

Übersicht

In dieser Diplomarbeit wurde ein bildbasiertes 3-dim. Rekonstruktionssystem für alltägliche Tischszenen entwickelt. Die Umgebungsbilder werden mit einer Mikro-Kopf Kamera an der mobilen Hand (Abb. 2) eines Roboterarms mit sechs Freiheitsgraden aus zwei verschiedenen Betrachtungswinkeln zeitnah nacheinander akquiriert (Abb. 1). Durch interaktive Selektion von 10 korrespondierenden Punkten beider Ansichten wird die Epipolargeometrie, mathematisch ausgedrückt durch die Fundamentalmatrix F , berechnet. Die Korrespondenzpunkte können aus einer Menge von zuvor mit einem Harris-Kanten-Detektor ermittelten Ecken ausgewählt werden. Anschließend werden mit Hilfe der Epipolargeometrie die Eingabebilder rektifiziert um korrespondierende Epipolarlinien in neuen virtuellen Ansichten auf die selbe vertikale Bildkoordinate abzubilden. Zur Disparitätsanalyse und der Erstellung einer 2.5-dim. Tiefenkarte (Abb. 3) wird ein schneller Stereoalgorithmus nach Birchfield eingesetzt. Aus der Tiefenkarte wird ein 3-dim. interaktives OpenGL Modell berechnet, das dem Nutzer als realistische, texturierte Punktwolke dargestellt wird.

Aufgabenstellung

Für robuste Greifplanungen wird ein 3-dim. Modell potentieller Interaktionsobjekte des Tisches benötigt. Durch den Einsatz einer monokularen mobilen Handkamera werden unnötige Bewegung der mobile Roboter-Plattform und Verdeckungen in den Bildern durch andere Roboterkomponenten (bspw. Roboterarm) vermieden.



Abbildung 1: Originalaufnahmen der Tischszene aus verschiedenen Betrachtungswinkeln

Systemarchitektur

Das Rekonstruktionssystem lässt sich in folgende Arbeitsstufen unterteilen:

- Kamerakalibration:** Vorab werden einmalig die Kameraparameter mit einem Kalibrationsalgorithmus nach (Tsai, 1986) ermittelt.
- Bildakquise:** Die Akquisition der Bilddaten wird mit einer Mikro-Kopf Kamera an der mobilen Hand eines Roboterarms mit sechs Freiheitsgraden realisiert (Abb. 2). Dabei werden die Bilder aus zwei verschiedenen Positionen zeitnah nacheinander aufgenommen (Abb. 1).
- Entzerrung:** Mit dem Verzerrungskoeffizienten κ aus der Kalibration wird die radiale Verzerrung über Gl. (1) korrigiert. Dabei ist $r^2 = u^2 + v^2$.

$$u = u_d(1 + \kappa_1 r^2) \quad \text{und} \quad v = v_d(1 + \kappa_1 r^2) \quad (1)$$

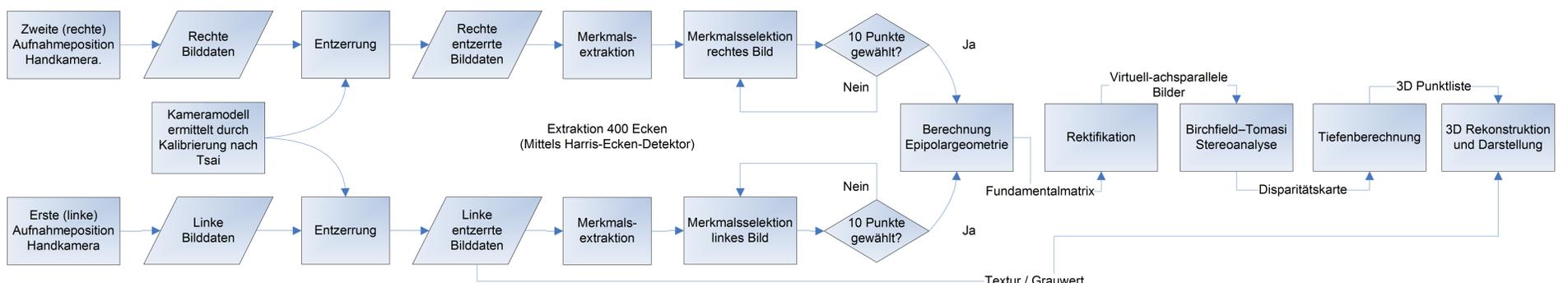


Abbildung 5: Flussdiagramm

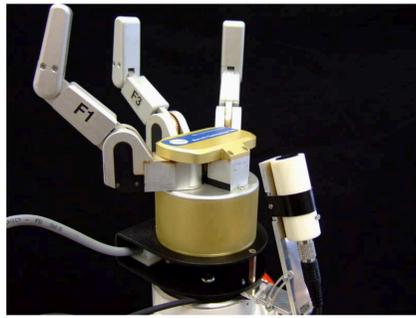


Abbildung 2: Mikro-Kopf Handkamera

- Merkmalsextraktion:** Über einen Harris-Kanten-Detektor wird eine variable Anzahl Ecken extrahiert.
- Berechnung Epipolargeometrie:** Durch interaktive Auswahl 10 korrespondierender Merkmale beider Ansichten wird über den Acht-Punkt-Algorithmus (Longuet-Higgins, 1981) die Epipolargeometrie berechnet. Diese wird nachfolgend für eine Dimensionsreduzierung des Suchraums der Korrespondenzanalyse verwendet.
- Rektifikation:** Ein Art Rückprojektion, wobei das konvergente Kamerasystem durch eine Rotation um die Projektionszentren der originalen Kameras virtuell in ein achsparalleles Stereosystem überführt wird. Praktische Umsetzung der zuvor erwähnten Reduzierung des Suchraums von zwei auf eine Dimension.
- Stereoanalyse:** Lösung des Korrespondenzproblems und Erstellung einer 2.5-dim. Tiefenkarte durch Nutzung eines Algorithmus nach (Birchfield und Tomasi, 1996). Je heller ein Objekt, desto näher ist es der Kamera (Abb. 3).



Abbildung 3: 2.5-dim. Disparitätskarte

- 3D Modell:** Über die Distanz beider Kamerazentren (Basislinie b) und dem Disparitätswert d der Tiefenkarte lässt sich für jeden Bildpunkt über Gl. (2) ein 3-dim. Raumpunkt ermitteln der den zugehörigen Grauwert aus den Originalaufnahmen erhält. Die Menge aller Raumpunkte lässt sich mittels OpenGL als Punktwolke darstellen (Abb. 4).

$$\begin{pmatrix} x_c \\ y_c \\ z_c \end{pmatrix} = \frac{b}{d} \begin{pmatrix} u_l \\ v_l \\ f \end{pmatrix} \quad (2)$$

Abbildung 5 zeigt in einem Flussdiagramm die einzelnen Verarbeitungsschritte des Rekonstruktionssystems.

Ergebnisse

Im Rahmen der Diplomarbeit wurde eine analytische Auswertung durchgeführt. Es sollte eine angemessene Basislinie für eine Tischszene und eine qualitative Bewertung der jeweils berechneten Disparitätsbilder des Stereoalgorithmus erlangt werden. Es konnte eine lineare Relation der wahren Objektentfernung zu den nach Gl. (2) ermittelten Tiefenwerten des Disparitätsbildes gezeigt werden. Aufgrund dieses Verhältnisses ließ sich ein Streckungsfaktor bestimmen, durch den die reale Objektiefe hinreichend genau berechnet werden kann. Abb. 4 zeigt zwei Ansichten des 3-dim. Modells der Tischszene rekonstruiert aus den Eingabebildern aus Abb. 1. Es kann deutlich zwischen Vorder- und Hintergrundobjekten unterschieden werden. „Löcher“ entstehen durch Verdeckungen in den Originalaufnahmen.

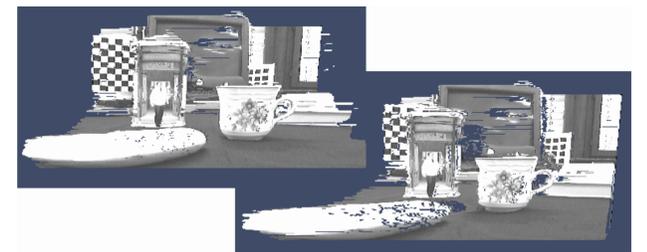


Abbildung 4: Verschieden 3-dim. Modellansichten

Über die eigentliche Aufgabenstellung hinaus ließen sich auch Modelle ganzer Räume erstellen.

Zusammenfassung & Ausblick

Durch Nutzung der Handkamera kann die Szene in Ihrer Ganzheit – unverdeckt durch Roboterkomponenten, wie es beim Stereokopf der Fall wäre – abgebildet werden. Wegen der Mobilität der Kamera am Roboterarm kann die Tischszene aus beliebigen Blickwinkeln aufgenommen werden ohne den Roboter neu zu positionieren und bietet somit eine hohe Flexibilität. Die Verwendung speziell auf Intel® Chipsätze optimierter Algorithmen der OpenCV Library äußert sich in einer guten Gesamtperformance des Rekonstruktionssystems.

Konzeptionelle Erweiterungsmöglichkeiten wären:

- **3-dim. Kollisionsvermeidung:** In dem ermittelten 3-dim. Modell könnten mögliche Kollisionen des Roboterarms mit Szenenobjekten für reale Greifaktionen vorab simuliert werden.
- **Multi view geometry:** Berechnung des 3-dim. Modells aus mehr als zwei Ansichten. Dies führe zu viel genaueren 3-dim. Modellen. Bereits genauer untersucht in (Hartley, 2003).
- **Next best view:** Algorithmen zur Vorhersage für die nächste Aufnahmeposition, die möglichst viele Informationen für das zu berechnende 3-dim. Modell beitragen.
- **Omnidirektionale Rekonstruktion:** Übertragung des Rekonstruktionssystems auf Bilder zweier unterschiedlicher Aufnahmepositionen der Omni-Kamera. Bessere Rekonstruktion ganzer Räume.