

# 3D-Deskriptoren für Aufgaben der Rekonstruktion und Objekterkennung

*Christian Lengert* — 15. Juni 2017

Sommersemester 2017 · Computervision Group · FSU Jena

## Zusammenfassung

Im Rahmen des Seminars Rechnersehen der Friedrich-Schiller-Universität entsteht dieser Text als Ausarbeitung des Themas *3D-Deskriptoren für Aufgaben der Rekonstruktion und Objekterkennung* und soll zusammen mit dem Vortrag einen Einblick in den Aufbau und die Einsatzbereiche von 3D-Deskriptoren geben. Abschnitt 1 beleuchtet das Konzept hinter dem Begriff 3D-Deskriptor, erklärt das für den Themenbereich notwendige Vokabular und liefert Einblick in die Einsatzbereiche. Dann folgt in 2 ein Überblick über den Entwicklungsprozess und die Bewertung von Deskriptoren. In 3 wird auf die in Tombari et. al. [1] entwickelte Methode (SHOT) eingegangen. Hier sollen die Unterschiede zu den vorherigen Methoden herausgearbeitet werden, wohingegen dann in ?? der Fokus ganz auf den Komponenten der **SHOT**-Methode liegen soll.

## 1 Einführung

**3D-Deskriptor** Ein 3D-Deskriptor ist eine Methode um Punkte  $p$ , welche zu einer Fläche im  $\mathbb{R}^3$  gehören, anhand der Punkte in Ihrer Umgebung zu beschreiben. Diese Beschreibung soll dann der Identifikation dienen, sodass man die Beschreibung eines Punktes mit einer anderen Beschreibung vergleichen kann. So ist es möglich, aus beschriebenen Punkten zusammengesetzte Flächen miteinander zu vergleichen und den Grad der Ähnlichkeit von Objekten festzustellen. Das Problem des Oberflächenvergleichs ist allgemein als **Surface-Matching** bezeichnet.

Die Eingabe für einen Deskriptor ist ein Teil einer Punktwolke, zum einen der Punkt, welcher später als Merkmal dienen soll, sowie die Punkte in seinem Umfeld, ab jetzt als **Nachbarschaft** bezeichnet. Um mehrere Features aus einem Bild zu extrahieren werden die Deskriptoren für viele Punkte im Bild berechnet. Diese können zum Beispiel zufällig gewählt sein [1] oder durch die Berechnung und Auswertung von Gradientenbildern an interessante Punkte im Bild, wie zum Beispiel Ecken gelegt werden [2]. Der Auswahlprozess wird als **Feature Selection** bezeichnet.

### 1.1 Einsatzbereiche

Die Einsatzbereiche sind sehr vielseitig, Beispiele sind die Lokalisation von Robotern anhand der Auswertung von Daten aus den an Ihnen angebrachten optischen Sensoren [3], sowie die Konstruktion von 3D-Modellen aus zweidimensionalen Bildern [4]. Desweiteren verbessert der Einsatz von dreidimensionalen Abbildern und Methoden im Bereich der Biometrie die Erkennungsrate signifikant, da Sie robuster gegenüber Variation in

Objektposition und Beleuchtung sind [5]. Bereits vor dem Erkennen von Objekten 1.2.3 wurden Deskriptoren für die Detektion von Bewegungen durch das Erkennen von Punktverschiebungen im sogenannten Motion-Tracking eingesetzt.

## **1.2 Struktur**

Nach [1] kann man die Art der Betrachtung der Nachbarschaft in zwei Gruppen einteilen. Der Unterschied besteht in der Form des Nachbarschaftsbereichs und in der Berechnung, die auf ihn angewandt wird.

### **1.2.1 Signatur**

signatures are potentially highly descriptive thanks to the use of spatially well localized information

The first category, that includes earliest works on the subject, describes the 3D surface neighborhood of a given point (hereinafter support) by defining an invariant local Reference Frame (RF) and encoding, according to the local coordinates, one or more geometric measurements computed individually on each point of a subset of the support

### **1.2.2 Histogramm**

Wird das Histogramm als Beschreibungsmethode gewählt, so wird eine diskrete Anzahl an Merkmalen benötigt, welche dann durch Auszählung der Ausprägungen als Wahrscheinlichkeitsdichtefunktion verwendet wird.

Die Reduktion des Umfeldes auf ein zählbares Merkmal

## **Surface-Matching**

### **1.2.3 Objekterkennung**

### **1.2.4 Rekonstruktion**

Die Rekonstruktion von

## **2 Was bisher geschah**

Es wird ein Überblick über die bisherige Entwicklung von 3D-Deskriptoren gegeben um unterschiedliche Ansätze aufzuzeigen und eine Reihe an Parametern definiert, anhand denen eine Bewertung stattfinden kann.

### **2.1 Probleme**

In der Vergangenheit wurde versucht, dieses Problem durch eine Beschreibung der gesamten betrachteten Oberfläche zu lösen.

Schon lange ist klar, dass die Betrachtung von einzelnen Punkten innerhalb einer Oberfläche

## 2.2 Bewertung

Die Güte eines solchen Deskriptors ist daran zu messen wie gut er folgende Eigenschaften erfüllt, bzw. wie gut er mit dem Auftreten folgender Effekte umgeht:

**Effizienter Vergleich** Der Vergleich von Oberflächen soll möglichst wenig Rechenaufwand erfordern, in hochaufgelösten Daten muss hochfrequent verglichen werden um im Optimalfall Verarbeitung in Echtzeit zu ermöglichen.

**Invarianz** Der Deskriptor soll Punkte und Ihre Umgebung auch dann noch als korrespondierend identifizieren, wenn sie Translation oder Rotation, also geometrischen Transformationen, ausgesetzt wurden.

**Rauschen** Der Einfluss von unterschiedlichen Rauscharten soll möglichst gering sein.

**Beleuchtung** Ein Wechsel der Beleuchtung, bzw. die Verschiebung eines Objektes relativ zu einer Lichtquelle soll das Ergebnis nicht verändern.

**Clutter** In realen Szenen sind die zu findenden Objekte meist nicht alleine, viele andere Objekte befinden sich im Bild und müssen ausgeschlossen werden.

**Verdeckung/Occlusion** Wenn andere Objekte im Bild sind kann es auch zu Verdeckungen kommen. Nicht alle bekannte Punkte eines Objekts sind erkennbar, anhand von unvollständigen Informationen muss eine Klassifizierungsentscheidung möglich sein.

## 2.3 Bisherige Ansätze

Nun soll eine Auswahl der in [1] verglichenen und erprobten Ansätze kurz vorgestellt werden.

**Spin Images [4]** Die Methode der Spin-Images wurde 1999 von Andrew Johnson und Martial Herbert vorgestellt. Sie wird in die Klasse der histogrammbasierten Methoden eingeordnet. Johnson und Herbert berechnen für ausgewählte Punkte die Oberflächennormale, indem Sie eine Ebene an die Nachbarschaft des Punktes einpassen. Die Nachbarschaft definieren sie als alle Punkte, welche durch ein Oberflächennetz verbunden direkte Nachbarn sind. Nun spannen sie einen Zylinder auf, dessen Radius die Distanz zwischen Punkt und Flächennormale ist, und dessen Höhe durch die vorzeichenbehaftete Distanz zwischen Punkt und Ebene bestimmt wird.

## 3 Fortschritt

Welche Probleme sehen Tombari et. al, wie lösen Sie diese

- Einfluss des Reference-Frames auf den Deskriptor
  - Ein oder mehrere Reference-Frames (uniqueness/ambiguity)
  - Ein guter Reference-Frame durch eigenvalue-decomposition und das Erzeugen von Eindeutigkeit durch Berechnung eines Vorzeichens.

## 4 Signaturen von Histogrammen SHOT

?? Tombari et. al. sehe

- Von 2D nach 3D: Warum funktioniert SIFT und wie kann man diese Prinzipien auf dreidimensionale Probleme anwenden
- Trotz Ableitungen robust gegenüber Rauschen?
- Kombination von Histogrammen über Winkel zwischen Punkt- und Umfeldnormalenvektoren als Deskriptor an der Schnittstelle zwischen Signatur und Histogramm
- Wahl des Koordinatensystems für die Signatur
- Robustheit gegenüber Schwankungen in der Punktdichte

## 5 Validierung im Experiment

- Wie wird Wiederholbarkeit garantiert?
- Wie werden faire Wettbewerbsbedingungen garantiert?
- Ist SHOT besser als die gewählten Gegner?
- Wenn ja, wie viel in welchen Bereichen.

## Literatur

- [1] Luigi Di Stefano Federico Tombari, Samuele Salti. Unique signatures of histograms for local surface description. 2010.
- [2] Mike Stephens Chris Harris. A combined corner and edge detector. 1988.
- [3] Se et al. Mobile robot localization and mapping with uncertainty using scale-invariant visual landmarks. 2002.
- [4] Martial Hebert Andrew E. Johnson. Using spin images for efficient object recognition in cluttered 3d scenes. 1999.
- [5] Licesio J. Rodríguez-Aragón Cristina Conde and Enrique Cabello. Automatic 3d face feature points extraction with spin images.