

1 Einleitung

Objekte in Bildern zu erkennen, ist für einen Menschen eine komplexe Aufgabe, die er auf Grund seiner Erfahrung und seines Wissens über den Kontext des Bildes intuitiv bewältigt. Wie stark der Kontext dabei die Wahrnehmung beeinflusst, ist an Abbildung 1.1 sichtbar.

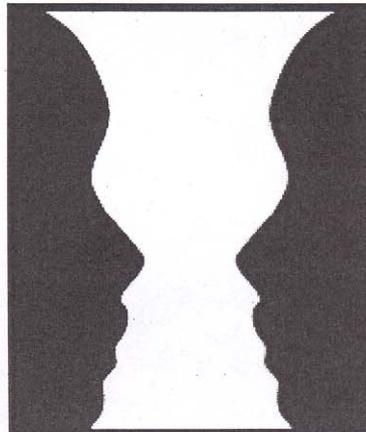


Abbildung 1.1: Kippbild

Zu erkennen sind entweder zwei dunkle Gesichter, die sich gegenseitig anschauen, oder die Konturen einer hellen Vase vor einem dunklen Hintergrund. Solche „optischen Täuschungen“ zeigen, dass Menschen trotz vieler Erfahrungen Objekte nicht immer ohne Schwierigkeiten erkennen können.

1.1 Motivation

Die computergestützte Objekterkennung versucht, den menschlichen Prozess der Objekterkennung mittels Algorithmen möglichst gut zu automatisieren. Die Informationen, die den verwendeten Algorithmen zu Grunde liegen, sind Pixel. Die Pixel beschreiben die Lichtintensitäten des aufgenommenen Bildes.

Ausgehend davon, dass die zu detektierenden Objekte sich durch Intensitäten von der Umgebung abheben, können wir Algorithmen zur Beschreibung der im Bild vorhandenen Objekte und dem Detektieren der gelernten Objekte in Bildern entwickeln.

Typische Anwendungsfälle für den Einsatz von Detektoren und Deskriptoren sind dabei die Kameraorientierung für 3D-Rekonstruktionen als auch Bildinterpretationsaufgaben wie Objektdetektion oder -erkennung.

Im Bereich der Objektdetektion bieten sich vielseitige Anwendungsfälle, wie z.B. die automatische Erkennung von Handschriften, Personenkontrollsysteme, automatische Erkennung von Tumoren und optische Qualitätskontrolle. Der Mensch ist meist in der Lage, Objekte auch dann zu erkennen, wenn sie durch Bildtransformationen, wie z.B. Änderung des Betrachtungswinkels, der Drehung und des Maßstabes, der Bildschärfe und der Belichtung verändert wurden. Die Untersuchung, in wie weit die automatische Detektion sich invariant gegenüber diesen Veränderungen verhält, ist Bestandteil dieser Arbeit.

1.2 Aufgabenstellung

Es existiert eine Vielzahl von Detektoren und Deskriptoren und Kombinationen, die sich aus beiden ergeben. Ziel dieser Arbeit ist es, eine Einordnung des neu entwickelten *Scale-invariant Feature Operator* [7] (Abk. SFOP) Detektors vorzunehmen. Dazu wird er im Rahmen der von Mikolajczyk et al. [16] vorgestellten Testumgebung mit etablierten Detektoren auf unterschiedlichen Bilderserien auf seine Sensitivität hin beurteilt. Ebenso werden die auf der Grundlage der Detektoren berechneten Deskriptoren in einer weiteren Testumgebung von Mikolajczyk und Schmid [14] beurteilt.

In einem weiteren Experiment wird der SFOP Detektor zur Objektdetektion mithilfe eines *Implicit Shape Models* genutzt. Anschließend erfolgt eine Beurteilung der Leistungsfähigkeit durch den Vergleich mit den Ergebnissen anderer Detektoren.

1.3 Verwandte Arbeiten

Die Beurteilung von Detektoren wurde in den Arbeiten von [11] [16] [3] versucht zu bewerten. Bei allen diesen Ansätzen wird die Sensitivität anhand der Replizierbarkeit gemessen.

Ansätze zur Bewertung der Sensitivität von Deskriptoren gibt es nur wenige. Eine Einordnung der Deskriptoren bezüglich der Verwendung in Zuordnungsverfahren (engl. Matching) und zur Objektdetektion bietet Carneiro und Jepson [4], dabei wird die *Receiver Operating Characteristics* des Deskriptors bewertet. Einen anderen Ansatz verwendet Mikolajczyk et al. [16], hier werden die Deskriptoren mithilfe der Anzahl richtiger und falscher Zuordnungen

zwischen zwei Bildern bewertet. Diesen Ansatz verwenden wir auch in unserer Arbeit. Im Kontext der *Texture Classification* bestehen Vorschläge von [18] [21] [19].

Es gibt viele Ansätze zur Objektdetektion, neuere Methoden verwenden dabei die sogenannte *top-down segmentation* [23] [5]. Dazu benötigen wir allerdings vorab Informationen darüber, ob ein Objekt im Bild ist und welche Größe es hat.

Der Ansatz von Borenstein und Ullman [2] ist dadurch eingeschränkt, dass nur ein Objekt in dem jeweiligen Bild auftreten darf.

Besonders flexibel ist der Ansatz von Leibe et al. [8]. Dabei werden die Detektion und die Segmentierung in einen gemeinsamen Ablauf zusammengefasst. Das Training erfolgt dabei automatisch durch ein Bilderset der zu detektierenden Klasse. Dieser Ansatz ist sehr flexibel und benötigt im Gegensatz zu Methoden von [17] [20] [22] weniger Trainingsdaten.

1.4 Aufbau der Arbeit

In Kapitel 2 beschreiben wir die Grundlagen der verwendeten Detektoren und Deskriptoren. Zusätzlich stellen wir die Bewertungsmaße für die Sensitivitätsanalyse vor und geben einen Einblick in die Objekterkennung mit Hilfe eines Implicit Shape Models.

In Kapitel 3 untersuchen wir, wie sich Deskriptoren und Detektoren gegenüber Bildtransformationen verhalten und wie gut die Leistungsfähigkeit des SFOP Detektors für die Objektdetektion ist.

Kapitel 4 bietet eine Zusammenfassung der Ergebnisse und einen Ausblick auf mögliche weiterführende Aufgabenstellungen.