

BACHELOR THESIS
Niklas Timo Kück

Optische Positionsbestimmung eines Videolaryngoskops im Rachen einer Reanimationspuppe

FAKULTÄT TECHNIK UND INFORMATIK
Department Informatik

Faculty of Engineering and Computer Science
Department Computer Science

Niklas Timo Kück

Optische Positionsbestimmung eines Videolaryngoskops im Rachen einer Reanimationspuppe

Bachelorarbeit eingereicht im Rahmen der Bachelorprüfung
im Studiengang *Bachelor of Science Informatik Technischer Systeme*
am Department Informatik
der Fakultät Technik und Informatik
der Hochschule für Angewandte Wissenschaften Hamburg

Betreuender Prüfer: Prof. Dr. Thomas Lehmann
Zweitgutachter: Prof. Dr.-Ing. Roman Kusche

Eingereicht am: 23. August 2024

Niklas Timo Kück

Thema der Arbeit

Optische Positionsbestimmung eines Videolaryngoskops im Rachen einer Reanimationspuppe

Stichworte

Videolaryngoskopie, Anästhesie, Markererkennung, ArUco, Pose Estimation

Kurzzusammenfassung

Diese Bachelorarbeit versucht unter Verwendung eines Videolaryngoskop die Position der Kamera im Rachen einer Reanimationspuppe zu bestimmen. Hierzu sollen optische Mustererkennungsverfahren verwendet werden. Die Arbeit soll einen Algorithmus zur Positions- und Ausrichtungsbestimmung liefern, der im Endeffekt Realbilder anstelle von Bildern aus der Reanimationspuppe anzeigen kann.

Niklas Timo Kück

Title of Thesis

Optical position determination of a video laryngoscope in the throat of a resuscitation manikin

Keywords

videolaryngoscopy, anesthesia, marker detection, ArUco, pose estimation

Abstract

This bachelor thesis attempted to determine the position of the camera in the throat of a resuscitation manikin using a video laryngoscope. For this purpose, optical pattern recognition methods should be used. The work aims to provide an algorithm for position and orientation determination, which can ultimately display real images instead of images from the resuscitation manikin.

Inhaltsverzeichnis

Abbildungsverzeichnis	vi
Abkürzungen	vii
1 Einleitung	1
1.1 Motivation	1
1.2 Zielsetzung der Arbeit	2
1.3 Spezifikation des Entwicklungsziels	3
1.4 Das Simulationszentrum im UKE	4
2 Medizinischer Hintergrund	5
2.1 Reanimationspuppen und Kopfphantome	5
2.2 Laryngoskopie	6
3 Analyse	8
3.1 Idee des Projekts	8
3.2 Videolaryngoskopie im Simulationszentrum	9
3.3 Spezielle Zielgruppen	12
3.4 Funktionale Anforderungen	12
3.5 Teilziele und Beweis der Usability von einzelnen Techniken	13
4 Technische Grundlagen	16
4.1 Kamera	16
4.2 Optische Erkennungstechniken	17
4.2.1 Vorteile	19
4.2.2 Nachteile	19
4.3 Marker und Pose Estimation	19

5	Prototyp	21
5.1	Physischer Prototyp und Randbedingungen	21
5.1.1	Laryngoskop	21
5.1.2	Endoskopkamera	23
5.1.3	Entwicklung am 2D-Modell	24
5.2	Lösungsansatz	26
5.3	Systementscheidungen	27
5.3.1	Minimum Viable Product	27
5.3.2	ArUco-Marker	27
5.3.3	Algorithmen	30
5.4	Implementierung	31
5.4.1	Entwurf Marker in Realumgebung	31
5.4.2	Detektion	34
5.4.3	Analyse	36
5.4.4	Kartierung	38
5.4.5	Darstellung im 3D-Programm	40
5.4.6	Entwicklung 3D-Karte	42
5.4.7	Kamerapositionsberechnung	43
5.4.8	Probleme und Lösungsansätze	48
5.5	Anpassungen nach erster Prototypvorstellung	48
6	Ergebnisse	50
7	Ausblick	52
8	Fazit	54
	Literaturverzeichnis	55
A	Anhang	57
A.1	Digitaler Anhang	57
A.2	Verwendete Hilfsmittel	58
A.3	Bedienungsanleitung Videolaryngoskopsoftware	58
A.4	Bedienungsanleitung Kamerakalibrierung	59
	Glossar	60
	Selbstständigkeitserklärung	61

Abbildungsverzeichnis

1.1	Darstellung des Einsatzes eines Videolaryngoskop	2
5.1	3D-Druck Prototyp des Videolaryngoskop	22
5.2	Vollständiges Prototyp-Videolaryngoskop	24
5.3	Durchsichtiges Kopfschnittmodell	25
5.4	Ablaufdiagramm zur Systemkonzeption: Berechnung der Position der Kameralinse und 3D-Darstellungsform (eigene Darstellung)	26
5.5	Erstentwurf der Anordnung von Markern	32
5.6	Endentwurf der Anordnung von Markern	33
5.7	Markeroutput im Graubild	34
5.8	Markeroutput im Binärbild	35
5.9	Markererkennung ohne Closing-Effekt	36
5.10	Markererkennung mit Closing-Effekt	38
5.11	Erster Darstellungsversuch der Marker im Weltkoordinatensystem ohne Kartierung im 3D-Programm	40
5.12	Versuch einer Nachbildung des relevanten Bereichs des Kopfschnittmodells	42
5.13	Test der Winkel des Marker 7: Linksdrehung, Normalstellung, Rechtsdrehung	44
5.14	Entfernungen von Markern zur Kamera im Durchlauf des Rachens	46
5.15	Entfernungen von Marker zur Kamera im Durchlauf des Rachens, charakteristischer Graph	47
6.1	Demonstration der Funktionalität des Prototyps: Bildinjektion und Positionsdarstellung	51
6.2	Demonstration der Funktionalität des Prototyps an einer Reanimationspuppe	51

Abkürzungen

AR Augmented Reality.

FPS Frames per Second.

ID Identifikator.

MVP Minimum Viable Product.

SLAM Simultaneous Localization and Mapping.

TCP Transmission Control Protocol.

UDP User Datagram Protocol.

UKE Universitätsklinikum Hamburg-Eppendorf.

VR Virtual Reality.

1 Einleitung

1.1 Motivation

In der Medizin funktioniert nichts ohne das richtige Werkzeug und die Kompetenz der ausführenden Fachkraft. Um Kompetenzen in den entsprechenden Fachbereichen der Medizin zu erlangen, ist eine gute Ausbildung und ein wiederkehrendes Training nötig.

So wird diese Ausbildung oder das Training dazu in Simulationszentren durchgeführt. In diesen werden praktische und theoretische Aktivitäten an analogen, aber mittlerweile auch immer mehr digitalen menschlichen Phantomen durchgeführt.

Anästhesisten trainieren beispielsweise das korrekte Setzen eines Endotrachealtubus oder die richtige Durchführung des Beginns einer Anästhesie. Im Universitätsklinikum Hamburg-Eppendorf (UKE) werden im dortigen Simulationszentrum für diese Aufgabe lebensgroße Reanimationspuppen oder Kopfphantome eingesetzt. Solche Puppen oder Phantome können im Training für die unterschiedlichsten Zwecke im konkreten realitätsnahen Umgang mit einem Patienten verwendet werden. Meist besitzen diese Trainingsgeräte viele unterschiedliche Spezialfunktionen, die die Immersion, das Gefühl des Arbeitens mit einem echten Patienten in realer Umgebung perfektionieren. Unterstützend gibt es häufig für die spätere Analyse digitale Ergebnisauslesemöglichkeiten oder direktes realitätsnahes Feedback.

Die Ausstattung einer solchen Reanimationspuppe oder eines solchen Kopfphantoms ist, so gut es simulativ umsetzbar ist, einer echten Person nachempfunden. Dies schließt hierbei auch die Nachbildung des Rachens und der Atemwege mit ein. Entsprechend wurden standardisierte Trainingsobjekte für dieses Einsatzgebiet entworfen.

Bei der Durchführung einer Anästhesie muss nicht nur das Setzen eines Endotrachealtubus trainiert werden, sondern vorab auch der Einsatz eines Videolaryngoskops zur Kontrolle des Rachens. Hierbei soll der Patient auf eventuelle Verletzungen oder Erkrankungen im Rachen geprüft werden, die bei der Durchführung der Anästhesie beachtet werden müssen. Das Videolaryngoskop wird dabei über den Mund, dem Rachen und der



Abbildung 1.1: Darstellung des Einsatzes eines Videolaryngoskop

Speise- und Luftröhre folgend eingeführt und liefert an einen angeschlossenen Monitor ein Livebild aus dem Rachen der Puppe. Eine Durchführung dessen ist in Abbildung 1.1 dargestellt. Aufgrund der Standardisierung der Puppen ist es verständlich, dass bei jeder Puppe und jedem Trainingseinsatz immer dieselben Bilder aus dem Rachen zu sehen sind. Diese sind jeweils die aktuellen Bilder von unterschiedlichen Rachennachbildungen aus den Puppen. So lassen sich leider mit diesen Phantomen keine angepassten Verhaltensweisen bei Abweichungen von der Normalität trainieren, die von den Anästhesisten und Rettungskräften bei Verletzungen oder Erkrankungen erwartet werden. Es ist wie beschrieben nur ein spezifisches Bild pro Puppe vorhanden. Um aktuell dennoch Variationen des Rachens zu erreichen, werden von Trainern Puppen modifiziert, die dann teilweise anderes Aussehen haben und ermöglichen, eine angepasste Arbeitsweise trainieren zu können.

1.2 Zielsetzung der Arbeit

Diesem Problem, der immergleichen Trainingssequenzen durch die Uniformität der Puppen, möchte diese Bachelorarbeit entgegenreten. Es soll Abhilfe verschaffen werden, ein größeres und noch immersiveres Trainingserlebnis zu gestalten.

Das Ziel dieser Arbeit wird es nicht sein, physisch an der Reanimationspuppe oder dem Kopfphantom eine Verbesserung oder Anpassung zu entwickeln, beispielsweise im Sinne

von der Erweiterung von mehreren Trainingsobjekten, die gegeneinander austauschbar sein können, sondern das bestehende Trainingssetup in der aktuellen Ausführung zu nutzen und mit digitalen Möglichkeiten eine Erweiterung und Verbesserung der Immersion zu erreichen.

Der digitale Ansatz kann unter anderem hierzu das Kamerabild verwenden. Die grundsätzliche Idee besteht darin, den Videolivestream des Videolaryngoskops zu nutzen und statt der Bilder aus der Reanimationspuppe Realbilder von unterschiedlichen Situationen und Veränderungen von echten Menschen aus dem Rachen als Ausgabe zu liefern. Weiterführend kann dann beispielsweise bei jedem Training das UKE diverse unterschiedliche Bildmaterialien von Realfällen einspielen. Diese Anpassungen lassen sich folglich für jedes Training spezifisch erstellen und nutzen. Die konkrete Idee besteht nun darin, über das Kamerabild des Videolaryngoskops eine Positionserkennung durchzuführen. Dies soll im Sinne möglicher Augmented Reality (AR)-Techniken stattfinden können. Weiterhin soll dann abschließend die Möglichkeit bestehen, in Abhängigkeit der Position ein manipuliertes Bild des Videolivestreams dem Trainee anzuzeigen.

Vorteil dieses Lösungsansatzes ist, eine kompetenz- und praxisorientierte Aus- und Fortbildung von Fachkräften im Umgang mit Patienten zu schaffen. Ziel hierbei ist, die häufig problematische und eingeschränkte Übungskünstlichkeit zu verringern.

1.3 Spezifikation des Entwicklungsziels

Die Umsetzung dieses Ansatzes besteht aus zwei Kernprozessen, um ein solches „AR-Videolaryngoskop“ in einer möglichen Form des praktischen Einsatzes verwenden zu können. Diese gliedern sich in einen Abschnitt der Lokalisation beziehungsweise der Positionsbestimmung eines Videolaryngoskops im Rachen der Reanimationspuppe oder des Kopfphantoms und der aus dieser Information zu gestaltenden 3D-Karte zur Darstellung und Übersetzung der Realbilder auf das dann veränderte Livebild des Videolaryngoskops.

Diese Bachelorarbeit wird sich ausschließlich mit der Entwicklung des ersten Kernprozesses der Positionsbestimmung im Rachen beschäftigen. Hierbei wird es sich vor allem darum drehen, einen möglichen Prototyp¹ für diesen Abschnitt zu erstellen. Fokus der

¹Wenn in der Arbeit von einem Prototyp gesprochen wird, ist damit das Gesamtprodukt gemeint. Dies schließt Teile der Software und auch physischen Entwicklungen mit ein.

Arbeit soll auf der optischen Positionsbestimmung liegen, weil die Idee ist, das bestehende Videolaryngoskop in seiner realen Form zu verwenden. Speziell ist der Versuch geplant, eine geeignete Technik zur Positionsbestimmung zu finden, sodass in einem weiteren Projekt aus den 3D-Informationen das tatsächliche, manipulierte Videobild zur Ideeervollständigung produziert werden kann. Trotzdem soll eine Kurzdemonstration dieses Entwicklungsschritts zur möglichen Gesamtfunktionalität auch in dieser Arbeit erfolgen.

1.4 Das Simulationszentrum im UKE

Diese Bachelorarbeit findet in Kooperation mit dem Simulationszentrum der Klinik für Anästhesiologie des UKE mit den Ansprechpartnern Dr. med. Philipp Breitfeld und Dr. med. Martin Petzoldt statt. Das UKE tritt in dieser Arbeit in Form eines Kunden auf, da die konkreten Anforderungen und Anpassungen in einem teilweisen interdisziplinären Austausch stattfinden. Das Simulationszentrum bietet neben vollausgestatteten Trainingsoperationsräumen, Trainingsrettungswagen bereits auch schon eine erprobte Virtual Reality Umgebung, die durch ein Projekt der HAW Hamburg entstanden ist, weshalb diese Arbeit sich gut in dieses Portfolio einreicht.

2 Medizinischer Hintergrund

Diese Bachelorarbeit ist stark an der Entwicklung von medizinischem Gerät und spezieller Software für eine Weiterentwicklung dessen orientiert. Für einige dieser Elemente ist dennoch eine Einführung in medizinische Aspekte nötig, um die Zusammenhänge bezüglich der Entwicklung eines möglichen Prototyps und der gewählten Techniken und Einschränkungen analysieren und anwenden zu können. Dieser Abschnitt beleuchtet diese Aspekte und führt in einige dieser Themengebiete ein.

2.1 Reanimationspuppen und Kopfphantome

Die Ausbildung von Medizinern, Rettungskräften oder Pflegepersonal benötigt neben theoretischen Inhalten insbesondere viel Training und Übung bei praktischen Aktivitäten im Zusammenhang mit Menschen. Eine lebensrettende Maßnahme nur theoretisch zu beherrschen, ist in kritischen Momenten nicht ausreichend, weshalb für das Training Puppen und Phantom des menschlichen Körpers herangezogen werden.

Solche Modelle gibt es in verschiedenster Ausführung aufgrund unterschiedlicher Trainingsziele. Reanimationspuppen bilden in der Regel den gesamten Körper ab, sodass eine vollständige und authentische Handhabung eines Patienten im Trainingsfall realisiert werden kann. Teilweise gibt es auch Varianten dieser Modelle, die für spezielle Trainingszwecke reduzierte Körperteile bereitstellen. Diese Modelle bieten neben der anatomischen Nachbildung des menschlichen Körpers häufig auch weitere technische und mechanische Modifikationen, die den realistischen Gegebenheiten und Verhaltensweisen nachempfunden sind. Mögliche Modifikationen reichen von digitalem Feedback bis hin zu Versionen, mit denen ein Zugang für Infusionen oder Medikamente gesetzt werden kann.¹

¹Die Begrifflichkeiten Reanimationspuppe, Puppe, Kopfphantom oder Phantom werden in dieser Arbeit teilweise als Synonyme verwendet, überwiegend jedoch mit dem Einsatz der spezifischen Einzelbedeutung des Begriffs.

2.2 Laryngoskopie

Atemwegssicherung spielt in der Medizin immer dann eine Rolle, wenn ein Patient sich durch unterschiedliche Umstände in einem bewusstlosen Zustand befindet. Dieser kann durch diverse Ursachen wie Gifte, Herz-Kreislauf-Probleme oder andere vom Körper gesteuerte Dinge Prozesse ausgelöst werden. Des Weiteren ist es möglich, den bewusstlosen Zustand auch bewusst in der Anästhesie zu erreichen, insbesondere dann, wenn bei einem Patienten bei einer Operation eine Vollnarkose eingeleitet werden muss. Bewusstlosigkeit hat immer das Problem, dass die entsprechende Person die Körperspannung verliert und damit alle Muskeln erschlaffen. Der Bereich der oberen Atemwege birgt hierbei ein großes Problem. Da die Zunge ebenfalls ein Muskel ist, kann im einfachsten und auch lebensbedrohlichen Fall die Zunge des Patienten im erschlafften Zustand in den Rachen rutschen und die Luftzufuhr und Verbindung zur Luftröhre verschließen. Zur notwendigen Atemwegssicherung wird daher die Intubation des Patienten angewandt. Einfach formuliert wird über den Mund, den Rachen und dann in die Luftröhre ein Schlauch eingeführt und damit dafür gesorgt, dass es einen durchgängigen Weg für die Beatmung der Person gibt. Dies findet entweder supraglottisch, also oberhalb der Stimmlippen, oder endotracheal, in der Luftröhre, statt. Bei der endotrachealen Variante, die primär in der Anästhesie verwendet wird, werden Hilfsmittel benötigt, um den sogenannten Endotrachealtubus problemlos an die richtige Stelle legen zu können.

Ein Laryngoskop wird verwendet, um den Tubus an die richtige Stelle zu positionieren und die Luftröhre damit freizuhalten. Der Einsatz eines Laryngoskops zur Atemwegssicherung bei der Endotrachealtubusanlage ist essenziell notwendig. Die Aufgabe dieses Werkzeugs besteht darin, dem Benutzer einen besseren Blick in den Rachen und damit auch den Kehlkopf bieten zu können. Dies geschieht dadurch, dass durch das Anheben der Zunge oder Strecken des Rachenraums der relevante Bereich für den Benutzer sichtbar gemacht und freigelegt wird.

Laryngoskope bestehen grundsätzlich aus einem Griff und angeschlossenem Spatel. Durch unterschiedliche Hersteller können sich die Spatelformen und Typen teilweise stark unterscheiden, sodass es Varianten gibt, die wiederverwendbar oder nur einmal zu gebrauchen sind. Der Aufbau eines Spatels ist dennoch dahingehend gleich, dass diese halbrund geformt und an der Spitze zusätzlich verstärkt sind. An dem Spatel befindet sich eine einseitige Erhöhung, welche als Führungsschiene für den Endotrachealtubus verwendet wird. An der vorderen Seite dieser Erhöhung befindet sich zudem eine Lichtquelle zur

Beleuchtung des Rachens. Ein weit verbreiteter Spateltyp in diesem Einsatzbereich ist der Macintosh.[5, 6]

Seit einigen Jahren hat sich die Medizin auch in diesem Fachbereich der Atemwegssicherung weiterentwickelt, sodass festgestellt wurde, dass bei dieser sogenannten direkten Laryngoskopie nicht alle Bereiche des Rachens immer gut einsehbar sind und der Einsatz eines klassischen Laryngoskops viel Übung und Training benötigt. Zur Behebung dieses Problems wurden Videolaryngoskope entwickelt. Jene erhöhen durch die Visualisierung der Kehlkopföffnung signifikant den Erfolg der Intubation. Durch die Erweiterung einer Endoskopkamera, die sich an der Spitze des Spatels befindet, kann an einen externen Monitor ein Livebild des Rachens geliefert werden. Die Unterstützung eines Livebilds erleichtert die korrekte Positionierung des Endotrachealtubus, da ein geringeres Training zur sicheren Handhabung vonnöten ist. Zusätzlich ermöglicht eine Visualisierung eine erweiterte Untersuchung des Rachens auf Verletzungen oder Erkrankungen. [5, 6]

3 Analyse

3.1 Idee des Projekts

Dadurch, dass die Idee des Projekts aus Erfahrungen und Schwierigkeiten aus der Praxis stammt, ist in Teilen ein interdisziplinärer Austausch für die Entwicklung eines Prototyps nötig. Neben technischen Eigenheiten und Einschränkungen müssen in der Entwicklung dennoch zwingend die realitätsnahen Haptiken im Umgang mit der Test- und Trainingsumgebung beibehalten werden, sodass die gelernten Handlungsweisen problemlos in Real-einsätzen eingesetzt werden können.

Dieser Herausforderung und Idee nimmt sich diese Bachelorarbeit nun an, ein verbessertes Intubationstraining mit einem Videolaryngoskop zu ermöglichen. Für die Umsetzung sind allerdings einige Analysen der Machbarkeit und Umsetzungsmöglichkeiten notwendig, da der geforderte Einsatz des Videolaryngoskops und der Reanimationspuppen einige Schwierigkeiten in der Kombination mit erweiterter Technik mitbringt. Die Grundidee des Projekts besteht nun deshalb darin, dass das reale Videolaryngoskop verwendet werden soll. An dem direkt angeschlossenen Monitor sollen dann in der Folge die entsprechenden Realbilder eines Menschen abhängig von der Position im Rachen der Trainingspuppe abgespielt werden. Für diesen Einsatz besteht nun die Idee darin, in erster Linie die direkt zur Verfügung stehenden Sensoren des Videolaryngoskops zu verwenden. Eine Erweiterung mit zusätzlichen Modifikationen an den Sensoren und auch des Videolaryngoskops soll verhindert werden, weil das oberste Ziel ist, das Gefühl geben zu können, nur mit dem realen, echten Objekt zu arbeiten.

Die Untersuchung des Videolaryngoskops stellt folglich nur eine Kamera als Sensor mit dem Livebild aus dem Rachen der Puppe oder dem Phantom zur Verfügung. Dieser Umstand sorgt also für eine Einschränkung dadurch, dass nur eine Form von Daten zur Verarbeitung und Berechnung genutzt werden. Mehrere Sensoren könnten jedoch sich teilweise gegenseitig unterstützen und auch Rückfallebenen bilden. Dies ist hier aber

nicht möglich, weshalb dieser Teil dazu führt, dass bei dem Videolaryngoskop nur Live-bilder der Kamera zur Analyse der möglichen Positionierung im Rachen nutzbar sind. Das Projekt wird aufgrund dieser Erkenntnis und Einschränkung auf einen optisch basierten Lösungsansatz festgelegt. Entsprechend spielen die Kamera des Videolaryngoskops und die physische Gestaltung des Rachens in der Reanimationspuppe eine bedeutsame Rolle zur Positionsbestimmung.

Weitere spezielle Anforderungen für einen möglichen Prototyp entstehen nun auch aus dem Umgang mit dem Videolaryngoskop und dem Arbeiten an einer Reanimationspuppe oder in einem Menschen. Allerdings kann direkt angemerkt werden, dass diese Arbeit nicht für einen Einsatz an Menschen, sondern nur an Trainingsgeräten gedacht ist. Dies würde bei einer möglichen Zertifizierung als Medizinprodukt eine Rolle spielen [10]. Medizinprodukte müssten dementsprechend zusätzliche Regularien und Vorgaben zum sicheren Einsatz einhalten. Da dieses Projekt jedoch keinen Einsatz bei oder in Menschen anstrebt, sondern nur in einem abgetrennten Trainingsbereich genutzt werden soll, kann auf diese spezifischen Vorgaben verzichtet werden. Trotzdem sollten die Regularien teilweise beachtet und umgesetzt werden, im Sinne des Einhaltens der realitätsnahen Aspekte.

3.2 Videolaryngoskopie im Simulationszentrum

Im Simulationszentrum des UKE stehen mehrere Videolaryngoskope, Kopfphantome und Puppen unterschiedlichster Hersteller zur Verfügung. Für die Entwicklung des „AR-Videolaryngoskop“ wird ein Ganzkörperphantom, das heißt eine Puppe mit intubierbarem Kopf und vollständigem Körper, bereitgestellt, sodass eine Analyse der physischen Gegebenheiten an diesem durchgeführt werden kann. Als Videolaryngoskop steht ein C-MAC[®] Videolaryngoskop 8401 BX mit einem C-MAC[®] Monitor 8403 ZX von dem Hersteller Storz zur Verfügung. Die gewählten Produkte sind bereits in Abbildung 1.1 zu sehen. Die Wahl des Ganzkörperphantoms ist dadurch bedingt, dass der Prototyp in erster Linie für ein Training im Übungsoperationsraum entwickelt und die grundsätzliche Idee des Projekts in dieser Umgebung angesiedelt wird. Zudem hat diese Puppe einen vollständig der Realität nachempfundenen Rachen. Weiterhin standen auch Puppen zur Verfügung, die nur aus dem Kopf bestanden.

Das gewählte Videolaryngoskop von Storz ist ein wiederverwendbares Produkt, welches in der Ausstattung des Simulationszentrums vorhanden ist. Die korrekte Handhabung eines Videolaryngoskops muss trainiert sein, weshalb der Einsatz eines Videolaryngoskops

an Menschen nur von Anwendern eingesetzt wird, die über eine entsprechende fachliche Qualifikation und Einweisung auf das jeweilige Gerät verfügen. Das gewählte Produkt ist vor Ort im Trainingsgebrauch eingesetzt. Die Funktionsweise und Anwendung wird nun zur Analyse einmal dargestellt.

Ein Videolaryngoskop wird einhändig gegriffen und über dem Mund zum senkrechten Einführen in den Rachen positioniert. Vorab wird sich hinter den Kopf positioniert. Nach dem Öffnen des Mundes, meist mit einem zusätzlichen Auseinanderdrücken der Schneidezähne, wird der Spatel des Videolaryngoskops mit leichtem Druck entlang der Zunge eingeführt. Bei der hier genutzten Reanimationspuppe ist zusätzlich dafür zu sorgen, dass die Zähne stärker auseinander gedrückt werden müssen, da die für diese Phantome verwendeten Materialien in diesen Bereichen aus festeren Gummiarten produziert sind. Durch das Einführen mit einer nach vorne gerichteten Bewegung wird dafür gesorgt, dass im Kamerabild im Verlauf der Einführungsbewegung der Rachen entlang des Gaumens über das Zäpfchen bis hin zur Trachea (Lufttröhre) sichtbar ist. Durch die fortlaufende Bewegung werden weiter vorne im Rachen befindliche Abschnitte über die Zeit aus dem Sichtfeld verschwinden. Durch ein Kippen und Drehen der Hand am Videolaryngoskop kann in einigen Abschnitten der Betrachtungsfokus auch auf Bereiche außerhalb der geradlinigen Bewegung des Videolaryngoskops erweitert werden, wenn der Platz im Rachen diese zulässt. [5, 6]

Das verwendete Videolaryngoskop von Storz besteht aus einem Laryngoskopspatel des Typs Macintosh mit einer Halterung für das Elektronikmodul. Zusätzlich ist eine Führung für einen Saugkatheter vorhanden. Das Elektronikmodul ist mit einem Anschlusskabel für einen externen Videomonitor versehen. Dieses ist nur mit dem dazugehörigen Monitor von Storz kompatibel. Folgende technische Eigenschaften des C-MAC® Videolaryngoskop mit Monitor sind von Bedeutung [7]:

- Die Kamera und eine LED an der Spitze des Spatels sind übereinander angeordnet.
- Die Software des Videolaryngoskops stellt das Bild automatisch für den angeschlossenen HD-Monitor scharf.
- Für einen adäquaten Einsatz ist eine standardmäßig ausreichende Bildqualität nötig.
- Das Videolaryngoskop kann über eine in den Monitor eingesteckte SD-Karte Videos und Bilder aufnehmen, sodass diese dann auf einem anderen Gerät angesehen wer-

den können. Weiterhin ist ein HDMI-Ausgang für die Verarbeitung von Livebildern vorhanden.

Für die Entwicklung des Prototyps in Verbindung des veränderten Videolaryngoskop und der Kombination der Reanimationspuppe leiten sich nun aus den Nutzungsanweisungen und den Spezifikationen folgende Einschränkungen, Anforderungen und Vorgaben ab:

- Aufgrund der variablen Gummifestigkeit der Materialien im Rachen der Reanimationspuppe, die nur in Teilen der Realität entspricht, können keine dauerhaften, vollständig, absolut im Raum definierten Markierungsstellen zur Positionsbestimmung der Kameralinse genutzt werden. Positionsänderungen müssen fortlaufend die gleichen Ergebnisse liefern. Bewegungen im Rachen verändern dauerhaft die Form, sodass keine feste Anpassung des Bereichs vorhanden ist. Entsprechend stellen weiche Bereiche ein Problem dar. Eine Berechnung der Position in Abhängigkeit von Linse und tatsächlichem Bild muss anhand dieser Teile umgesetzt werden. Eine jeweils aktuelle Messung dieser Positionen ist nötig.
- Die Problematik der auseinander zu drückenden Zähne kann dazu genutzt werden, die Zähne als Beginn des Übergangs vom Livebild der Kamera zu dem später 3D-gerenderten Bild der Realbilder zu verwenden. Zusätzlich können markante Stellen im Rachen als Hilfe zur Positionsbestimmung genutzt werden. Gleichzeitig steht in diesem Fall eine Prüfung der Anzahl der nötigen Markierungsstellen aus.
- Der Einsatz eines Videolaryngoskops mit einem externen Bildschirm ist für die Entwicklung des Prototyps vorteilhafter, da hiermit eine Trennung von Aufgabenbereichen erreicht würde. Die Trennung bezeichnet konkret die Bildverarbeitung und das 3D-Rendering. Dennoch wird bei der Trennung eine Einschränkung hingenommen, da eine zusätzliche Hand-Augen-Koordination des Bedienenden bei dieser Art erforderlich ist, als wenn sich der Bildschirm direkt am Spatel befindet.
- Für eine Verarbeitung des Videobildes ist ein direkter Zugriff auf die entsprechende Kamera zur Auslese der Bilddaten notwendig. Ein Zugriff auf ein Livebild ist über den externen Monitor des Videolaryngoskops nicht möglich. Der Zugriff über den HDMI-Anschluss ist nicht ohne zusätzliches Videocapturing mit der Verwendung von speziellen Geräten möglich. Damit eine Kamera hier verwendet werden kann, muss diese als aktives, angeschlossenes Gerät an einem Computer erkannt werden. Ein Einsatz des C-MAC® Videolaryngoskops für das Projekt ist damit nicht möglich.

- Eine Kamera mit relativ hoher Bildqualität ist notwendig, um relevante Bereiche für die Positionsbestimmung scharf stellen zu können.

Entsprechend dieser Analyse ergeben sich einige Grundanforderungen und Festlegungen für die Entwicklung des Prototyps. Jene beziehen sich damit auf physische Gegebenheiten für das Videolaryngoskop, den Algorithmus und die Software zur Positionsbestimmung.

Das Ziel der grundsätzlichen Idee, ein bestehendes Videolaryngoskop wie das C-MAC® Videolaryngoskop 8401 BX als Basis zu verwenden, muss aufgrund der Schwierigkeit der Livevideobildübertragung und Auslese der Kamera verworfen werden. Für die Prototypentwicklung wird auch in Absprache mit den Beteiligten im UKE ein eigenes Videolaryngoskop gebaut und verwendet werden müssen. Dabei sind die haptischen und technischen Eigenschaften des realen Objekts einzuhalten. Dies führt weiterhin dazu, dass zur Positionsbestimmung als Sensoren nur eine Kamera mit entsprechender Beleuchtung an der Spitze des Laryngoskopspatels zu verwenden ist. Mit dieser Einschränkung stehen damit absolut zur Positionsbestimmung nur Methoden auf Basis von Bildern zur Verfügung. Der Einsatz von Objekterkennung anhand von Mustern, Markern und markanten Stellen soll verfolgt werden. Für diese Umsetzung sind Veränderungen und Modifikationen an den genutzten Puppen oder Phantomern notwendig. Dies kann beispielsweise die Platzierung von Markierungsstellen in den genannten Puppen sein. [5, 6, 7]

3.3 Spezielle Zielgruppen

Dieses „AR-Videolaryngoskop“ richtet sich aufgrund des speziellen Einsatzgebietes an Personen, die im medizinischen Bereich arbeiten und mit dem Fachgebiet der Anästhesie oder der endotrachealen Intubation in Kontakt stehen. Dies sind unter anderem Studierende der Medizin, Rettungskräfte, Pflegepersonal und ärztliches Personal. Neben dieser Gruppe kann die Arbeit auch für Personen interessant sein, die sich für Ideen zur Positionsbestimmung in engen und nahen Bereichen interessieren.

3.4 Funktionale Anforderungen

Neben den technischen Anforderungen und Vorgaben aus dem Umgang und der Handhabung mit dem Videolaryngoskop, sowie der nötigen Kopfphantome, ergeben sich weitere spezielle funktionale Anforderungen aus der Idee und der tatsächlichen Zielsetzung des

Projekts. Diese spezifischen funktionalen Anforderungen, bezogen auf die Funktionalität der Software, werden im Folgenden definiert:

- Die Software soll aus dem Kameralivebild des Videolaryngoskops die genaue und spezifische Position der Kameralinse im Rachen exakt berechnen.
- Die Berechnung muss neben der Position auch eine Ausrichtung des Kamerablickfelds zurückgeben. Beide Werte, die Position und Drehung, sollen eine hohe Genauigkeit aufweisen, sodass eine einwandfreie Hand-Augen-Koordination des Bedienenden am Laryngoskop sichergestellt ist.
- Die Hand-Augen-Koordination muss eine Echtzeitreaktion bei Bewegung und Drehung des Videolaryngoskops ermöglichen. Die Reaktionszeiten zur erneuten Berechnung einer Position und Ausrichtung dürfen nur so lang benötigen, wie der Bediende braucht, die aktuelle Position bei Benutzung auf Relevanz¹ der Intuition zu analysieren.
- In Abhängigkeit der Berechnung der jeweiligen Position und Drehung (Ausrichtung) im Rachen muss eine Darstellung in Form einer Koordinatenrepräsentation in einem Weltkoordinatensystem erstellt werden. Diese Darstellungsform soll für die Umsetzung in einem 3D-Rendering verwendet werden.
- Wenn das Laryngoskop nicht bewegt wird, soll keine Veränderung der Position erfolgen.
- Die Software soll für einen einfachen Bedienungseinsatz im Training optimiert sein.

3.5 Teilziele und Beweis der Usability von einzelnen Techniken

Mit den Anforderungen, Vorgaben und Einschränkungen wird in der folgenden Arbeit nun versucht werden, einen funktionsfähigen Prototyp in Form von einem Videolaryngoskop und dazu programmierter Software zu entwickeln. Entstehen soll ein verbessertes Trainingslaryngoskop mit digitaler und technischer Unterstützung zur Anpassung der Darstellungsinhalte im Rachen einer Reanimationspuppe. Aus den definierten Vorstellungen und Anforderungen ergeben sich für den Ablauf der Arbeit und Entwicklung

¹Gemeint ist hiermit die Möglichkeit eine Verletzung, Erkrankung oder Anpassung der Nutzung des Laryngoskops im Rachen an der jeweiligen Position erkannt zu haben.

einige Teilziele, die erreicht und getestet werden sollen. Diese leiten sich hauptsächlich aus den funktionalen Anforderungen ab. Bei der Entwicklung des Prototyps steht hierbei neben der Softwareentwicklung zur Positionsbestimmung zusätzlich die Gestaltung der physischen Komponenten wie des speziell angepassten Videolaryngoskops und der nötigen Anpassungen für Markierungsstellen in der Reanimationspuppe an.

Bei der Entwicklung werden folgende Ziele der Reihe nach bearbeitet und auf deren Erfolg hin überprüft. Hierbei wird im Entwicklungsprozess viel Experimentelles mit mehreren Iterationsstufen nach Beweis einzelner Einsatzmöglichkeiten der unterschiedlichen Techniken zu finden sein.

Nutzbarkeit der Kamera

Aufgrund der nicht einsetzbaren Kamera des C-MAC[®] Videolaryngoskop 8401 BX wird der Prototyp mit einem Laryngoskopspatel und externen Endoskop gebaut. Dies führt dazu, dass die Verwendung der entsprechenden Kamera speziell getestet und auf das Projekt angepasst werden muss. Die Nutzbarkeit der Kamera für die Positionserkennung soll geprüft werden. Hierbei können, wie bereits unter Abschnitt 3.1 beschrieben, Regularien zur Medizinproduktanerkennung in Teilen außer Acht gelassen werden.

Funktionalität der Erkennungstechnik

Es existieren diverse Techniken zur Bestimmung einer Position in einem Bild. Eine gute und einfach einzusetzende Technik zu finden und zu implementieren, wird das Ziel der Arbeit sein. Zudem muss die Funktionalität der Erkennungstechnik im Zusammenhang des Projektes bewiesen werden. Weiterhin ist hierbei auch zu prüfen, welche Technik mit welcher Art Marker und Menge zu nutzen ist.

Funktionalität möglicher Marker in Software und Puppe

Der Einsatz von entsprechend sinnvollen Markern in Form von markanten Stellen oder spezifischen Lokalisierungsstellen in Verbindung der Software und der Reanimationspuppe muss als ein Teilziel auf Einsetzbarkeit überprüft werden.

Berechnung der Position und Drehung

Die Bestimmung von Position und Drehung der Kameralinse aus dem Kameralivebild als Kernelement der Arbeit erhält einen besonderen Fokus bei der Erarbeitung. Dieser

Abschnitt als ein Kernteilziel wird mit zusätzlicher Sorgfalt beachtet werden.

Einfache Darstellung des 3D-Modells

Damit die Arbeit im Nachhinein einfach adaptiert werden kann, muss als ein Ziel auf die einfache Einbindung von den Daten in 3D-Modelle geachtet werden. Als Beweis der Funktionalitäten und für die Einbindung in ein 3D-Programm soll eine simple Darstellungsform präsentiert werden.

Echtzeit

Neben der grundsätzlichen Funktionalität der Positionsbestimmung muss speziell auf den Teil des Ziels der Echtzeitverwendung geachtet werden. Auch wenn keine direkte Einbindung einer Virtual Reality (VR)-Technologie vorherrscht, spielt bei der Entwicklung dieser Arbeit das Thema Motion Sickness, also ein Missverhältnis der realen Bewegung am Laryngoskop und der übertragenen und übersetzten digitalen Version am Monitor, eine Rolle. Wichtig ist hierbei, das Gefühl der echten Bewegungen und Wahrnehmungen ohne Probleme direkt darstellen zu können und keine Verschiebungen zu haben.

Umsetzung der Idee

An oberster Stelle steht außerdem das Ziel zu zeigen, dass es überhaupt mit diesem optischen Positionsbestimmungsansatz möglich ist, diese Art der Verbesserung eines Trainingsvideolaryngoskops zu erreichen und dass die Idee einen Mehrwert für das Training von ärztlichem Personal und Rettungskräften bieten kann.

4 Technische Grundlagen

Entsprechend der Analyse der Projektidee ergaben sich einige Vorgaben für die Entwicklung des Prototyps. Viele dieser identifizierten Techniken haben spezifische Eigenschaften und Hintergründe für ihren Einsatz in den entsprechenden Verwendungsgebieten. Für ein besseres Verständnis und zur Folgerung für den tatsächlichen Einsatz am Prototypen führt dieser Abschnitt in die Arbeit mit optischen Erkennungstechniken als Kernelement dieser Bachelorarbeit ein.

4.1 Kamera

Kameras werden verwendet, um die Umwelt in interpretierbare Datensätze zu übersetzen, welche dann einen Informationsgehalt in Form eines Bildes darstellen. Die grundsätzliche Funktion besteht darin, dass entsprechend mit unterschiedlicher Wellenlänge und Reflexion Lichtstrahlen auf ein kleines Loch treffen und von dahinter befindlichen Photodioden wahrgenommen werden. Mehrere Photodioden sind als Kamerasensor zusammengeschlossen. Bilder werden deswegen in einer Matrixrepräsentation abhängig von den Bildpunkten aufgenommen. Objektive werden häufig zusätzlich vor das Loch gesetzt, um damit spezielle perspektivische Projektionen oder Kissen- oder Tonnenverzeichnungen zu erzeugen oder zu reduzieren. Mit diesen Kissen- oder Tonnenverzeichnungen lassen sich meistens beispielsweise Fish-Eye-Kameras bauen. Die Basistechnik ist dennoch das beschriebene Lochkameramodell zur Bildaufnahme und Darstellung. Der Kamerateyp kann sich zwischen einer Flächenkamera und Zeilenkamera unterscheiden. In Videokameras ist in der Regel der Flächenkamerateyp verbaut. Dies begründet sich vor allem daran, dass meistens geringere Auflösungen vorhanden sind, aber eine zeitliche Bildfolge gefordert ist. [9, 11]

Die Auswahl einer geeigneten Kamera für das entsprechende Aufgabengebiet wird daher anhand mehrerer Faktoren getroffen. Zu diesen Faktoren zählen neben Auflösung,

Verzeichnung, Objektiven auch die Möglichkeit einer Beleuchtungsart. So können allein Beleuchtungsentscheidungen die Qualität der aufgenommenen Bilder in entscheidender Weise bestimmen. Beleuchtungsarten gliedern sich von direktem und indirektem Licht bis Auflicht und Durchlicht. [14]

4.2 Optische Erkennungstechniken

Die Detektion und Erkennung von Objekten ist ein eigenes Fachgebiet in der Informatik, welches sich in der Bildverarbeitung eingliedert. Einfach formuliert wird versucht, aus Bilddatensätzen Informationen zu gewinnen, die zur Spezifikation oder einer Merkmalsbestimmung verwendet werden können. Diese sind dann zur Identifikation der geforderten Auskunft über ein Bild einzusetzen.

Möglichkeiten der Erkennung von Mustern in Bildern gibt es einige. Für diese Arbeit werden nun einige Relevante kurz betrachtet und analysiert.

Vorverarbeitung

Die meisten Erkennungstechniken in der Bildverarbeitung benötigen vorab eine entsprechende Vorverarbeitung, damit die Bilder erfolgreich nach den festgelegten Merkmalen analysiert werden können. Der Einsatz der Vorverarbeitung ist in der Regel obligatorisch durchzuführen, dennoch unterscheiden sich die benötigten Vorverarbeitungsschritte von der später gewählten Mustererkennungstechnik.

- Für die meisten Mustererkennungsverfahren sind keine Farbinformationen notwendig, sodass einer der häufigsten ersten Schritte die Umstellung auf eine Grauwertdarstellung des Bildes ist.
- Teilweise sind die Grauwerte der Bilder auch noch zu viel Informationsgehalt, sodass weiter reduziert wird. Dies findet auf eine schwarz-weiße, also binäre, Darstellungsform statt.
- Der Einsatz von Filtern wird in diesem Zusammenhang zur Verbesserung der Bildqualität verwendet. Bei Mustererkennungsverfahren könnten beispielsweise die Verstärkung von Linien, morphologische Filter, relevante Punkte oder Ähnliches genutzt werden.

Mustererkennung

Mustererkennung ist ein Verfahren zur Detektion von definierten Informationen in einem spezifizierten Bildbereich anhand von unterschiedlichen Merkmalen. Diese Merkmale legen fest, welche die Erkennungsmarken für die gesuchte Information sind. Je nach Typus liefern diese einen zusätzlichen Informationsgehalt zum Bildaufbau oder gefundenen Musterwert. [14]

- Kantendetektion kann als eine Variante zur Berechnung von markanten Stellen verwendet werden. Diese lassen sich dann einfach herausrechnen, beispielsweise wenn die Längen der errechneten Kanten gemessen werden, um zusammenhängende Bereiche zu identifizieren. Einsetzbar wäre die Kantendetektion dennoch auch für eine weitere Vorverarbeitung zur Qualitätserhöhung der Bilder.
- Die Subtraktion von zwei Bildern im Zeitverlauf kann verwendet werden, um relevante Bereiche zu lokalisieren, da in dem Ergebnisbild nur die Differenzen hervorgehoben werden. In diesem Zusammenhang ist auch die teilweise Segmentierung von einzelnen Bildbereichen nützlich, um einzelne spezifische Bereiche von anderen zu differenzieren.
- Genauso lassen sich durch die Analyse von Farbhäufigkeiten Bereiche in Bildern feststellen, die einem gesuchten Muster entsprechen.
- Das Template Matching als Technik zum Finden von konkreten Bildausschnitten in einem Bild wird häufig in der Qualitätsanalyse oder Bestimmung von absoluten definierten Elementen in einem Bild genutzt. Der Einsatz dieses Verfahrens fordert für einen Erfolg dennoch gute Vorbereitung und hat auch schnell keine positiven Ergebnisse parat, da nur exakt gleiche Templates und Interessenspunkte im Bild verwertet werden können.
- Das Feature Based Matching verbessert diesen Ansatz der Templatesuche und vergleicht stattdessen markante einzelne Punkte zur Mustererkennung. Dieser Ansatz wird meist mit zusätzlicher Bilddatenanalyse durchgeführt.

Während alle bisher aufgeführten Techniken nach natürlichen Merkmalen oder Texturen in einem Bild suchen, gibt es auch Varianten wie Markererkennungen beispielsweise mit ArUco-Markern, sogenannte Passermarken, die eine vereinfachte Detektion von Mustern ermöglichen. Diese Ansätze nutzen speziell vordefinierte Muster, die in die entsprechenden Bilder vorab eingesetzt werden, sodass an diesen dann mit hoher Effizienz und Genauigkeit die geforderte Information extrahiert werden kann. [3]

4.2.1 Vorteile

Optische Erkennungstechniken vor allem in Zusammenhang mit Mustererkennung bringen damit die Vorteile, aus einem Bild oder auch nur einem Bildausschnitt durch entsprechende Verfahren ein Großteil an Information über eine Lokalität zu erlangen. Ein Kamerasensor ist kostengünstig und kann für die jeweiligen Aufgabengebiete weiter angepasst werden. Der Einsatz von Vorverarbeitung ist für die Mustererkennung hierbei elementar. Marker beispielsweise stellen für dieses Projekt eine gute Grundlage und Anpassungsfähigkeit dar.

4.2.2 Nachteile

Dennoch gibt es einige Probleme, die die Erfolge der optischen Erkennungstechniken stark mindern können. Diese fangen bei der Bildaufnahme an, wenn ein schlecht beleuchtetes Bild verarbeitet wird. Zwar gibt es Möglichkeiten, diese Bilder nachzubearbeiten, leider verlieren sie dann jedoch einiges an Auflösung, also der verwendbaren Pixel, sodass Mustererkennungsverfahren Schwierigkeiten haben, noch Informationen zu extrahieren.

4.3 Marker und Pose Estimation

Bei der Verwendung von Mustererkennungsverfahren spielen Marker mittlerweile immer häufiger eine Rolle, vor allem wenn diese in Kombination von AR-Anwendungen verwendet werden. Für diese ist es essenziell, reale und genaue Informationen über die Bildinhalte zu erlangen. Sogenannte Passermarken, meist aus Kombinationen mit schwarz-weißen Punkten und Kästchen, bilden optische Referenzpunkte, die in einem Bild integriert sind. Die Identifizierung der jeweiligen Information aus dem Marker geschieht anhand von Regeln, die in einem komplexen Analyseprozess diese Auskunft aus den Referenzpunkten in der Bildposition berechnen. Nach diesen Schritten lassen sich einfach Informationen über die Position im Bild und auch im Verhältnis zur realen Welt interpretieren. Einzige Abhängigkeit dieses Verfahrens ist, dass diese Marker vorab in der Realität positioniert werden müssen.

Anhand der Detektion von Markern lassen sich Objekte leicht lokalisieren, beispielsweise wenn die jeweiligen Auftritte der Objekte direkt an den Markern zu finden sind. Nebenbei ist auch die Bestimmung von Positionen von Bildpunkten zu ihrer Repräsentation in der

realen Welt möglich. Es findet eine Orientierung an den Raumdimensionen statt. Eine einfache Detektion, klare Bestimmungsregeln und technische Erweiterungen ermöglichen sogenannte Pose Estimation. Hierbei werden die Marker untereinander in Verbindung gesetzt und mit neuer Bedeutung versehen.

Pose Estimation ist das Verfahren der Bildverarbeitung zur Bestimmung von Positionen, Ausrichtungen, Orientierungen und auch Größen von Objekten oder Personen. Diese Technik nutzt definierte Punkte, als Beispiel Marker, zur präzisen Schätzung der Position und Orientierung in einem Bild. Verwendet werden hierbei zur Berechnung zusätzliche Entsprechungen der realen Welt. [3, 17]

5 Prototyp

Das folgende Kapitel beschreibt nun anhand der Analysen und notwendigen thematischen Hintergründen den Entwicklungsweg des Prototyps für das Videolaryngoskop mit der Software der Positionsbestimmung für die 3D-Kartenerstellung. Für die Konzeption des Prototyps sind Entwicklungsschritte aus physischer Betrachtung und aus softwareseitiger Entstehung notwendig. Die Entwicklung wird sich hierzu zu Beginn auf einige physische Gestaltungsaspekte fokussieren. Erst nach einer fertigen Konzeption der Abbildung des realen Videolaryngoskopsetup wird anhand dieser dann gültigen Spezifikationen eine softwaretechnische Seite angefangen zu entwerfen und zu programmieren. Wichtig hierbei ist trotzdem, dass eine andauernde Anpassung von Software und auch der physischen Begebenheiten beachtet und umgesetzt wird.

5.1 Physischer Prototyp und Randbedingungen

Für die Entwicklung des physischen Prototyps konnte keine direkte Umsetzung im Simulationszentrum erfolgen. Einerseits aufgrund der Gegebenheit, dass keine Reanimationspuppe vom UKE zur dauerhaften Ausleihe bereitgestellt werden konnte und eine direkte Arbeit an dieser auch nicht zielführend zur generellen Testung der Idee gewesen wäre. Andererseits, da das Videolaryngoskop des Simulationszentrums nicht direkt transportierbar ist und nicht die nötigen Funktionen für das Projekt bereitstellen kann.

Dementsprechend werden diese Teile einzeln betrachtet und analysiert, sodass ideale Produkte als Alternativen und Vorteile genutzt werden können. Ziel ist, einen einfachen Weg zur Integration in das Realprodukt zu erreichen.

5.1.1 Laryngoskop

Die Verwendung eines eigenen neuen Laryngoskops bringt für die Gestaltung des Prototyps einige Vorteile mit sich. Durch die frei wählbaren Spezifikationen im Rahmen der



Abbildung 5.1: 3D-Druck Prototyp des Videolaryngoskop

Anforderungen sind mehrere Verbesserungen und Anpassungen in der Zusammensetzung des Laryngoskops und der später nötigen Endoskopkamera möglich.

Um der freien Gestaltung gerecht zu werden, wird nun ein Prototyp aus einem 3D-Drucker verwendet. Ein Einscannen des realen Laryngoskops zur Produktion einer Druckvorlage musste nicht umgesetzt werden, da das UKE eine bereits entwickelte prototypische Version¹ empfehlen konnte. Die Vorlage dieses Laryngoskops ist das eines mit einem Macintosh-Spatel, sodass es den Gegebenheiten des ursprünglichen realen Videolaryngoskop direkt entsprach und, wie in Abbildung 5.1 dargestellt, 3D gedruckt nutzbar ist. Die Entscheidung, dieses Modell zu nutzen und in den Prozess der Prototypentwicklung zu integrieren, begründet sich mit folgenden Argumenten.

- Die Vorlage muss eine Version sein, die die Integration einer separaten Endoskopkamera unterstützt. Die gewählte Form des sogenannten AirAngel Project ist für diesen Anwendungsfall speziell entwickelt worden. Das Projekt wirbt explizit für diesen Einsatz in Simulationstrainings.
- Zusätzlich steht die Haptik und das Gefühl, ein reales Laryngoskop zu verwenden, bei dieser Arbeit an erster Stelle. Da das nun Gewählte explizit für diesen Bereich nachempfunden ist, ist diese Vorgabe gewährleistet.

¹<https://www.airangelblade.org>

- Weiterhin musste ein Laryngoskoptyp gefunden werden, der sich nach Abschluss der Bachelorarbeit leicht von Bedienenden reproduzieren und in das Training einfügen lässt.

Dennoch sind ein paar Dinge bei dem Einsatz dieses Produkts zu bedenken. Es ist keine spezielle Position für den Einbau einer externen LED-Beleuchtung vorhanden, sodass die getrennten Teile von Kamera und Beleuchtung des C-MAC[®] Videolaryngoskop nicht exakt nachgebildet werden können. Bei der Auswahl der zugehörigen Endoskopkamera muss dementsprechend eine LED-Beleuchtung mit integriert sein.

5.1.2 Endoskopkamera

Die Kamera als Kernelement der Bachelorarbeit hat einige Vorgaben zu erfüllen, so dass sich die Suche und Entscheidung nach einem geeigneten Produkt für den Prototyp schwieriger herausstellte als erwartet.

Eine Endoskopkamera wird zur Beobachtung von innenliegenden Bereichen in unterschiedlichen Fachgebieten verwendet. Die Zusammensetzung eines Endoskops besteht aus einer Kamera mit einer kleinen Linse, häufig einer Beleuchtungs-LED und einem langen Kabel. Dieses Kabel sorgt für den Anschluss an einen Monitor oder Computer und ist zusätzlich halbstarr, dass es verformt und in einer Form fixiert in enge Bereiche eingeführt werden kann.

Für den Prototyp ist eine Kamera mit guten Spezifikationen gesucht, da die Mustererkennungsverfahren hohe Auflösungen benötigen, um präzise Informationen bezüglich der Linsenposition zu berechnen. Unabhängig davon steht die Lösung des Kompatibilitätsproblems aus der Analyse an erster Stelle. Für den Einsatz der Endoskopkamera in Verbindung eines Computers, perspektivisch eines Mikrocontrollers, muss ein USB-A-Anschluss vorhanden sein. Dieser Umstand ist vor allem dadurch bedingt, dass eine Kompatibilität zu Windows, Mac oder Linux-Betriebssystemen gewährleistet werden kann. Eine Recherche mit dem Fokus dieser Spezifikationen stellte sich als schwierig heraus. Die meisten auf dem Markt zu findenden Endoskopkameras haben nur USB-C-Anschlüsse, die mit speziellen Restriktionen versehen sind, welche nur Kompatibilitäten zu Android oder IOS-Betriebssystemen unterstützen. Hinderlich ist zudem, dass bei einigen Kameras zumeist nur mit speziellen Apps auf das Videosignal direkt zugegriffen werden kann. Für eine zielgerichtetere und auch technologieoffene Einsatzmöglichkeit des endgültigen Prototyps muss daher die direkte Aufnahme des Kamerabilds ohne spezielle



Abbildung 5.2: Vollständiges Prototyp-Videolaryngoskop

Protokolle möglich sein. Neben den genannten Faktoren liefert das nun gewählte Laryngoskop 3D-Modell ein Loch für die Linse von maximal 8 mm Größe, was zusätzlich die letzte Vorgabe der Kameraspezifikation ist.

Die Endoskopkamera TENG-END-31UB der Marke Lightswim² erfüllte am besten alle geforderten Vorgaben. Neben einer Auflösung von Super Full HD 1200P (4:3) 2MP des CMOS-Sensors ist die Plug-and-Play-Funktionalität der Kamera für Windows und Mac-Betriebssysteme entscheidend gewesen. Anhand einiger Tests ist zudem die nötige Schärfe der Bilder für die relevanten Entfernungen von 3 bis 10 cm im Rachen einer Testpuppe belegt worden.

Das gewählte Produkt ist nun, wie in Abbildung 5.2 zu sehen, in den Laryngoskopprototyp eingesetzt. Ein Ringlicht mit einer Helligkeitssteuerung ist um die Linse herum positioniert und über die Kamera steuerbar.

5.1.3 Entwicklung am 2D-Modell

Die Entwicklung an der realen Reanimationspuppe bringt, wie bereits beschrieben, einige Schwierigkeiten mit sich. Angefangen bei Unhandlichkeit und Größe, da der für das

²<https://www.amazon.de/dp/B089Y3RJ4K>



Abbildung 5.3: Durchsichtiges Kopfschnittmodell

Projekt relevante Abschnitt sich nur auf den Kopf bezieht. Die Problematik der weichen beweglichen Bereiche im Rachen stellt sich im Erstentwurf des Prototyps mit dem Positionsberechnungssystem als eine Verkomplizierung des geforderten Systems dar. In Abstimmung aller beteiligten Parteien wird die Entwicklung deshalb an einem Testrachen durchgeführt. Dieser Testrachen weist einige Reduktionen der realen Eigenschaften auf, ist aber trotzdem so gestaltet, dass eine einfache Übersetzung des fertigen Produktes auf eine reale Anwendung möglich ist.

Die Wahl fällt damit auf ein Kopfschnittmodell des Herstellers VBM, welches über den Fachbereich Breitenausbildung der Johanniter-Unfall-Hilfe e.V. Regionalverband Hamburg bereitgestellt werden konnte. Dieses spezielle Modell wird normalerweise zur Demonstration und dem Training der korrekten Intubation benutzt. Auch Videolaryngoskopie kann damit geübt werden und ist dadurch ideal für dieses Projekt geeignet. Wie in Abbildung 5.3 zusehen, ist der transparente Kopf durch sein statisches Design in einer Position festgelegt, was die Reduzierung von Komplexität in Hinsicht auf den 3D-Raum mit sich bringt, da keine Veränderungen möglich sind. Ein einfaches Design der nötigen Markierungen ist ebenfalls möglich, weil die Transparenz des Modells ein Platzieren von Markern vereinfacht.

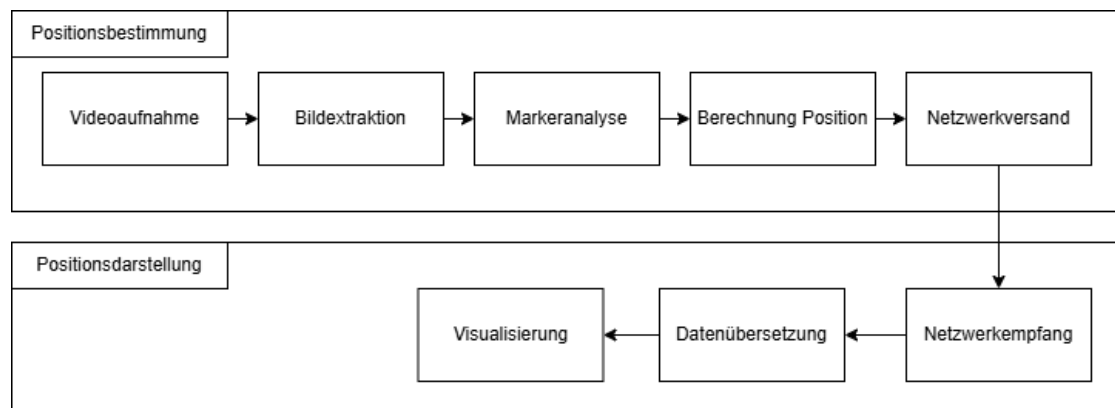


Abbildung 5.4: Ablaufdiagramm zur Systemkonzeption: Berechnung der Position der Kameralinse und 3D-Darstellungsform (eigene Darstellung)

5.2 Lösungsansatz

Nach den physischen Entwicklungsschritten zur Erstellung eines verwendbaren Videolaryngoskops und Kopfphantoms folgt nun der Lösungsansatz zur digitalen Verarbeitung der Livebilder und der Umwandlung in eine 3D-Karte.

Wie in Abbildung 5.4 erkennbar, gliedert sich der Lösungsansatz in zwei Softwareteile. Die Hauptentwicklung liegt auf der Mustererkennungsseite mithilfe der Kamerabilder und der daraus folgenden Positionsbestimmung. Das Konzept für diesen Abschnitt basiert auf einfacher Bildverarbeitungsideologie. Dies heißt, dass ein Bild aufgenommen oder aus dem Video extrahiert wird und nach einem Vorverarbeitungsschritt der Daten an die nächste Ebene zur Analyse weitergegeben wird. Jede weitere Ebene der Analyse und Verarbeitung der Daten findet getrennt und nur mit den für diesen Aufgabenbereich benötigten Daten statt. Die Trennung dieser Informationsebenen soll dafür genutzt werden, bei einer späteren Auseinandersetzung zur Weiterentwicklung dieses Systems mit dem Lösungsansatz nur einzelne Teile anpassen zu müssen.

Der zweite Softwareteil wird zur Verifikation des ersten Schritts und Demonstration des Hauptziels der Darstellung der 3D-Position der Kamera umgesetzt. Da diese Trennung eine Kommunikation benötigt, wird diese über das Netzwerk abgewickelt.

5.3 Systementscheidungen

Neben den physischen Entscheidungen für den Prototyp sind zusätzlich Entscheidungen für den Systemaufbau und Softwareentwurf zu treffen. Festlegungen dieser Art sind an die Erfüllung der Entwurfsziele gebunden. Bei diesen sind trotzdem die Testung der Einsatzmöglichkeiten auch in der Verifikation der Funktionalitäten für dieses Projekt vorgesehen.

5.3.1 Minimum Viable Product

Da das gesamte Projekt in einigen Teilen experimentellen Charakter besitzt, ist als Artefakt der Software nur ein Minimum Viable Product (MVP) vorgesehen. MVPs sind Produkte, die ein Maximum an Funktionalität bei geringstmöglichem Einsatz in der Testung mitbringen [8]. Sie dienen vor allem dazu, die grundsätzliche Funktionalität der Idee für entsprechendes Feedback zeigen zu können. Bedeutsam ist hierbei zusätzlich, dass in kleinen Schritten das Gesamtprodukt entworfen wird. Aus diesem Grund wird für die Positionsbestimmung ein Markersystem genutzt, welches sich einfach und schnell umsetzen und integrieren lässt.

Folgende kleine Arbeitsschritte ergeben sich für diesen MVP:

- Erkennung und Berechnung der Markerposition
- Erstellung einer 3D-Kartenrepräsentation mit Abspeicherung der Karteninformation
- Berechnung einer spezifischen Kameraposition und Kameraausrichtung
- Transport von 3D-Dateninformationen in ein 3D-Programm
- Einfache Darstellung der Realweltinformationen in einem 3D-Programm mit Erstdarstellung von Echtbildern

5.3.2 ArUco-Marker

Posenschätzungen und Positionsbestimmungen verwenden in den meisten Fällen sogenannte Passermarken. Eine der beliebtesten und einfach zu nutzenden Bibliotheken für dieses Einsatzgebiet ist die ArUco-Bibliothek von Rafael Muñoz und Sergio Garrido

[3, 12]. Die Integration des zugehörigen `arcuo`-Moduls von OpenCV in Python öffnet die einfache Entwicklung für jedermann.

ArUco-Marker sind quadratische Marker, die aus einer schwarzen Umrandung und einer inneren binären Codierung, abhängig von der gewählten Markergröße, der genutzten Marker-Identifikator (ID)-en und der Größe des Abstands der binären Codierung, bestehen. Sie besitzen die Eigenschaften, durch die vier Ecken des Quadrats eine markerabhängige und eindeutige Kennzeichnung bestimmen zu können. Diese Marker ermöglichen durch eine robuste Implementation eine schnelle Erkennung im Bild unter Verwendung mehrerer Fehlererkennungs- und Korrekturtechniken. Eine präzise Kameraposeschätzung ist über diese Eigenschaften reibungslos umsetzbar. Durch die Identifikation von spezifischen IDs lassen sich zudem mehrere Marker unabhängig zueinander einsetzen, ohne die Funktionalität der anderen einzuschränken [2]. Dabei ist eine direkte Korrespondenz auf die Position der Kamera und des jeweiligen Markers möglich, da sie eine festgelegte Drehung in der realen Welt besitzen. Anhand dieser Informationen können dann Berechnungen zu Entfernungen, Positionen, Rotationen und anderen geometrischen Zusammenhängen getroffen werden.

Die Marker müssen für einen Einsatz beispielsweise nur auf Papier gedruckt vor der Kamera positioniert werden und sind dann direkt durch die Bibliothek im Bild einsatzbereit. Funktionen zur Generierung der Marker werden bereitgestellt. Doch mit speziellen Onlinegeneratoren³ geht dieser Schritt zielgerichteter und einfacher. Wichtig bei der Verwendung ist, zwischen den physisch gedruckten und den in der Software ausgewählten Markern eine Konformität zu gewährleisten, da die binäre Codierung spezifisch auf die Größe und die sogenannte Wörterbuchgröße, Anzahl der IDs, berechnet wird.

Der Markererkennungsprozess findet durch die Bibliothek schnell in zwei Schritten anhand von Bildverarbeitungsverfahren statt. Die Posenschätzung ist ein darauffolgender dritter Schritt. [2]

Markeridentifizierung

Zu Beginn wird das Bild anhand von adaptiven Schwellwerten segmentiert, sodass versucht wird, quadratische Formen zu extrahieren. Durch weitere Filter werden Konturen verworfen, die teilweise konvex, zu groß oder klein oder nicht für eine weitere Analyse aufgrund zu geringer Trennungen verwendet werden können. [2]

³<https://chev.me/arucogen/>

Berechnung der Codierung

Über eine Anwendung einer perspektivischen Transformation, weshalb eine Kamerakalibrierung vorab benötigt wird, werden folgend die einzelnen Markerbits zur Berechnung der Codierung extrahiert. Die Anwendung von Otsu zur binären Übersetzung des Bildes ist hierbei obligatorisch. Durch das Zählen der übriggebliebenen Pixel und ihrer entsprechenden Anordnungen im Marker Kandidaten werden echte Marker identifiziert. [2]

Positionierung und Ausrichtung im Bild

Über eine Posenschätzung werden durch den Einsatz der Kamerakalibrierungswerte die Position und Drehung des Markers bestimmt. Diese entsprechen einer 3D-Transformation vom Marker-Koordinatensystem zum Kamera-Koordinatensystem. Die Rückgabe dazu sind ein Translationsvektor und ein Rotationsvektor in der Achsendarstellung von x , y und z . Die Ausrichtung des Translationsvektors ist in der Kameralinsenmitte in Richtung des Markers ausgerichtet. Der Rotationsvektor in der Form einer Rotationsmatrix muss zur weiteren Nutzung manuell transformiert werden.

Durch bereitgestellte Methoden der Bibliothek lassen sich die Parameter der einzelnen Grenzwerte zur Identifikation oder Bestimmung der Marker anpassen. Diese Funktionalitäten sind optional einsetzbar und wurden aufgrund der MVP Ausrichtung des Projekts in der Verwendung darauf hier verzichtet. [2]

Die Möglichkeit, direkt nutzbare Daten in Form von Vektoren für den Prototyp verwenden zu können, soll verfolgt werden, da eine 3D-Raumdarstellung viele Vorteile mit sich bringt. Als ideale Variante erweist sich, dass die Position der Kamera im Inneren des Rachens eines Kopfphantoms durch einfache mathematische Vektorrechnung durchzuführen ist. Komplizierte Neuschaffungen von Markererkennungsverfahren bleiben damit aus, die mitunter durch die zusätzlichen experimentellen Schritte nicht genau genug sein könnten. Die Bereitstellung von vielen vorgefertigten Informationen durch die ArUco-Marker vereinfacht zudem den zu entwickelnden Positionsberechnungsalgorithmus.

Mit dem Durchführen erster Tests analog zu den Tutorials der ArUco-Marker-Anwendung bestätigt sich die Eignung der Nutzung von ArUco-Markern für den Einsatz in diesem Projekt. Die Marker sind gut und lange bezogen auf die Entfernung detektierbar. Dennoch sind weitere Vorverarbeitungen für noch bessere Ergebnisse nötig, da Randbereiche oder wie um 90 Grad gekippte oder gebogene Marker Schwierigkeiten bei der Erkennung darstellen.

5.3.3 Algorithmen

Dieser Abschnitt beschreibt ein paar Verfahren und Algorithmen, die für die Implementierung des Systems von Bedeutung sein können. Dies soll jedoch lediglich auf die Verwendung und mögliche Komplikationen aufmerksam machen und die nötigen Vorgaben für den Gebrauch definieren.

Kamera-Kalibrierung

Eine Kalibrierung der Kamera ist für die Verwendung der Methoden aus der ArUco-Bibliothek notwendig. Die Posenschätzung findet unter Einsatz der Verrechnung der Kameramatrixwerte und der Verzerrungskoeffizienten statt. Jede Verwendung einer neuen Kamera oder die Veränderung des Fokus derer benötigt zur korrekten Funktion eine erneute Kalibrierung. Der benötigte Algorithmus zur Durchführung ist relativ simpel und liefert sehr genaue Messwerte. Angelehnt an ein Videotutorial zur Kalibrierung von Kameras mittels OpenCV und Python von Nicolai Nielsen⁴ und dem OpenCV Dokumentationseintrag [11] wird die Kalibrierung über eine Berechnung der intrinsischen und extrinsischen Kameraparameter über die Realgrößen-Entsprechungen eines Schachbretts⁵ vollzogen. Wenn die Werte für eine Kameraeinstellung einmal bestimmt sind, sind diese sofort verwendbar und müssen nicht erneut vermessen werden.

Vorverarbeitung

Bei der Bildverarbeitung sind neben den Analyseverfahren für die Markererkennung teilweise kleinere Vorbereitungsverfahren zur Verbesserung der Bildqualität notwendig. Bedingt durch die Kalibrierung der Kamera stellt sich eine Verzerrungsrausrechnung als Möglichkeit bereit. Diese wird durchgeführt, um das Eingabebild mehr an die realitätsnahen Bedingungen anzupassen. Dies erleichtert später die Detektion. Die zusätzliche Verzerrungsrausrechnung muss allerdings in Verbindung der Nutzung der ArUco-Bibliothek nicht umgesetzt werden, da die verwendeten Methoden die Verzerrungen bereits selber kompensieren.

Daneben können Bildmanipulationsverfahren wie eine Graubildwandlung, Binärisierung, Closing oder Opening zusätzliche Verbesserungen bringen. Die Möglichkeiten werden direkt bei der Implementierung getestet.

Positionsbestimmungen

Für den Hauptabschnitt der Software, der Berechnung der tatsächlichen Position der

⁴https://www.youtube.com/watch?v=_-BTKiamRTg

⁵<https://markhedleyjones.com/projects/calibration-checkerboard-collection>

Kameralinse, ist kein bestehender Algorithmus vorhanden und kann direkt verwendet und angepasst werden. Ein eigener Algorithmus ist zur Verarbeitung der Vektorinformationen und der daraus folgenden Lokalisation nötig. Für den Erstentwurf der Bestimmung der Position und Ausrichtung dienen zwei Videotutorials von Nicolai Nielsen⁶ und Asadullah Dal⁷ als Inspirationen zum Umgang mit ArUco-Markern in OpenCV und Python.

5.4 Implementierung

Die Implementierung des Detektionsalgorithmus und der Positionsbestimmung orientiert sich an den Eigenschaften der ArUco-Marker. Die sehr einfache Detektion eines einzigartigen Markierungspunkts mit den dazugehörigen Daten, bezogen auf Drehung zur Kameralinse und der realen Entfernung in Form von Vektoren, lässt eine Nutzung möglichst weniger zu platzierender Marker im Rachen des Kopfphantoms zu. Die präzisen Messwerte ermöglichen eine Positionsbestimmung anhand weniger Berechnungen.

Als Datenverarbeitungsprogramm zur Bildverarbeitung und Integration der OpenCV ArUco-Bibliothek wird Python als Programmiersprache verwendet. Da neben Python auch die OpenCV ArUco-Bibliothek umfangreiche Open-Source-Dokumentationen bereitstellt, ist die Implementierung effizient durchzuführen. Die Umsetzung der Software soll im Endeffekt einen Frame einer Kamera annehmen und als Rückgabe die Kameraposition und Drehung zurückgeben. Realisiert wird dies durch die Weitergabe von jeweils relevanten Datensätzen von einer Verarbeitungsebene zur nächsten. Grundsätzlich sind zwei Implementationen für das Gesamtsystem nötig. Neben der Berechnung und Entdeckung einer 3D-Karte soll zudem gleichzeitig die Möglichkeit der Positionsbestimmung bestehen. Dieses Problem einer Kartierung und gleichzeitigen Lokalisation bezeichnet sich als Simultaneous Localization and Mapping (SLAM) und ist ein Forschungsgebiet der Robotik [15]. Eine mögliche Lösung der gleichzeitigen Bestimmung soll entworfen werden.

5.4.1 Entwurf Marker in Realumgebung

Neben der Software müssen die Marker für eine ideale Erkennung und Verarbeitung in den Rachen des Kopfphantoms eingesetzt werden. Die Marker dürfen für die Verwendung

⁶<https://www.youtube.com/watch?v=GEWoGDdj1Sc>

⁷<https://www.youtube.com/watch?v=mn-M6Qzx6SE>

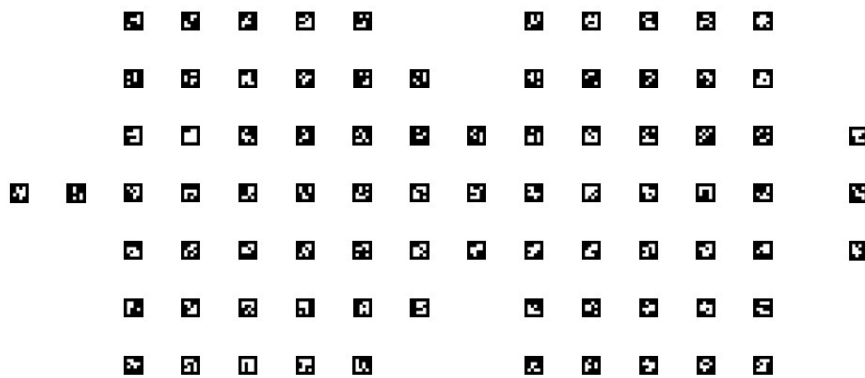


Abbildung 5.5: Erstentwurf der Anordnung von Markern

im Rachen die Größe von 5x5 mm nicht über- und unterschreiten. Dieses Maß definiert sich durch die Breite von etwa 3 cm des Rachens des durchsichtigen Kopfmodells. Zudem müssen sich die Marker in einem Gleichverhältnis der Größen in diesem definierten Raum befinden. Überdimensionierte Marker erschweren die Notwendigkeit, von mehreren Markern die Positionsdaten erlangen zu können, um idealerweise die Bestimmung derer triangulierend durchzuführen. Die Verwendung von wenigen Markern erschwert zusätzlich die Kartierung des Rachens anhand der Markierungspunkte, da wenige Referenzpunkte die Festlegung der Messergebnisse zueinander einschränken. Entsprechend der definierten Größe wurden mehrere Marker, Markergröße „4x4_250“ (4x4 Quadrat der Binärcodierung mit 250 definierten IDs) auf Softwareebene für das Wörterbuch, so verteilt, wie in Abbildung 5.5 erkennbar, dass der gesamte sichtbare Bereich bei der Bewegung des Laryngoskops mit Markern schachbrettartig ausgekleidet ist. Die sichtbaren Bereiche beginnen auf Höhe der Zähne, wo ein Marker als Startpunkt gesetzt ist. Die weiteren Bereiche befinden sich dann durchgängig am Gaumen, mit einer Ausweitung nach links und rechts des Laryngoskopsichtfelds verteilt bis zum Kehlkopf, wo sich Luftröhre und Speiseröhre trennen und der Kehildeckel zum Verschließen sichtbar ist. In diesem ersten Design sind die Marker noch ohne eine spezielle Position verteilt und ausgerichtet.

Der Detektionsalgorithmus von OpenCV kann mit jeglicher Menge an Markern im Sichtfeld umgehen, da die Analyse über den gesamten Bildausschnitt stattfindet. Wichtig ist hierbei nur, dass der Marker zur Detektion vollständig im Bild zu sehen sein muss. Weiterhin ergaben Tests im Rachen mit dieser großen Menge an Markern, dass viele und vor allem spezifische Marker in der Mitte der Kameraausrichtung teilweise dauerhaft verfügbar waren. Diese und spätere Ergebnisse der Entwicklungsschritte fokussieren auf

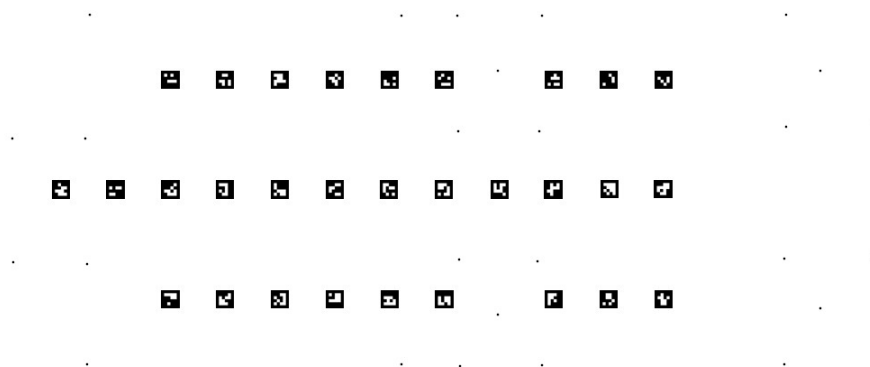


Abbildung 5.6: Endentwurf der Anordnung von Markern

die Erkenntnis, dass eine geringere Menge notwendiger Marker zur Detektion genutzt werden kann. Dies spielt deshalb eine Rolle, da es dann einfacher ist, diesen Entwurf in Reanimationspuppen zu integrieren.

Anhand der Ergebnisse ergibt sich die Positionierung der Marker, wie in Abbildung 5.6 zu sehen. Dabei dürfen die Marker entsprechend nur eine bestimmte maximale Ausrichtung besitzen, weil die Detektion und Wertbestimmung unmöglich wird. Dies kann als Beispiel ein Nick-Winkel zur Kamera von 90 Grad sein. Außerdem gibt es in diesem Markerentwurf eine Reduktion der Markeranzahl. Die Marker sind nun auf das Mindeste reduziert. Nur der mittlere Streifen wird in der Positionsbestimmung verwendet, während die äußeren Streifen bei der Kartierung zusätzlich eingesetzt werden. Die Ausrichtungen der Marker sind hierin für einfachere Vektorrechnungen optimiert. Zusätzlich sind in Abbildung 5.6 Hilfsmarkierungen für das Zurechtschneiden und Falten auf einem Papierausdruck zu sehen.

Der Ausdruck der Marker wird durch einen Tintenstrahldrucker auf normalem, mattem Papier ausgeführt. Ein Laserstrahldrucker hatte bei den Drucktests in den Bereichen der kleinen Binärcodierungen leider nicht die benötigte Genauigkeit erreicht. Das Problem dabei war, dass obwohl diese meist eine hohe Genauigkeit bei scharfen zu druckenden Kanten aufweisen können, die Kanten minimal gezackt gedruckt wurden. Die Wahl fiel dann zusätzlich auf unlaminierendes Papier, da die Plastiksicht durch die Kamerabeleuchtung für ein Glänzen der Marker und damit eine nicht mögliche Detektion sorgte.



Abbildung 5.7: Markeroutput im Graubild

5.4.2 Detektion

Die Detektion von Markern findet zu jedem erhaltenen Frame der Kamera statt. Maximal 27 Frames können aufgrund der Spezifikationen von der Endoskopkamera TENG-END-31UB in einer Sekunde gelesen werden, sodass die maximale Echtzeitberechnung hier nach oben diese Grenze aufweist. Durch die Verarbeitungsschritte zum Erhalt der Messwerte reduziert sich diese Zeit jedoch noch etwas weiter.

Das Programm ist nun so aufgebaut, dass in einer Schleife Frames von der Kamera gelesen und an die Berechnungsmethoden weitergegeben werden. Im ersten Schritt erfolgt dabei die Vorverarbeitung des Bildes zur Verbesserung der Bildqualität. Aufgrund dessen, dass, wie unter Abschnitt 5.3.3 und 5.3.2 bereits beschrieben, die ArUco-Bibliothek bereits viele Bildqualitätsverbesserungsmaßnahmen automatisch über die Verwendung der Methoden abdeckt, bleiben nur wenige zusätzliche selbstgesteuerte Maßnahmen übrig.

Die Anwendung der Reduktion der benötigten Farbkanäle ist dennoch hilfreich und verringert gleichermaßen die zu transportierende Datenmenge des Bildes. Insofern wird der Frame von einem Farbbild in ein Grauwertbild transformiert. Das Ergebnis, wie in Abbildung 5.7 zu sehen, fokussiert erfolgreich in Verbindung mit der Beleuchtung der Kamera den relevanten Sicht- und Arbeitsbereich. Durch die Verwendung sind gleichzeitig die Konturen der schwarz-weißen Bereiche der Marker besser voneinander getrennt. In der Abbildung ist die Kamera aktuell noch um 180 Grad verdreht. Zur Lösung dieses Pro-



Abbildung 5.8: Markeroutput im Binärbild

blems wird sie physisch gedreht.

Normalerweise kann ein Binärbild mit der zusätzlichen Trennung und Differenzierung der Konturen weitere Trennung der zu suchenden Elemente bringen. Eine Anwendung der Binärisierung des Frames ist in Abbildung 5.8 zusehen. Die explizite Trennung der dunkleren und helleren Bereiche ist gut zu erkennen. Problematisch ist allerdings, dass einige Marker stark verwaschen werden. Viele der Marker verlieren ihre Binärcodierung. Zudem wendet der Detektionsalgorithmus der ArUco-Bibliothek dieses Verfahren selbstständig erneut zur Segmentierung des Bildes an. Dadurch entsteht eine doppelte Anwendung des Binärisieren und die möglichen Marker werden stark reduziert. Eine Demonstration dessen ist in den Abbildungen 5.7 und 5.8 in den unteren Abbildungsabschnitten dargestellt. Jede dortige Zeile stellt die erkannten Marker in diesem jeweiligen Frame dar. Während in Abbildung 5.7 auch oder noch Marker an den Rändern erkannt werden, bleiben in Abbildung 5.8 nur eindeutige mittlere Marker als Ergebnis übrig. Zur Verhinderung dieser Einschränkung wird die Binärisierung in der Vorverarbeitung damit ausgelassen.

Konkret nimmt in der Implementation die Methode `findMarker` nun das vorbearbeitete Frame an und gibt als Rückgabe eine Liste der gefundenen Marker-IDs und zugehörigen Eckpunkte im Kamerabild zurück. Zur tatsächlichen Detektion wird das vordefinierte Markerwörterbuch mit 250 möglichen IDs in der Markergröße 4x4 verwendet. Die Entscheidung dieser großen Marker-ID-Reichweite spielt in der späteren Kartierung und einfacheren Unterscheidung der Marker eine Rolle. Dies ermöglicht, mathematische Un-

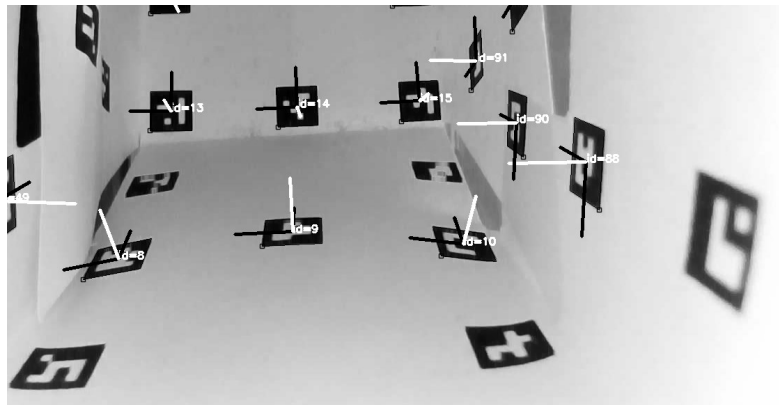


Abbildung 5.9: Markererkennung ohne Closing-Effekt

terscheidungen zwischen Markern treffen zu können, beispielsweise in Sortierverfahren. Im Endentwurf (Abbildung 5.6) sind die Markerreihen nach Ziffern sortiert.

5.4.3 Analyse

Sobald mindestens ein Marker von der Detektion bestimmt wurde, schließt sich im Verlauf die Methode `calculatePose` zur Bestimmung der spezifischen Lokalisationswerte eines Markers an. Über die Parameter der IDs, Markereckpunkte und der Kamerakalibrierungswerte (Kameramatrix und Verzerrungskoeffizient) berechnet sich der Rückgabewert dieser Methode als eine Liste mit diversen Messinformationen zu den jeweiligen Markern. Die Methode `estimatePoseSingleMarkers` des `aruco`-Moduls aus Python liefert hierbei über die Posenschätzung der einzelnen Marker den `rvec` (Rotationsvektor) und den `tvec` (Translationsvektor). Diese beiden Werte sind immer aus dem Sichtbereich und der Ausrichtung der Kamera definiert und lassen sich darauf zurückführen. Dabei bildet der Bildmittelpunkt den Koordinatenursprung. `rvec` stellt in diesem Kontext die Verdrehung des Markers in Relation zum Kamerasichtfeld dar.

`tvec` ist ein sogenannter Verbindungsvektor, der vom Koordinatenursprung aus die Positionierung des Markermittelpunkts definiert. Bei der Verwendung der Messwerte entsprechen die Entfernungen des Vektors selbstdefinierter Einheiten. Wenn also bei der Methode `estimatePoseSingleMarkers` Entfernungen von Millimetern als Referenzwerte verwendet werden, beträgt das Ergebnis der Posenschätzung weiterhin Millimeter. Das Lochkameramodell [2, 11] liegt der Berechnung als Basis zugrunde. Die Koordinatensystemausrichtungen sind dementsprechend auch teilweise nach diesem Modell ausgerichtet und folgen nicht der ersten bekannten Intuition. Die positive y-Koordinatenachse geht nach

unten, die positive x-Koordinatenachse nach rechts und die positive z-Koordinatenachse nach vorne, alles immer aus der Betrachtung des Bildes gesehen. Dieser Fakt, dass die Koordinatenachsen anders als die der Mathematik ausgerichtet sind, muss bei der Weiterverarbeitung bedacht werden. Zudem sind die Vektoren immer in Abhängigkeit des aktuellen Kamerakoordinatensystems zu verstehen. Eine Umsetzung auf ein Weltkoordinatensystem für die Positionsbestimmung ist in jedem Fall unabdingbar.

`rvec` stellt die Verdrehung des Markers aus Sicht der Kamera in der Rotationsmatrixdarstellung dar [11]. Diese Darstellungsform ist für eine Verwendung im Kontext von komplexen Vektorrotationen ideal. Da bei diesem Projekt allerdings nur einfache Beschreibungen benötigt werden, ist eine andere Rotationsbeschreibung des Nickens, Gierens und Rollens für diesen Einsatz idealer. Wie Addison [13] beschreibt, müssen die Darstellungsformen für diese Veränderung über mehrere Schritte transformiert werden. Die Methode `rodRigesToQuaternionToEuler` nimmt eine Rotationsmatrix an, wandelt sie über die Rodriguesformel und die Quaternionform in die Eulerdarstellung mit Winkeln in Grad um. Im Gegensatz zur Rotationsmatrixdarstellung lassen sich Winkel in Eulerform einfacher verrechnen und darstellen, da hier keine zusätzlichen Hilfsformen genutzt werden, sondern direkte reale Winkelwerte vorhanden sind.

Neben den Rotationsvektoren und Translationsvektoren stellt `calculatePose` noch die absolute Entfernung des Kameramittelpunkts zum Markermittelpunkt als Berechnung⁸ des Betrags eines Vektors bereit. Die Bildmittelpunktkoordinaten des Markers werden für weitere Verarbeitungsschritte zusätzlich bereitgestellt.

An dieser Stelle der Software ist eine Verwandlung der Rohdaten in ein JSON-Format zur Weitergabe an weitere Akteure notwendig. In Abbildung 5.9 ist zudem eine Visualisierung der detektierten Marker mit den jeweiligen Rotationsmarkierungen zu sehen.

Betrachtungen der Ergebnisse der Posenschätzungen zeigen auf, dass die `rvec` und `tvec` Werte zwischen den einzelnen Frames über die Zeit teilweise stark schwanken, auch wenn die Kamera stabil und ruhig an einer Stelle gehalten wird. Die Unterschiede bei den Translationsvektoren sind nicht so stark ausgeprägt wie bei den Rotationsvektoren. In der Visualisierung der Marker (Abbildung 5.9) lässt sich ein teilweises starkes Zittern der visualisierten Vektordarstellungen ausmachen. Dadurch, dass der Detektionsalgorithmus in jedem Bild die Bereiche, wo Marker zu finden sind, neu bestimmt und die Beleuchtung teilweise leicht abweichen kann, hat die Posenschätzung Schwierigkeiten, die immer exakt gleichen Daten zu bestimmen. Der Versuch einer Anwendung des Closingfilters auf das

⁸<https://github.com/Asadullah-Dal17/Basic-Augmented-reality-course-opencv/tree/master/DISTANCE-ESTIMATION>

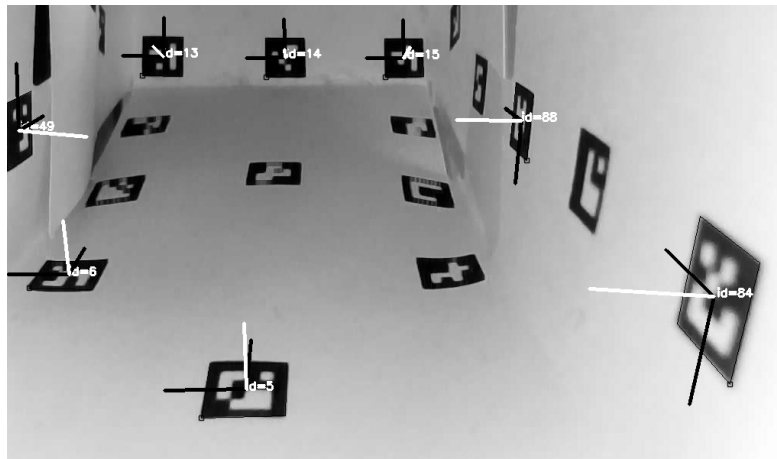


Abbildung 5.10: Markererkennung mit Closing-Effekt

Grauwertbild zur Entfernung von Pfeffer-Rauschen und Beibehaltung der Helligkeit [14] als Lösungsansatz erzeugt, wie in Abbildung 5.10 erkennbar, nur unbefriedigende Ergebnisse. Das Closing reduziert zwar die Rauschbereiche um die Marker herum, entfernt aber zusätzlich teilweise die Binärcodierung. Im Vergleich der erkannten Markermengen, die in Abbildung 5.9 und 5.10 zu sehen und durch die Beschriftungen markiert sind, sind die für eine Weiterverarbeitung verfügbaren Marker reduziert. Da eine erhöhte Menge an Markern jedoch wichtiger ist, um genauere Ergebnisse erlangen zu können, wird diese Lösungsidee verworfen.

5.4.4 Kartierung

Die SLAM Problematik bezeichnet sich als gleichzeitige Positionsbestimmung und Erkundung der Umgebung [15]. Dieses Problem spielt in diesem Projekt teilweise eine Rolle. Dies passiert immer dann, wenn ein neues Kopfphantom verwendet wird und beim erstmaligen Einsatz noch keine Karte mit Positionsinformationen dieses spezifischen Kopfphantoms existent ist. Dadurch, dass dieses Problem der nicht existenten Karte nur bei der erstmaligen Verwendung auftritt, fällt die Entscheidung zur Vereinfachung des SLAM Problems auf einen getrennten Arbeitsschritt dieser beiden Aufgaben. Eine Positionierung muss daher erst nach einer Erstellung einer Karte stattfinden und nicht bereits währenddessen. Außerdem muss dieser Schritt der Kartenerkundung nur einmalig durchgeführt werden, da eine bestehende Karte bei erneuter Benutzung geladen werden kann.

Für die Kartierung werden nun die Translationsvektoren der Marker über eine Koordinatentransformation miteinander verrechnet. Relevant und notwendig für diesen Schritt ist eine Reduktion des starken Zitterns der Marker, weil nur Marker verwendet werden können, die einigermaßen stabil über eine längere Zeit detektiert werden konnten. Die Generierung dieses stabilen Markerabbilds findet über die jeweils letzten 3 Bildframes der Kamera statt. Die Menge der stabilen Marker über diesen Zeitraum von 3 Frames wird im Anschluss zur Kartierung der Methode `calculateCoordtoMarker` übergeben. Die grundsätzliche Funktion dieser Methode ist, aus den Translationsvektoren des Kamerasisichtfelds Koordinaten in einem Weltkoordinatensystem zu bestimmen. Anhand eines definierten Startmarkers, welcher auf den Schneidezähnen positioniert ist, wird der Koordinatensystemursprung gestaltet. Die Wahl fällt hierbei aufgrund mehrerer Einflüsse wie der mathematischen Vorteile des ID-Spektrums und des gewählten Markerwörterbuch auf den Marker 31. Unter Anwendung normaler Koordinatentransformationen und Vektorrechnungen werden nun die jeweiligen Koordinaten durch die im Frame sichtbaren Marker bestimmt. Die Verwendung von Wissen über die Positionierung der Marker in der realen Welt wird zur Verbesserung der Kartierungsergebnisse herangezogen. So können die Marker auf einer Achse alle denselben Wert bekommen, wenn sie in einer Reihe liegen. Je weiter das Laryngoskop in den Rachen vordringt, desto schneller verschwindet der Startmarker. Sobald dies passiert, wird ein bekannter Marker mit Koordinate für die weitere Bestimmung verwendet. Hierbei kann die fortlaufende Bewegung einen Kartierungsfehler erzeugen, weil die reale Position durch den neuen Festmarker abweichen kann. Aus diesem Grund werden für die Kartierung auch alle drei Markerlinien verwendet, um diese Abweichung gering halten zu können.

Unerlässlich ist bei der Kartierung die Verwandlung des Kamerakoordinatensystemvektors in einen für die Arbeit im Weltkoordinatensystem fähigen Vektor. Die Methode `picVecToWksVec` transformiert die Vektoren, sobald sie in die Vektorberechnung integriert werden. Transformiert werden folgende Teile für eine Implementation in ein Weltkoordinatensystem:

- Die Achsen z und y werden miteinander vertauscht, da im Kamerakoordinatensystem die z -Achse als Achse gerade aus dem Bild kommend ausgerichtet definiert ist. Für die Weltkoordinatensystemdarstellung muss diese allerdings als y -Achse für die Höhe betrachtet werden. Dies ist dadurch bedingt, dass die Normalstellung der Kamera des Videolaryngoskops nach unten auf die Zähne gerichtet ist (siehe Abbildung 1.1).

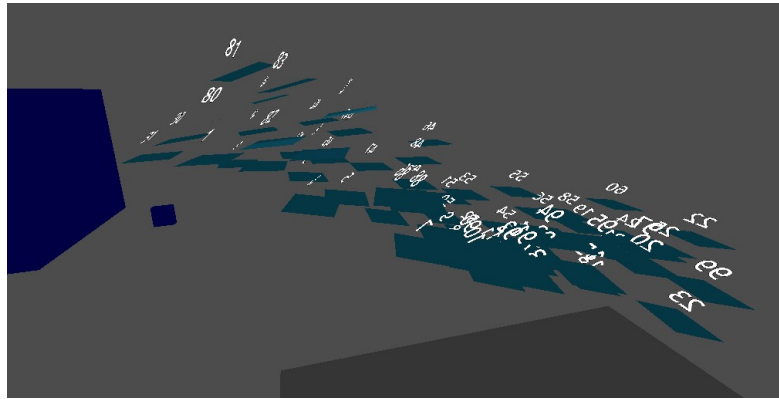


Abbildung 5.11: Erster Darstellungsversuch der Marker im Weltkoordinatensystem ohne Kartierung im 3D-Programm

- Die neue y-Achse muss mit -1 multipliziert werden, da die Ausrichtung durch das Kamerakoordinatensystem verdreht ist.

5.4.5 Darstellung im 3D-Programm

Die Markerkoordinaten und Positionswerte in Form von Kamerakoordinaten und Ausrichtungswinkeln stehen nach außen durch einen Versand ins Netzwerk zur weiteren Nutzung zur Verfügung. Neben den Positionswerten, die fortlaufend bei jeder Aktualisierung des Frames versandt werden, lassen sich auch die Karteninformationen abrufen. Der Zugriff auf diese Daten läuft über eine User Datagram Protocol (UDP) Verbindung des 3D-Programms und der Positionsbestimmungssoftware des Videolaryngoskops. Die Entscheidung auf diesen Verbindungstyp begründet sich darin, dass ein Datenübertragungsverlust der Koordinaten weniger problematisch ist, als wenn die Übertragung einer Transmission Control Protocol (TCP) Verbindung fehlschlägt und die Positionsbestimmungssoftware blockiert wird. Die unterschiedlichen Verbindungsproblematiken dieser Verfahren spielten bei der Wahl des besten Verbindungstyps eine entscheidende Rolle.

Dieses Projekt verwendet zur einfachen Visualisierung der errechneten Informationen die OpenSource Spiel-Engine Godot⁹. Markerkoordinaten, Kameraposition und die Präsentation der Funktion der Kernidee der Bilderinjektion analog zur bestimmten Position werden damit hier umgesetzt. In Godot können in einfachen Schritten 2D- und 3D-Spiele gestaltet und programmiert werden. Neben einer grafischen Benutzeroberfläche

⁹<https://godotengine.org/>

zur Gestaltung der Objekte bietet Godot zusätzlich eine eigene Skriptsprache zur Implementierung von Verhalten. C# und C++ können ebenfalls für diesen Schritt ausgewählt werden. In GDScript dagegen lässt sich einfach einarbeiten, weil die Syntax dieser Sprache an Python angelehnt ist.

Die zugehörige Entwicklungsumgebung stellt einen Szenenmanager, 2D- oder 3D-Ansichten und einen Skripteditor bereit. Die Spiel-Engine stellt ein 2D- oder 3D-Spiel als einen Baum aus mehreren Nodes dar. Kombiniert stellt dies dann eine Szene zur Interaktion bereit [4]. Über einen Szenenbaum wird der Aufbau des Spiels gemanagt. Nodes können in einem 3D-Editor in der 3D-Welt positioniert und angepasst werden.

Speziell die Gestaltung eines 3D-Spiels, wie zur Visualisierung der Dateninformationen oder des Rendering der Realbilder in diesem Projekt, benötigt einige elementare Nodes. Notwendig ist eine Kamera, die die positionierten Nodes und Szenen aufnimmt, da in einer 3D-Welt für die Spieldarstellung auf eine 2D-Darstellung fokussiert wird. Zudem wird eine Lichtquelle benötigt, welche die sichtbaren Bereiche der Kamera bestimmt. Sollten neue Objekte eingefügt werden, passiert dies am einfachsten über eigene Szenen, die in diese Umgebung eingespielt werden.

Die Darstellung der Marker ohne einen angewandten Kartierungsalgorithmus ist in Abbildung 5.11 aufgeführt. Marker sind dort als ebene Flächen mit ihrer ID in Abhängigkeit ihres aktuellen untransformierten Translationsvektor in Form einer Szene dargestellt. Durch die Situation, dass die Marker zu Beginn der Kartierung des Rachens unbekannt sind, ist in Godot dieser Weg der schrittweisen Darstellung zu gehen. Der Koordinatenursprung ist hier aktuell als großer Würfel und eine Darstellungsidee der Kameralinse als kleiner Würfel umgesetzt. Wie bereits in Abschnitt 5.4.4 beschrieben, demonstriert die Abbildung 5.11 zudem die Problematik, dass sich die Achsen in den unterschiedlichsten 3D- und Vektor-Programmen unterscheiden. Godot verwendet in diesem Bereich ein rechtshändiges Koordinatensystem.

Zusätzlich lassen sich in Godot vorab fertig generierte 3D-Modelle einfügen. In Kombination der Testung der Erweiterung zur Darstellung der Marker an einer festgelegten Stelle im 3D-Spiel, wurden einige Entwürfe des Rachens des Kopfphantoms in Blender¹⁰ einer freien 3D-Grafiksoftware zur Modellierung von 3D-Modellen angefertigt. Das Modell ist in Abbildung 5.12 zu sehen und wurde im Endeffekt zum besseren Verständnis von Positionierung im Raum genutzt. Ein Einsatz dieses Modells wurde im 3D-Spiel verworfen. Die Begrifflichkeit 3D-Spiel bezeichnet in dieser Arbeit nun die Umsetzung der Darstellung der Marker, Kamera und Darstellungsform des 3D-Rendering-Versuchs.

¹⁰<https://www.blender.org/>

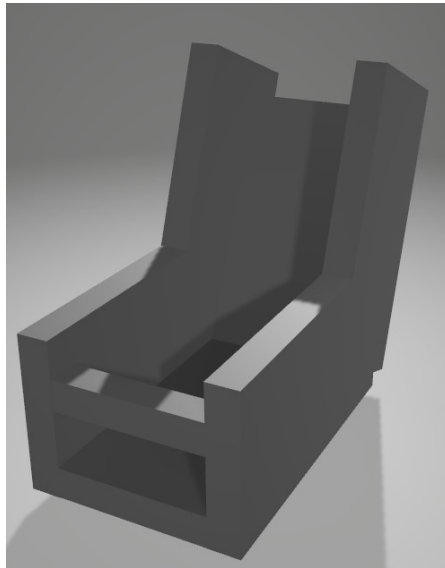


Abbildung 5.12: Versuch einer Nachbildung des relevanten Bereichs des Kopfschnittmodells

5.4.6 Entwicklung 3D-Karte

Die Umsetzung der Darstellung der Demonstration der digitalen Funktionalität der Positionsbestimmung und des 3D-Renderingversuchs wird in Godot sehr einfach gehalten. Neben der 3D-Darstellung wird zudem die Möglichkeit der Bildinjektion eingefügt. Die 3D-Karte stellt sich über eine Einblendung von den Markern in Form von Szenen dar. Zusätzlich wird die Kamera des Videolaryngoskop für eine Außensicht als blauer Würfel eingefügt. Das Bild dieser Kamera im 3D-Spiel, welche zu Beginn in einer Neutralposition steht, verändert sich beim Empfang eines neuen Datensatzes der Positionsbestimmungssoftware über die UDP Verbindung. Die erhaltenen Werte werden zur Veränderung der Position und Ausrichtung in den xyz-Achsen der Weltkoordinatenposition im 3D-Spiel eingesetzt. Weitere Kameras für die Darstellungen im 3D-Spiel, zwischen denen gewechselt werden kann, stehen zur Betrachtung der Außenansicht der 3D-Umsetzung des Kopfphantoms bereit. Die nötigen Marker können ebenso über die UDP Verbindung angezeigt werden. Das Bild der Kamera des Videolaryngoskop ist außerdem in einen Bildinjektionsmodus zur Anzeige von Realbildern an bestimmten Punkten schaltbar. Die zum Testen verwendeten Bilder stammen aus einer Reanimationspuppe.

5.4.7 Kamerapositionsberechnung

Der zweite Aufgabenteil der SLAM Problematik ist die gleichzeitige Positionsbestimmung relativ zu einer gerade neu bestimmten und erkundeten Karte. Diese Aufgabe der absoluten gleichzeitigen Erkundung und Berechnung lässt dieses Projekt außen vor. Die Entscheidung für die Positionsbestimmung, eine bereits vorab berechnete Karte des aktuell ausgewählten Kopfphantoms zu produzieren und dann zu nutzen, lässt sich einfach umsetzen. Dies reduziert und blendet das Kernproblem der Gleichzeitigkeit aus. Der einzige Moment, wo es eine gefühlte Gleichzeitigkeit dieser Aufgabenbereiche geben könnte, ist, wenn die Karte gerade erst neu erkundet wird.

Die Positionsbestimmung benötigt dennoch zur guten Funktionalität dieselben Voraussetzungen wie die Kartierung. Stabile Marker sind hierbei ebenfalls gefordert. Wie bei der Kartierung werden in diesem Zusammenhang die Marker der letzten 3 Frames der Kamera analysiert und diejenigen Marker, die in allen sichtbar waren, zur Verarbeitung weitergegeben. Die Verwendung von 3 Frames beruht auf der Entscheidung und Testung, was die Endoskopkamera aufgrund ihrer Spezifikationen leisten kann. Dadurch, dass die Kamera maximal etwa 27 Frames per Second (FPS) generieren kann, ist der Spielraum für die Menge der nutzbaren Frames nicht sonderlich ausgeprägt, wenn als Teilziel eine Echtzeitverwendung möglich sein soll. Es bleibt damit eine Überprüfung der Markerwerte von 9 FPS konkret 9 Positionsberechnungen pro Sekunde übrig. Dies entspricht einer Aktualisierungshäufigkeit der Positionen in der 3D-Karte von etwa 111 ms. Das ist etwa 2,6 mal länger als bei dem Idealwert von 24 FPS, wo der Mensch eine Bildfolge als flüssige Bewegung wahrnimmt. Trotzdem reicht dieser Wert in diesem Anwendungsgebiet gerade aus, weil neben der Echtzeitanforderung beim Umgang mit dem Videolaryngoskop vor allem die gewissenhafte Anwendung zur Verletzungsreduktion eine Rolle spielt und die Bewegungen dadurch deutlich langsamer durchgeführt werden. Eine langsamere Aktualisierung sensibilisiert hierbei den korrekten Umgang. Das avisierte Einsatzgebiet dieses Prototyps ermöglicht einen größeren Spielraum.

Unabhängig dessen schwankt die reale Aktualisierungszeit der Anwendung zusätzlich um eine Zeitspanne von etwa 20 ms. Die tatsächliche Dauer beträgt etwa zwischen 95 ms und 120 ms. Dieser Fakt hängt allerdings mit der Positionsbestimmung zusammen, die durch unterschiedlich lange Schleifendurchläufe schwanken kann. Diese Designentscheidung, eine Art der Superloop zu nutzen, ist vor allem durch das kontinuierliche Lesen des nächsten Frames der Kamera bedingt. Diese Tests liefen unter zwei Rechnern mit den Prozessorspezifikationen „12th Gen Intel(R) Core(TM) i7-12700K @ 3.61 GHz“ und „11th Gen Intel(R) Core(TM) i5-1135G7 @ 2.40 GHz“ relativ gleich ab.

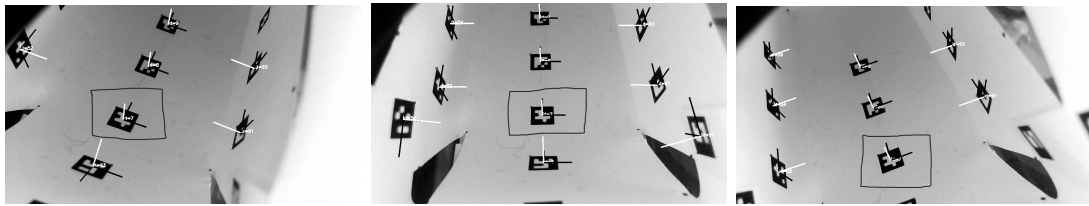


Abbildung 5.13: Test der Winkel des Marker 7: Linksdrehung, Normalstellung, Rechtsdrehung

Alle 3 Frames wird entsprechend in der Software die Methode `calculateCamPosition` als ein Kernprodukt dieser Bachelorarbeit aufgerufen. In dieser Methode wird der letzte verfügbare Markerdatensatz der Frames verwendet, um daraus die Position und Drehung der Kamera des Videolaryngoskops zu berechnen. Diese beiden Berechnungen laufen allerdings getrennt voneinander ab, sodass eine Neuberechnung derjenigen Information nur bei einer tatsächlichen Veränderung umgesetzt wird.

Position

Wie bereits beschrieben, schwankt die Aktualisierungsdauer. Außerdem können die berechneten Markermesswerte ebenfalls zwischen den einzelnen Frames durch das Pfeffer-Rauschen am Markerrand schwanken. Entsprechend wird für die Positionsbestimmung ein Algorithmus entworfen, welcher es trotzdem schafft, stabile Weltkoordinaten für eine Lokalisation der Position zu bestimmen.

Dadurch, dass jeder Translationsvektor eines verfügbaren Markers die Beschreibung einer Position ist, kann durch die bekannte Karte mit den Weltkoordinaten der Marker über einfache Vektoroperationen wie Addition, Subtraktion und Multiplikation die Kameraposition für diesen speziellen Marker rückgerechnet werden [1]. Beginnend mit einer Transformation des Translationsvektors in das Weltkoordinatensystem, werden diese durch eine Multiplikation in der Richtung gedreht und zum Richtungsvektor des Ortsvektors, der Weltkoordinate des aktuellen Markers, gemacht. Die Spitze dieses Vektors zeigt nun auf die Position der Kameralinse. Im Anschluss werden alle so berechneten Kamerakoordinaten trianguliert, sodass eine spezifische Kameraweltkoordinate bestimmt ist. Hierzu werden alle aktuell im Kamerabild verfügbaren Marker der mittleren Markerlinie verwendet, um eine hohe Genauigkeit der Position zu erlangen. Zur Reduktion von Schwankungen dieser Koordinate wird eine neue Koordinate nur dann zum Versand ins Netzwerk geschickt, wenn es signifikante Bewegungsänderungen auf einer der drei Achsen gibt.

Drehung

Anders als die Position ist die Drehung nicht möglich über eine direkte Transformation der Rotationsvektoren zu berechnen. Die Eigenschaft, dass die Rotationsvektoren nur eine Information der aktuellen Kameraposition bereitstellen können, stellt eine Schwierigkeit bei der Berechnung der Roll-, Gier- und Nickwinkel der Kameraausrichtung dar. Es gibt für die Berechnung der Drehung aus den Markern keine für alle Winkel vollständig festgelegten und eindeutigen Referenzstellen, an denen sich für eine Verrechnung orientiert werden kann. Referenzstellen sind in diesem Fall Punkte, die dauerhaft aus Sicht des Kamerabildes im Zusammenhang der Weltausrichtung gleich bleiben, egal wie die Kamera positioniert oder gedreht ist. Dies liegt daran, dass die Rotationsvektoren von einer sich durch die Position dauernd ändernden Berechnungsgrundlage durch die Bewegung entstehen.

Wie in Abbildung 5.13 zu sehen, lassen sich für zwei der drei Winkel trotzdem direkte Eigenschaften aus den Vektoren ableiten, die bei der kombinierten Verbindung von spezieller Platzierung der Marker in der realen Welt und der Kamerabildberechnungseigenschaften zu finden sind. In diesen Bildern wurde der Rotationsvektor des Markers 7, der schwarz umrandet ist, auf das Verhalten der Drehung des Videolaryngoskops betrachtet. Durch die spezielle Platzierung der Marker in der Realumgebung lässt sich im mittleren Bild bei einer Normalstellung des Videolaryngoskops der Gierwinkel auf 0,62 Grad und der Rollwinkel auf 1,11 Grad bestimmen. Eine Linksdrehung und Rechtsdrehung ergeben jeweils fast dieselben Werte, links Gierwinkel bei 1,16 Grad und Rollwinkel bei 20,93 Grad und rechts Gierwinkel bei -1,39 Grad und Rollwinkel bei -20,63 Grad. Diese Winkel sind bei den Drehungen jeweils etwa gleich weit gedreht und haben nur um das Vorzeichen gedrehte Änderungen in Bezug auf die Richtung. Durch dieses Verhalten, dass hier eine feste Definition von Winkeln für einen kleinen Bereich festzustellen ist, ist für die Rotationsbestimmung auf zwei Winkelachsen diese Eigenschaft ideal zu verwenden.

Der Nickwinkel kann dieses Verhalten nicht widerspiegeln, sodass eine ähnliche Bestimmung nicht stattfinden kann. Dies liegt daran, dass dieser Winkel sich nicht an den Rotationsreferenzpunkten orientieren kann. Ein anderer Zusammenhang von Messwerten und beobachteten Eigenschaften des Videolaryngoskops kann aber für die Nickwinkelbestimmung verwendet werden. Eine Betrachtung der sich im Verlauf der Bewegung durch den Rachen verändernden Markerentfernungen lässt sich hierfür anwenden. Die Abbildung 5.14 zeigt den Verlauf aller Entfernungen der einzelnen Marker. Eine Analyse dieser Werte zeigt, dass auf ein Muster in dem Datensatz zuschließen ist, welches eine Abhängigkeit der Entfernungen von einem spezifischen Nickwinkel an der jeweiligen Position darstellt.

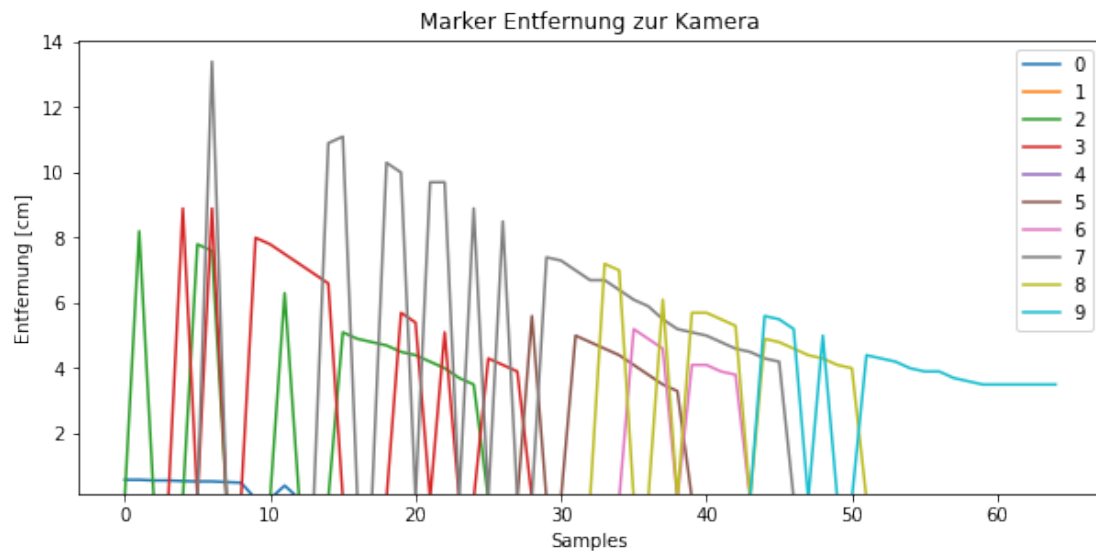


Abbildung 5.14: Entfernungen von Markern zur Kamera im Durchlauf des Rachens

Abbildung 5.15 demonstriert den Zusammenhang besser, da durch eine Verschiebung der Messpunkte die Kurve und das Muster besser erkennbar ist. Anhand der Grafik lassen sich nun Stellen ausmachen, wo spezifische Winkeländerungen auftreten. So ist bei Sample 12, konkret an den Punkten des Marker 2, der Sprung der Kamera in den Mund sichtbar und ab Sample 28 bis 42 der Schritt des aktiven Kippens des Videolaryngoskops zur Streckung der Zunge erkennbar. Dieser Zusammenhang von der Entfernung und dem Winkel kann in der Implementation über einen LookUp-Table umgesetzt werden, dass ein korrespondierender Winkel zur Entfernung ausgegeben wird.

Der Drehungsbestimmungsalgorithmus startet unabhängig der Winkelberechnungsschwierigkeit mit einer Analyse zur Wahl eines verwendbaren Markers für diesen Schritt. Diese Aufgabe teilt sich in zwei Schritte:

- Zu Beginn findet die Identifikation eines stabilen Markers aus der mittleren der drei Markerlinien im Rachen statt. Gefiltert wird dieser Marker nach Intensität des Zitterns. Dies bezeichnet Marker, die in den letzten zwei Aktualisierungen keine 180 Grad Drehung vollzogen haben. Das Pfeffer-Rauschen lässt solche Abweichungen entstehen. Weiterhin wird geprüft, ob eine signifikante Drehungsänderung stattgefunden hat. Dieser Schritt kann ebenfalls die Schwierigkeit des Pfeffer-Rauschens reduzieren.

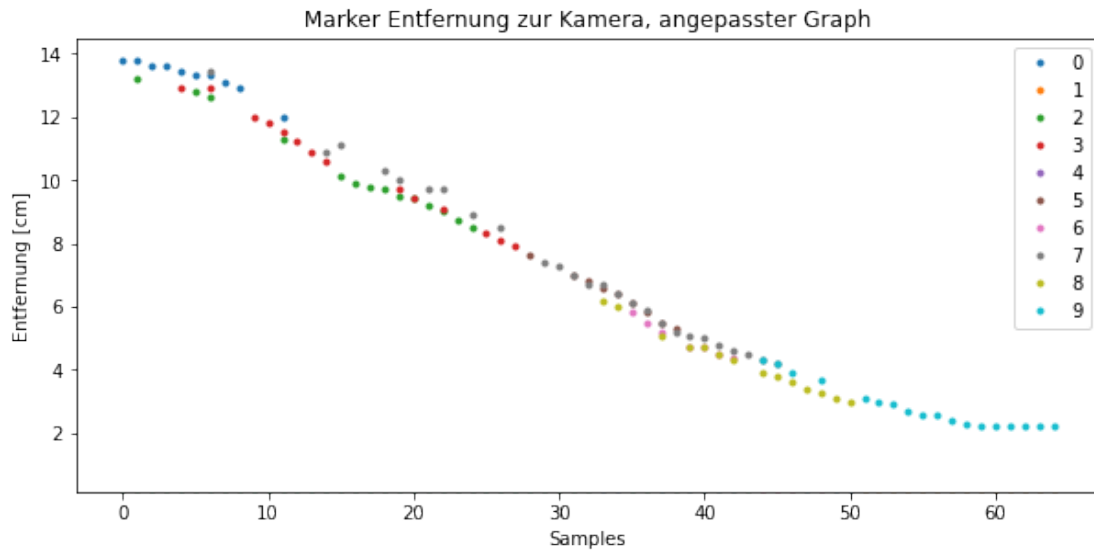


Abbildung 5.15: Entfernungen von Marker zur Kamera im Durchlauf des Rachens, charakteristischer Graph

- Im weiteren Schritt folgt aus den übriggebliebenen Markern die Suche nach demjenigen, der sich am mittigsten im Kamerabild befindet. Dies ist dadurch begründet, dass häufig durch den ersten Filterschritt eine Auswahl zwischen mehreren potenziellen Markern getroffen werden muss. Dabei stellt der mittigste Marker meistens die beste Möglichkeit dar. Dieser ist meist einer, der über einen längeren Zeitraum sichtbar ist und wenig als Referenzpunkt gewechselt werden muss.

Die endgültige Rotationsbestimmung übernimmt die Methode `calculateCamRotation`. In dieser werden die Roll- und Gierwinkel des ursprünglichen Rotationsvektors mit dem Nickwinkel aus der LookUp-Table kombiniert. Sollte keine gültige Rotation berechnet worden sein, wird die Ursprungsrotation vom Einführen des Videolaryngoskops verwendet. Dies ist nötig, um kontinuierliche valide Werte an das Netzwerk weitergeben zu können, wenn bei der Filterung der potenziellen Marker Fehler auftreten. Die Verwendung dieser LookUp-Table muss allerdings im Idealfall für jedes Kopfphantom speziell angefertigt und ausgemessen werden, um die bestmöglichen Ergebnisse zu erlangen. Die für das durchsichtige Kopfschnittmodell angefertigten Werte stellen trotzdem eine gute Universalität dar. Dabei ist es dennoch möglich, da nicht jeder Winkel abgebildet werden kann, dass die Rotation Sprünge in der Drehung produziert.

5.4.8 Probleme und Lösungsansätze

Der entwickelte Algorithmus funktioniert mit der Detektion und Positionsbestimmung relativ stabil. Auf Datenfehler und Bildausfälle ist der Entwurf angepasst, sodass selbst das 3D-Spiel bei solchen Fehlern weitere Rückgaben liefert.

Ein Problem steht dennoch bereits am Anfang der Verarbeitungskaskade. Mit der Qualität des Eingangsbilds aus dem aktuellen Kameraframe steht und fällt der Erfolg der Verarbeitung. Einflüsse nehmen hierauf vor allem die äußeren Umstände der Marker. Probleme machen speziell die starken Verzerrungen und Unschärfen bei den Markern, die sich an den Rändern des Rachens des Kopfphantoms befinden. Weiterhin hat die Beleuchtung einen sehr großen Effekt auf die Detektionsfähigkeit der Marker. Durch die Verschiebung des Lichtkegels entstehen zusätzliche Schwierigkeiten in Bezug auf Pfeffer-Rauschen.

Lösungen, vor allem für einen späteren Einsatz in einer echten Reanimationspuppe, könnten folgende sein:

- Eine starke und großflächige Ausleuchtung des gesamten Rachens durch weitere LEDs kann das in Abschnitt 5.4.3 beschriebene Pfeffer-Rauschen reduzieren und die Kanten der schwarz-weißen Codierungen der ArUco-Marker mehr differenzieren.
- Der Druck der Marker muss auf kleinen Platten aus Aluminium oder ähnlichem Material für formstabile Bereiche stattfinden, damit die Marker als quadratische Flächen identifiziert werden können.

5.5 Anpassungen nach erster Prototypvorstellung

Der Prototyp bestehend aus Kopfphantom und Videolaryngoskop wurde im Rahmen der Entwicklungsschritte einmal den Experten und Ansprechpartnern Dr. med. Philipp Breitfeld und Dr. med. Martin Petzoldt im Simulationszentrum des UKE zur Ansicht und Austausch vorgestellt. Die präsentierte Version hatte zu dem Zeitpunkt noch keine fertiggestellte, vollständig funktionierende Rotationsbestimmung der Kamera. Weiterhin war keine Funktionalität der Bildinjektion im 3D-Spiel implementiert. Die grundsätzlich geplante und angestrebte Funktionsidee konnte trotzdem voll demonstriert werden.

Durch die Präsentation dieses ersten Entwurfs konnten vor allem Schwierigkeiten vorgeführt werden, die teilweise durch die geforderten Spezifikationen entstanden sind. Präsentiert wurde die Möglichkeit, wie ein Videolaryngoskop mit der Trainingsverbesserung

aussehen könnte. Dabei wurde der experimentelle Entwurf dargestellt, der anhand dieser Vorgaben zu Beginn des Projekts nur über die grundsätzliche Idee des Projekts anzuwenden war.

Eine Rückmeldung dieser ersten Präsentation ist, dass das Videolaryngoskop nicht in vollständiger Länge in den Rachen des Kopfphantoms eingeführt werden muss. Eine Tiefe, sodass die Spatelspitze das Zungenende gerade erreicht, ist ausreichend. Die speziellen Anpassungen durch medizinische Vorgaben sorgen dazu, dass die Platzierung des Markers 7 ideal ist und erweitert genutzt werden kann. Vorteil an diesem ist, dass er bereits zu Beginn des Einführens des Laryngoskops sichtbar ist und erst sehr spät aus dem Sichtfeld verschwindet. Diese Eigenschaft verringert zudem das Problem, dass je weiter die Kamera in den Rachen geführt wird, ein Positionsfehler bei der Erkundung entsteht und auch erweitert wird. Die Abweichung der Position von der Realität kann durch diese Anpassung reduziert werden. Die relevanten Bereiche sind damit weiter vorne, oben im Rachen lokalisiert.

Weiterhin hat sich herausgestellt, dass eine vollständige Reanimationspuppe nur bedingt verwendbar ist, da die Positionierung auch bei der geringen Anzahl an Markern schwierig umgesetzt werden kann. Eine Puppe müsste vermutlich in jedem Fall aufgeschnitten werden, damit die nötigen Marker platziert werden können. Als Idee für eine weitere Entwicklung zur Lösung in einer anderen Arbeit wurde der Vorschlag getätigt, eventuell einen eigenen Kopf für dieses Aufgabengebiet zu gestalten. Dieser würde dann die statischen Eigenschaften des durchsichtigen Kopfschnittmodells von VBM beibehalten und in den Bereichen der Zunge mit weichem und biegsamem Material zur Schaffung der realen Haptik ausgetauscht werden. Als erste Idee für diesen Teil wurde die Verwendung eines Knautschballs vorgeschlagen, weil dieser die ähnlichen Eigenschaften der Gummifestigkeit wie eine Puppe in dem Bereich aufweist. Neben diesen Vorschlägen wurde auch von der Möglichkeit einer Gamifizierung des gesamten Einsatzes des Laryngoskops gesprochen, damit eine Art Blackbox entworfen wird, die oben das Laryngoskop in Form eines Joysticks bereitstellt. Die nötigen Positionen und Drehungen könnten dann anhand der Bewegung des Joysticks abgeleitet und berechnet werden. Hierzu ist allerdings anzumerken, dass dies die ursprüngliche Idee der Verwendung eines realen Videolaryngoskops außer Acht lässt.

6 Ergebnisse

Der vollständige Prototyp ist in seiner Endform in Abbildung 6.1 zu finden. Die Abbildung stellt den kombinierten Einsatz der einzelnen physischen und softwaretechnischen Elemente dar. Die beiden Abbildungen zeigen die Verwendung des durchsichtigen Kopfschnittmodells mit den eingesetzten benötigten Markern. Die Marker sind in diesem Modell auf einem Papierstreifen aufgedruckt, der zugeschnitten und gefaltet im Rachen eingelegt ist. Das Videolaryngoskop ist zur Hälfte in den Rachen eingeführt. Im hinteren linken Bereich ist jeweils das Livebild der Kamera zusehen. Detektierte Marker sind mit ihren Achsen gekennzeichnet.

Das linke Bild zeigt die Verwendung des Bildinjektionsmodus des 3D-Spiels, während das rechte Bild die Positionierung anhand der Weltkoordinaten im 3D-Raum präsentiert.

Die Veränderung der Position oder Anpassung des aktuellen 3D-Renderingbildes funktioniert im geforderten Umfang. Das haptische Gefühl eines realen Videolaryngoskops ist vorhanden. Wie in Abschnitt 5.4.7 beschrieben, ist eine leichte Verzögerung der Bildaktualisierung festzustellen. Beim Testen stellte dies jedoch keine große Problematik dar, weil das Gefühl der realen Haptik überwog. Die Schwierigkeiten der ungenauen Positionsbestimmung im Sinne der zitternden und instabilen Marker sind wenig bis kaum bei diesem Entwurf zu bemerken. Die Verwendung des Videolaryngoskops im Rachen des durchsichtigen Kopfschnittmodells lässt zudem keine hektischen Bewegungen zu. Festzustellen sind dann keine Änderungen der Position oder Drehung, allerdings sind diese hektischen Bewegungen auch nicht im Sinne der Bedienung.

Zum Abschluss wurde der Entwurf an einer vollständigen Reanimationspuppe, welche durch die Johanniter-Unfall-Hilfe e.V. Regionalverband Hamburg bereitgestellt wurde, getestet. Dazu wurde ein spezieller Papierstreifen mit den für die Positionsbestimmung nötigen Markern vorbereitet, in den Mund eingeführt und am Gaumen befestigt. Das Ende des Streifens an den Lippen der Puppe ist in Abbildung 6.2 zu erkennen. Das Videolaryngoskop wird identisch zum durchsichtigen Kopfschnittmodell bei dieser Puppe angewandt. Das Ergebnis der Positionsbestimmung ist in Abbildung 6.2 im hinteren

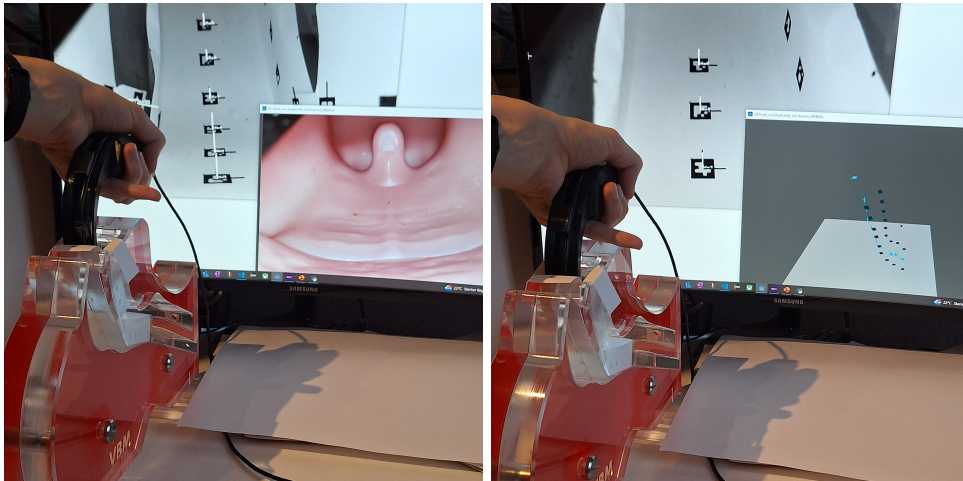


Abbildung 6.1: Demonstration der Funktionalität des Prototyps: Bildinjektion und Positionsdarstellung

Monitor zu sehen. Die Funktionalität zwischen den Kopfphantomen ist dieselbe. Als Kartengrundlage konnte problemlos die des durchsichtigen Kopfschnittmodells verwendet werden.

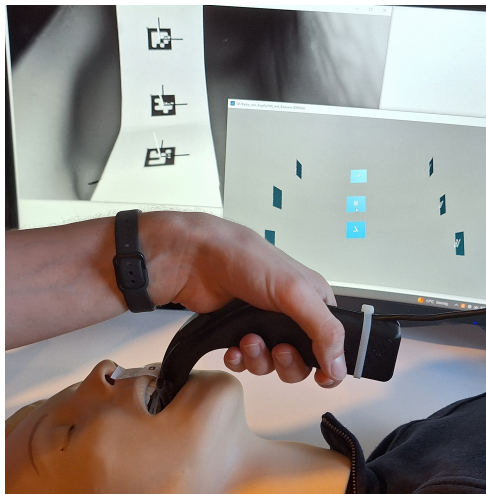


Abbildung 6.2: Demonstration der Funktionalität des Prototyps an einer Reanimationspuppe

7 Ausblick

Ziel dieser Bachelorarbeit war es, einen möglichen Positionsbestimmungsalgorithmus mit entsprechender Erkennungstechnologie in Form eines Prototyps zu entwickeln. Hierbei lag der Fokus auf den experimentellen Schritten zur Zusammensetzung des gesamten Produkts, bestehend aus Software und notwendigen physischen Teilen. Der Prototyp hat dabei eine Entwurfsrichtung für die Einsatzgebiete der Technologien und die Verwendung dieses Ideenspektrums im Simulationszentrum des UKE geliefert. Aus den Erkenntnissen dieser ersten Umsetzung lassen sich einige Vorschläge für weiterführende Arbeiten in diesem Fachgebiet ableiten.

Rendering der Realbilder

Unabhängig neuer Erkenntnisse stellt diese Bachelorarbeit nun mit den Koordinaten der Kamera und Marker und der Rotation alle nötigen Informationen bereit, um ein Rendering der Realbilder aus dem Rachen einer Person zu ermöglichen. Dieser Arbeitsschritt ist für eine Vervollständigung der gesamten Projektidee und für einen Produktivbetriebseinsatz unabdingbar.

Erweiterbarkeit des Einsatzbereichs

Der Entwurf der optischen Positionsbestimmung ist so designt, dass eine Integration in andere in der Medizin verwendete Videoskopien einfach möglich ist. Für die jeweilige Videoskopie, also der genauen Beobachtung mithilfe eines Endoskops, sei es die Bronchoskopie, Gastroskopie oder eine andere, ist der Einsatz theoretisch einfach umsetzbar. Softwareseitig muss dafür vermutlich fast nichts angepasst werden. Die Verwendung der Markererkennung müsste aufgrund der äußeren Umstände wie Beleuchtungsmöglichkeiten und Platz für die Marker in einer solchen Arbeit analysiert und angepasst werden.

Verbesserung der Detektion

Neben der Erweiterung der Einsatzgebiete können weiterführende Arbeiten konkret die Detektionstechnologien verbessern, sodass eine höhere Genauigkeit der Kameraposition

und Ausrichtung erreicht werden. Ein anderer Schritt wäre die Erhöhung der Echtzeitaktualisierung. Mögliche Ansätze wären hierzu die Verwendung anderer Marker. Markertechniken mit einer genaueren Detektion vor allem in den Kantenbereichen zur geringeren Verschiebung der tatsächlichen Positionierung in den Bildern könnten hilfreich sein. Herr Prof. Dr. Peer Stelldinger der HAW Hamburg hat eine Markertechnik entwickelt, die sich denen der ArUco-Marker ähneln. Die Technik besteht aus einer Art Schachbrettmuster mit Einbuchtungen an den geraden Stücken. Dadurch kann die Position selbst an abgerundeten Bereichen eindeutig bestimmt werden. Diese Eigenschaft könnte für dieses Projekt hilfreich sein.

Die Anpassung der verwendeten Sensoren zur Lokalisierung kann ebenso Abhilfe verschaffen. Eine veränderte Beleuchtung, aber auch die Wahl einer höherwertigen Kamera ist denkbar.

Zusätzlich könnte es auch Erfolg versprechend sein, weitere Sensortypen miteinander zu kombinieren. Dies wäre zwar eine Abkehr des Ziels dieser Arbeit, nur die optischen Möglichkeiten einer Kamera zu nutzen, aber trotzdem mit viel Potenzial versehen. Ein Beispiel dazu ist in einer Arbeit von Seyed Shahim Vedaie und Khan A. Wahid [16] zu finden. Diese verwenden zu einer Lokalisation eines Kapselendoskops im Dünndarm eine Kombination von Kameras und Magnetsensoren. Für die Erweiterung dieser Bachelorarbeit könnten sich Geschwindigkeitssensoren, Gyroskope oder Magnetismussensoren bei der Rotationsbestimmung der Kamera eignen.

8 Fazit

Im Rahmen dieser Bachelorarbeit wurde ein optischer Positionsbestimmungsalgorithmus für die Verwendung eines Videolaryngoskops im Rachen einer Reanimationspuppe entworfen. Dabei wurden neben softwaretechnischen Schritten auch notwendige physische Entwicklungen des Prototyps umgesetzt. Die verwendete Positionierungstechnik über ArUco-Marker konnte für einen Einsatz in diesem medizinischen Zusammenhang für Trainingsgeräte erprobt und erfolgreich eingesetzt werden. Die Anwendung dieser Marker vereinfachte zudem aufwendige Positionierungsverfahren, die sich an unspezifischen Markierungen orientieren. Die Verwendung eines optischen Mustererkennungsverfahrens stellte sich als teilweise robust, einfach umsetzbar und erweiterbar dar.

Das Endprodukt des Prototyps aus Kombination des eigens zusammengestellten Videolaryngoskops und verändertem Kopfphantom in Bezug auf das innere Aussehen funktioniert für den geforderten Einsatzbereich einwandfrei. Es konnte in erster Linie ein erweitertes und verbessertes Anwendungserlebnis der Videolaryngoskopie präsentiert werden. Die dafür nötigen Anforderungen der Idee des Projekts im Sinne von Echtzeitübersetzung, Verwendung der ursprünglichen Kamera eines Videolaryngoskops und der Koordinatenerstellung für ein 3D-Rendering der Realbilder konnten für eine erfolgreiche Funktionalität umgesetzt werden. Das ursprüngliche Gefühl der Verwendung eines Videolaryngoskops konnte eingehalten werden.

Die Verwendung von Python und Godot als Programmiersprachen machte eine Umsetzung schnell und einfach möglich. Diese Entscheidung ermöglicht zudem die unkomplizierte Adaptierbarkeit zur Weiterentwicklung des Prototyps.

Die Herausforderungen des Projekts bestanden vor allem darin, im Rahmen der durch das Videolaryngoskop erzeugten Einschränkungen, die bestmögliche Genauigkeit für die Positionierung und Beibehaltung des realen Gefühls zu erlangen.

Literaturverzeichnis

- [1] FELSCH, Mathias ; MARTIN, Karlheinz: *Das große Tafelwerk interaktiv - Formelsammlung für die Sekundarstufen I und II*. Cornelsen Verlag, 2009. – 62–67 S. – ISBN 978-3-464-57144-6
- [2] GARRIDO, Sergio ; PANOV, Alexander: *Detection of ArUco Markers*. docs.opencv.org, 2015. – URL https://docs.opencv.org/3.1.0/d5/dae/tutorial_aruco_detection.html
- [3] GARRIDO-JURADO, S. ; MUÑOZ-SALINAS, R. ; MADRID-CUEVAS, F.J. ; MARÍN-JIMÉNEZ, M.J.: Automatic generation and detection of highly reliable fiducial markers under occlusion. In: *Pattern Recognition* 47 (2014), Nr. 6, S. 2280–2292. – URL <https://www.sciencedirect.com/science/article/pii/S0031320314000235>. – ISSN 0031-3203
- [4] Godot (Veranst.): *Godot Engine Dokumentation*. docs.godotengine.org. 2014. – URL <https://docs.godotengine.org/de/4.x/>
- [5] GORANOVIĆ, Tatjana: Videolaryngoscopy, the Current Role in Airway Management. In: SHALLIK, Nabil A. (Hrsg.): *Special Considerations in Human Airway Management*. Rijeka : IntechOpen, 2020, Kap. 13. – URL <https://doi.org/10.5772/intechopen.93490>
- [6] GÖTZ, Eva ; THEILER, Lorenz ; KNAPP, Jürgen: Videolaryngoskopie – Pro und Kontra. In: *Notfall + Rettungsmedizin* 27 (2023), S. 100–104. – URL <https://link.springer.com/article/10.1007/s10049-023-01261-7>
- [7] Karl Storz (Veranst.): *Gebrauchsanweisung C-MAC Videolaryngoskop 8401xxx, Elektronikmodul 8401/8402 X, C-MAC PM 8401 XD*. 2019

- [8] KUSAY-MERKLE, Ursula: *Minimum Viable Product & Minimum Marketable Features – Warum es so wichtig ist, in „Wellen“ oder „Häppchen“ zu denken.* S. 171–175. In: *Agiles Projektmanagement im Berufsalltag: Für mittlere und kleine Projekte.* Berlin, Heidelberg : Springer Berlin Heidelberg, 2018. – URL https://doi.org/10.1007/978-3-662-56800-2_14. – ISBN 978-3-662-56800-2
- [9] LEHMANN, Prof. Dr. T.: *Vorlesungunterlagen Intelligente Sensorsysteme.* HAW Hamburg, SoSe 22
- [10] LINS, Prof. Dr. C.: *Vorlesungunterlagen Health Informatics.* HAW Hamburg, WiSe 22
- [11] OpenCV (Veranst.): *Camera Calibration and 3D Reconstruction.* docs.opencv.org. 2017. – URL https://docs.opencv.org/3.3.1/d9/d0c/group__calib3d.html
- [12] ROMERO-RAMIREZ, Francisco J. ; MUÑOZ-SALINAS, Rafael ; MEDINA-CARNICER, Rafael: Speeded up detection of squared fiducial markers. In: *Image and Vision Computing* 76 (2018), S. 38–47. – URL <https://www.sciencedirect.com/science/article/pii/S0262885618300799>. – ISSN 0262-8856
- [13] SEARS-COLLINS, Addison: *How to Perform Pose Estimation Using an ArUco Marker.* automaticaddison.com, 2021. – URL <https://automaticaddison.com/how-to-perform-pose-estimation-using-an-aruco-marker/>
- [14] SEPPKE, Dr. B.: *Vorlesungunterlagen Mustererkennung und Maschinelles Lernen / Bildverarbeitung.* HAW Hamburg, WiSe 22
- [15] THRUN, Sebastian ; LEONARD, John J.: *Simultaneous Localization and Mapping.* S. 871–889. In: SICILIANO, Bruno (Hrsg.) ; KHATIB, Oussama (Hrsg.): *Springer Handbook of Robotics.* Berlin, Heidelberg : Springer Berlin Heidelberg, 2008. – URL https://doi.org/10.1007/978-3-540-30301-5_38. – ISBN 978-3-540-30301-5
- [16] VEDAEI, Seyed S. ; WAHID, Khan A.: MagnetOFuse: A Hybrid Tracking Algorithm for Wireless Capsule Endoscopy Within the GI Track. In: *IEEE Transactions on Instrumentation and Measurement* 71 (2022), S. 1–11
- [17] *Augmented Reality Marker.* worldofvr.de. – URL <https://worldofvr.de/augmented-reality-marker/>. – Accessed 19-08-2024

A Anhang

A.1 Digitaler Anhang

Auf dem Datenträger befinden sich neben dieser Bachelorarbeit in PDF-Form folgende für diese Arbeit relevante Dokumente:

- Ordner **3D-Modell**: Godot-Projektordner der Visualisierung der Positionierung und Ausrichtung des Videolaryngoskop
- Datei **calibrate_cam.py**: Code zur Durchführung der Kamerakalibrierung
- Datei **videolaryngoscope_aruco_position.py**: Code zur Positionierung und Ausrichtung des Videolaryngoskop anhand von Livebildern einer Kamera
- Datei **Kalibrierungsschachbrett.svg**: Kalibrierungsschachbrett zum Ausdrucken
- Datei **Kopfphantom-Marker.pdf**: Vorlage zum Ausdrucken der ArUco-Marker

A.2 Verwendete Hilfsmittel

In der Tabelle A.1 sind die im Rahmen der Bearbeitung des Themas der Bachelorarbeit verwendeten Werkzeuge und Hilfsmittel aufgelistet.

Tabelle A.1: Verwendete Hilfsmittel und Werkzeuge

Tool	Verwendung
L ^A T _E X	Textsatz- und Layout-Werkzeug verwendet zur Erstellung dieses Dokuments
Overleaf	Editor und Compiler für L ^A T _E X
LanguageTool	Werkzeug zur Überprüfung der Rechtschreibung
OpenCV	Python-Bibliothek zur Verarbeitung von Bildern und Erkennung der ArUco-Marker
Godot	Spiel-Engine zur Darstellung von 3D-Inhalten

A.3 Bedienungsanleitung Videolaryngoskopsoftware

In den Tabellen A.2 und A.3 sind die Tastaturbelegungen der Python- und Godot-Programme beschrieben. Zum Starten der Software sollte das Godot 3D-Spiel als erstes gestartet werden. Dies nimmt nach dem Start der Python-Analysesoftware direkt über die UDP-Verbindung die Kartendaten und dann fortlaufend die Kameraposition an.

Das Python-Programm wird mit dem Parameter `estimate` zur Positionsbestimmung gestartet. Dies setzt allerdings eine bestehende Koordinatendatei voraus. Durch das Anhängen des Parameters `create` wird der Kartierungsalgorithmus mit ausgeführt und eine Koordinatendatei erstellt.

Tabelle A.2: Tastaturbelegung Python

Taste	Funktion
q	Programm schließen
d	Aktuelle Kartendarstellung löschen
n	Bestehende Karte erneut via UDP versenden

Tabelle A.3: Tastaturbelegung Godot

Taste	Funktion
1, 2, 3, 4, 5	Kamera auf gewählte Ziffern setzen
Tab	Nächste Kamera auswählen
r	Wechsel zwischen Positionsdarstellung und Realbildern
Pfeiltasten, w, a, s, d, Shift, Leertaste	Bewegung der freien Kamera 5

A.4 Bedienungsanleitung Kamerakalibrierung

Das Python-Programm wird mit dem Parameter `getImage` zur Erstellung der nötigen Schachbrettfotos gestartet. Es werden zwischen 20 und 30 Stück für ein gutes Ergebnis benötigt. Mit dem Parameter `calibrate` werden die Kalibrierungsparameter bestimmt und gespeichert. Tabelle A.4 zeigt die Tastaturbelegungen des Python-Programms.

Tabelle A.4: Tastaturbelegung Python

Taste	Funktion
q	Programm schließen
s	Bild speichern

Glossar

Anästhesie Zustand des Menschen, in dem es einen vollständigen Verlust der Steuerung und Empfindung der Sinnesorgane gibt, Umgangssprachlich auch "Narkose" genannt.

endotracheal in der Trachea (Luftröhre).

Endotrachealtubus Eine Art Schlauch, der über den Rachen in die Trachea eingeführt wird und mit kleinen aufgeblasenen Ballons ein verrutschen verhindert und damit die Atmung eines Patienten sicherstellt.

Hand-Augen-Koordination Koordination von visueller Wahrnehmung und einer dazugehörigen physischen Bewegung, meist die eines Menschen.

Immersion Eintauchen in ein Setting, welches der Realität so gut wie möglich in der Nachempfindung entspricht.

supraglottisch oberhalb der Stimmlippen.

Trachea Luftröhre.

Erklärung zur selbständigen Bearbeitung

Hiermit versichere ich, dass ich die vorliegende Arbeit ohne fremde Hilfe selbständig verfasst und nur die angegebenen Hilfsmittel benutzt habe. Wörtlich oder dem Sinn nach aus anderen Werken entnommene Stellen sind unter Angabe der Quellen kenntlich gemacht.

Ort

Datum

Unterschrift im Original