



TECHNISCHE UNIVERSITÄT MÜNCHEN
Ingenieur fakultät Bau Geo Umwelt
Photogrammetrie und Fernerkundung

Direct Plenoptic Odometry –
Robust Tracking and Mapping with
a Light Field Camera

Niclas Zeller

Dissertation

Abstract

Research and industry proceed to build autonomous driving cars, self navigating unmanned aerial vehicles, and intelligent, mobile robots. Furthermore, due to the demographic change intelligent assistance devices for visually impaired an elderly people are in demand. For these tasks systems are needed which reliably map the 3D surrounding and are able to self-localize in these created maps. Such systems or methods can be summarized under the term simultaneous localization and mapping (SLAM) or visual odometry (VO). The SLAM problem was broadly studied for different visual sensors like monocular, stereo and RGB-D cameras. During the last decade plenoptic cameras (or light field cameras) became available as commercial products. They capture 4D light field information in a single image. This light field information can be used for instance to retrieve 3D structure.

This dissertation deals with the task of VO for plenoptic cameras. For this purpose a new model for micro lens array (MLA) based plenoptic cameras was developed. On the basis of this model a multiple view geometry (MVG) for MLA based plenoptic cameras is derived.

An efficient probabilistic depth estimation approach for MLA based light field cameras is presented. The method establishes semi-dense depth maps directly from the micro images recorded by the camera. Multiple depth observations are merged in a probabilistic depth hypotheses. Disparity uncertainties resulting e.g. from differently focused micro images and sensor noise are taken into account. This algorithm is developed on the basis of single light field images and will be extended by the introduced MVG afterwards.

Furthermore, calibration approaches for focused plenoptic cameras at different levels of complexity are presented. Starting from depth conversion functions which convert the virtual depth estimated by the camera into metric distances we continue to derive a plenoptic camera model on the basis of the estimated virtual depth map and the corresponding synthesized totally focused image. This model considers depth distortion and a sensor tilted with respect to the main lens. Finally it leads us to a plenoptic camera model which define the complete projection of an object point to multiple micro images on the sensor. Based on this model we emphasize the importance of modeling squinting micro lenses. The depth conversion functions are estimated on a series of range measurements while the parameters of all other models are estimated in a bundle adjustment using a 3D calibration target. Our bundle adjustment based methods significantly outperform existing approaches.

The plenoptic camera based MVG, the depth estimation approach, and the model obtained from the calibration are combined in a VO algorithm called Direct Plenoptic Odometry (DPO). DPO works directly on the recorded micro images. Therefore, it has not to deal with aliasing effects in the spatial domain. The algorithm generates semi-dense 3D point clouds on the basis of correspondences in subsequent light field frames. A scale optimization framework is used to adjust scale drifts and wrong absolute scale initializations. Up to our knowledge, it is the first method that performs tracking and mapping for plenoptic cameras directly on the micro images. DPO is tested on a variety of indoor and outdoor sequences. With respect to the scale drift it outperforms state-of-the-art monocular VO and SLAM algorithms. Regarding the absolute accuracy DPO is competitive to existing monocular and stereo approaches.

Kurzfassung

Sowohl in der Forschung als auch der Industrie werden autonome Fahrzeuge, selbstfliegende Drohnen und intelligente, mobile Roboter entwickelt. Außerdem wird aufgrund des demografischen Wandels die Nachfrage nach intelligenten Hilfsmitteln für sehbeeinträchtigte und ältere Menschen immer größer. Für die beschriebenen Anwendungen werden Systeme benötigt, welche zuverlässig die 3D Umgebung erfassen und sich selbst in dieser Umgebung lokalisieren können. Diese Systeme können unter den Begriffen Simultaneous Localization and Mapping (SLAM) oder Visuelle Odometrie (VO) zusammengefasst werden. Bisher wurden SLAM-Methoden hauptsächlich basierend auf monokularen, Stereo- und RGB-D Kameras untersucht. In der letzten Jahren kamen plenoptische Kameras auf den Markt. Diese sind in der Lage, anhand des aufgenommenen 4D Lichtfelds, z.B. 3D Information zu erlangen.

In dieser Arbeit werden VO-Methoden basierend auf einer plenoptische Kamera untersucht. Dafür wurde ein neues Modell sowie eine Multi-View-Geometrie (MVG) für Lichtfeldkameras mit Mikrolinsengittern (engl.: micro lens arrays, MLA) hergeleitet.

Es wurde eine Methode zur Tiefenschätzung für MLA basierte Lichtfeldkameras entwickelt. Diese Methode berechnet eine teildichte Tiefenkarte direkt aus den Mikrolinsenbildern der Kamera. Mehrere Tiefenbeobachtungen werden hierbei in einem stochastischen Modell vereint. Das Modell berücksichtigt Disparitätsunsicherheiten welche aus z.B. unterschiedlich fokussierten Mikrolinsenbildern und Sensorrauschen resultieren.

Weiterhin werden verschiedene Kalibriermethoden für fokussierte plenoptische Kameras vorgestellt. Zunächst werden Funktionen zur Umrechnung der geschätzten virtuellen Tiefe in metrische Abstände definiert. Anschließend leiten wir Modelle basierend auf der geschätzten virtuellen Tiefenkarte und dem zugehörigen totalfokussierten Bild her. Diese Modelle berücksichtigen Tiefenverzerrung und eine Sensorverkipfung bezüglich der Hauptlinse. Schließlich wird ein Modell definiert, welches die komplette Projektion eines Objektpunkts in mehrere Mikrolinsenbilder auf dem Sensor beschreibt. Hier wird die Bedeutung der Modellierung von schielenden Mikrolinsen, wie sie in einer plenoptischen Kamera vorkommen, herausgearbeitet. Die Funktionen zur Tiefenumrechnung werden anhand einer Reihe von Abstandsmessungen bestimmt, während alle weiteren Modelle mit einem 3D Kalibrierobjekt im Bündelausgleich geschätzt werden. Unsere auf Bündelausgleich basierende Kalibriermethoden erzielen robustere Ergebnisse als bisherige Methoden.

Die MVG für plenoptische Kameras, das Verfahren zur Tiefenschätzung und das für die Kalibrierung entwickelte Modell werden in einem VO-Algorithmus namens Direct Plenoptic Odometry (DPO) vereint. DPO arbeitet direkt auf den Mikrolinsenbildern und vermeidet damit Unterabtastungseffekte bei der Tiefenschätzung. Der Algorithmus erzeugt teildichte 3D Punktwolken basierend auf Korrespondenzen in aufeinanderfolgenden Lichtfeldaufnahmen. Nach unserem Wissen ist DPO die erste Methode welche die Kameraposition sowie eine 3D Karte aus den Aufnahmen einer plenoptischen Kamera erstellt. DPO wurde mit verschiedenen Innenraum- und Außenbereichsequenzen getestet. Bezüglich des Skalierungsdrifts übertrifft DPO aktuellste monokulare VO und SLAM Algorithmen. Die absolute Genauigkeit liegt in der Größenordnung von aktuellen monokularen und stereobasierten Algorithmen.