

Structure from Motion (SfM)

Studiengang: BSc in Informatik | Vertiefung: Computer Perception and Virtual Reality

Betreuer: Prof. Marcus Hudritsch

Experte: Dr. Harald Studer (Optimo Medical AG)

Menschen und andere Lebewesen können die räumliche Struktur einer Umgebung erfassen, indem sie sich durch diese bewegen. Diese für uns Menschen nebenläufige Aufgabe wird im Bereich der Computer-Vision Structure from Motion (SfM) genannt und ist ein Prozess, bei dem räumliche Strukturen (3D) anhand von Bildern (2D) berechnet werden. Im Rahmen dieser Bachelor-Thesis wurde eine Applikation entwickelt, die aus Bild-Sequenzen die jeweilige räumliche Struktur berechnet.

Zielsetzung

Ziel dieser Bachelor-Thesis war es die technischen Grundlagen des SfM-Prozesses zu verstehen und in einer Applikation mit einem geeigneten Framework umzusetzen. Diese soll aus einer Bildsequenz (z. B. Video) mittels Merkmalerkennung die Position der Kamera ermitteln und anschliessend daraus die räumliche Struktur berechnen. Im Vergleich zu anderen Methoden ist SfM kostengünstig, da ausser dem Computer nur eine handelsübliche Kamera vonnöten ist. Unter anderem deshalb finden sich für SfM viele Anwendungen. Es sei beispielsweise die Erfassung der räumlichen Struktur eines Gebäudes oder Geländes mit Hilfe einer Drohne genannt.

Implementierung

In der entwickelten Anwendung wird die Rekonstruktion des Objekts in einem mehrstufigen Verfahren umgesetzt. Diese wurde als C++ Applikation basierend auf den Frameworks Qt, OpenCV und OpenGL implementiert.

In einem ersten Schritt werden markante Bildmerkmale wie Kanten oder Ecken in den einzelnen Bildern der Sequenz mit dem SURF Algorithmus identifiziert. Zu diesen werden anschliessend auf den Nachbarbildern die korrespondierenden Punkte gesucht.

Für die Korrespondenzen des ersten Bildpaares wird ermittelt, wo sich die zwei Kameras während den Aufnahmen räumlich zueinander befanden. Aus diesen Po-

sitionen können nun die Punkte des Objekts im Raum berechnet werden. Mit den berechneten Objektpunkten und den Korrespondenzen des nächsten Bildpaares, kann die nächste Kameraposition berechnet werden. Damit können nun weitere Objektpunkte ermittelt werden. Durch diesen iterativen Prozess erhält man eine massstabsgetreue (up-to-scale) Punktwolke, wobei der Massstab selbst jedoch unbekannt ist.

Die 3D-Struktur wird in einem OpenGL Viewer mit den berechneten Kameras visualisiert. Dieser ermöglicht eine Rotation des Objekts, sodass dieses von allen Seiten betrachtet werden kann.

Die Applikation ermöglicht zudem einen Export der 3D-Punkte im verbreiteten ply-Format, das in weiteren Anwendungen genutzt werden kann.

Fazit

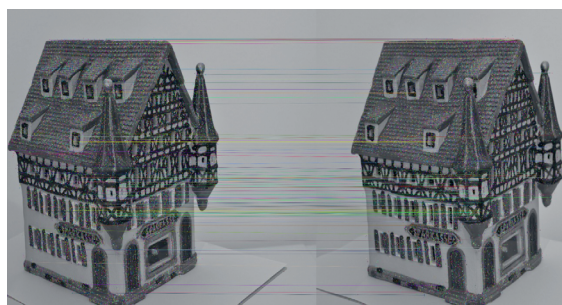
Die entstandene Arbeit und die Implementation der dazugehörigen Applikation ist eine ideale Einführung in das Thema und bietet eine Grundlage für weitere Projekte in diesem Bereich. Der Export der Rohdaten eröffnet viele Möglichkeiten für deren Weiterverwendung in anderen Applikationen. Diese können beispielsweise als Basis für die Generierung eines Meshs dienen, das für einen 3D-Druck genutzt werden kann.



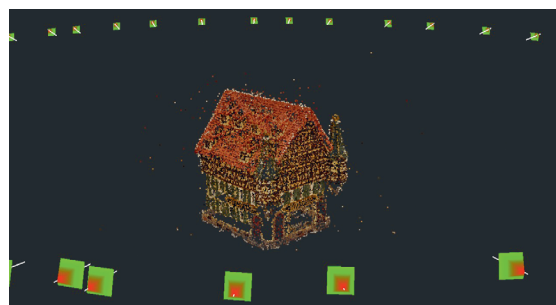
Marco Füllemann



Janosch Rohdewald



Einige korrespondierende Bildmerkmale zwischen einem Bildpaar



Berechnete Punktwolke mit einigen korrespondierenden Kameras