiefe berechnet werden. Am Ende steht ein plastisches Geometriemodell mit der Optik einer „Knetmassen-Kreation“, das Fachleute wie Laien in seiner Detailgenauigkeit beeindruckt. Dennoch – und darauf legt Goesele Wert – betreibt er Grundlagenforschung. „Sicher würden sich neben den kooperierenden Kollegen von Microsoft auch Unternehmen wie Google für die Technik interessieren“, sagt Goesele, „doch bis zur Marktreife dürfte noch einige Zeit vergehen.“

„Du sollst Dir ein Bild machen“: Bisher haben die Darmstädter Pixelmacher vor allem die Geometrie herangezogen, während andere Informationen ungenutzt blieben. So hat die Zahl der Eingabefotos bisher einige Hundert nicht überschritten. Und Kenndaten zu optischen Effek-

ten wie Schatten, Farben und Reflexe konnten erst in Ansätzen berücksichtigt werden. Deshalb wirken die Modelle auf den Betrachter wie Plastilin-Objekte. Um Kontur- und Oberflächenmodellierung voranzubringen, könnten „photometrische Verfahren“ erprobt werden. Sicher ist: Um zu realitätsnäheren Ergebnissen zu kommen, müssen weitere Informationen in die Modelle einfließen. An dieser Mammutaufgabe arbeitet die Emmy Noether-Nachwuchsgruppe.

Die beiden Wissenschaftlichen Mitarbeiter der Arbeitsgruppe, Jens Ackermann und Simon Fuhrmann, teilen sich ein Büro im Fraunhofer-Institut für Graphische Datenverar-



Links: Simon Fuhrmann (stehend) und Jens Ackermann bei der Simulation eines Keramik-Laubfrosches. Unten: „Exploded Views – Remapping Firenze“. Die Betrachter „erlaufen“ sich die virtuelle Stadtlandschaft.

beitung, mit dessen Forschern die Gruppe eng zusammenarbeitet. Die beiden Nachwuchsforscher haben Abgabetermine für Konferenzpapiere im Nacken – und präsentieren dennoch ihre Projekte.

Den Diplom-Mathematiker Jens Ackermann beschäftigt die Frage, wie die Reflexionseigenschaft von Objekten besser abgebildet werden könnte – eine ebenso grundlegende wie schwierige Frage. Denn häufig hat ein Objekt matte und glänzende Oberflächenbereiche, die einfallendes Licht unterschiedlich reflektieren. Ackermann hat sich dieser Herausforderung bei der Simulation einiger Objekte, wie einem dickbäuchigen Keramik-Laubfrosch, angenommen; dabei kombiniert er „Multi View Stereo“- mit einem photometrischen Verfahren. So gewann er Daten, die die Raumorientierung der betrachteten Oberfläche verraten, was hilft, Oberflächen präziser zu rekonstruieren.

Auch Diplom-Informatiker Simon Fuhrmann geht es um Fortschritte bei der Oberflächenrekonstruktion, die er mit der Frage verbindet, wie mit Bilddaten unterschiedlicher Auflösung (Skalen) produktiver umgegangen werden kann. Dafür spielt die Kombination von Tiefenkarten – Bilder, die an jedem Pixel den Abstand zur Kamera speichern – eine entscheidende Rolle. „3-D ist unsere Computerwelt“, sagt Fuhrmann, „die vierte Dimension ist die Auflösung.“ Von dieser vierten Dimension profitieren – vereinfachend gesagt – seine Rekonstruktionsversuche, zum Beispiel an der Fassade von Notre Dame. Auf der Konferenz

SIGGRAPH Asia der Association for Computing Machinery konnte Fuhrmann im Dezember 2011 sein neues Verfahren in Hongkong vorstellen.

Es ist ein verlockender Wunsch, dreidimensionale Visualisierungen von touristischen Hotspots, die auch Goesele im Blick hat, wie in einem Computerspiel durchwandern zu können. Wie das unter künstlerisch-spielerischen Vorzeichen aussehen kann, zeigt die Multimediainstallation „Exploded Views – Remapping Firenze“ des niederländischen Künstlers Marnix de Nijs (www.marnixdenijs.nl/exploded-views.htm). Ihre Besonderheit: Der auf einem Laufband stehende Besucher erläuft sich buchstäblich die virtuelle Stadtlandschaft von Florenz, indem er auf die Projektionsleinwand zugeht. Die dreidimensionalen Florentiner Welten für diese ungewöhnliche Erfahrungsmaschine haben Goesele und Kollegen erarbeitet.

Ob für künstlerische oder technische Aufgaben – Schritt für Schritt könnte die Realitätsnähe dreidimensionaler Rekonstruktionen durch computergraphische Grundlagenforschung verbessert werden. Damit geraten neue und vielversprechende Anwendungen in den Blick. Die Grenzen dieses Ansatzes bleiben noch auszuloten. Doch dass Internetbilder ein potenter Datenschatz für das Visual Computing sind und die Studien zum 3-D-Mapping voranbringen werden – dafür spricht schon jetzt viel.

Dr. Rembert Unterstell
ist Chef vom Dienst der „forschung“.

* Das Aufmacherbild ist eine Collage der Emmy Noether-Gruppe Goesele; die Fotos stammen von den Flickr-Usern John Bailey, Holger Baschleben, Raymond Chan, D-HausBT (2x), Bert Kaufmann, Frank Kovalchek, Angela Marie, Will Palmer, Francois Philipp, Christian Prade, Paul Sableman (2x), Ralf Schulze (2x), Frank Steiner (2x), Ruben Vique (2x) und yeowatzup.