

DISS. ETH NO. 22933

**3D RECONSTRUCTION AND RENDERING  
FROM HIGH RESOLUTION LIGHT FIELDS**

A thesis submitted to attain the degree of  
DOCTOR OF SCIENCES of ETH ZURICH  
(Dr. sc. ETH Zurich)

presented by

CHANGIL KIM

MSc in Computer Science, ETH Zurich

born on  
June 19, 1979

citizen of  
The Republic of Korea

accepted on the recommendation of

Prof. Dr. Markus Gross  
Dr. Alexander Sorkine-Hornung  
Prof. Dr. Brian Curless

2015

# Abstract

This thesis presents a complete processing pipeline of densely sampled, high resolution light fields, from acquisition to rendering. The key components of the pipeline include 3D scene reconstruction, geometry-driven sampling analysis, and controllable multiscopic 3D rendering.

The thesis first addresses 3D geometry reconstruction from light fields. We show that dense sampling of a scene attained in light fields allows for more robust and accurate depth estimation without resorting to patch matching and costly global optimization processes. Our algorithm estimates the depth for each and every light ray in the light field with great accuracy, and its pixel-wise depth computation results in particularly favorable quality around depth discontinuities. In fact, most operations are kept localized over small portions of the light field, which by itself is crucial to scalability for higher resolution input and also well suited for efficient parallelized implementations. Resulting reconstructions retain fine details of the scene and exhibit precise localization of object boundaries.

While it is the key to the success of our reconstruction algorithm, the dense sampling of light fields entails difficulties when it comes to the acquisition and processing of light fields. This raises a question of optimal sampling density required for faithful geometry reconstruction. Existing works focus more on the alias-free rendering of light fields, and geometry-driven analysis has seen much less research effort. We propose an analysis model for determining sampling locations that are optimal in the sense of high quality geometry reconstruction. This is achieved by analyzing the visibility of scene points and the resolvability of depth and estimating the distribution of reliable estimates over potential sampling locations.

A light field with accurate depth information enables an entirely new approach to flexible and controllable 3D rendering. We develop a novel algorithm for multiscopic rendering of light fields which provides great controllability over the perceived depth conveyed in the output. The algorithm synthesizes a pair of stereoscopic images directly from light fields and allows us to control stereoscopic and artistic constraints on a per-pixel basis. It computes non-planar 2D cuts over a light field volume that best meet described constraints by minimizing an energy functional. The output images are synthesized by sampling light rays on the cut surfaces. The algorithm generalizes for multiscopic 3D displays by computing multiple cuts.

The resulting algorithms are highly relevant to many application scenarios. It can readily be applied to 3D scene reconstruction and object scanning, depth-assisted segmentation, image-based rendering, and stereoscopic content creation and post-processing, and can also be used to improve the quality of light field rendering that requires depth information such as super-resolution and extended depth of field.

# Zusammenfassung

Diese Dissertation präsentiert eine komplette Verarbeitungspipeline für dicht abgetastete, hochauflösende Lichtfelder, von der Akquisition bis zu deren Rendering. Die wichtigsten Komponenten dieser Pipeline umfassen 3D-Szenenrekonstruktion, Geometrie-gesteuerte Abtastanalyse, und kontrollierbares multiskopisches 3D-Rendering.

Diese Dissertation beschäftigt sich zunächst mit 3D-Geometrierekonstruktion von Lichtfeldern. Wir zeigen, dass dichte Abtastung einer Szene in Form von Lichtfeldern eine robuste und genaue Tiefenmessung ermöglicht, und zwar ohne Nachbarschaften von Pixeln zu vergleichen, und ohne auf kostspielige globale Optimierungsprozesse zurückzugreifen. Unser Algorithmus schätzt die Tiefe für jeden Lichtstrahl im Lichtfeld mit grosser Genauigkeit, und die resultierenden pixelweisen Tiefenwerte sind von besonders hoher Qualität in der Nähe von Tiefendiskontinuitäten. In der Tat sind die meisten Operationen für kleine Abschnitte des Lichtfeldes lokalisiert, was entscheidend für die Skalierbarkeit für höher aufgelöste Eingabedaten ist und auch eine effiziente parallele Implementierung erlaubt. Die resultierenden Rekonstruktionen erhalten feine Seznendetails mit einer präzisen Lokalisierung von Objektkanten.

Während es der Schlüssel zum Erfolg unseres Rekonstruktionsalgorithmus ist, führt die hohe Abtastdichte von Lichtfeldern zu Problemen bei deren Aufnahme und Verarbeitung. Dies wirft die Frage der optimalen Abtastdichte auf welche zur genauen Geometrierekonstruktion erforderlich ist. Bestehende Arbeiten konzentrieren sich mehr auf das Alias-freie Rendering von Lichtfeldern und Geometrie-getriebene Analyse hat deutlich weniger Forschungsanstrengungen

gesehen. Wir schlagen ein Analysemodell zur Bestimmung von Aufnahmepositionen vor, die im Sinne von qualitativ hochwertigen Geometrierekonstruktion optimal sind. Dies wird durch die Analyse der Sichtbarkeit der Szenenpunkte und der Tiefenauflösbarkeit, sowie der Bestimmung der Verteilung der zuverlässigen Schätzungen für potenzielle Aufnahmepositionen erreicht.

Ein Lichtfeld mit genauer Tiefeninformation ermöglicht einen völlig neuen Ansatz für flexibles und steuerbares 3D-Rendering. Wir entwickeln einen neuartigen Algorithmus für multiskopisches Rendering von Lichtfeldern, der grosse Steuerbarkeit über die wahrgenommene Tiefe in der Ausgabe erlaubt. Der Algorithmus synthetisiert ein Paar von stereoskopischen Bildern direkt von Lichtfeldern und ermöglicht es, stereoskopische und künstlerische Vorgaben auf Pixelebene zu steuern. Er berechnet nicht-planare 2D-Schnitte in einem Lichtfeldvolumen, welche die vorgegebenen Vorgaben bestmöglich erfüllen durch die Minimierung eines Energiefunktionals. Die Ausgangsbilder werden durch Abtasten von Lichtstrahlen auf den Schnittflächen synthetisiert. Der Algorithmus kann für multiskopische 3D-Displays verallgemeinert werden indem mehrere Schnitte berechnet werden.

Die resultierenden Algorithmen sind von grosser Bedeutung für viele Anwendungsszenarien. Sie sind leicht anwendbar für 3D-Szenenrekonstruktion und Objektscanning, Tiefen-gestützte Segmentierung, bildbasiertes Rendering sowie stereoskopische Inhaltsgenerierung und Nachbearbeitung, und können auch verwendet werden, um die Qualität von Lichtfeld-Rendering zu verbessern, welches Tiefeninformation erfordert, wie zum Beispiel Super-Resolution und erweiterte Tiefenschärfe.