

Institut für Visualisierung und Interaktive Systeme

Universität Stuttgart
Universitätsstraße 38
D-70569 Stuttgart

Bachelorarbeit Nr. 094

Kamera-basierte Oberflächenanalyse zu Erkennung von geeigneten Projektionsflächen

Paul Brombosch

Studiengang: Softwaretechnik
Prüfer/in: Prof. Dr. Albrecht Schmidt
Betreuer/in: Dipl.-Inf. Markus Funk

Beginn am: 4. November 2013

Beendet am: 6. Mai 2014

CR-Nummer: H.5.2, H.5.m

Kurzfassung

Die visuelle Aufbereitung von Informationen begleitet uns durch unseren Alltag, dabei werden vermehrt Projektoren eingesetzt, beispielsweise an U-Bahnhaltestellen. Früher waren Projektoren auf Grund ihrer Größe, Kosten und dem Einsatz in abgedunkelten Räumen nur beruflichen oder öffentlichen Organisation vorbehalten. Durch die Weiterentwicklung der verbauten Technologie passen solche Geräte bereits in Hosentaschen oder sind in Handys verbaut, damit eignen sie sich auch für den privaten Gebrauch. Ein großes Manko dieser Projektoren bleiben die statisch, definierten Projektionsflächen in qualitativer Abhängigkeit zum Hintergrund. Aus diesem Grund werden in der vorliegenden Arbeit Flächen auf Projektionen untersucht. Hierfür wird zunächst ein System zur Erfassung von Gegenständen, Freiräumen und Metadaten der Oberflächen entwickelt. Dabei kommt eine duale Abtastung der Fläche durch eine RGB-Kamera und einen 3D-Tiefensensor zum Einsatz. Mittels einer Benutzerstudie werden verschiedene Oberflächen auf ihre Projektionsfähigkeit hin eingestuft. Durch diese Informationen in Verbindung mit den Metadaten der Oberflächen wird ein Modell, zur Echtzeitklassifizierung, entworfen. Ziel dieser Arbeit ist die Klassifizierung von Oberflächen in Abhängigkeit von ihrer Projizierbarkeit.

Abstract

The visual presentation of information accompanies us through our daily lives, although projectors are increasingly used, for example at subway stations. In the past, projectors were reserved for professional or public organization, because of their size, cost and use in darkened rooms. Through the further development of the technology such devices already fit in pockets or are installed in mobile phones, so they are also suitable for home use. A major shortcoming of this projector are the static, defined projection in qualitative dependence to the background. For this reason, surfaces are tested on projections in the present work. For this purpose, a system for detection of objects, open spaces and metadata of the surfaces were developed. The scanning of the surfaces used a RGB camera and a 3D depth sensor. By means of a user study different surfaces are graded on their ability towards projection. With this informations and the metadata of the surfaces a model for classification were design. The aim of this work is to develop a system for the classification of surfaces, depending on their projectability.

Inhaltsverzeichnis

1. Einleitung	9
1.1. Zielsetzung der Arbeit	10
1.2. Aufbau der Arbeit	10
2. Grundlagen	11
2.1. Verwandte Arbeiten	11
2.1.1. Grundlagen	11
2.1.2. Statische Projektion	12
2.1.3. Mobile Projektion	14
2.1.4. Fazit und Abgrenzung	15
2.2. Technische Grundlagen	15
2.2.1. Rahmenbedingungen	15
2.2.2. Microsoft Kinect	16
2.2.3. OpenCV	18
3. Beschreibung des Prototypen	19
3.1. Begrenzung der Arbeitsfläche	19
3.2. Erfassung von Gegenständen	20
3.2.1. Canny-Algorithmus	21
3.3. Freie Flächen	21
3.3.1. Algorithmus zur Flächenberechnung	21
3.4. Projektionsflächen auf Gegenständen	22
3.5. Wiedererkennung von Gegenständen	23
3.5.1. SURF Algorithmus	23
3.6. Rotationsänderung von Gegenständen	26
3.7. Erfassung der Oberflächenmerkmale	27
3.7.1. RGB-Kamera	27
3.7.2. Tiefenkamera	28
4. Evaluation	29
4.1. Benutzerstudie	29
4.1.1. Design	29
4.1.2. Teilnehmer	29
4.1.3. Apparaturen	30
4.1.4. Vorgehen	31
4.2. Ergebnisse	33
4.2.1. Auswertung der Fragebögen	33

4.2.2. Erfasste Daten	35
4.2.3. Auswertung der Daten	37
4.3. Diskussion	37
5. Fazit und Ausblick	39
5.1. Fazit	39
5.2. Ausblick und weitere Arbeiten	39
A. Appendix	41
A.1. Reihenfolge der Projektionsobjekte	41
A.2. Fragebogen	42
Literaturverzeichnis	45

Abbildungsverzeichnis

2.1. motionEAP Montagetisch	16
2.2. Kinect-Bestandteile	17
3.1. Ansicht der GUI während der Objekterfassung	20
3.2. Ansicht der GUI während der Freiflächenberechnung	22
3.3. Beispiel für ein Summed Area Table	24
3.4. Bildpyramide	25
3.5. SURF-Deskriptor	26
3.6. Beispiel einer SURF Anwendung	27
4.1. Skizzen der Projektionsobjekts	30
4.2. Gesamtbild aller Projektionsobjekte	32
4.3. Gesamtüberblick der Studienergebnisse	33
4.4. Bewertung der Strukturen	34
4.5. Bewertung auf Basis der Formen	35
4.6. Metadaten der Struktur	36
4.7. Metadaten der Form	36

Tabellenverzeichnis

4.1. Studienergebnisse	33
----------------------------------	----

1. Einleitung

Die visuelle Aufbereitung von Informationen begleitet uns durch unseren Alltag, angefangen von der Uhrzeit unseres Weckers, über LED-Tafeln an Haltestellen, Ampeln im Straßenverkehr, endend mit dem gemütlichen Fernsehabend. So werden klassischen Abfahrtstafel mittlerweile digital aufbereitet oder der Nummernblock an Fahrscheinautomaten durch Touch-Displays ersetzt.

Unter anderem setzten sich Videoprojektoren, kurz Beamer, mehr und mehr durch. In früheren Zeiten begrenzte sich der Einsatz von Projektoren auf Kinos oder Meetingräume. Doch durch die Weiterentwicklung der Technologie in den Gerät, könnten sie in Alltagssituationen vermehrt anklang finden, beispielsweise werden Beamer an U-Bahnhaltestellen eingesetzt um aktuelle Nachrichten wiederzugeben.

Wichtige Faktoren, die das steigende Interesse an Projektoren auszeichnen, sind die großen Bildflächen, ihre steigende Auflösung, die immer kleiner werdenden Geräte und der angemessene Preis. Auch das Abdunkeln von Räumlichkeiten kann durch die heutige Belichtungsstärke der Beamer vernachlässigt werden.

Diese Eigenschaften führen auch dazu, dass Projektoren im privaten Umfeld eingesetzt werden, dies wird verdeutlicht durch das Erscheinen erster Handys mit integrierten Projektoren. Auch in der Arbeitswelt ist diese Entwicklung zusehen, so werden beispielsweise Informationen in das Blickfeld der Arbeitskräfte projiziert, dies kann sowohl in der Fertigung als auch in der Büroumgebung der Fall sein. Hierbei wird für gewöhnlich ein top-mounted Projektor oberhalb eines Arbeitsplatzes platziert, der den Arbeitsplatz von oben mit Informationen bestrahlt. Auch hier zeigen sich die Vorteile von Projektoren gegenüber Displays, so kann die Bildfläche nicht verkratzen oder Informationen werden nicht durch herumliegende Gegenstände verdeckt.

Eine Erweiterung der Informationswiedergabe ist die Interaktion mit diesen, so bestehen Systeme, die die Projektionsfläche zusätzlich als Touchfläche nutzen können. Durch die Mobilität der Beamer sind solche Anwendungen auf jeder Oberfläche möglich, hierbei muss es sich nicht mal um planare Flächen handeln. Momentan werden Projektionen an zuvor definierten Plätzen projiziert. In der Realität kann zum Beispiel ein Arbeitsplatz aber sehr unordentlich sein und von Tag zu Tag anders aussehen. Gegenstände können sich auf den definierten Projektionsflächen befinden und somit die Darstellung beeinträchtigen.

Aus Nutzersicht wäre es wünschenswert diese Problematik automatisiert zu umgehen. Abhängig von der momentan Situation auf der Arbeitsfläche sollten die Informationen auf Flächen projiziert werden, die dafür am Besten geeignet sind. Beispielsweise könnte das System die Arbeitsfläche abtasten um eine nicht bedeckte Fläche zu finden oder auf Gegenstände projizieren, die sich dafür besonders gut eignen. Alternativ sollte der Nutzer die Projektionsfläche fest an einen Gegenstand binden können, so dass der Gegenstand als bewegliche Bildfläche genutzt werden kann.

1.1. Zielsetzung der Arbeit

Das Ziel der vorliegenden Arbeit ist die Entwicklung eines System zur automatischen Lokalisierung und Klassifizierung von Oberflächen in Abhängigkeit von ihrer Projizierbarkeit zu entwickeln. Dabei werden auf die Daten einer RGB-Kamera und einer Tiefenkamera zurückgegriffen. Bei der Lokalisierung wird auf eine duale Abtastung der Fläche durch beide Kameras gebaut. Im Bereich der Klassifizierung decken die RGB-Daten primär die Erfassung der Oberflächenstruktur ab, wie zum Beispiel Farben und Muster. Wiederrum erfasst der Tiefensensor die Form der Oberfläche, dies beinhaltet Unebenheiten und Konturen. Für die Realisierung werden empirische Daten mittels einer Benutzerstudie gesammelt, hierbei wird ermittelt welche Oberflächen für Projektionen bevorzugt werden. An Hand der gesammelten Daten wird ein Modell erstellt, welches die Klassifizierung von Oberflächen in Echtzeit ermöglicht. Abschließend wird dieses Modell evaluiert.

Als Nebenprodukt zeigt sich das Erfassen von Gegenständen. Daraus bilden sich Teilziele, so wird an Hand der Position und der Grundfläche des Gegenstands die größte, rechtwinklige, freie Fläche berechnet. Des weiteren verfolgt diese Arbeit das Ziel Gegenstände wieder zu erkennen, hierbei werden optische Merkmale von bereits bekannten Gegenständen verglichen. Das letztes Ziel dieser Arbeit ist die Bestimmung von waagerechten Drehungen eines Gegenstandes, auch in diesem Fall wird von einem Ausgangszustand ausgegangen um anschließend die Veränderung zu vergleichen und zu berechnen.

1.2. Aufbau der Arbeit

Die vorliegende Arbeit beginnt in Kapitel 2 mit einem Überblick über verwandte Arbeiten der Bereiche Objekterkennung, Projektionen und Interaktionen mit Oberflächen. Ebenso wird in diesem Kapitel eine Einblick in die technischen Grundlagen verschafft.

Im Kapitel 3 wird auf die Entwicklung des Prototypen eingegangen. Hierbei werden der Funktionsumfang, die Implementierung und die verwendeten Algorithmen beschrieben, beginnend mit der Objekterfassung über die Freiflächenberechnung und Wiedererkennung bis hin zur Oberflächenanalyse.

In Verbindung mit den, vom Prototypen validieren Metadaten, wurden ein Benutzerstudie durchgeführt, bei der die subjektive Bewertung zur Qualität von Projektionen auf unterschiedlichen Oberflächen erarbeitet wurde. Des weiteren wird im Kapitel 4 auch auf die Evaluation der Daten eingegangen. Ebenso wird hier die Entwicklung und Überprüfung eines Modells zur Klassifizierung der Oberflächen beschrieben.

Die Arbeit endet im letzten Kapitel mit dem Fazit der kompletten Arbeit sowie einem Ausblick auf mögliche Verfeinerungen und Weiterentwicklungen.

2. Grundlagen

In diesem Kapitel werden Einblicke über verwandte Arbeiten und Technische Grundlagen gegeben. Die hierbei gewonnen Erkenntnisse dienen der gesamten Arbeit sowohl als Grundgerüst wie auch als äußerer Rahmen.

2.1. Verwandte Arbeiten

In den folgenden Abschnitten werden bereits existierende wissenschaftliche Arbeiten aufgeführt, die in drei Bereiche gegliedert werden. Der erste Bereich spiegelt die grundlegenden Arbeiten wieder, welche in gewisser Form in allen weiteren Arbeiten enthalten sind. Anschließend werden die weiteren Arbeiten in fest montierte Apparaturen und tragbare Systeme unterteilt. Im Fazit wird ein Gesamtüberblick über die vorgestellten Arbeiten sowie die Unterscheidung zur vorliegenden Arbeit gegeben.

2.1.1. Grundlagen

Als erste grundlegende Arbeit ist das „Everywhere Display Projector“-Projekt von Claudio Pinhanez[Pin01] zu erwähnen. Diese Entwicklung ermöglichte es Räume mit nur einem montierten Instrument als Projektionsflächen zu nutzen. Durch die Kombination eines Beamers und eines drehbaren Spiegels würde eine Abdeckung von 60° vertikal und 230° horizontal erreicht. Als Projektionsflächen könnten sämtliche planare Flächen genutzt werden, welche vom Benutzer definiert werden mussten. Zur Behebung der Verzerrung projiziert das System ein Muster auf die Fläche und passt anschließend das Bild an. Dem System war es somit möglich dynamisch die Projektionsfläche zu wechseln.

Pierre Wellner präsentiert in seiner Arbeit „The DigitalDesk Calculator: Tangible Manipulation on a Desk Top Display“[Wel91] eine Taschenrechneranwendung die mittels Projektion auf einer Arbeitsfläche verwendet werden konnte. Durch die Integration einer Kamera wurden Interaktionen mit Fingern ermöglicht. Einerseits könnten Zahlen mittels Nummernblock eingegeben werden, doch auch das Zeigen auf Zahlen auf Dokumenten wurde akzeptiert. Beispielsweise reichte es aus auf den Betrag einer Rechnung zu zeigen und das System erfasste den vor dem Finger befindlichen Bereich und übernahm diese Zahl. Dies ermöglicht die Fusion von herkömmlichen und elektronischen Dokumenten.

Bereits im Jahre 2005 veröffentlicht Andrew D. Wilson unter dem Titel „PlayAnywhere: A Compact Interactive Tabletop Projection-Vision System“[Wil05] ein interaktives Tischsystem. Der Prototyp bestand aus einem IR-Strahler, einer Kamera mit IR Pass Filter und einem Projektor. Hierbei wurden

2. Grundlagen

Touch Events durch den vom Finger erzeugten Schatten berechnet. Des Weiteren wurde Bildbearbeitungstechnik für die Manipulation der Projektionsbilder und zur kontinuierlichen Verfolgung von Papierblättern eingesetzt. Somit ermöglichte er eine bewegliche Projektionsfläche.

In Verbindung mit der Manipulation von Projektionsbildern steht auch die Arbeit von Bimber et. al., so wird im Artikel „The Visual Computing of Projector-Camera Systems“ [BIWG07] beschrieben wie Projektionen auf geometrisch komplexen, farbigen und strukturierten Oberflächen je nach Hintergrund angepasst werden. Unter anderem ist dies auf eine ständig Abtastung der Projektion durch eine Kamera zurückzuführen, dabei gemessene Lichtverhältnisse werden mit dem Originalbild verglichen und angepasst.

Andrew D. Wilson präsentiert 2010 erneut eine Touch Anwendung, diesmal allerdings auf Grundlage von Tiefendaten [Wil10]. Somit stellte er einen neuartigen Algorithmus vor, mit dem jede Oberfläche touch-sensitive werden kann. Dabei spielen Unebenheiten oder die Lage der Oberfläche keine Rolle mehr, diese können sich sogar an der Decke befinden. Das Touch Event wird durch einen Offset oberhalb einer Ebene erzeugt, sobald ein Finger in diesen Wertebereich eintaucht wird dies als Interaktion wahr genommen. Abschließend erwähnter auch die Grenzen solcher Anwendungen, die in der Auflösung der Tiefensensoren, der benötigten freien Sichtlinie und der exakten Bestimmung des Berührungspunktes liegen.

Als letzte, in diesem Zusammenhang, grundlegende Arbeit im Umgang mit Tiefendaten ist der Artikel „KinectFusion: Real-time 3D Reconstruction and Interaction Using a Moving Depth Camera“ [NIH⁺11] zu erwähnen. Schwerpunkt dieser Arbeit liegt in der 3D-Rekonstruktion von Objekt und ganzen Innenraumszene. Hierbei werden nur die Tiefendaten des Kinect-Sensors verwendet. Ein weiterer Aspekt ist die interaktive Simulation der rekonstruierten 3D-Modelle, so können digitale Bälle auf das Modell geworfen werden und diese zeigen ein realitätsnahes Verhalten auf den Oberflächen des Modells.

2.1.2. Statische Projektion

Erneut war es Andrew D. Wilson in Zusammenarbeit mit Hrvoje Benko der 2010 eine tiefense-norbasierte Anwendung unter dem Titel „Combining Multiple Depth Cameras and Projectors for Interactions On, Above, and Between Surfaces“ [WB10] präsentierte. Diese Anwendung ermöglicht Interaktionen zwischen zwei getrennten Projektionen, so können zum Beispiel Informationen von einem Display in die Handfläche gelegt werden und durch Berührung des anderen Display wiederum dort angezeigt werden. Ähnliche Interaktionen sind auch durch das zeitgleiche Berühren beider Displays möglich. Für die Berechnung einer solchen Anwendung müssen Tiefendaten sowohl von beiden Projektionsflächen als auch dem Bereich zwischen ihnen erfasst werden.

Im Jahr 2012 waren es John Hardy und Jason Alexander die unter dem Titel „Toolkit Support for Interactive Projected Displays“ [HA12] ein umfangreiches Toolkit veröffentlichten. Dieses Toolkit ermöglichte es komplette multi-touch interaktive Displays auf allen beliebigen Oberflächen zu erzeugen und das alles nur durch eine Kinect und einen Beamer. Das Toolkit zeichnet sich durch seine einfache Kalibrierung der Oberflächen und das Einbinden von Interfaces dieser Flächen aus. Auch die Arbeiten von Klompaker et. al. [KNF12], Funk et. al. [FKS14] sowie Echtler und Wimmer [EW13] basieren auch auf der Interpretation von Tiefendaten innerhalb einer definierten Fläche. „dSensingNI – A Framework

for Advanced Tangible Interaction using a Depth Camera“[KNF12] umfasst die Multitouch und Objekt-Interaktion. Durch die Tiefendaten werden Hände, Finger und Gegenstände erkannt, durch die Kombination dieser Erkenntnisse werden Aktionen wie das Greifen und Stapeln von Gegenständen wahrnehmen. Auch beim Tracking der Objekte werden diese Interaktionen berücksichtigt. Im Artikel „The Interactive Dining Table“[EW13] wird durch dreier Szenarien die Nützlichkeit von interaktiven Oberflächen im privat Haushalt verdeutlicht. So werden Brettspiel durch visuelle Reize ergänzt und Informationen wiedergeben, wie zum Beispiel während des Frühstücks die Wettervorhersage. Das letzte, erwähnte Konzept deckt eine Art Augmented-Restaurant ab, hierbei werden die Bestandteile der Speisen auf dem Tisch angezeigt. Hingegen beschreibt Funk et. al.[FKS14] die Möglichkeit der Interaktion durch Alltagsgegenstände, durch verschieben oder drehen beliebiger Gegenstände können beispielsweise Regler simuliert werden. Auch in der Arbeit „iCon: Utilizing Everyday Objects as Additional, Auxiliary and Instant Tabletop Controllers[CLC⁺10]“wird der Aspekt der Interaktion mit Gegenständen betrachtet. Doch im Unterschied zu [FKS14] lediglich durch die Verwendung von einer Kamera und Aufklebern mit unterschiedlichen Mustern. Drei Interaktionen werden dabei genauer betrachte, Klick, Rotation und Verschieben. Diese zwei Arbeiten zielen auf eine Erweiterung der Eingabegeräte durch beliebige Gegenstände, so lassen sich beispielsweise Gegenstände und ihre Aktionen bestimmten Anwendungen zuordnen.

Bei „LuminAR: Portable Robotic Augmented Reality Interface Design and Prototype[LM10]“handelt es sich um einen Prototyp für ein Projektor-Kamera-System, welches das traditionelle Glühbirnengewinde als Energiequelle nutzt. Somit ergänzt der Prototyp jeder Lampe zu einer interaktiven Schnittstelle. In der erwähnten Arbeit kommt ergänzend ein Roboterarm zum Einsatz. Mittels Multi-Touch und Gestensteuerung lassen sich das projizierte Interface, die Projektionsfläche so wieder der Arm steuern. Auch in der Arbeit von Molyneaux et.al [MG09] werden durch optische Merkmale Gegenstände erkannt und Interfaces aus sie projiziert. An Hand mehrerer Bilder eines Objekts wird ein 3D-Modell erstellt. Diese 3D-Modell ermöglicht die Wiedererkennung der Objekte und die damit verbundene Anwendung, beispielsweise das abrufen einer Projektion.

Auch in Freizeitaktivitäten kommen solchen Anwendungen vermehrt zum Einsatz, so wird beispielsweise im Artikel „PoolLiveAid: Augmented reality pool table to assist in experienced players [ASR13]“ein Billardtisch präsentiert. Durch die Ermittlung der Kugeln, die Ausrichtung Queues so wie deren Treffpunkt wird die Laufrichtungen der Kugeln auf den Tisch projiziert. An andres Beispiel kommt aus dem Bereich der Musik, so wird in der Arbeit von Jorgà et. al.[JGAK07] ein Tisch zur Erstellung von Musik gezeigt. Die Tischplatte wird von unten bestrahlt und abgescannt. Zur Erkennung der einzelnen Objekte werden RFID Marken verwendet. Durch die Position und Ausrichtung der Marken wird die Musik beeinflusst, das somit erstellte „Musikinstrument“beinhaltet sämtliche Möglichkeiten um live elektronische Musik zu erzeugen. Ein weiterer Aspekt im Bereich der Freizeitaktivitäten deckt der Beitrag „PingPongPlus: Design of an Athletic-Tangible Interface for Computer-Supported Cooperative Play“[IWO⁺99] von Ishii et. al. ab. Hierbei handelt es sich um eine Anwendung in Verbindung mit der Sportart Tischtennis. In Abhängigkeit zum erfassten Ball werden Sounds, Bilder und Animationen wiedergegeben. So wird zum Beispiel der Rhythmus der Sound von der Geschwindigkeit der Ballwechsel beeinflusst.

Die nächsten zwei Veröffentlichungen zeichnen sich durch die Modifikation ihrer projektionsbasierte Bildfläche aus. In „Flexpad: Highly Flexible Bending Interactions for Projected Handheld

2. Grundlagen

Displays“[SJM13] werden mittels Tiefendaten Oberflächenverformungen rekonstruierten und entsprechend Informationen präsentiert. So kann ein Blattpapier als Projektionsfläche dienen, Veränderungen der Höhe oder Krümmung dieser Oberfläche werden vom System erfasst und interpretiert. So kann ein durch die Veränderung der Höhen ein Zoom-Effekt auf der Projektion erzeugt werden. Die zweite Veröffentlichung, „FoldMe: Interacting with Double-sided Foldable Displays“[KLKS12], beschäftigt sich mit einem doppelseitigen Display. Durch Scharniere können Größe und Form dynamisch geändert werden, auch die beim Faltern erzeugten Rückseiten werden erkannt. Das System basiert auf dem „Optitrack Motion Capture“¹-System bestehend aus sechs Infrarot-Kameras und mehrere Infrarot-reflektierenden Markierungen auf dem faltbaren Prototypen. Durch das Tracking dieser Markierungen (Position, Orientierung, Klapp Stand der Prototyp) stehen Informationen zur Verfügung, die das Anpassen der projizierten Bilder auf den Prototyp in Echtzeit ermöglichen.

2.1.3. Mobile Projektion

Projektionen werden auch zur Erleichterung von Lernvorgängen verwendet, dies wird auch am Beispiel von „guitAR – Supporting Guitar Learning through Mobile Projection“[LGJK11] verdeutlicht. Da es sich bei der Gitarre ist eines der am weitesten verbreiteten Instrumente unter Autodidakten handelt, setzen Löchtersfeld et. al hier an. Bei ihrem Prototypen wurde ein Projektor am Gitarrenkopf befestigt, dieser projizierte Griffe für Akkorde oder ganze Lieder in einstellbarer Geschwindigkeiten auf die jeweiligen Saiten. Bei der Arbeit von Toshikazu Karitsuka and Kosuke Sato „Wearable Mixed Reality with an On-board Projector“[KS03] handelt es sich um eine der ersten tragbaren Anwendung zur Erweiterung der Realität auf Projektionsbasis. Zur Erkennung der Position der Projektionsflächen wurden reflektierende Marker eingesetzt. Jede mit diesen Markern ausgestattet Fläche dient als Projektionsfläche so lang sie sich in Sichtfeld der Anwendung befindet, zur Interaktion mit der Fläche kommt eine Kappe für den Finger mit einer IR-LED zum Einsatz. Bei dieser Arbeit sowie denen von Chris Harrison et. al[HBW11] und Winkler et. al.[WSDR14] handelt es sich um Lösungsansätze, bei dem sich der Projektor über der Schulter befindet. Durch diese Positionierung ist mindestens eine Hand für eine Interaktion zur Verfügung, wo bei die andere Hand in Verbindung mit der Projektionsfläche steht. Dadurch ermöglicht OmniTouch[HBW11] Multitouch-Interaktionen auch sämtlichen in Reichweite befindlichen Oberflächen, selbst auf dem eigenen Arm ist dies möglich. In Gegensatz dazu verfolgt der AMP-D-Prototyp[WSDR14] den Ansatz der Gestensteuerung und eine Informationswiedergabe in Abhängigkeit zur Umgebung. Zum Beispiel kann bei ein Navigation durch Pfeile auf dem Gehweg realisiert werden. Bei der Verwendung von tragbaren Projektoren muss der Anwendungskontext berücksichtigt werden, so wirken unterschiedliche Faktoren auf das projizierte Bild, wie zum Beispiel die Position des Projektor, der Präsentationsinhalt und ob der Anwender in Bewegung ist oder steht. Diese Problematik wird unter dem Titel „A Method for Wearable Projector Selection that Considers the Viewability of Projected Images“[OTTT10] von Ota et. al. behandelt, dabei gehen sie von am Körper befestigten Projektoren aus. Es stellt sich heraus, dass je auch Situation unterschiedliche Positionen geeignet sind. Daraus entwickelten sie einen Prototypen, bestehende aus drei Projektoren, der je nach Situation dynamisch den geeigneten Projektor zur Wiedergabe wählt.

Im weiteren wird auf Anwendungen basieren auf Pico-Projektoren² eingegangen. Im Artikel „Light-

¹<http://www.naturalpoint.com/optitrack/>

²<http://www.picoprojektor.com/>

Beam: Interacting with Augmented Real-World Objects in Pico Projections“[HSL⁺12]verfolgt man den Ansatz, Informationen auf Alltagsgegenständen wie Tassen oder ähnliches wieder zu geben. Die auch schon von Funk et. al. erwähnte Interaktion durch Gegenstände ist hier von Bestand. So wird in diesem Kontext die Rotation der Kaffeetasse durch einen RGB-Kamera registriert.

In Anbindung an Projektionen mit Pico Projektoren ist die Interaktion mit der Projektionsfläche recht umständlich, da man den Fokus von der Projektion auf das Gerät wenden muss. Mit dieser Problematik setzt sich der Artikel „ShadowPuppets: Supporting Collocated Interactionwith Mobile Projector Phones Using Hand Shadows[CL11]“auseinander. Um eine bessere Interaktion zu gewährleisten, wird dem Nutzer die Möglichkeit geboten durch Schattenspiele die Projektion zu beeinflussen. Dieser Schattenbilder lassen sich auch von nebenstehenden Nutzern erzeugen, wodurch eine Interaktionen mehrere Anwender möglich ist. Dies wird ebenfalls mittels einen RGB-Kamera registriert. Auch Willis et. al. befassten sich in „HideOut: Mobile Projector Interaction with Tangible Objects and Surface“[WSM13] mit Pico Projektoren. In ihrer Arbeit nutzten sie Infrarot Markierungen um Projektionen aufzurufen. Dies Markierungen sind für das bloße Auge nicht sichtbar. So wird zum Beispiel aus einem Kinderbuch ein interaktives Erlebnis.

2.1.4. Fazit und Abgrenzung

Im Gegensatz zu vorliegender Arbeit unterscheiden sich die verwandten Arbeiten dadurch, dass sie jede Oberfläche als ganzes sehen und nicht auf Teilflächen eingehen. Ebenso wird keine Untersuchung der Projektionstauglichkeit durchgeführt, statt dessen werden Projektionen umgehend angepasst. Sämtliche Arbeiten behandeln die Erfassung oder Projektion auf Objekte, doch nicht den Nebeneffekt der Objekterkennung nämlich die Berechnung des freien Raums. Somit lassen sich aus allen verwandten Arbeiten Teilaspekte gewinnen, doch in der Gesamtheit handelt es sich hierbei um eine separate Arbeit.

2.2. Technische Grundlagen

In diesem Unterkapitel werden die grundlegenden Bestandteile dieser Arbeit beschrieben, als zentrales Elemente ist der Kinect-Sensor zusehen. Darauf folgend wird auf die verwendete Bibliothek eingegangen.

2.2.1. Rahmenbedingungen

Diese Arbeit steht in Verbindung mit dem motionEAP³ Projekt der Universität Stuttgart. Das Projekt verfolgt das Ziel ein System zur Unterstützung und Effizienzsteigerung bei Produktionsprozessen in Unternehmen zu entwerfen. Das Assistenzsystem soll vor allem leistungsminderte Menschen prozessorientiert unterstützen und motivieren. Als Grundgerüst für den Aufbau des Assistenzsystems dient ein Aluminiumrahmen (sieh Abbildung 2.1), an dem eine Microsoft Kinect und ein Projektor

³<http://www.motioneap.de>

2. Grundlagen

befestigt sind. Diese Konstruktion wird auf einen beliebigen Tisch positioniert. Die Anbringung der Kinect und des Projektors lag in etwa bei einem Meter Höhe. Insgesamt standen drei dieser Montagetische zur Verfügung wobei die Höhe etwas variierte. Die vom Aluminiumrahmen umschlossene Fläche wird als Arbeitsfläche bezeichnet und spiegelt die überwachte und projizierbare Fläche wieder. Ein weiterer Bestandteil des motionEAP-Systems ist eine auf dem Tisch montierte Leap-Motion.

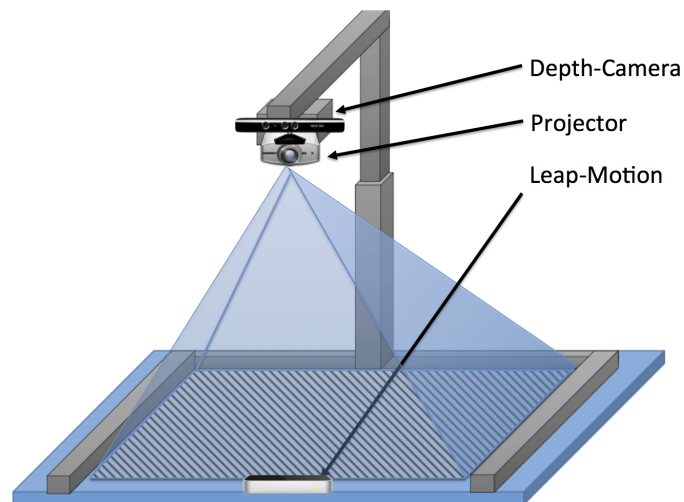


Abbildung 2.1.: motionEAP Montagetisch [FKS14]. Die Tiefenkamera und der Projektor sind in etwa einem Meter Höhe an einem Aluminiumrahmen befestigt.

2.2.2. Microsoft Kinect

Beim Kinect⁴ Sensor handelt es sich ursprünglich um ein Bedienelement der Videospielekonsole Xbox 360 von Microsoft. Kinect lässt sich aus dem englischen kinetic connect (deutsch Kinetische Verbindung) ableiten und wurde erstmals auf der E3-Spielemesse am 1. Juni 2009, unter dem Codenamen Project Natal, der Öffentlichkeit präsentiert. Hiermit war es nun möglich mit Hilfe der Sprache und durch Bewegungen des Körpers mit der Spielkonsole zu interagieren. Bereits Tage nach dem Verkaufsstart im November 2010 erschienen die ersten Treiber, welche das Auslesen der unterschiedlichen Sensoren am PC ermöglichten. Als Folge der ersten Treiber wurden von der Open Source Community auch ein Framework zur Verfügung gestellt, dieses erleichterte vielen Entwicklern den Einstieg. Microsoft zog im Juni 2011 nach und veröffentlichte das Kinect Software Development Kit für Windows 7. Im Februar 2012 erschien schließlich eine offizielle Version des Kinect Sensors für Windows. Bis Februar 2013 wurde der Kinect Sensor bereits 24 Millionen mal verkauft. Mittlerweile ist die zweite Generation des Sensors erhältlich.

⁴<http://www.microsoft.com/en-us/kinectforwindows/>

Technische Daten

Wesentliche Bestandteile der Kinect sind der PrimeSense-Tiefensensor⁵, 3D-Mikrofon und Farbkamera, wodurch eine Interaktion ohne herkömmliche Hardware möglich ist. Nachfolgend ein kleiner Blick ins Innere einer Kinect⁶.

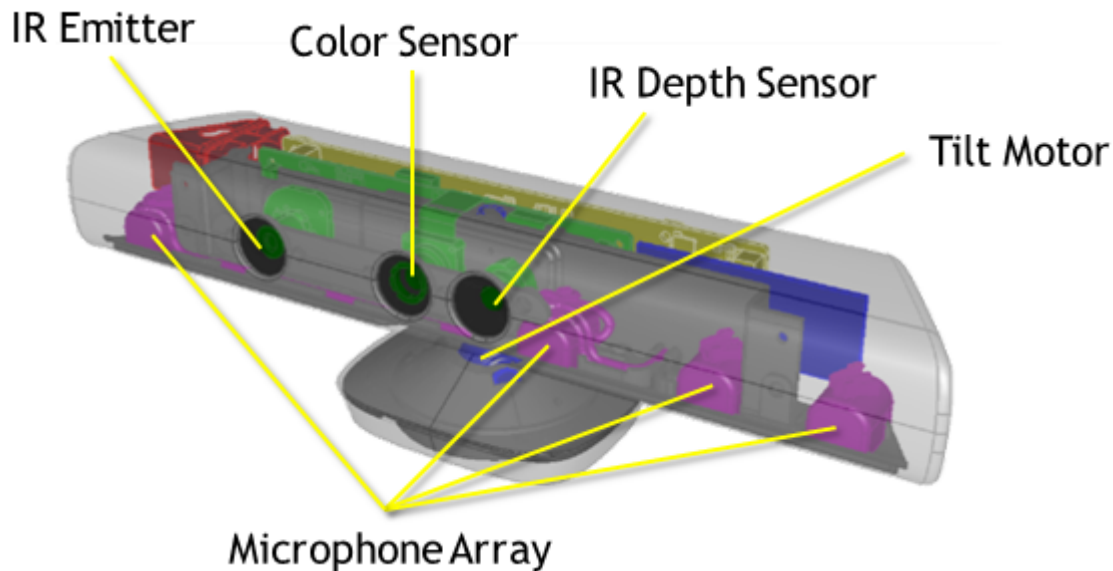


Abbildung 2.2.: Dies Abbildung zeigt das Innenleben einer Microsoft Kinect

Quelle: <http://msdn.microsoft.com/en-us/library/jj131033.aspx>

Tiefensensor Beim Tiefensensor handelt es sich um die größte und zentrale Innovation des Kinect. Dieser Sensor setzt sich aus einer Kombination eines Infrarot Laser Projektor und eines CMOS-Sensor zusammen. Diese Kombination liefert eine Auflösung 640x480 Pixels bei einer Bildrate von 30 FPS. Laut Microsoft werden Daten in einem Tiefenbereich von 0.4m - 4m erfasst, wobei im Bereich von 0.4m-3.5m bessere Ergebnisse erzielt werden. Dies wird allerdings in zwei Modi, dem Near Mode und dem Default Mode, unterteilt[Mir]. Das Tiefenbild verhält sich ähnlich wie ein Farbbild, dass zu jedem Pixel statt eines Farbwertes eine Tiefeninformation zur Verfügung steht.

Kamerasensor Auch die RGB-Kamera der Kinect liefert eine Auflösung 640x480 Pixeln. Sowohl RGB- als auch Tiefenkamera decken einen Sichtbereich von Horizontal 57,5° und Vertikal 43,5° ab.

Audiosensoren Die akustische Aufnahme erfolgt über 4 Mikrofonen, hiermit wird eine 3D-Lokalisierung ermöglicht.

⁵<http://www.freepatentsonline.com/20100118123.pdf>

⁶<http://msdn.microsoft.com/en-us/library/jj131033.aspx>

2. Grundlagen

Kippmotor Mit Hilfe des Kippmotors im Sockel der Kinect ist eine Veränderung des Neigungswinkels um -27° und 27° möglich, ausgehend von der 0° Stellung.

Kinect SDK

Beim Kinect SDK handelt es sich um das offizielle Entwicklungskit von Microsoft, mit dem der Zugriff auf sämtliche von der Kinect gelieferten Rohdaten wesentlich erleichtert wurde. Des Weiteren beinhaltet es Methoden, die es möglich machen, Gliedmaßen und Bewegungen nachzuverfolgen. Mittlerweile ist die Version 1.8 erhältlich.

2.2.3. OpenCV

Bei OpenCV⁷ handelt es sich um eine freie Programmbibliothek, welche unter den BSD-Lizenzbedingungen steht und somit kostenlos für akademische und kommerzielle Nutzungen einsetzbar ist. Das „CV“ im Namen steht für englisch "Computer Vision" und wurde 1999 als Forschungsprojekt von Intel gestartet. OpenCV beinhaltet Schnittstellen für C++, C, Python sowie JAVA und unterstützt Windows, Linux, Mac OS, iOS und Android. OpenCV[BK08] wurde mit einem starken Fokus auf Echtzeit-Anwendungen konzipiert. Die Bibliothek ist in der Lage, unter der Nutzung von C++ bzw. C, die Vorteile der Mehrkernprozessoren anzuwenden. Die Bibliothek verfügt über mehr als 2500 optimierte Algorithmen. Das Einsatzspektrum reicht von interaktiver Kunst bis hin zur fortschrittlichen Robotik.

EmguCV Da es sich beim motionEAP-Projekt (siehe 2.2.1) um eine Entwicklung in C# handelt, musste EmguCV⁸ zum Einsatz kommen. Bei EmguCV handelt es sich um einen plattformübergreifenden .Net Wrapper, wo mit OpenCV Funktionen in .Net-kompatiblen Sprachen wie C#, VB, VC++ und IronPython aufgerufen werden können.

⁷<http://opencv.org>

⁸http://www.emgu.com/wiki/index.php/Main_Page

3. Beschreibung des Prototypen

Das folgende Kapitel beschäftigt sich mit den grundlegenden Funktionen und ihrer Implementierung. Um Oberflächen analysieren zu können, müssen diese erst einmal erkannt werden. Daher setzt sich der erste Abschnitt mit der Bestimmung der Arbeitsfläche auseinander und anschließend Gegenstände innerhalb dieser Fläche zu erkennen. Da nun die Position der Gegenstände erkannt wurden, kann in Abschnitt 3.3 die größte, freie Fläche bestimmt werden. Im Anschluss daran wird auf die Projektion auf einem Gegenstand eingegangen. In Abschnitt 3.5 und 3.6 steht der SURF-Algorithmus im Mittelpunkt, hier wird die Wiedererkennung und Rotationsberechnung von Gegenständen erläutert. Im letzten Abschnitt des Kapitels wird auf die Oberflächen der Gegenstände eingegangen. Hierbei wird ein Einblick über die fundamentalen Charakteristiken der Oberflächen erlangt.

3.1. Begrenzung der Arbeitsfläche

Diese Funktion gilt als grundlegend für alle weiteren Interaktionen. Um dem weiteren System nur relevante Informationen zu liefern, ist zu Beginn eine Kalibrierung nötig. In diesem Schritt werden die Bildbereiche der Tiefenkamera und Farbkamera an die Arbeitsfläche angepasst, hierfür werden die Koordinaten zweier Eckpunkte (oben-links und unten-rechts) verwendet. Durch diese Punkte wird ein Rechteck erzeugt, alle außerhalb dieser Fläche liegenden Punkte werden verworfen (siehe Abbildung 3.1). Zusätzlich wird für den Tiefensensor die maximale Tiefe definiert, so dass Tiefeninformation, die von der Arbeitsfläche erzeugt werden, vernachlässigt werden. Im selben Schritt werden nach der Akzeptierung der innen-liegenden Pixel, diese in einen Grauwert umgewandelt. Graustufen haben einen Wertebereich von 0 bis 255, dies entspricht auch einem Byte, dieser Umstand wird bei der Umrechnung der Tiefendaten verwendet. Die Tiefendaten beinhalten für jeden Pixel eine Distanz, für die Graustufendarstellung wird das erste Byte dieser Distanz verwendet. Die Umwandlung in Graustufenbilder zur Visualisierung und Weiterverarbeitung notwendig.

3. Beschreibung des Prototypen

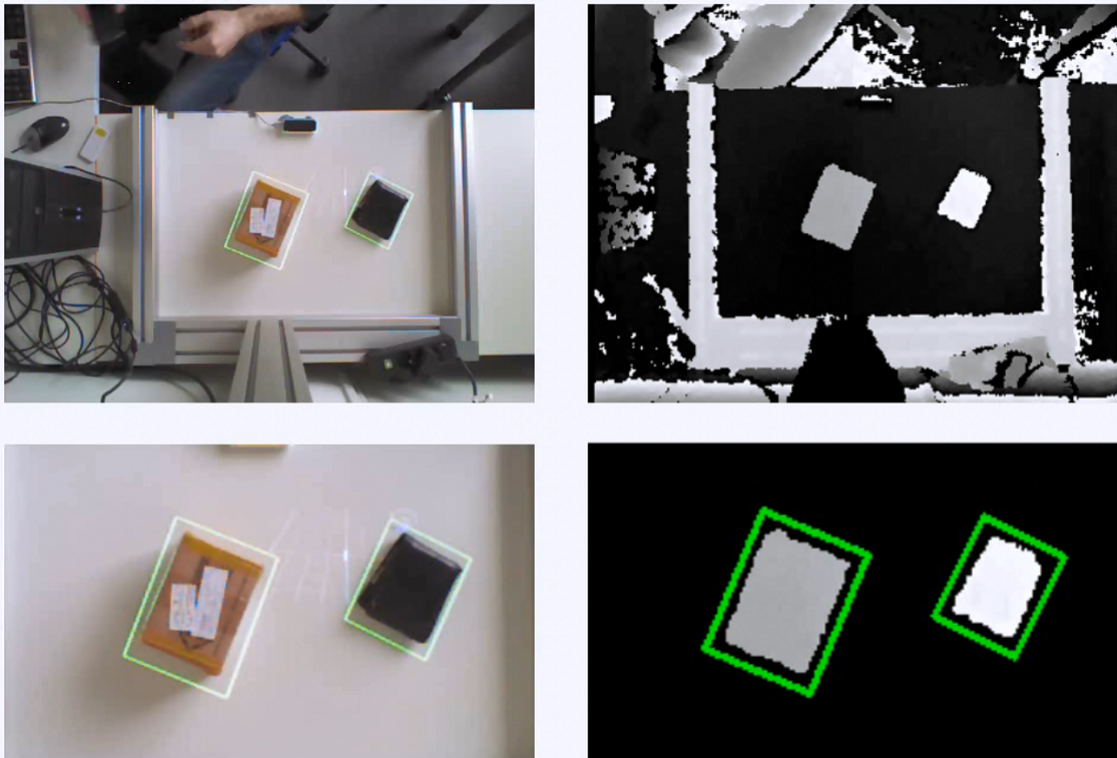


Abbildung 3.1.: Die oberen Bilder zeigen die verfügbaren Bildflächen, die unteren nur den Arbeitsbereich. Links ist die RGB- und rechts die Tiefenansicht zu sehen. Die grüne Umrandung wird durch die Erfassung der Objekte erzeugt.

3.2. Erfassung von Gegenständen

Zur Erfassung von Gegenständen wird die FindContour-Methode von EmguCV verwendet.

```
Contour<System.Drawing.Point> contours = gray_image.FindContours(  
    Emgu.CV.CvEnum.CHAIN_APPROX_METHOD.CV_CHAIN_APPROX_SIMPLE,  
    Emgu.CV.CvEnum.RETR_TYPE.CV_RETR_EXTERNAL,  
    storage);
```

Beim FindContour-Algorithmus handelt es sich um einen einfachen Algorithmus, der systematisch das gesamte Bild abtastet. FindContour arbeitet mit Binärbildern, also Bilder die nur aus 0 und 1 bestehen. Der Canny-Algorithmus erzeugt solche Bilder. Sobald eine Komponente getroffen wird, gilt dies als Ausgangspunkt und seine Kontur wird verfolgt, existieren keine weiteren Kontaktpunkte ist die Kontur abgeschlossen. Der Algorithmus beginnt nun erneut mit der Suche nach neuen Komponenten. Eine Kontur wird wiedergegeben durch eine Ansammlung von Punkten.

Bei der Implementierung von EmguCV besteht die Möglichkeit ein Rechteck um dies Konturen zu legen, somit ist eine Bestimmung der Position und Größe gewährleistet. Dieses Rechteck wird als Bildausschnitt zum Objekt vermerkt.

Diese Methode tastet sowohl auf das Tiefenbild sowie das RGB-Bild ab. Anschließend werden die Konturen durch ihre Positionen und Größe auf Überlappung und Duplizität verglichen. Abschließend werden diese Konturen als Gegenstand definiert und mit ihrem Bildausschnitt abgespeichert.

3.2.1. Canny-Algorithmus

Es handelt sich hierbei um einen Algorithmus der Bildbearbeitung zur Filterung von Kanten, welcher im Jahre 1986 von Canny[Can86] entwickelt wurde.

Dabei legt er seinen Fokus auf eine gute Lokalisierung, sowie eine geringe Fehlerrate, in Bezug auf die Erkennung und die Mehrfacherkennung von einzelnen Kanten. Der Algorithmus tastet die Grauwerte der Bilder ab und untersucht diese auf Unstetigkeit.

Zu Beginn des Verfahrens wird das Bild durch einen zweidimensionalen Gaußfilter geglättet, um Störungen zu beseitigen. Anschließend werden die gerichteten Ableitungen ermittelt, zur Faltung eignen sich hier der Sobel-Operator oder aber auch der Laplace-Operator. Darauf aufbauend wird nun die Kantenrichtung ermittelt, womit auch eine Aussage über die Steigung der Kante getroffen werden kann. Als Kantensteigung wird die Wertedifferenz zwischen dem Bereich vor der Kante und der Kante selbst bezeichnet. Im vierten Schritt kommt die so genannte „Non Maximum Suppression“ Technik zum Einsatz. Bei dieser Technik soll sicher gestellt werden, dass nur Pixel markiert wurden, die sich nahe der maximalen Kantensteigung befinden, dies führt zur Ausdünnung der Linien. Jedes Pixel wird mit seinen 8 Nachbarn verglichen, keines der benachbarten Pixel darf eine höhere Steigung aufweisen, außer es handelt sich um ein Pixel in Kantenrichtung.

Für Pixel mittlerer Steigung muss entschieden werden, ob es sich um eine Kante oder Rauschen handelt. Durch den Einsatz von zwei Schwellenwerten, der oberen und unteren Schranke, können so Fehlinformationen eliminiert werden. Dieser Schritt wird als Hysterese bezeichnet.

3.3. Freie Flächen

Nach der Erkennung von Gegenständen können nun auch freie Flächen berechnet werden. Durch die bereits bekannten Positionen der Gegenstände werden diese in einer Matrix festgehalten. Bildlich lässt sich dies durch ein Raster, welches auf die Arbeitsfläche gelegt wird verdeutlichen. Die Größe der Rasterrechtecke ist dabei variabel. Befindet sich nun eine Kante eines Objekts innerhalb eines Rasterrechtecks, so wird dieses mit 1 als belegt markiert. Somit ergibt sich eine Matrix aus 0 und 1. Durch den anschließend beschriebenen Algorithmus wird nun mittels der Matrix das größte unmarkierte Rechteck berechnet.

3.3.1. Algorithmus zur Flächenberechnung

Dieser Algorithmus¹ untersucht jede Position innerhalb der Matrix auf seine räumliche Einschränkung durch Einsen. Er arbeitet sich zeilenweise von oben nach unten durch, dabei werden pro Spalte

¹<http://stuckinananinfiniteloop.blogspot.de/2011/09/finding-maximum-zero-submatrix-with-c.html>

3. Beschreibung des Prototypen

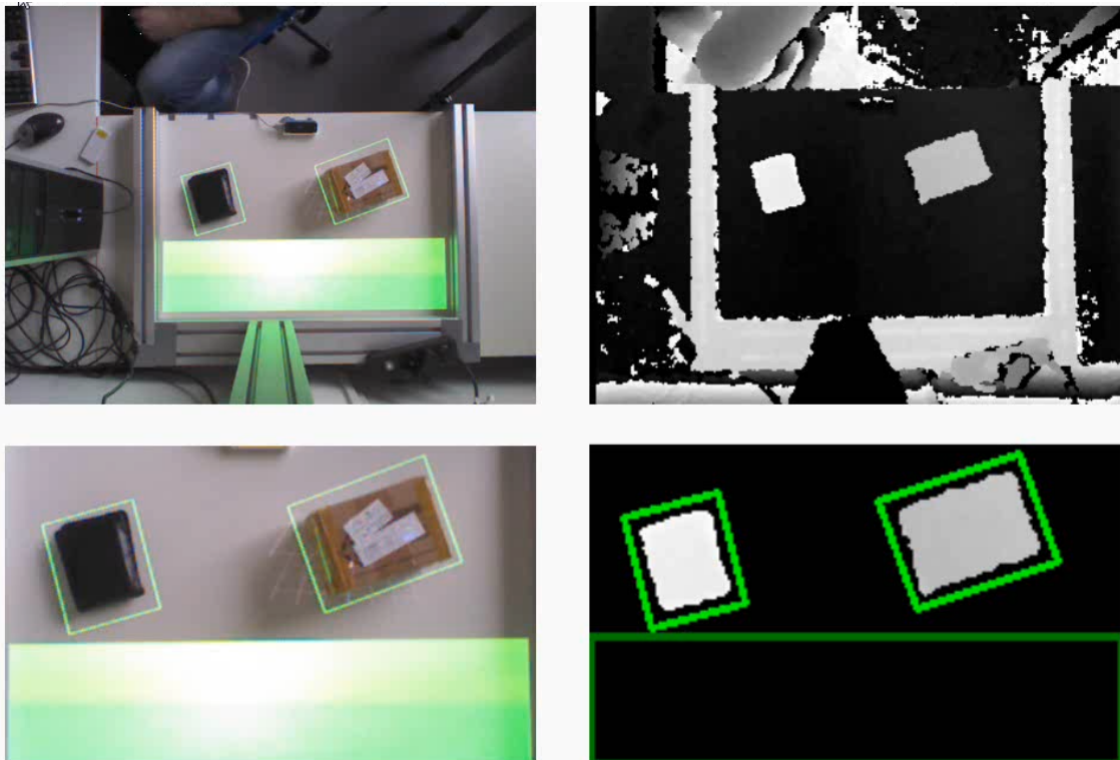


Abbildung 3.2.: Zusätzlich zur Abbildung 3.1 wird hier das größte freie Rechteck angezeigt.

die oberen und seitlichen Abgrenzungen gespeichert. Nach jeder Zeile wird durch diese Informationen die größte Fläche ermittelt und gegebenenfalls vermerkt. Mit absteigenden Zeilen können die seitlichen Abgrenzung nur enger werden, es sei denn an der aktuellen Position befindet sich eine Eins, in diesem Fall wird dies als obere Grenze markiert und die seitlichen Abgrenzungen auf das Maximum gesetzt.

3.4. Projektionsflächen auf Gegenständen

Als Projektionsfläche auf einem Gegenstand wird das größtmögliche Rechteck innerhalb dessen bezeichnet. Da Gegenstände auch komplexe Grundflächen besitzen können ist die Bestimmung der Projektionsfläche nicht immer eindeutig.

Diese Funktion basiert auf der Berechnung von freien Flächen (siehe Kapitel 3.3), doch im Gegensatz dazu Pixel genau. Für Berechnung der Fläche wird der bei der Erfassung des Gegenstands (siehe Kapitel 3.2) erzeugte Bildausschnitt verwendet. Jene Pixel die außerhalb der Objektkontur liegen werden als markiert belegt. Im Weiteren wird, wie bei der Freiflächenberechnung, der Algorithmus angewandt und somit ergibt sich eine rechteckige Projektionsfläche in Abhängigkeit zum Objekt.

3.5. Wiedererkennung von Gegenständen

Dieses Kapitel befasst sich mit der Wiedererkennungen von Gegenständen, dazu wird der SURF-Algorithmus verwendet. Dabei werden dem Algorithmus eine Liste mit bekannten Gegenständen und das zu vergleichende Objekt übergeben. Sobald der das Objekt in der Liste erkannt wird, wird seine ID ausgegeben.

3.5.1. SURF Algorithmus

Das Akronym SURF[BTG06] steht hier bei für „Speeded Up Robust Features“, frei übersetzt bedeutet es so viel wie „Beschleunigte, robuste Merkmale“. Dieser Algorithmus wurde 2006 veröffentlicht und umfasst die Detektion und Beschreibung von Bildmerkmalen, wie Ecken und charakteristisch Bildbereiche. Bei seiner Entwicklung wurde der Fokus darauf gelegt einen schnellen aber in Bezug auf die erfassten Daten robusten Algorithmus zu entwerfen.

Der Algorithmus arbeitet zudem mit Intensitätswerten, d.h. mit Bildern mit nur einem Kanal. Meistens werden diese Bilder Schwarzweiß dargestellt. SURF setzt sich aus zwei Bestandteilen zusammen, zum Einen der SURF-Detektor, welcher hervorstechende Bildstrukturen erkennt, und zum Anderen der SURF-Deskriptor. Der SURF-Deskriptor erstellt zu jedem erkannten Bildbereich einen Vektor, der die Umgebung beschreibt.

Integralbild

Sowohl der Detektor als auch der Deskriptor verwenden zur Faltung einen Box-Filter. Das Verfahren basiert auf dem Konzept der Summed Area Tables citeSummed und dienen der schnellen Summenbildung über die Pixel einer rechteckigen Bildfläche. Pro Pixel des Integralbildes steht die Summe aller Pixel, die innerhalb des Rechtecks zwischen dem aktuellen Punkt und des Bildursprungs liegen.

$$I_{\Sigma}(x, y) = \sum_{i=0}^{i \leq x} \sum_{j=0}^{j \leq y} I(i, j)$$

3. Beschreibung des Prototypen

16	9	21	1	10	10	28
7	22	5	9	16	18	19
2	25	30	11	2	4	20
3	26	10	19	12	13	7
6	16	14	5	17	28	11
7	5	13	6	2	17	18
18	13	26	29	4	13	13

16	25	46	47	57	67	95
23	54	80	90	116	144	191
25	81	137	158	186	218	285
28	110	176	216	256	301	375
34	132	212	257	314	387	472
41	144	237	288	347	437	540
59	175	294	374	437	540	656

Abbildung 3.3.: Links ist die Ausgangstabelle zusehen, rechts die Bereits berechnete Summed Area Table

Mit nur drei Additionen, und somit in konstanter Zeit, lassen sich nun Summen von Pixelwerten in beliebig große Rechtecke berechnen.

$$I_{\Sigma}(Fläche_D) = I_{\Sigma}(P_1) + I_{\Sigma}(P_4) - I_{\Sigma}(P_2) - I_{\Sigma}(P_3)$$

Dies führt zu einer großen Effizienzsteigerung des Algorithmus.

SURF-Detektor

Als Grundlage der Erkennung eines interessanten Punkte dient die Berechnung der Determinante einer Hesse-Matrix. Hierbei wird ein Box-Filter, zur Glättung des Bildes, verwendet, welcher sich auch für große Filtermasken effizient verhält.

Die Determinante der Hesse-Matrix zu einem beliebigen Punkt kann dadurch über

$$(\det(H_{approx})) = D_{xx} * D_{yy} - (0.9 * D_{xy})^2$$

berechnet werden, wobei D_{xx} die durch Box-Filter approximierte zweite partielle Ableitung darstellt. Um Skalierungsinvarianz zu erreichen wird die Größe der Box-Filters angepasst um unterschiedlich große Teile des Bildes abzudecken und so unterschiedliche Skalierungen zu simulieren. Durch die Skalierungsinvarianz kann ein Skalenraum aufgezogen werden, Bildlich wird dieser meist als Bildpyramide dargestellt. In jeder Ebene befindet sich das selbe Bild mit abnehmendem Detailgrad und Skalierung. Ein Punkt wird als interessant markiert, sobald es sich um ein Maximum auf allen Ebenen handelt.

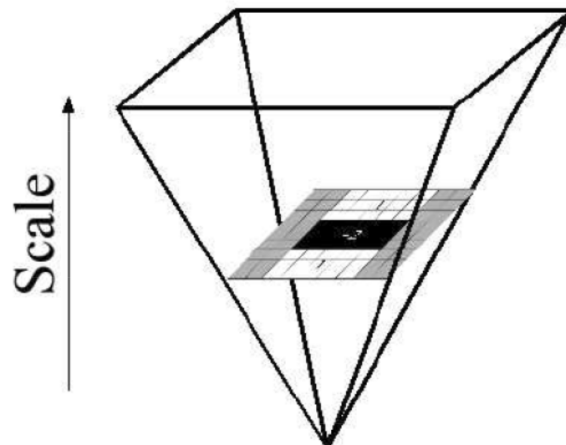


Abbildung 3.4.: Der SURF-Ansatz verläuft in entgegengesetzter Richtung zur traditionellen Skalenraum-Pyramide.

Quelle: http://www.informatik.hu-berlin.de/forschung/gebiete/viscom/teaching/media/cphoto10/cphoto10_04.pdf

SURF-Deskriptor

In diesem Teil des Algorithmus werden die erfassten Punkte, unter Verwendung von der Integralbilder in Verbindung mit Haar-Wavelets Filtern, genauer beschrieben. Durch Einsatz der Haar-Wavelets Filtern lassen sich Vektoren in X und Y- Richtung bestimmen, die die Helligkeitsveränderungen beschreiben. Nach Gewichtung und Aufsummierung der einzelnen Vektoren lässt sich zu jedem erfassten Punkt eine Richtung bestimmen.

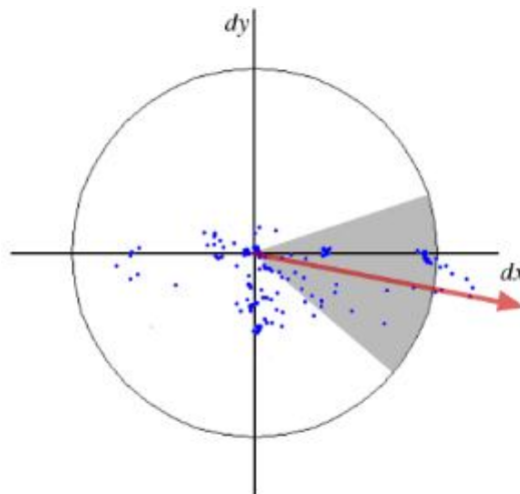


Abbildung 3.5.: Bestimmung des roten Deskriptors. Die Ergebnisse der Haar-Wavelets Filter werden blau dargestellt.

Quelle: http://docs.opencv.org/trunk/doc/py_tutorials/py_feature2d/py_surf_intro/py_surf_intro.html

Nun wird ein Rechteck um den erfassten Punkt gelegt und anhand der zuvor berechneten Richtung ausgerichtet. Das Rechteck wird nun in 16 gleichgroße Subregionen unterteilt, worin erneut die Haar Wavelets berechnet werden. Durch die 16 Subregionen ist eine Eindeutigkeit sowie eine Invarianz gegenüber der Rotation gegeben.

3.6. Rotationsänderung von Gegenständen

Bei Interaktion und Projektionen mit Gegenständen ist die Erkennung von Rotationen nützlich, so können die Drehungen als Eingabeimpulse aufgefasst werden. In Verbindung mit Projektionen können diese Informationen genutzt werden, um sie anzupassen oder auszurichten.

In dieser Arbeit wird die Rotationsänderung durch den Vergleich von zwei Bildern realisiert. Da es sich beim SURF-Algorithmus um einen rotationsinvarianten Algorithmus handelt, kommt er auch bei der Rotationsbrechung zum Einsatz. Der SURF-Algorithmus erkennt den Gegenstand auf dem Bild wieder, nun werden die KeyPoints beider Bilder auf Analogie untersucht. Robert Langanieri beschreibt in seinem Buch *OpenCV 2* [Lag11] die Möglichkeit eines Symmetrie Tests von KeyPoints. Durch diesen Test ergeben sich Punktpaare zwischen beiden Bildern (siehe Abbildung 3.6). Es genügen zwei dieser Punktpaare um die Rotation zu berechnen. Es werden die Steigungen der zwei Punkte im jeweiligen Bild berechnet, um diese anschließend zu vergleichen, dies ergibt die Rotation des Gegenstands.

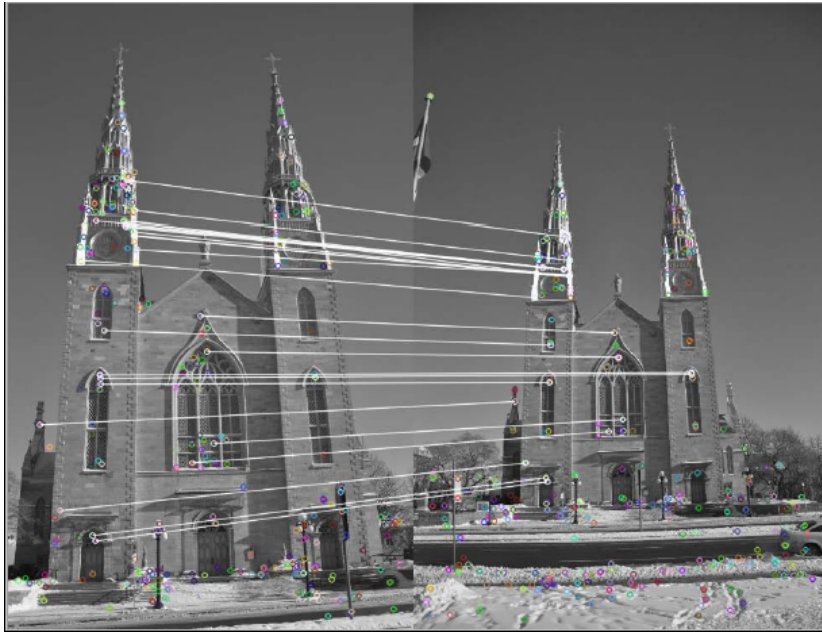


Abbildung 3.6.: Bei den farbigen Kreisen handelt es sich um FeaturePoints, identische Punkte werden durch eine Linie zu Punktpaaren verbunden.
Quelle: [Lag11]

3.7. Erfassung der Oberflächenmerkmale

Die Oberflächenbeschaffenheit lässt sich, für Projektionswiedergabe, in drei Merkmale zerlegen. Zum einen die Helligkeit, ein weiteres Merkmal ist die Struktur, hierbei sind Farbwechsel und Muster gemeint und zu Letzt die Form der Oberflächen, sowohl im zweidimensionalen als auch im dreidimensionalen Raum. Im weiteren Verlauf wird die Implementierung zur Erfassung dieser Oberflächenmerkmale durch die RGB- und Tiefenkamera beschrieben. Zu Untersuchung der Oberflächen wird erneut nur der verkleinerte Bildausschnitt eines Objekts verwendet (siehe 3.2)

3.7.1. RGB-Kamera

Wie schon in Abschnitt 3.1 erwähnt, wird das RGB-Bild in Graustufenwerte umgewandelt. Für Projektionen spielt der Kontrast eine wichtige Rolle, je stärker dies ist um so deutlicher wird eine Projektion von ihrem Umfeld hervorgehoben. Da es sich beim Kontrast um ein Unterscheidungsmerkmal im Helligkeitsverlauf des Bildes handelt, werden nun die gesamte Helligkeit und die Farbwechsel betrachtet.

Helligkeit Um die Helligkeit eines Gegenstandes zu berechnen wird der durchschnittliche Grauwert der Objektfläche ermittelt. Der so erfasste Wert liegt zwischen 0 für Schwarz und 255 für weiß.

```
intensity = (int) rgbImage.GetAverage().Intensity;
```

3. Beschreibung des Prototypen

Struktur Um die Struktur der Oberfläche zu analysieren wird es aufwendiger. Hierbei werden farblich zusammenhängende Flächen mit Hilfe der findContour Methode (siehe 3.2) von EmguCV als solche erkannt. Um weitere Informationen zu erlangen werden FeaturePoints des SURF-Algorithmus abgezählt. Je mehr Farbwechsel sich auf einer Oberfläche befinden um so mehr FeaturePoints sind zu erkennen, auch Reflexionen werden so erkannt. Für die Analyse von Oberflächen ist an dieser Stelle lediglich die Anzahl der Konturen und FeaturePoints von Relevanz. Die erfassten Werte der Struktur lagen im Laufe der gesamten Arbeit im Bereich zwischen 0 und 100. Mit diesen Werten wird die Anzahl der störenden Einflüsse wiedergegeben.

```
//FeaturePoint
MKeyPoint[] featurePoint = surfCPU.DetectKeyPoints(rgbImage, null);
featurePointCount = featurePoint.Count();
```

3.7.2. Tiefenkamera

Bei der Erfassung der Form durch die Tiefenkamera wird ähnlich wie bei der Erfassung von Gegenständen vorgegangen.

Form Anhand der bereits erkannten Kontur wird die Grundform kategorisiert, es bestehen die Gruppen der Kreise, Dreiecke, Vierecke und Vielecke. Anschließend werden Höhenunterschiede innerhalb dieser Form zu untersuchen. Hierbei kommt erneut der FindContour-Algorithmus (siehe 3.2 zum Einsatz, die nun gelieferten Konturen beschreiben abrupte Höhenunterschiede die mindestens eine Grundfläche von 0,5 x 0,5 cm und einer Tiefe von 1 cm besitzen. Mit diesen Informationen lassen sich die Anzahlen der Kanten und Löcher auf den Oberflächen erkennen, welche zur Minderung der Projektionsfähigkeit eines Gegenstandes führen können.

Zusätzlich wird die Anzahl von Linien innerhalb des Tiefenbildes festgehalten, dies dient zur Mittlung der tatsächlichen Kantenzahl.

```
LineSegment2D[] lines = cannyEdges1.HoughLinesBinary(
    // 1, //Distance resolution in pixel-related units
    // Math.PI / 45.0, //Angle resolution measured in radians.
    // 10, //threshold
    // 5, //min Line width
    // 10 //gap between lines
    // )[0]; //Get the lines from the first channel
```

4. Evaluation

Zentraler Bestandteil dieses Kapitels ist die durchgeführte Benutzerstudie. Die hierbei gewonnenen Aussagen werden daraufhin analysiert und bewertet. Anschließend wurde aus diesen Erkenntnissen ein Modell zur Klassifizierung entworfen und geprüft. Abschließend wird die gesamte Evaluation nochmals auf ihre guten und weniger guten Aspekte hin betrachtet.

4.1. Benutzerstudie

Zur Erfassung des Einflusses der unterschiedlichen Oberflächen auf die Projektionsfähigkeit wurde eine Benutzerstudie durchgeführt. Unter dem Titel „Oberflächenanalyse zur Erkennung von geeigneten Projektionsflächen“ wurden Probanden nach ihrer subjektiven Meinung zur Projektionsfähigkeit unterschiedlicher Objekte befragt. Dabei zeigen sich glatte, helle Oberflächen als besonders geeignet, während kantige, schwarze oder glänzende Oberflächen schlecht abschnitten.

4.1.1. Design

Diese Studie verwendet eine „repeated-measures“-Design, um eine abwechselnde, sich nicht wiederholende Reihenfolge zu erstellen wurde auf das Latin-Square Verfahren (siehe Anhang A.1) zurückgegriffen. Die unabhängigen Variablen setzen sich aus der Form und Struktur der Projektionsobjekte zusammen. Wiederrum unterteilt sich die Form in glatt, gelocht und Kanten durchzogen. Die Struktur setzt sich aus weiß, schwarz, beschriftet, glänzend und matt zusammen. Bei den abhängigen Variablen handelt es sich um alle erfassten Werte, beginnend mit der Bewertung der Probanden zu den einzelnen Objekten. Ebenso gehören die durch das System festgehaltenen Daten dazu, wie die Helligkeit, die Anzahlen der FeaturePoints, der Konturen aus dem Farb- und Tiefenbild und der Linien im Tiefenbild.

4.1.2. Teilnehmer

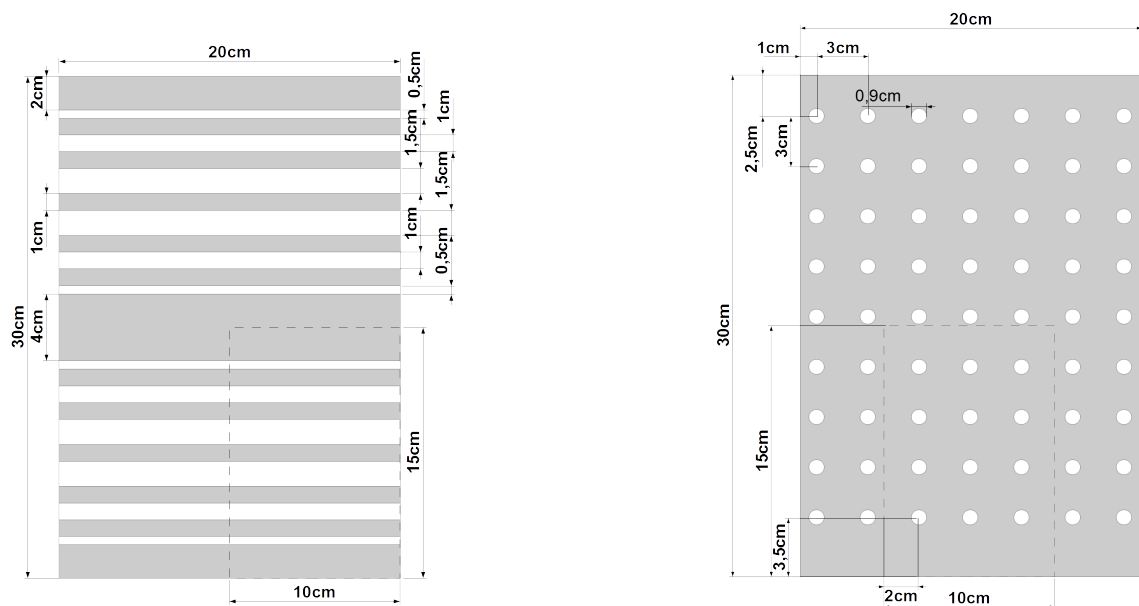
Die Rekrutierung der Probanden erfolgte mündlich. Die Studie stellte keinerlei Voraussetzungen an die Probanden. An der Studie nahmen 15 Probanden teil, die zwischen 10 und 31 Jahren alt waren. Das Berufsspektrum reichte von Schüler über Krankenschwestern bis hin zu Diplom Informatikern. Hiervon waren 4 Teilnehmer weiblich und 11 männlich. Als Vergütung erhielten die Teilnehmer Süßigkeiten.

4.1.3. Apparaturen

Die Studie wurde am Montagetisch im Labor des Institutes in Verbindung mit dem bereits beschriebenen System (siehe 3) durchgeführt. Für die Dokumentierung der Meinungen wurde ein Fragebogen erstellt (siehe A.2), dieser beinhaltet neben demographischen Fragen zu Geschlecht, Alter und Berufsbild, 15 Fragen zur Projektionsfähigkeit der jeweiligen Oberflächeneigenschaften sowie drei weitere allgemein gehaltene Fragen. Als Projektionsfläche dienten bearbeitete Holzklötze, im weiteren Verlauf werden diese als Projektionsobjekte bezeichnet.

Projektionsobjekte

Diese Projektionsobjekte wurden unterteilt in fünf Oberflächenstrukturen und jeweils in drei Oberflächenformen (siehe 4.2). Zur Erfassung der Daten wurden die 15 Objekte in zwei Größen angefertigt. Die zwei Größen dienten lediglich zur Mittlung der erfassten Daten über unterschiedliche Größen hinweg. Zur Erläuterung der Bemessung der Formen siehe Abbildung 4.1.



(a) Skizze des kanten durchgezogenen Projektionsobjekts. Die Abstände der oberen Hälfte wiederholen sich im unteren Teil

(b) Skizze des gelochten Projektionsobjekts. Der Abstand zwischen den Löchern beträgt jeweils 3 cm.

Abbildung 4.1.: Die Masse der Grundfläche gelten auch für die glatten Projektionsobjekte. Die gestrichelte Fläche innerhalb der Skizzen kennzeichnet die kleinen Projektionsobjekte, die weißen Flächen kennzeichnen die Vertiefungen

4.1.4. Vorgehen

Die Studie wurde persönlich mit den Probanden durchgeführt, um äußere Einflüsse ausschließen zu können wurde stets gleiche Belichtungsverhältnisse erzeugt. Zu Beginn der Studie wurde jedem Probanden, das Testbild präsentiert. Anschließend mussten die Probanden nach der jeweiligen Reihenfolge A.1 die Projektionsobjekte auf die Arbeitsfläche legen. Diese wurden daraufhin vom System mit dem Testbild belichtet, die somit entstanden Projektionen wurden von den Probanden mit Hilfe des Fragebogens in eine sieben Punkte Likert-Skala eingestuft. Pro Projektionsobjekt wurden vor und während der Projektion sämtliche abhängige Variablen festgehalten. Abschließend wurden noch drei Fragen in Bezug auf die gesamte Studie bearbeitet. Durch Abbildung 4.2 lässt ein kleiner Einblick auf die unterschiedlichen Projektionsverhalten wiedergeben.

4. Evaluation


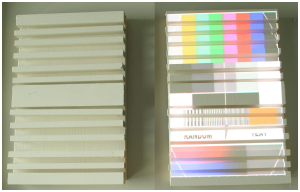
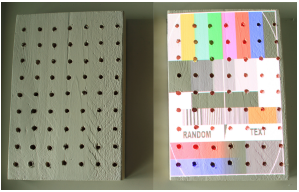

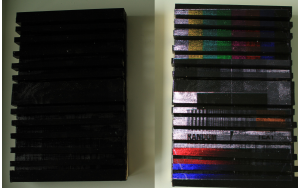
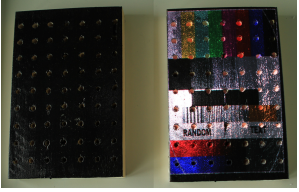
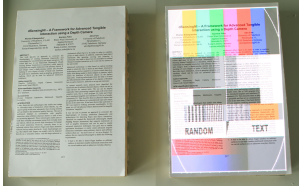

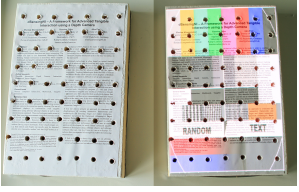
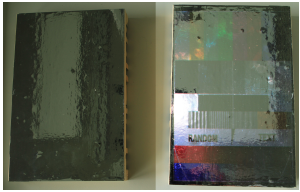
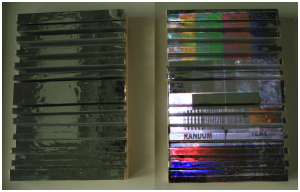
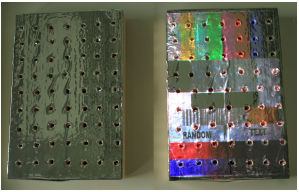


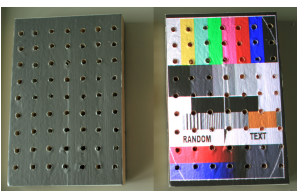
	Glatt	Kanten	Löcher
Weiß	 <p>(a) Weißes glattes Objekt</p>	 <p>(b) Weißes kantiges Objekt</p>	 <p>(c) Weißes gelochtes Objekt</p>
Schwarz	 <p>(d) Schwarzes glattes Objekt</p>	 <p>(e) Schwarzes kantiges Objekt</p>	 <p>(f) Schwarzes gelochtes Objekt</p>
Text	 <p>(g) Beschriftetes glattes Objekt</p>	 <p>(h) Beschriftetes kantiges Objekt</p>	 <p>(i) Beschriftetes gelochtes Objekt</p>
Glanz	 <p>(j) Glänzend glattes Objekt</p>	 <p>(k) Glänzendes kantiges Objekt</p>	 <p>(l) Glänzendes gelochtes Objekt</p>
Matt	 <p>(m) Mattes glattes Objekt</p>	 <p>(n) Mattes kantiges Objekt</p>	 <p>(o) Mattes gelochtes Objekt</p>

Abbildung 4.2.: Das Gesamtbild setzt sich aus allen Projektionsobjekten zusammen. Jeweils zeigt der linke Teil das Projektionsobjekt und der rechte Teil das selbe Objekt mit Projektion.

4.2. Ergebnisse

Im folgende Abschnitt werden die Ergebnisse der Fragebögen und die vom System erfassten Metadaten ausgewertet.

4.2.1. Auswertung der Fragebögen

Nach dem Prinzip „Größer ist Besser“ wurden die Werte des Fragebogens invertiert, folglich handelt es sich bei einem Wert von 7 um eine sehr gute Oberfläche hingegen 1 eine sehr schlechte bezeichnet. Um allgemeingültige Aussagen über die einzelnen Projektionsoberflächen zu erhalten wurden der Median aller Probanden bezüglich der einzelnen Projektionsobjekte berechnet. Die Bewertung der Objekte lässt sich in Abbildung 4.3 und exakter aus Tabelle 4.1 abzulesen.

	Weiß	Schwarz	Schrift	Glanz	Matt
Glatt	7	3	6	2	6
Kanten	5	2	4	1	4
Löcher	6	2	5	2	5

Tabelle 4.1.: Gemittelte Werte aller Projektionsobjekte. In den Spalten werden die Strukturen und in den Zeilen die Formen aufgeführt.

Aus Abbildung 4.3 ist ersichtlich, dass die Struktur einen größeren Einfluss auf die Bewertung der Wahrnehmung hat als die Form der Oberfläche. Aus diesem Grund wurden die strukturellen Gemeinsamkeiten farblich hervorgehoben, beispielsweise werden alle weißen Projektionsobjekte als weißer Balken dargestellt. Im weiteren werden Struktur und Form erstmals getrennt betrachtet.

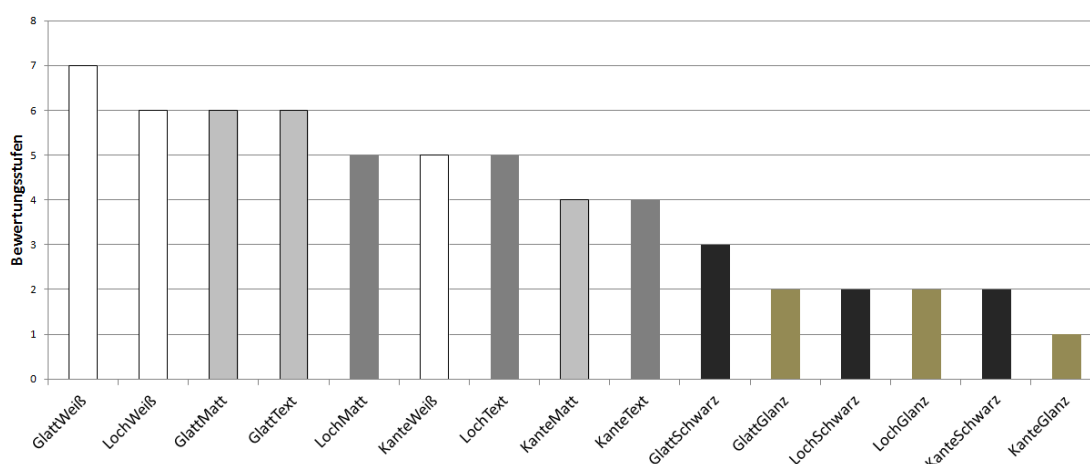


Abbildung 4.3.: Gemittelte Werte aller Projektionsobjekte

Betrachtung der Struktur

Zunächst wird die Form der Projektionsobjekte vernachlässigt, somit lässt sich die Auswertung der Fragebögen wie in Abbildung 4.4 komprimieren.

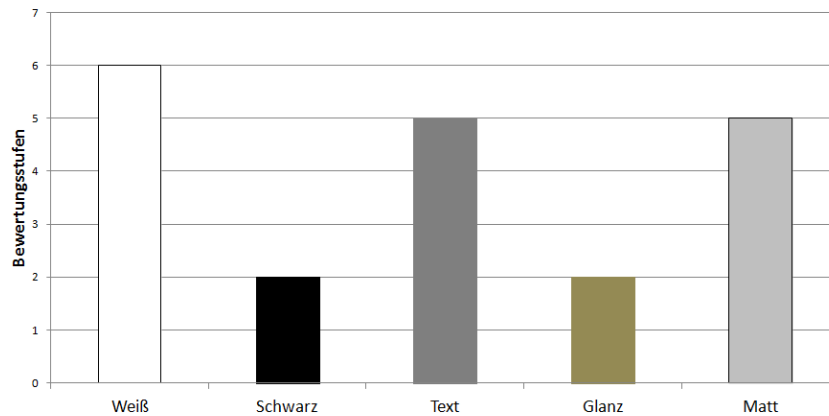


Abbildung 4.4.: Bewertung der Projektionsobjekte auf Grund ihrer Struktur

Durch die Ergebnisse lässt sich erkennen, dass Glänzende Projektionsobjekte als ungünstig erweisen, als Grund dafür wurde von den Probanden die Reflexionseigenschaft genannt. Abhängig vom Blickwinkel unterschied sich das Testbild gravieren, da der sichtbare Bildausschnitt und der Kontrast variierten (siehe z.B. Abbildung 4.2(j)). Ebenso zeigte sich, dass die Wiedergabequalität von Projektionen abhängig der Helligkeit ihrer Oberfläche ist, verdeutlicht durch den Vergleich von „Weiß“ und „Schwarz“. Ein weiterer Aspekt zeigte sich bei der Betrachtung von „Matt“ und „Text“, zwar besitzen beide eine vergleichbare, durchschnittliche Oberflächenhelligkeit, doch wird der durch die Beschriftung erzeugte Kontrast als ungeeigneter betrachtet.

Betrachtung der Form

Abbildung 4.5 verdeutliche den in Abbildung 4.3 angedeutete Einfluss der Formen auf die Projektionsfähigkeit. Logischerweise zeigen sich glatte Oberflächen besser als gelocht, den Abschluss bildeten kantigen überzogene Formen. Der Unterschied zwischen den einzelnen Formen erweist sich nicht so deutlich wie die der Strukturen.

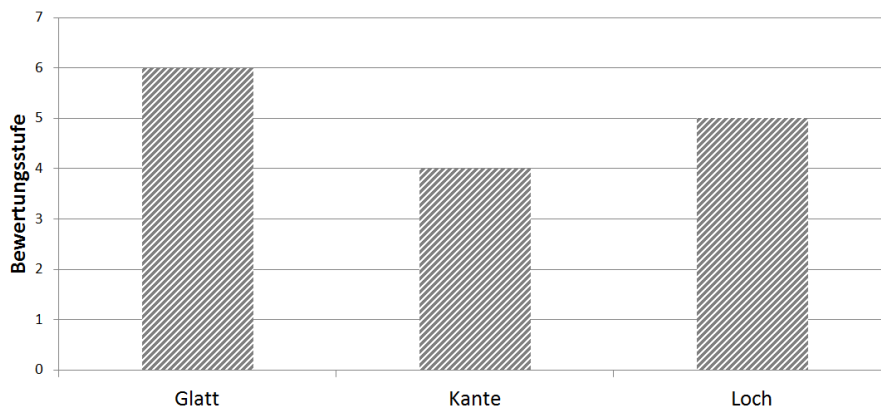


Abbildung 4.5.: Bewertung der Projektionsobjekte auf Grund ihrer Form

Insgesamt kann die Form als Manipulationsfaktor der Struktur gesehen werden. Ferner scheint durch sinken der Helligkeit dieser Faktor geringer zu werden.

4.2.2. Erfasste Daten

Im folgenden Abschnitt werden die vom System erfassten Metadaten betrachtet. Wie in Kapitel 3.7 erwähnt teilen sich dies in fünf Werte auf, die Helligkeit und die Anzahlen der FeaturePoints, der Farbkonturen, der Linien und der Tiefenkonturen. Bei den verwendeten Werten handelt es sich um Durchschnittswerte.

Analyse der erfassten Strukturdaten

Zur Analyse der Struktur, werden die Informationen zur Form vernachlässigt, da sich diese für alle Strukturen als identisch erweisen. In Abbildung 4.6 werden die durchschnittlichen erfassten Werte über alle Formen dargestellt. Die ersten drei Strukturen lassen sich relativ einfach einstufen, da ein Anstieg der störenden FeaturePoints ersichtliche ist. Auch der von den Probanden erwähnte Kontrastverlust kann durch die Helligkeitswerte der schwarzen im Vergleich zu weißen Oberfläche nachvollzogen werden. Als interessant erweist die glänzende Oberfläche, da sie eine große Anzahl an FeaturePoints anzeigt, hierbei handelt es sich um die Reflexionen auf der Oberfläche.

4. Evaluation

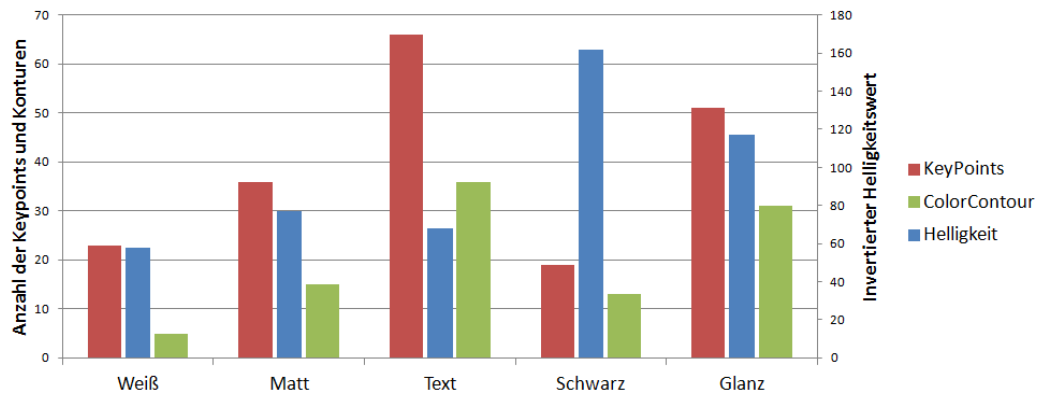


Abbildung 4.6.: Metadaten der Struktur. Die Helligkeit wird in Abhängigkeit zur rechten, sekundär Achse angezeigt.

Aus den erfassten Daten während der Projektionsphase konnten keine nützliche Schlussfolgerung gezogen werden.

Analyse der erfassten Daten zur Form

Die Helligkeit und die FeaturePoints sind in diesem Teil zu vernachlässigen. Beim Erkennen der Löcher zeigte sich die Schwäche des Systems im Bereich der Tiefenwahrnehmung, da der Durchmesser der Löcher zu gering war. Dies wurde mit Hilfe der optischen Aufnahme von Konturen reguliert. Die erfassten Anzahlen (siehe Abbildung 4.7) können nicht als absolute Werte gesehen werden, da beispielsweise glatte Oberflächen maximale eine Tiefenkontur und vier Linien enthalten können. Aus diesem Grund können diese Werte nur in Relation zu einander betrachtet werden.

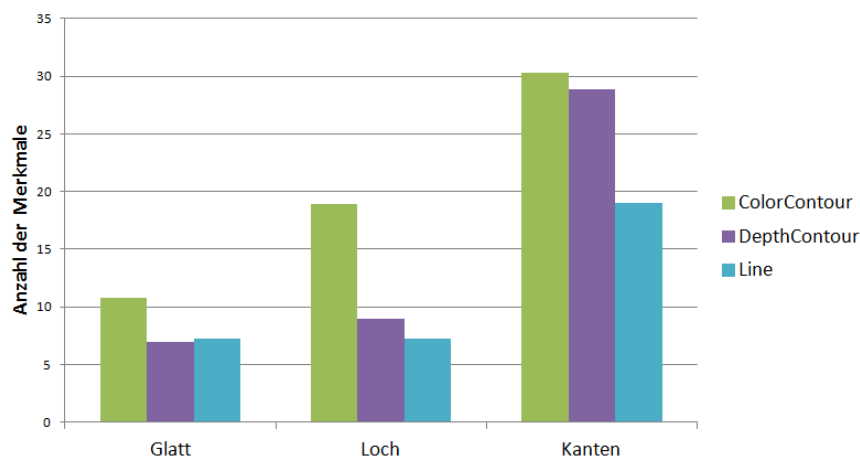


Abbildung 4.7.: Metadaten der Form. Die Werte verhalten sich proportional zur Unebenheit der Oberfläche.

4.2.3. Auswertung der Daten

Die Auswertung der Daten wurden nach Kendalls Tall, Spearman und Pearson durchgeführt, hierbei handelt es sich um Rangkorrelationsanalyse-Verfahren. Da sich die Ergebnisse nicht gravierend unterschieden, wird des weiteren auf die Berechnung nach Spearman eingegangen. Alle fünf Oberflächenmerkmale wurden in Relation zur Bewertung der Probanden stellt und somit ihr Korrelationskoeffizient berechnet. Der Korrelationskoeffizient kann die Werte von -1 bis 1 annehmen. Je näher sich der Wert an seinen Randbereichen befindet um so größer ist der lineare Zusammenhang, hingegen besteht bei einem Korrelationskoeffizient von 0 kein linearer Zusammenhang. Hierbei zeigt sich wie bereits vermutet, dass die Helligkeit der Oberfläche den Größten Einfluss auf die Bewertung hat, sein Korrelationskoeffizient liegt bei 0,85. Diese Korrelation ist bei einem Niveau von 0,01 signifikant. An Hand der jeweiligen Korrelationskoeffizient lässt sich folgendes lineares Modell aufstellen:

$$\begin{aligned} \text{Bewertung} = & 0,849 * \text{Helligkeit} - 0,165 * \text{FeaturePoints} - 0,509 * \text{ColorContour} \\ & - 0,581 * \text{DepthContour} - 0,19 * \text{Lines} \end{aligned} \quad (4.1)$$

Bei der Evaluation des Modells zeigt sich, dass diese lineare Annäherung das Verhalten der Oberflächenmerkmale zur Bewertung angemessen simuliert, bei der Berechnung des Korrelationskoeffizient zwischen dem Modell und der Bewertung ergab sich ein Wert von 0,94 (Die Korrelation ist ebenfalls bei einem Niveau von 0,01 signifikant).

4.3. Diskussion

In dieser Benutzerstudie wurde ein eindeutiger Effekt von Oberflächeneigenschaften auf die Projektionseigenschaft festgestellt. So erwiesen sich Projektion auf weiße Oberflächen besser als auf beschrifteten oder gar schwarzen Oberflächen. Innerhalb jeder Strukturen wurden glatte Flächen bevorzugt. Durch das somit erstellte Modell lassen sich tendenzielle Einstufungen von Oberflächen festhalten, doch auf Grund des kleinen Studenumfangs lässt sich keine genau Evaluierung von Oberflächen durchführen. Bei der Betrachtung einzelner Oberflächeneigenschaften erwiesen sich nur die Kontur sowie und die Helligkeit als stetige Faktoren zur Projektion, da alle weiteren eine höhere Signifikanz als 5% besaßen. Laut den Ergebnissen wird die Bewertung der Projektion zu 37% von der Helligkeit der Oberfläche beeinflusst, doch ist dieser Punkt auch am einfachsten auszugleichen. Hierbei reichen einfach Verfahren zur Aufhellung des projizierten Bildes aus.

Bei der Verwendung der Tiefendaten zeigten sich Schwierigkeiten, denn durch den Drang möglichst kleiner Unebenheit und Konturen zuerkennen zeige sich die Schwäche der verwendeten Technik. Beim ständige Abtasten der Oberfläche traten Fehlinformation bzw. Schwankungen der Distanzen auf. Hinsichtlich der Ambition ein echtzeitfähiges System zu entwickeln konnte dies nur bedingt unterdrückt werden. Daher wurde bei der Erfassung von Linen und Konturen durch Tiefendaten auf gemittelte statt der exakten Werte zurückgegriffen. Dies lässt nun Vermuten, dass die erfassten Tiefendaten nicht verwendbar sind und somit weggelassen werden können. Dies ist nicht der Fall, da sie dennoch einen Einfluss von 33,6 % auf die Bewertung haben.

Bei der Oberflächenanalyse von glänzenden Flächen erwiesen sich diese als durchaus tückisch, da bei Betrachtung der Oberfläche der Blickwinkel eine richtige Rolle spielt, je nach Winkel ist die

4. Evaluation

Projektion sehr gut bis gar nicht erkennbar. Dies konnte durch den Aufbau nicht in seiner ganzen Komplexität berücksichtigt werden, dennoch stimmt das entworfene Modell mit der Tendenz der Bewertung überein.

Abschließend lässt sich sagen, dass die Bewertung der Projektionen eine subjektive Entscheidung ist und durch die Komplexität der Projektionsoptik nur eine grundlegende Modellierung der Problemstellung möglich ist. Somit trifft das Modell für viele Projektionsflächen zu, doch ist sie kein Garant sämtliche Flächen richtig zu klassifizieren.

5. Fazit und Ausblick

Das letzte Kapitel fasst die theoretischen und praktischen Erkenntnisse der gesamten Arbeit in einem kurzen Fazit zusammen. Abschließend wird ein Ausblick über weitere, ergänzende Arbeiten gegeben.

5.1. Fazit

Ziel der vorliegenden Arbeit war es ein System zur Lokalisierung und Klassifizierung von Oberflächen in Abhängigkeit von ihrer Projizierbarkeit zu entwickeln. In diesem Rahmen wurde ein Prototyp zur Erfassung von Gegenständen sowie freien Flächen entworfen, darauf aufbauend wurden unterschiedliche Oberflächen in den Fokus gezogen. Hierfür wurden Einstufungen verschiedener Oberflächeneigenschaften auf Grund ihrer Projizierbarkeit erhoben. Durch diese gewonnenen Erkenntnisse wurde abschließend ein Modell zur Klassifizierung der Oberflächen erstellt.

Bei der Evaluation der Projektionsflächen zeigte sich, dass ein Zusammenhang zwischen der Projizierbarkeit und den unterschiedlichen Oberflächeneigenschaften zu erkennen ist. Es erwiesen sich helle, ebene Flächen als gute Grundlage für Projektionen. Mit zunehmender Abdunklung und steigender Unstetigkeit der Farbe und Tiefe verschlechterte sich auch die Qualität der Projektionen. Abschließend lässt sich zur Projektionsfläche frei nach der Redewendung „weniger ist mehr“ festhalten, je heller und je weniger Impulse von einer Oberfläche ausgehen, um so besser ist für eine Fläche geeignet. Bei der Betrachtung des vorgestellten Prototypen in Verbindung mit dem erarbeiteten Modell ist ein möglicher Ansatz zur Klassifizierung von potenziellen Projektionsflächen zu erkennen. Zwar erwiesen sich die einzelnen, erfassten Komponenten als nicht exakt und aussagekräftig, doch durch das Zusammenspiel aller Komponenten zeigte es sich im Laufe der Evaluationen als akzeptable Umsetzung.

Bei der Projizierbarkeit handelt es sich um eine subjektive Wahrnehmung. Meines Erachtens stellt das System, aus diesem Grund, einen guten Klassifikator im vierstufigen Bereich dar. Im Anwendungsfall erweist sich dies auch als sinnvoll, da somit stabile, grundlegende und allgemeingültige Aussagen getroffen werden.

5.2. Ausblick und weitere Arbeiten

Die Lokalisierung und Klassifizierung von Oberflächen in Verbindung mit Projektionen erwies sich als eine sehr umfangreiche Thematik. Im Bereich Lokalisierung werden momentan rechtwinklige Flächen

5. Fazit und Ausblick

in Betracht gezogen, es liegt nahe sich in weiteren Arbeiten mit anderen Flächenformen auseinander zu setzen. Hierbei könnten die Position des Betrachter und die damit verbundene Ausrichtung der Informationen berücksichtigt werden. Des weiteren könnte die Positionierung der Projektionen nicht nur in Abhängigkeit zur verfügbaren Fläche sondern auch zur benötigte Fläche stehen. Zum Beispiel benötigen die Projektion eine L-förmige Fläche mit bestimmten ausmaßen, darauf hin sollte das System sämtliche Oberflächen auf die entstehend Abhängigkeiten untersuchen.

Bei der Klassifizierung der Oberflächen erwiesen Reflexionen und Formen als sehr interessant. In der vorliegenden Arbeit wurden drei grundlegende Formen betrachtet, dabei wurde nicht auf das Verhalten der Projektion auf ansteigende Komplexität, Hochunterschiede oder Flächengröße eingegangen. Beispielsweise kann eine Oberfläche mit einem kleinen oder großen Loch gleichermaßen klassifiziert werden, doch bei mittlerer Größe des Loches als schlechter erweisen. Grundlegende hierfür ist die Verbesserung der erfassten Tiefendaten durch das ermitteln exakter und stabiler Distanzen. Bei einer weiteren Untersuchung der Reflexionen wäre eine Klassifizierung der Objekte aus dem Blickwinkel der Benutzer wünschenswert, da diese Flächen je nach Winkel anders wirken.

Ein anderer Bereich wäre die Visualisierung, so könnten nach der Bestimmung einer Projektionsfläche(siehe [BIWG07]) die Projektion bezüglich der Unebenheit bzw. der Färbung des Hintergrunds und des Projektionswinkels angepasst werden. Abschließend wäre die Entwicklung eines kompakten, semi-autonomen Projektionssystems möglich, mit dem innerhalb von einen Klick Informationen auf den bestmöglichen Flächen wiedergegeben werden.

A. Appendix

A.1. Reihenfolge der Projektionsobjekte

Reihenfolge der Projektionsobjekte pro Proband

Proband

1	Loch Weiß	Glatt Text	Kante Weiß	Kante Glanz	Loch Matt	Kante Schwarz	Kante Text	Loch Text	Loch Schwarz	Glatt Schwarz	Glatt Glanz	Glatt Matt	Glatt Weiß	Kante Matt	Loch Glanz
2	Loch Text	Kante Glanz	Glatt Schwarz	Kante Weiß	Loch Schwarz	Kante Text	Kante Schwarz	Loch Weiß	Loch Glanz	Glatt Glanz	Glatt Matt	Loch Matt	Glatt Text	Glatt Weiß	Kante Matt
3	Kante Matt	Glatt Glanz	Loch Glanz	Loch Schwarz	Kante Schwarz	Glatt Weiß	Loch Matt	Glatt Matt	Kante Text	Loch Text	Loch Weiß	Kante Weiß	Kante Glanz	Glatt Text	Glatt Schwarz
4	Loch Matt	Loch Weiß	Kante Schwarz	Kante Matt	Glatt Weiß	Glatt Glanz	Loch Text	Loch Schwarz	Glatt Text	Loch Glanz	Kante Glanz	Glatt Schwarz	Glatt Matt	Kante Weiß	Kante Text
5	Kante Glanz	Glatt Schwarz	Kante Text	Glatt Glanz	Kante Weiß	Loch Text	Loch Glanz	Kante Matt	Loch Weiß	Glatt Text	Loch Schwarz	Glatt Weiß	Kante Schwarz	Glatt Matt	Loch Matt
6	Loch Glanz	Loch Schwarz	Loch Text	Glatt Weiß	Glatt Glanz	Kante Weiß	Glatt Text	Kante Schwarz	Kante Glanz	Loch Matt	Glatt Schwarz	Kante Matt	Kante Text	Loch Weiß	Glatt Matt
7	Kante Schwarz	Kante Text	Kante Matt	Loch Matt	Glatt Schwarz	Loch Schwarz	Kante Weiß	Kante Glanz	Glatt Glanz	Glatt Matt	Glatt Weiß	Loch Text	Loch Weiß	Loch Glanz	Glatt Text
8	Kante Text	Loch Glanz	Kante Glanz	Loch Weiß	Kante Matt	Glatt Matt	Glatt Schwarz	Glatt Weiß	Loch Text	Loch Schwarz	Kante Schwarz	Glatt Text	Loch Matt	Glatt Glanz	Kante Weiß
9	Glatt Schwarz	Kante Weiß	Glatt Weiß	Kante Text	Loch Text	Glatt Text	Glatt Matt	Loch Glanz	Kante Matt	Kante Schwarz	Loch Matt	Loch Weiß	Loch Schwarz	Kante Glanz	Glatt Glanz
10	Glatt Matt	Loch Matt	Glatt Text	Kante Schwarz	Kante Text	Kante Matt	Kante Glanz	Glatt Schwarz	Glatt Weiß	Loch Weiß	Kante Weiß	Loch Glanz	Glatt Glanz	Loch Schwarz	Loch Text
11	Glatt Weiß	Kante Matt	Loch Matt	Loch Text	Loch Glanz	Kante Glanz	Loch Weiß	Kante Text	Glatt Matt	Kante Weiß	Glatt Text	Glatt Glanz	Glatt Schwarz	Kante Schwarz	Loch Schwarz
12	Kante Weiß	Loch Text	Loch Weiß	Glatt Schwarz	Glatt Matt	Loch Glanz	Glatt Glanz	Glatt Text	Kante Schwarz	Glatt Weiß	Kante Text	Loch Schwarz	Kante Matt	Loch Matt	Kante Glanz
13	Glatt Text	Glatt Weiß	Glatt Matt	Loch Glanz	Loch Weiß	Loch Matt	Loch Schwarz	Glatt Glanz	Glatt Schwarz	Kante Matt	Loch Text	Kante Glanz	Kante Weiß	Kante Text	Kante Schwarz
14	Loch Schwarz	Glatt Matt	Glatt Glanz	Glatt Text	Kante Glanz	Loch Weiß	Kante Matt	Loch Matt	Kante Weiß	Kante Text	Loch Glanz	Kante Schwarz	Loch Text	Glatt Schwarz	Glatt Weiß
15	Glatt Glanz	Kante Schwarz	Loch Schwarz	Glatt Matt	Glatt Text	Glatt Schwarz	Glatt Weiß	Kante Weiß	Loch Matt	Kante Glanz	Kante Matt	Kante Text	Loch Glanz	Loch Text	Loch Weiß

A.2. Fragebogen

Oberflächenanalyse zur Erkennung von geeigneten Projektionsflächen

1. Geschlecht

Welches Geschlecht haben Sie?

Wählen Sie alle zutreffenden Antworten aus.

- Weiblich
 Männlich

2. Alter

Wie alt sind Sie?

.....

3. Beruf/Studiengang

In welchem Bereich sind Sie tätig bzw. studieren Sie?

.....

Glatte Oberfläche

4. Weißes Objekt (Glatt)

Wie gut konnten Sie das Testbild erkennen?

Markieren Sie nur ein Oval.

	1	2	3	4	5	6	7	
sehr gut	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	sehr schlecht

5. Schwarzes Objekt (Glatt)

Wie gut konnten Sie das Testbild erkennen?

Markieren Sie nur ein Oval.

	1	2	3	4	5	6	7	
sehr gut	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	sehr schlecht

6. Beschriftetes Objekt (Glatt)

Wie gut konnten Sie das Testbild erkennen?

Markieren Sie nur ein Oval.

	1	2	3	4	5	6	7	
sehr gut	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	sehr schlecht

7. Glänzendes Objekt (Glatt)

Wie gut konnten Sie das Testbild erkennen?

Markieren Sie nur ein Oval.

	1	2	3	4	5	6	7	
sehr gut	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	sehr schlecht

8. Mattes Objekt (Glatt)

Wie gut konnten Sie das Testbild erkennen?

Markieren Sie nur ein Oval.

	1	2	3	4	5	6	7	
sehr gut	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	sehr schlecht

Kantenüberzogene Oberfläche

9. Weißes Objekt (Kanten)

Wie gut konnten Sie das Testbild erkennen?

Markieren Sie nur ein Oval.

	1	2	3	4	5	6	7	
sehr gut	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	sehr schlecht

10. Schwarzes Objekt (Kanten)

Wie gut konnten Sie das Testbild erkennen?

Markieren Sie nur ein Oval.

	1	2	3	4	5	6	7	
sehr gut	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	sehr schlecht

11. Beschriftetes Objekt (Kanten)

Wie gut konnten Sie das Testbild erkennen?

Markieren Sie nur ein Oval.

	1	2	3	4	5	6	7	
sehr gut	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	sehr schlecht

12. Glänzendes Objekt (Kanten)

Wie gut konnten Sie das Testbild erkennen?

Markieren Sie nur ein Oval.

	1	2	3	4	5	6	7	
sehr gut	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	sehr schlecht

13. Mattes Objekt (Kanten)

Wie gut konnten Sie das Testbild erkennen?

Markieren Sie nur ein Oval.

	1	2	3	4	5	6	7	
sehr gut	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	sehr schlecht

Gelochte Oberfläche

14. Weißes Objekt (Löcher)

Wie gut konnten Sie das Testbild erkennen?

Markieren Sie nur ein Oval.

	1	2	3	4	5	6	7	
sehr gut	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	sehr schlecht

15. Schwarzes Objekt (Löcher)

Wie gut konnten Sie das Testbild erkennen?

Markieren Sie nur ein Oval.

	1	2	3	4	5	6	7	
sehr gut	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	sehr schlecht

A. Appendix

16. Beschriftetes Objekt (Löcher)

Wie gut konnten Sie das Testbild erkennen?

Markieren Sie nur ein Oval.

	1	2	3	4	5	6	7	
sehr gut	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	sehr schlecht

17. Glänzendes Objekt (Löcher)

Wie gut konnten Sie das Testbild erkennen?

Markieren Sie nur ein Oval.

	1	2	3	4	5	6	7	
sehr gut	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	sehr schlecht

18. Mattes Objekt (Löcher)

Wie gut konnten Sie das Testbild erkennen?

Markieren Sie nur ein Oval.

	1	2	3	4	5	6	7	
sehr gut	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	sehr schlecht

Allgemeine Fragen

19. Welches Objekt schnitt besonders schlecht ab?

Bitte begründen Sie.

20. Welches Objekt schnitt besonders gut ab?

Bitte begründen Sie.

21. Würden Sie die Möglichkeit nutzen auf Objekt zu projizieren, falls keine freie Fläche zur Verfügung steht?

Die Alternative wäre das Schaffen von freiem Platz. Bitte begründen Sie.

Literaturverzeichnis

- [ASR13] ALVES, Ricardo ; SOUSA, Luis ; RODRIGUES, J.M.F.: PoolLiveAid: Augmented reality pool table to assist inexperienced players. In: *WSCG'13 21st International Conference on Computer Graphics, Visualization and Computer Vision 2013*, 2013 (Zitiert auf Seite 13)
- [BIWG07] BIMBER, Oliver ; IWAI, Daisuke ; WETZSTEIN, Gordon ; GRUNDHÖFER, Anselm: The Visual Computing of Projector-Camera Systems. In: *EUROGRAPHICS 2007*, 2007 (Zitiert auf den Seiten 12 und 40)
- [BK08] BRADSKI, Gary ; KAEHLER, Adrian: *Learning OpenCV*. O'Reilly, 2008 (Zitiert auf Seite 18)
- [BTG06] BAY, Herbert ; TUYTELAARS, Tinne ; GOOL, Luc V.: SURF: Speeded Up Robust Features. In: *Proceedings of the 9th European Conference on Computer Vision*, 2006 (Zitiert auf Seite 23)
- [Can86] CANNY, John: A Computational Approach to Edge Detection. (1986) (Zitiert auf Seite 21)
- [CL11] COWAN, Lisa G. ; LI, Kevin A.: ShadowPuppets: supporting collocated interaction with mobile projector phones using hand shadows. In: *CHI '11 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2011 (Zitiert auf Seite 15)
- [CLC⁺10] CHENG, Kai-Yin ; LIANG, Rong-Hao ; CHEN, Bing-Yu ; LIANG, Rung-Huei ; KUO, Sy-Yen: iCon: Utilizing Everyday Objects as Additional, Auxiliary and Instant Tabletop Controllers. In: *CHI '10 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2010 (Zitiert auf Seite 13)
- [EW13] ECHTLER, Florian ; WIMMER, Raphael: The Interactive Dining Table. In: *CHI'13 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2013 (Zitiert auf den Seiten 12 und 13)
- [FKS14] FUNK, Markus ; KORN, Oliver ; SCHMIDT, Albrecht: An Augmented Workplace for Enabling User-Defined Tangibles. In: *CHI'14 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2014 (Zitiert auf den Seiten 12, 13 und 16)
- [HA12] HARDY, John ; ALEXANDER, Jason: Toolkit Support for Interactive Projected Displays. In: *MUM '12 Proceedings of the 11th International Conference on Mobile and Ubiquitous Multimedia*, 2012 (Zitiert auf Seite 12)
- [HBW11] HARRISON, Chris ; BENKO, Hrvoje ; WILSON, Andrew D.: OmniTouch: wearable multitouch interaction everywhere. In: *UIST'11 Proceedings of the 24th annual ACM symposium on User interface software and technology*, 2011 (Zitiert auf Seite 14)

- [HSL⁺12] HUBER, Jochen ; STEIMLE, Jürgen ; LIAO, Chunyuan ; LIU, Qiong ; MÜHLHÄUSER, Max: Light-Beam: Interacting with Augmented Real-World Objects in Pico Projections. In: *MUM'12 Proceedings of the 11th International Conference on Mobile and Ubiquitous Multimedia*, 2012 (Zitiert auf Seite 15)
- [IWO⁺99] ISHII, Hiroshi ; WISNESKI, Craig ; ORBANES, Julian ; CHUN, Ben ; PARADISO, Joe: PingPong-Plus: design of an athletic-tangible interface for computer-supported cooperative play. In: *CHI '99 Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, 1999 (Zitiert auf Seite 13)
- [JGAK07] JORDE, Sergi ; GEIGER, Günter ; ALONSO, Marcos ; KALTENBRUNNER, Martin: The reacTable: exploring the synergy between live music performance and tabletop tangible interfaces. In: *TEI '07 Proceedings of the 1st international conference on Tangible and embedded interaction*, 2007 (Zitiert auf Seite 13)
- [KLKS12] KHALILBEIGI, Mohammadreza ; LISSERMANN, Roman ; KLEINE, Wolfgang ; STEIMLE, Jürgen: FoldMe: Interacting with Double-sided Foldable Displays. In: *TEI '12 Proceedings of the Sixth International Conference on Tangible, Embedded and Embodied Interaction*, 2012, S. 33–40 (Zitiert auf Seite 14)
- [KNF12] KLOMPMAKER, Florian ; NEBE, Karsten ; FAST, Alex: dSensingNI: a framework for advanced tangible interaction using a depth camera. In: *TEI '12 Proceedings of the Sixth International Conference on Tangible, Embedded and Embodied Interaction* Bd. 12 TEI, 2012, S. 217–224 (Zitiert auf den Seiten 12 und 13)
- [KS03] KARITSUKA, Toshikazu ; SATO, Kosuke: A Wearable Mixed Reality with an On-board Projector. In: *ISMAR '03 Proceedings of the 2nd IEEE/ACM International Symposium on Mixed and Augmented Reality*, 2003 (Zitiert auf Seite 14)
- [Lag11] LAGANIERE, Robert: *OpenCV 2 Computer Vision Application Programming Cookbook*. Packt Publishing, 2011 (Zitiert auf den Seiten 26 und 27)
- [LGJK11] LÖCHTEFELD, Markus ; GEHRING, Sven ; JUNG, Ralf ; KRÜGER, Antonio: guitAR: supporting guitar learning through mobile projection. In: *CHI '11 Extended Abstracts on Human Factors in Computing Systems*, 2011 (Zitiert auf Seite 14)
- [LM10] LINDER, Natan ; MAES, Pattie: LuminAR: Portable Robotic Augmented Reality Interface Design and Prototype. In: *UIST '10 Proceedings of the 23rd annual ACM symposium on User interface software and technology*, 2010 (Zitiert auf Seite 13)
- [MG09] MOLYNEAUX, David ; GELLERSEN, Hans: Projected interfaces: enabling serendipitous interaction with smart tangible objects. In: *TEI '09 Proceedings of the 3rd International Conference on Tangible and Embedded Interaction*, 2009 (Zitiert auf Seite 13)
- [Mir] MIRCOSOF (Hrsg.): *Kinect for Windows | Human Interface Guidelines v1.8*. Microsoft (Zitiert auf Seite 17)
- [NIH⁺11] NEWCOMBE, Richard ; IZADI, Shahram ; HILLIGES, Otmar ; MOLYNEAUX, David ; KIM, David ; DAVISON, Andrew ; KOHLI, Pushmeet ; SHOTTON, Jamie ; HODGES, Steve ; FITZGIBBON, Andrew: KinectFusion: Real-time 3D Reconstruction and Interaction Using a Moving

- Depth Camera. In: *UIST'11 Proceedings of the 24th annual ACM symposium on User interface software and technology*, 2011 (Zitiert auf Seite 12)
- [OTTT10] OTA, Shuhei ; TAKEGAWA, Yoshinari ; TERADA, Tsutomu ; TSUKAMOTO, Masahiko: A Method for Wearable Projector Selection that Considers the Viewability of Projected Images. In: *Computers in Entertainment (CIE) - Theoretical and Practical Computer Applications in Entertainment* (2010) (Zitiert auf Seite 14)
- [Pin01] PINHANEZ, Claudio S.: The Everywhere Displays Projector: A Device to Create Ubiquitous Graphical Interfaces. In: *Proceeding UbiComp '01 Proceedings of the 3rd international conference on Ubiquitous Computing*, 2001 (Zitiert auf Seite 11)
- [SJM13] STEIMLE, Jürgen ; JORDT, Andreas ; MAES, Pattie: Flexpad: Highly Flexible Bending Interactions for Projected Handheld Displays. In: *CHI '13 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2013 (Zitiert auf Seite 14)
- [WB10] WILSON, Andrew D. ; BENKO, Hrvoje: Combining Multiple Depth Cameras and Projectors for Interactions On, Above and Between Surfaces. In: *UIST '10 Proceedings of the 23rd annual ACM symposium on User interface software and technology*, 2010, S. 273–282 (Zitiert auf Seite 12)
- [Wel91] WELLNER, Pierre: The DigitalDesk calculator: tangible manipulation on a desk top display. In: *UIST '91 Proceedings of the 4th annual ACM symposium on User interface software and technology*, 1991 (Zitiert auf Seite 11)
- [Wil05] WILSON, Andrew D.: PlayAnywhere: A Compact Interactive Tabletop Projection-Vision System. In: *UIST'05 Proceedings of the 18th annual ACM symposium on User interface software and technology*, 2005 (Zitiert auf Seite 11)
- [Wil10] WILSON, Andrew D.: Using a Depth Camera as a Touch Sensor. In: *ITS'10 ACM International Conference on Interactive Tabletops and Surfaces*, 2010, S. 69–72 (Zitiert auf Seite 12)
- [WSDR14] WINKLER, Christian ; SEIFERT, Julian ; DOBBELSTEIN, David ; RUKZIO, Enrico: Pervasive Information through Constant Personal Projection: The Ambient Mobile Pervasive Display (AMP-D). In: *CHI'14 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2014 (Zitiert auf Seite 14)
- [WSM13] WILLIS, Karl D. D. ; SHIRATORI, Takaaki ; MAHLER, Moshe: HideOut: Mobile Projector Interaction with Tangible Objects and Surfaces. In: *TEI '13 Proceedings of the 7th International Conference on Tangible, Embedded and Embodied Interaction*, 2013 (Zitiert auf Seite 15)

Alle URLs wurden zuletzt am 27. 04. 2014 geprüft.

Erklärung

Ich versichere, diese Arbeit selbstständig verfasst zu haben. Ich habe keine anderen als die angegebenen Quellen benutzt und alle wörtlich oder sinngemäß aus anderen Werken übernommene Aussagen als solche gekennzeichnet. Weder diese Arbeit noch wesentliche Teile daraus waren bisher Gegenstand eines anderen Prüfungsverfahrens. Ich habe diese Arbeit bisher weder teilweise noch vollständig veröffentlicht. Das elektronische Exemplar stimmt mit allen eingereichten Exemplaren überein.

Ort, Datum, Unterschrift