

Master thesis

DANNY SCHOTT

—

^

○

Für Enno

Abschlussarbeit im Studiengang Interaction Design der Hochschule Magdeburg-Stendal zur Erlangung des akademischen Grades „Master of Arts“ (M.A.)

Hands-free Interaction

—

Entwicklung von multimodalen Interaktionstechniken zur handfreien Selektion und Manipulation von medizinischen Bilddaten.

Eingereicht von: Danny Schott

Erstgutachter: Prof. Steffi Hußlein

Zweitgutachter: Jun.-Prof. Dr. Christian Hansen

Betreuer: MA Benjamin Hatscher

Ort, Datum: Magdeburg, 15.07.2019



DANKSAGUNG

Diese Arbeit wäre nicht entstanden ohne die vielen lieben Menschen, die mich tagtäglich mit einem offenen Ohr umgeben und mir Unterstützung, Freude und Motivation geben.

Ich bedanke mich bei Prof. Steffi Hußlein für die gestalterischen Diskussionen, die konstruktive Kritik, sowie den anregenden kreativen Austausch. Im gleichen Maße danke ich Jun.-Prof. Dr. Christian Hansen für das Einbringen neuer Sichtweisen, die organisatorische Hilfe und die kritische Reflektion meiner Ideen. Ebenso bedanke ich mich bei den MitarbeiterInnen der Arbeitsgruppe, die für Fragen und Hilfe für mich da waren. Besonders danke ich Benjamin Hatscher und Mareike Gabele für die fachlichen Diskurse und die inhaltliche Betreuung. Ich danke Dr. Maria Luz für die wertvollen Empfehlungen zur Planung der Nutzerstudie und die Hilfe bei der Betrachtung psychologischer Aspekte. Danke an die Mitstreiter im Labor für den ein oder anderen Ratschlag bei der technischen Umsetzung. Vielen Dank an Thomas für das fotografische Dokumentieren von Teilen dieser Arbeit und Simon für die mathematische Nachhilfe. Den Probanden meiner Studie möchte ich für die Teilnahme und das Durchhaltevermögen einen Dank aussprechen. Für das medizinische Feedback und die Möglichkeit der Hospitation danke ich den Mitarbeitern des Universitätsklinikums Magdeburg.

Ohne euch wäre ich nicht der, der ich bin und wäre nicht da, wo ich bin: herzlichen Dank für die gesamte Unterstützung während meines Studiums liebe Eltern und Großeltern. Tausend Dank an meinen Bruder, mit dem ich stundenlang über persönliche, technische und theoretische Probleme und deren Lösungen philosophieren kann.

Abschließend richte ich das größte Dankschön an meine liebevolle Antje, die mir nicht nur immer zur Seite steht, sondern auch während der Entwicklung dieser Arbeit das größte Glück auf Erden schenkte.

Inhaltsverzeichnis

1	EINLEITUNG	13
1.1	ZIELSETZUNG UND VORGEHEN	16
2	GRUNDLAGEN	19
2.1	TERMINOLOGIE	20
2.1.1	MODALITÄTEN	20
2.1.2	MULTIPLE RESSOURCEN	22
2.1.3	MULTIMODALITÄT	23
2.1.4	INTERAKTION	24
2.2	BENUTZERSCHNITTSTELLEN	25
2.2.1	NATÜRLICHE BENUTZERSCHNITTSTELLEN	25
2.2.2	GESTISCHE BENUTZERSCHNITTSTELLEN	25
2.2.3	MULTIMODALE BENUTZERSCHNITTSTELLEN	26
2.3	EINGABEBEGERÄTE	28
2.4	VERWANDTE ARBEITEN	30
3	ANFORDERUNGSANALYSE	33
3.1	NUTZUNGSKONTEXT	34
3.1.1	AUFGABENGEBIET NEURORADIOLOGE	34
3.1.2	MEDIZINISCHE GRUNDLAGEN	36
3.1.3	ARBEITSUMFELD	38
3.1.4	MEDIZINISCHE EIN- UND AUSGABEBEGERÄTE	40
3.2	USER RESEARCH	44
3.2.1	INTERAKTIONSAUFGABEN	45
3.2.2	BEOBACHTUNGEN UND WORKFLOWANALYSE	46
3.2.3	EXPERTENINTERVIEW	47
3.3	KRITERIENKATALOG	52
4	KONZEPTION	55
4.1	VOREXPERIMENT	58
4.2	LIMITATION	60
4.2	SENSORIK	61
4.3	WORKFLOW	62
4.4	EINGABEMODALITÄTEN	66
4.4.1	INTERAKTIONSKONZEPTE	68
4.4.2	FEEDBACK	70
4.5	INTERFACE	71
4.6	INTERAKTIONSAUFGABEN	77
4.6.1	PRIMÄRAUFGABE	77
4.6.2	SEKUNDÄRAUFGABE	77
4.7	FINALES STUDIENKONZEPT	78
5	PROTOTYPING	83
5.1	SOFTWAREPROTOTYP	84
5.1.1	KOPFBEWEGUNGEN	86
5.1.2	KÖRPERPOSITION	88
5.1.3	SPRACHEINGABE	88
5.1.4	MIMIK	88
5.1.5	LOG-FUNKTION & INDIVIDUELLE PARAMETER	90
5.2	HARDWAREPROTOTYP	91
5.2.1	UMSETZUNG	91
5.3	SETUP	94
6	EVALUATION	97
6.1	FRAGESTELLUNGEN	98
6.2	PLANUNG	100
6.2.1	VERSUCHSPERSON	100
6.2.2	VERSUCHSUMGEBUNG	101
6.2.3	STUDIENDESIGN	102
6.2.4	ABLAUF	104
6.2.5	BEWERTUNGSKRITERIEN	106
6.2.6	NON-EXPERIMENT	108
6.3	DURCHFÜHRUNG	108
6.3.1	ZUSAMMENSETZUNG DER PROBANDEN	108
6.3.2	ERGEBNISSE	112
6.3.3	BEOBACHTUNG UND NACHBEFRAGUNG	118
6.4	DISKUSSION	120
6.4.1	OPTIMIERUNGSMÖGLICHKEITEN	123
7	ZUSAMMENFASSUNG	127
7.1	AUSBLICK	130
	LITERATURVERZEICHNIS	132
	ABBILDUNGSVERZEICHNIS	137
	ANHANG	138

KAPITEL 1

I < O

Ein- eitu- ng

Spätestens seit dem Film „Minority Report“ (2002) scheint ein neuer Begriff den Mainstream erreicht zu haben: Multimodalität. Dabei sind multimodale Interaktionen, wie im Film zu sehen, keine Zukunftsfantasien mehr, obwohl das Jahr, bei Erstellung dieses Textes, noch längst nicht 2054 schreibt. Viele Produkte werden heutzutage mit multimodaler Bedienbarkeit beworben. So kann das Smartphone, zusätzlich zur Berührung, Gesten und Sprachbefehle interpretieren. Im Forschungsbereich der Mensch-Computer-Interaktion ist sie bereits seit den frühen 1980er Jahren Gegenstand [1].

Die Idee mehrere Modalitäten zu kombinieren, scheint vielversprechend zu sein. Ziel ist eine „natürliche“ Kommunikation mit Maschinen und die intuitive Bedienung dieser. Diese Potenziale werden auch im medizinischen Kontext gesehen, wodurch auch dort berührungslose Interaktionsschnittstellen seit vielen Jahren Gegenstand der Forschung sind. So stellt eine vielversprechende Methode für die Interaktion in sterilen Umgebungen, unter anderem die Verwendung einer sprach- und gestenbasierten Steuerung dar [2]. Der Ideenreichtum an möglichen Modalitäten kann eingegrenzt werden. Denn bei der Entwicklung solcher Konzepte, ist neben den sterilen Standards in einer Operation, auch auf diverse Bewegungseinschränkungen und eine ergonomische Haltung des Nutzers zu achten.

Gerade bei radiologischen Interventionen spielt dies eine bedeutende Rolle, denn Angiographie-Systeme bieten wenig Bewegungsfreiraum. Bildgebende Verfahren wie die Angiographie, ermöglichen es, den Arzt medizinische Instrumente ohne Sichtverbindung, zur gewünschten Struktur im Inneren des Patienten zu bewegen und erfordern die Navigation auf Basis von Echtzeitbilddaten [3]. Die Augen des ausführenden Radiologen richten sich dabei meistens auf dem vor ihm befindlichen Bildschirm, welcher durch Darstellung der medizinischen Bilddaten Orientierung und Informationen bietet. Zur Manipulation dieser Daten müssen nun Arbeitsschritte unterbrochen werden, weil der Arzt zur Bedienung der Steuerungselemente häufig seine Position verändern muss [4] [5]. Auch verändert dieser seine Körperhaltung, weil durch eine kontrastarme Darstellung und ungünstige Positionierung des Monitors die Strukturen schlecht erkennbar sind.

Moderne Angiographie-Systeme bieten bereits multimodale Interaktionsschnittstellen mit vielen Funktionalitäten. Ein modulares Bedienfeld mit Joysticks, Tasten, Touchdisplays und Fußpedalen ermöglicht es dem Arzt die Anlage zu steuern, den Patienten zu positionieren, Darstellungsparameter zu verstellen und pathologische Aufnahmen zu tätigen. Die wichtigsten Bedienelemente sind für den geübten Arzt

schnell zu finden, obwohl sämtliche Oberflächen in Folie eingehüllt sind. Zudem bieten sie keine Individualisierungsmöglichkeiten, wodurch Funktionen, gerade in kritischen Situationen, nicht schnell genug zu finden sind [4]. Demnach bieten sich intuitive Systeme an, die so wenig Interaktionen wie möglich benötigen [6]. Eine Erweiterung bestehender Techniken können unter anderem die Kombination von Blick- und Fußgesten bieten [7].

Neben Recherchen in aktuellen Publikationen und kontextgebundener Literatur, ließen sich auch aufgrund eigener Beobachtungen in einer radiologischen Intervention in der Universitätsklinik Magdeburg diese Probleme bestätigen und durch zusätzliche Interviews neue Erkenntnisse gewinnen. In einer dieser Beobachtung hatte der dortige Radiologe teilweise Schwierigkeiten sich auf dem Display zu orientieren. Um Details besser zu erkennen, lehnte er sich nach vorne und fokussierte sich stark auf die Nadelspitze des Katheters.

Dies ruft nun die Frage auf, welche Interaktionsformen ihn während der Durchführung der Intervention beim Blick auf dem Bildschirm unterstützen, um eine bessere Orientierung in den medizinischen Bilddaten zu gewährleisten? Wie kann der Nutzer berührungslos durch das System navigieren? Welche Ein- und Abgabemodalitäten oder gar Kombinationen verschiedener Schnittstellen kommen für diesen spezifischen Anwendungsfall in Frage? Kann das System dabei automatisch und intuitiv auf den Nutzer reagieren, bspw. wenn dieser eine bestimmte Bewegung ausübt? Welche Technologien, Sensoren und Methoden kommen für eine präzise Erfassung des Nutzers in Frage? Befindet sich die Sensorik dafür am Nutzer oder am System? Wie lässt sich eine sichere, fehler- und ablenkungsfreie Interaktion gewährleisten? Welche Rolle spielt Feedback und welche Arten kommen in Frage? An welchen Stellen und wie lässt sich dadurch der Arbeitsablauf optimieren? Welche weiteren Umgebungseinflüsse spielen eine Rolle und lassen sich auch andere Nutzer für die Bedienung des Systems integrieren?

1.1 Zielsetzung und Vorgehen

Innerhalb eines iterativen Vorgehens sollen Chancen multimodaler Konzepte, im medizinischen Kontext einer radiologischen Intervention, erforscht werden. Dabei gilt es konkret herauszufinden, ob und welche handfreie Interaktionstechniken, den Arzt unterstützen können medizinische Bilddaten zu explorieren. Dabei sollen Modalitäten zur Selektion und Bildmanipulation bereitgestellt werden, die unter anderem eine berührungslose Bildvergrößerung ermöglichen.

Angelehnt ist das Vorgehen am menschenzentrierten Designprozess (Human-centred Design) der nach DIN EN ISO 9241-210 standardisierte Abläufe enthält. Demnach soll in vier Phasen ein Konzept entwickelt werden, womit eine hohe Usability und User Experience gewährleistet werden kann.

Auf theoretischer Grundlage zum Thema multimodaler und berührungsloser Interaktion, sowie einer Darstellung verwandter Arbeiten, werden durch Beobachtungen und Befragungen Nutzungsanforderungen am Beispiel einer radiologischen Intervention erschlossen. Darauf aufbauend werden Nutzer- und Systemanforderungen definiert, worauf der Entwurf von Gestaltungslösungen folgt. Iterativ werden diese Ideen analysiert und schrittweise prototypisch umgesetzt, getestet und optimiert. Eine abschließende Evaluation in einer Nutzerstudie soll Aufschluss darüber geben, welche Interaktionstechniken und Kombinationen verschiedener Modalitäten sich am besten in diesem Anwendungsfall eignen und wie sich diese im Arbeitsprozess integrieren lassen.

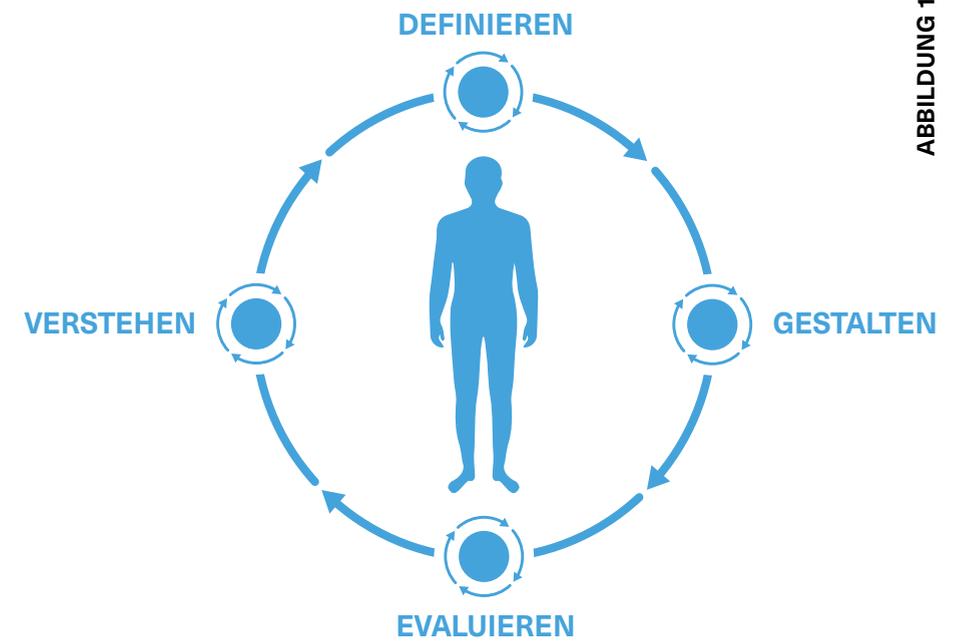


ABBILDUNG 1

Grundlagen

KAPITEL 2

IKO

In diesem Abschnitt werden Grundlagen eingeführt, die für das Verständnis dieser Arbeit von Bedeutung sind. Dabei werden Theorien zur menschlichen Wahrnehmung vorgestellt und wichtige Begriffe der Mensch-Computer-Interaktion erläutert. Ebenso werden verschiedene Benutzerschnittstellen und Eingabegeräte vorgestellt, ehe ein Abschluss mit verwandten Arbeiten und der Einordnung dieser Arbeit erfolgt.

2.1 Terminologie

2.1.1 MODALITÄTEN

Das Wort Modalität wird in vielen Kontexten, wie u.a. in Medizin, Sprachwissenschaft und Philosophie verwendet und ist nicht klar abzugrenzen. Als Modalität wird in der Psychologie häufig eine bestimmte Art von Reizen wahrzunehmen bezeichnet. In der Physiologie wird von Sinnesmodalitäten gesprochen. Diese gelten als der Empfindungskomplex aus Sehen, Hören, Riechen, Schmecken und Fühlen [8]. Menschen nehmen ihre Umwelt durch ihre Sinne wahr und handeln darauf, durch die motorische Steuerung ihrer Effektoren (Finger, Hände, Arme, Beine, Augen, (Körper-)Sprache, Mimik, ...). Hedicke [9] unterscheidet in diesem Zusammenhang zwischen Aktions- und Wahrnehmungsmodalitäten (s. Abbildung 2). Von weiteren Autoren werden dabei Aktionsmodalitäten als Eingabemodalitäten und Wahrnehmungsmodalitäten als Ausgabemodalitäten vereinfacht zusammengefasst [10].

Durch den Gebrauch der Sinne wird es ermöglicht, Erfahrungen zu machen, diese zu verstehen und zu handeln. Dabei ist die Art und Weise, wie Informationen genutzt und analysiert werden, auf die jeweilige Aktivität angepasst. Ein Wechselspiel aus sensorischen, kognitiven und physischen Fähigkeiten bestimmt den Einsatzzweck und lenkt die Aufmerksamkeit auf eine Tätigkeit. Dies geschieht bewusst und auch unbewusst, denn irrelevante Informationen werden durch einen natürlichen Fokus herausgefiltert. Die Fähigkeit, wichtige Informationen über Aufgaben und Aktivitäten zu selektieren und diese dann zu priorisieren, wird als modaler Fokus bezeichnet. Es kommt zu einem Rückkopplungs-Effekt, denn die Aufmerksamkeit schärft den Fokus und dieser wiederum die Sinne [11]. Während Geruch und Geschmack, als olfaktorische und geschmackliche Sinne, essentiell mit der

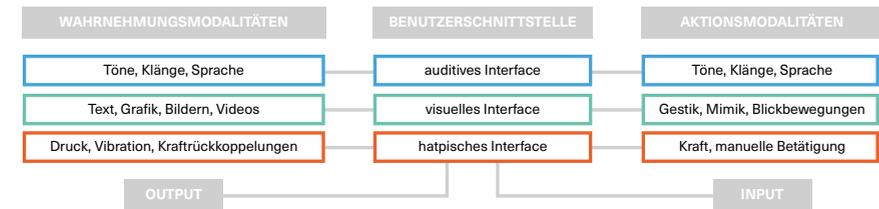


ABBILDUNG 2: Modalitäten einer Benutzerschnittstelle nach Hedicke.

menschlichen Wahrnehmung verknüpft sind, spielen sie hingegen in der Schaffung von Benutzerschnittstellen eine geringfügige Rolle. Park und Alderman [12] klassifizieren die dominanten Modalitäten für Nutzerinterfaces, basierend auf den jeweiligen Sinnen:

VISUELL – Basierend auf Sehvermögen

AUDITIV – Basierend auf Klangwahrnehmung

HAPTISCH – Basierend auf Tastsinn und Bewegung

PROPRIOKTIV – Basierend auf dem Sinn der eigenen Bewegung und Orientierung.

Modalitäten sind Muster der Wahrnehmung, Kognition und Handlung [...], die es ermöglichen Empfindungen zu konzentrieren, Wichtiges von Unwichtigem zu differenzieren und die Sinne so anzupassen, dass wir verstehen, was passiert, so Alderman [12]. Beuter [13] definiert Modalitäten wie folgt: „Eine Modalität bezeichnet ein kommunikatives System, das durch die Art und Weise, wie Informationen kodiert und interpretiert werden, gekennzeichnet ist“. Im technischen Sinn orientiert sich dabei der Begriff Modalität als konkreten Zusammenschluss eines Interaktionsgerätes mit einer Interaktionssprache [10].

2.1.2 MULTIPLE RESSOURCEN

Eine einzelne Modalität, wie bspw. die visuelle Aufmerksamkeit, kann nicht geteilt werden. Im Straßenverkehr kommt es so häufig zu einer Interferenz: Dem parallelen Bearbeiten von zwei Aufgaben, indem gleichzeitig der Verkehr beobachtet und ein Infotainment System bedient wird. Um Rücksicht in der Gestaltung von Mensch-Computer-Interaktionen auf diese Effekte zu nehmen und Probleme frühzeitig zu identifizieren, kann das Modell multipler Ressourcen von Wickens & Holland [14] herangezogen werden. Dabei ist die Grundidee, dass Nutzer eines Systems prinzipiell über verschiedene Ressourcen zur simultanen Durchführung kognitiver Prozesse verfügen (s. Abbildung 3).

Die Ressourcen werden dabei in drei Dimensionen eingeteilt: 1. Verarbeitungsphasen (Stages: perception, cognition, responding), 2. Wahrnehmungsmodalitäten (Modalities: visual vs. auditory), 3. Verarbeitungstyp (Codes: spatial vs. verbal). Demnach stören sich Ressourcen für die Wahrnehmung (perception) und Ressourcen für das Antwortverhalten (response) nicht, womit theoretisch beide Prozesse parallel ohne Interferenzeffekte ausgeführt werden. So besteht in Multitasking-Situationen die Möglichkeit, während der Antwort auf Aufgabe A bereits der Wahrnehmungsprozess für Aufgabe B laufen kann. Diese Möglichkeit spielt eine zentrale Rolle in einer Multitasking-Situation, da während der Antwort für eine Aufgabe A bereits der Wahrnehmungsprozess für eine andere Aufgabe B ablaufen kann [15]. Jedoch sind diese Modelle theoretischer Natur und sollten immer im Kontext der jeweiligen Situation betrachtet werden. Ebenso spielen eigene Erfahrungen und Expertisen eine triviale Rolle. So haben Nutzer, die besonders vertraut und geübt in einer Aktivität sind, prinzipiell einen Vorteil gegenüber denen, die diese zum ersten Mal ausüben.

2.1.3 MULTIMODALITÄT

Die menschliche Interaktion mit der Welt ist von Natur aus multimodal [16]. Multimodalität zeichnet sich durch eine Kombination von mindestens zwei Modalitäten aus und beschreibt komplexe Verhaltensweisen zwischen Menschen und ihrer Umwelt. Der Begriff der Multimodalität findet aktuell große Verbreitung in der Gestaltung von Mensch-Maschine-Interaktionen. Ziel ist stets eine optimierte Anpassung der menschlichen Systemsteuerung. Es mangelt an einer einheitlichen Definition des Begriffs der Multimodalität, sowie einem integrativen Modell, zur Gestaltung von multimodalen Systemen, die einen Interaktions- und Informationsaustausch zwischen Menschen und Computern bieten [17]. Im Alltag treffen wir oft auf die technologische Integration verschiedener Modalitäten: Bei der Bedienung von Smartphones, TV, Computer- und IoT-Geräten wird parallel zur visuellen oder haptischen Ebene, die Sprachsteuerung ermöglicht. Leistungsstarke und immer erschwinglichere Geräte und Sensoren, treiben die rasante Entwicklung multimodaler Forschung im Bereich der Sprache, Berührung, Sehen und Gestik voran [18].

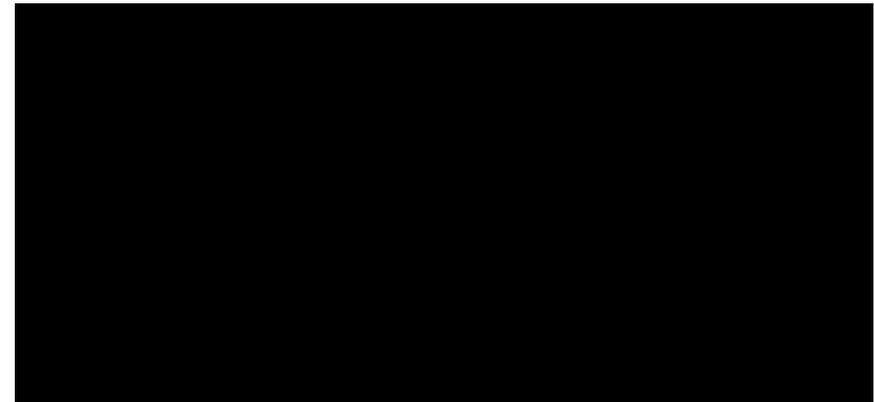


ABBILDUNG 3: Modell multipler Ressourcen nach Wickens [14].

2.1.4 INTERAKTION

Unter den Begriffen Mensch-Maschine-Interaktion oder auch Human-Computer-Interaction versteht sich der wechselseitige Austausch von Informationen zwischen einem menschlichen Nutzer und einem maschinellen System [19]. Es kann unterschieden werden zwischen Eingaben des Nutzers und den Auswirkungen auf ein System und die Ausgabe des Systems und dessen Auswirkung auf den Nutzer. Die Interaktion zwischen Menschen und Computern ist in Teilaspekten zu sehen und nach Boles et al. [20] in folgende Begrifflichkeiten zusammengefasst:

INTERAKTIONSFORM

Unter Interaktionsform werden Paradigmen in vier Klassen der Informationseingabe zusammengefasst: Kommandosprachen, Masken, User-Interface-Komponenten (UI-Elementen) und Direkte Manipulation [20].

INTERAKTIONSAUFGABE

Beschreibt den Typ der Information, mit dem ein Nutzer eine Eingabe tätigen kann. Nach Foley [21] werden Basisinteraktionsaufgaben unterschieden in Texteingabe, Positionierung, Bestätigung, Auswahl und Quantifizierung.

INTERAKTIONSTECHNIKEN

Unter dem Begriff Interaktionstechnik werden Möglichkeiten verstanden, wie Eingabegeräte eingesetzt werden können, um eine Interaktionsaufgabe zu erfüllen [21].

INTERAKTIONSZIEL

Die Auswirkung der Eingabe und die damit verbundene Erwartungshaltung der Nutzer ein Ziel zu erreichen, wird als Interaktionsziel beschrieben. Dabei wird die tatsächliche Ausgabe des Systems als Interaktionsauswirkung (Reaktion) bezeichnet (vgl. Boles [20]).

2.2 Benutzerschnittstellen

2.2.1 NATÜRLICHE BENUTZERSCHNITTSTELLEN

Eine natürliche Benutzerschnittstelle (engl. Natural User Interface, NUI) zeichnet sich nach Lee [22] durch eine intuitive Bedienung aus, in der die eigentliche Aufgabe des Computers in den Vordergrund tritt, während die Schnittstelle nicht sichtbar ist: „NUI generally refers to an interface that is highly intuitive and effectively becomes invisible to the user when performing a task“. Ergänzend dazu konkretisiert Blake [23], dass ein NUI Gestaltungsprinzip hinsichtlich einer einfachen Bedienung auf Basis alltäglich gebrauchter Fertigkeiten des Nutzers folgt: „A natural user interface is a user interface designed to reuse existing skills for interacting directly with content“. Denn die Fähigkeiten der zwischenmenschlichen Kommunikation und der Interaktion mit seiner Umgebung ist dem Menschen von Natur aus gegeben. NUIs erweitern den bisher künstlichen Umgang mit Computern (Maus und Tastatur) und bieten eine natürliche Benutzerschnittstelle. Bollhoefer [24] beschreibt sie als Schnittstelle, welche unmittelbar mit den Sinnen des Nutzers verknüpft ist. Dabei wird unterschieden in aktive Interfaces, indem direkt mit dem System interagiert wird, wie bspw. Multitouch Displays und passive Interfaces, wobei der Nutzer eine passive Rolle einnimmt und ohne motorisches Zutun eine Eingabe erfolgt, wie bei Brain-Computer-Interfaces (BCI), in dem die Synapsenaktivitäten des Gehirns zur Steuerung eines Systems gemessen werden.

2.2.2 GESTISCHE BENUTZERSCHNITTSTELLEN

Gestenbasierte Benutzerstellen spielen bei der berührungslosen Interaktion eine wichtige Rolle. Unterordnen lassen sich diese im Kontext von natürlichen Benutzerschnittstellen. Gesten sind Bewegungsmuster und stellen eine intuitive Form der non-verbalen Kommunikation dar und umfassen symbolische Körpersprache und expressive Bewegungen (Sprache, Blick, Handgesten, Körperbewegung, Mimik, Berührung). Saffer [25] unterscheidet zwei Typen gestischer Benutzerschnittstellen: Touch User Interfaces und Free-form Gestural Interfaces. Touch Interfaces erfassen Gesten aufgrund von Berührungen, wohingegen bei Free-form Gesten die Erkennung berührungsloser Interaktion im Raum möglich ist und kein direktes Eingabegerät erkennbar sein muss.

2.2.3 MULTIMODALE BENUTZERSCHNITTSTELLEN

Wie in vorherigen Abschnitten kenntlich gemacht, ist festzuhalten, dass die Wahrnehmung des Menschen ein komplexes Unterfangen aus unzähligen Wechselwirkungen verschiedener Modalitäten ist. Die multimodale Mensch-Computer-Interaktion ist seit jeher bestrebt, diese natürlichen Fähigkeiten auf computergestützte Systeme zu übertragen, um überzeugende, leistungsfähige und interaktive Erlebnisse zu schaffen. Bei der Konzeption multimodaler Benutzungsschnittstellen (engl. Multi-Modal-Userinterfaces, MMI) sollten nicht die technologischen Fragen im Vordergrund stehen, sondern die wahrnehmungspsychologischen und kognitiven Grundlagen des Nutzers [26]. Die am häufigsten vorzufindende Definition, multimodaler Systeme als parallel angebotene Wahrnehmungskanäle zur Abgrenzung der Begriffe multimodal und multimedial zu verstehen, bietet Oviatt [26]: "Multimodal system process combined natural input modes – such as speech, pen, touch, manual gestures, gaze, and head and body movements – in a coordinated manner with multimedia system output". Dabei unterscheidet Oviatt drei verschiedene Eingabemodi [27]:

ACTIVE INPUT MODES

Werden vom Benutzer absichtlich als expliziter Befehl eingesetzt (z.B. Sprache).

PASSIVE INPUT MODES

Durch natürlich vorkommendes Benutzerverhalten ausgelöste Eingaben, die passiv, d.h. ohne direkten Befehl reagieren (z.B. Gesichtsausdrücke, manuelle Gesten).

BLENDED MULTIMODAL INTERFACE

Die Systemerkennung beinhaltet mindestens eine passive und eine aktive Eingabe (z.B. Sprach- und Lippenbewegungssysteme).

Multimodale Systeme verteilen Informationen auf mehrere Sinne, um die Beanspruchung des Nutzers zu reduzieren. Die Ziele multimodaler Interaktionen lassen sich nach Rötting und Seifert in vier Aspekte aufteilen [28]:

NATÜRLICHKEIT

Die Kommunikation mit dem Computer basiert auf realen Nutzererfahrungen ähnlicher Situationen, wodurch für die Interaktion kein Fachwissen vorhanden sein muss.

INTERAKTIONSSICHERHEIT

Ein System sollte zuverlässig und robust sein, indem Mehrdeutigkeiten von Nutzereingaben vermieden werden.

FLEXIBILITÄT

Flexible örtliche, zeitliche, situative und auch leistungsabhängige Parameter der Interaktion sollten durch ein Angebot verschiedener Ein- und Ausgabemodalitäten gewährleistet werden.

LEISTUNGSFÄHIGKEIT

Durch Ausschöpfung des Potenzials der Informationskanäle, nimmt die Qualität von übermittelten Informationen zu, wodurch sich die Leistungsfähigkeit eines Systems erhöht.

Die Vorteile multimodaler Systeme scheinen offensichtlich zu sein. Jedoch merkt Hedicke [9] an, dass die Ausnutzung der gesamten Bandbreite der Wahrnehmungs- und Aktionsmodalitäten mit Bedacht gewählt werden sollte. Es besteht die Gefahr der Reizüberflutung und der Nutzer kann durch die Informationsdarbietung auf verschiedenen Kanälen überfordert sein.

2.3 Eingabegeräte

Neben haptischen Interfaces, welche sich auf direkte Berührungseingaben beschränken, ermöglichen visuelle und kamerabasierte Interfaces die räumliche Erfassung von Positionen und Bewegungen von Objekten im Raum und ermöglichen somit eine berührungslose Interaktion. Mit deren Hilfe lassen sich Körper, Gesichter und Gesten mehrerer Nutzer detektieren. Zur Erfassung menschlicher Aktivitäten lassen sich zwei Gruppen klassifizieren [29]:

BILDBASIERTES TRACKING

Erstreckt sich über einfache und stereoskopische Kamerasysteme zur Bildaufnahme, Tiefenkameras zur Tiefenbestimmung und Markersystemen zur Lokalisierung relevanter Punkte (vgl. Preim [30]). Vielversprechende Technologien dieser Art sind u.a. der Leap Motion Controller [31], die stereoskopische Kamera Zedd der Firma Stereolabs [32] und kompakte Time-of-Flight Sensoren zur 3D-Bildverarbeitung zum Beispiel von Terabee [33]. Die kommerziell erfolgreichste und maßgebende Hardware stellt Microsoft in Form der Kinect [34] zur Verfügung, dessen Produktion 2014 eingestellt wurde. Weitergeführt wird die Technologie der Kinect in der Hololens (1 und 2) [35] als IoT (engl. Internet of Things, Internet der Dinge) Gerät unter dem Namen Azure Kinect [36]. Die Analyse von Bild- und Sprachinformationen wird hierbei mithilfe künstlicher Intelligenz unterstützt. Die Komponenten werden immer erschwinglicher und kompakter, wodurch u.a. Gesichtsauffindung und Gestenerkennung in mobilen Geräten und auch in Spielekonsolen zum Einsatz kommen. Beispiele hierfür sind Apples „TrueDepth Face ID“, der „G8 ThinQ“-Sensor von LG [37] und der Controller der Nintendo Switch [38]. Insgesamt lässt sich hier auch noch unterscheiden zwischen tragbaren Geräten für immersive VR/AR/MR-Anwendungen (Hololens, Magic Leap [39] und Meta [40]) und Geräten, welche Nutzer im Raum erfassen, wie anhand vorher genannter Beispiele gezeigt wurde.

NICHT-BILDBASIERTES TRACKING

Sind unter anderem Band-basierte Sensoren, die am Arm getragen und über drahtlose Technologien und Elektromyogrammsensoren Bewegungen erfassen, sowie Datenhandschuhe, die mit Beschleunigung- und Gyroskopsensoren arbeiten und auch nicht-tragbare Sensoren, die über Hochfrequenzbänder ohne Kontakt Bewegungen verfolgen können. Aktuelle Technologien sind unter anderem der Datenhandschuh von Manus VR [41], das Myo Armband [42] und das Project Soli von Google [43].



ABBILDUNG 4:

Das Projekt Kinect Azure von Microsoft analysiert Tiefeninformationen von Bildern mit Unterstützung maschinellen Lernens.

2.4 Verwandte Arbeiten

Historisch gilt als eine der ersten Referenzen multimodaler Systeme das Werk des MIT-Wissenschaftlers Richard Bolt [1]. Im MIT-Media Lab entwickelte Bolt 1980 ein Demonstrationssystem unter der Publikation „Put that there“, welche die Interaktion von Sprache mit Zeigegesten ermöglicht (s. Abbildung 1). Auch weitere Autoren unterbreiteten danach Vorschläge, wie verschiedene Modalitäten sinnvoll aufeinander kombiniert werden können, um eine gebrauchstaugliche Benutzungsschnittstelle zu schaffen [17] [44].

Allgemeine Beispiele multimodaler Interaktionen wurden bereits erwähnt und sind ersichtlich aus dem Potenzial, was diese Systeme mit sich bringen. Es stellt sich heraus, dass aktuell viele Arbeiten im Bereich multimodaler Interaktion im Gebiet virtueller und erweiterter Realitäten verfasst werden [45] [46]. Xu et al. zeigen neuartige multimodale Konzepte zur berührungslosen Interaktion mit Head-Mounted-Displays [47]. Dem Nutzer ist es durch richtungsweisende Bewegungen möglich mit seiner Umwelt zu interagieren.

Da in dieser Arbeit jedoch die Interaktion an einem Monitor untersucht wird, werden die genannten Gebiete nicht mit einbegriffen, was keineswegs die Relevanz dieser Thematik untergraben soll. Zur Exploration geographischer Daten an einem Multitouch-Wanddisplay realisierten Schöning et al. verschiedene Ganzkörperinteraktionen [48]. Mittels eines Wii Balance Boards war es dem Nutzer möglich durch Vorlehnen eine Bildvergrößerung hervorzurufen.

Wie Eingangs festgehalten, bieten sich im Szenario einer OP laut O'Hara et al. Sprach- und Gestenbasierte Steuerung an [2]. Diese Interaktionen sind natürlich, denn sie finden alltäglich Verwendung in der menschlichen Kommunikation. Relevante Untersuchungen mit Schwerpunkt einer radiologischen Intervention liefert die Arbeit von Hatscher et al. [49], in der gestische und sprachbasierte Eingabemodalitäten hinsichtlich der Interaktion in sterile Umgebungen analysiert wurden. In der durchgeführten Studie wurde untersucht, welchen Einfluss die verschiedenen Modalitäten (Handgesten, Fußgesten und Sprache) auf die Tätigkeiten des Arztes haben, mit dem Ergebnis, dass die Eingabe über Fußpedale die beste Lösung bietet, jedoch Sprach- und Handeingaben nur geringfügig schlechter beurteilt wurden. In einer weiteren Arbeit wurden Fußinteraktionen auf ihre Bewegungsgrade und die Unterstützung der handfreien Steuerung eines Angiographie-Systems ge-

prüft [50]. Weiterhin untersuchte Hatscher das Zusammenwirken von Augen- und Fußinteraktionen [7]. Dabei wurden mit einer mobilen Eye-Tracking-Brille die Bewegungen der Augen erfasst und eine Selektion verschiedener Bereiche auf einem Display ermöglicht. Dabei stören sich Blick- und Fußinteraktionen bei gleichzeitiger Nutzung leicht. Eine Sensormatte als Eingabegerät von Fußinteraktionen fand bei Wagner et al. [51] Verwendung. Es wurden Gestensets zur Navigation in Bilddaten integriert, die unter anderem eine Eingabe über Gewichtsverlagerung und durch Schrittgesten betätigte Buttons ermöglichte.

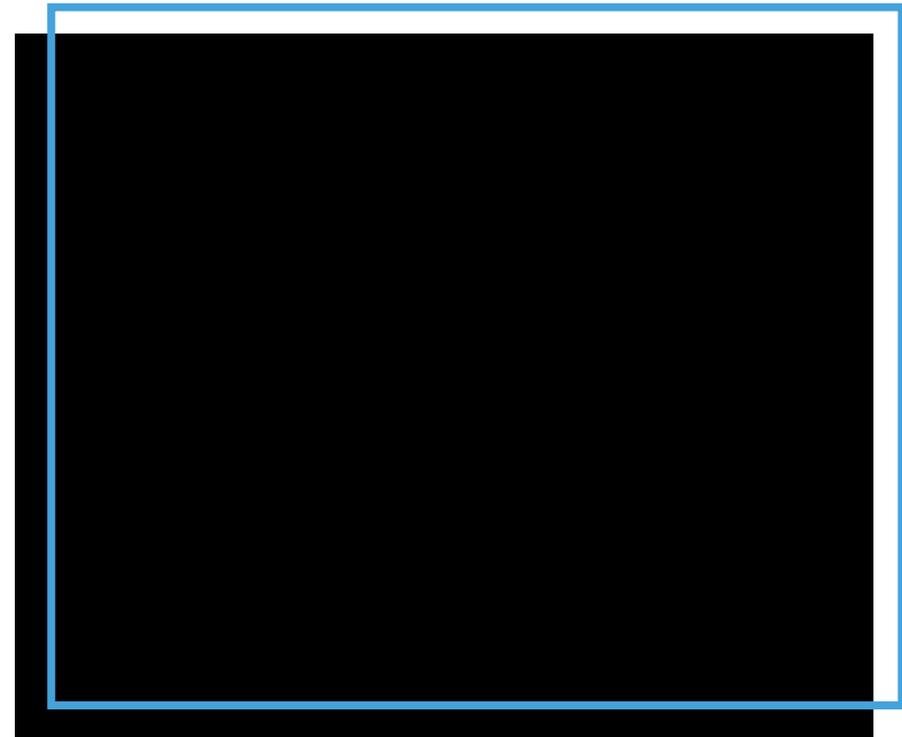


ABBILDUNG 5: *Doppelbelichtetes Foto aus dem Media Raum des MIT in den 70er Jahren zeigt die Interaktionen aus Zeigegeste und Spracheingabe an einem großen Bildschirm.*

IKO

Anfor derun gsana lyse

KAPITEL 3

In diesem Kapitel sollen Anforderungen erschlossen werden, welche der Entwicklung von späteren Konzepten der berührungslosen Interaktion im OP dienen. Grundlage bildet die strukturierte Literaturrecherche aus dem vorherigen Kapitel, sowie eigene Feldbeobachtungen und Interviews mit Experten und Personal des Instituts für Neuroradiologie des Universitätsklinikums Magdeburg. Im ersten Schritt wurde das allgemeine Aufgabengebiet und Arbeitsumfeld von Radiologen beschrieben und medizinische Grundlagen erläutert. Darauf folgt die Auswertung der Beobachtungen und Gespräche im klinischen Kontext, woraufhin typische Arbeitsabläufe einer neuroradiologischen Intervention und dessen Nutzern analysiert und in Form einer User Journey Map zusammengefasst wurde. Abschließend wurde anhand der Anforderungsanalyse ein Kriterienkatalog erstellt, der zur späteren Konzeption und Evaluierung des Systems herangezogen wird.

3.1 Nutzungskontext

Grundlage zur Entwicklung nutzerzentrierter Gestaltung bildet die Identifikation von Nutzerbedürfnissen durch die Analyse der Umgebung, ergonomischen Aspekten sowie spezifischen Handlungsabläufen. Hierbei wird die allgemeine Situation mit eigenen Beobachtungen ergänzt.

3.1.1 AUFGABENGEBIET NEURORADIOLOGE

Fachärzte für Radiologie arbeiten im Röntgen, der Nuklearmedizin sowie in der Strahlentherapie. Das Aufgabengebiet des Neuroradiologen umfasst dabei diagnostische und therapeutische Tätigkeiten zur Identifikation von Erkrankungen und Behandlungen des zentralen Nervensystems, d.h. von Gehirn und Rückenmark. Dabei schließen Neuroradiologen die Lücke zwischen verschiedenen Nachbardisziplinen wie der Neurologie, Neurochirurgie, Neuropädiatrie, Psychiatrie, Orthopädie und Augenheilkunde [52]. „Die Radiologie ist ein diagnostisches Fach. In der Praxis bedeutet das: Sie entscheiden mit. Ist der Arm gebrochen oder nicht? Das klärt der Radiologe – und damit hat er auch entscheidenden Einfluss auf die weitere Behandlung“ [53].

Mithilfe bildgebender Hochleistungsdiagnostik, wie Magnetresonanztomografie (MRT), Computertomographie (CT), Angiographie, Sonographie (Ultraschall), Myelographie (Untersuchung des Rückenmarkkanals) sowie Röntgenuntersuchungen, kann die Radiologie eine vielfältige Patientenversorgung gewährleisten [54]. Therapeutische Verfahren, auf welche der Fokus dieser Arbeit liegt, werden unter dem Begriff der Interventionellen Neuroradiologie zusammengefasst und lassen sich unterteilen in:

- gefäßeröffnende Maßnahmen (rekanalisierende Maßnahmen)
bei einem Gefäßverschluss wird das Gefäß mithilfe von Kathetern mechanisch wiederoeffnet
- gefäßverschießende Maßnahmen (embolisierende Maßnahmen)
bei Blutungen im Hirn wird die Gefäßwunde mit Klebstoff verschlossen
- schmerztherapeutische Maßnahmen bei chronischen Rückenschmerzen

Abhängig von der Art des Eingriffs steht eine ärztliche Assistenz oder speziell ausgebildetes Fachpersonal, Medizinisch-technischen Radiologieassistenten (MTRA), zur Seite. Diese bereiten u.a. die OP vor, in dem sie die Instrumente und Utensilien bereitstellen, das System einrichten und zusammen mit dem Radiologen bedienen [55]. In der Regel steht mindestens ein steriler MTRA direkt dem Operateur zur Seite, während eine oder mehrere „unsterile“ Assistenten im sogenannten Kontrollraum die Operation begleiten.

In dieser Arbeit werden Interaktionstechniken am Beispiel einer neuroradiologischen Intervention an einem Angiographiesystem untersucht wird gilt es aus diesem Grund wichtige Begriffe zunächst zu erläutern:

In der Medizin wird als Angiografie die radiologische Darstellung von Gefäßen mittels Bildgebungsverfahren bezeichnet, welche die Untersuchung von Blutgefäßen ohne operativen Eingriff ermöglicht. So können Gewebeschädigungen, Durchblutungsstörungen und Schlaganfälle, die durch Blutungen in Hirngefäßen verursacht werden, frühzeitig diagnostiziert werden. Zur Darstellung dieser Strukturen wird häufig ein Kontrastmittel in das Blut injiziert, wodurch in den mittels Röntgen oder Magnetresonanztomografie (MRT) erstellten 2D- und 3D-Aufnahmen der Gefäßinnenraum sichtbar wird [56].

Eine spezielle Form der Angiographie stellt die digitale Subtraktionsangiographie dar. Hierbei führt der Radiologe einen Katheter oder eine Injektionsnadel über den Lendenbereich des Patienten zur Zielstruktur – zum Beispiel im Gehirn. Das Instrument (Katheter) wird vor Abgang (Verzweigung) des Blutgefäßes positioniert und es werden zeitlich aufeinander abfolgende Röntgenaufnahmen erstellt, während das Kontrastmittel eingebracht wird. Das Ergebnis ist ein durch Subtraktion erschaffenes digitales Maskenbild der Gefäßstruktur aus Sequenzbildern mit und ohne Kontrastmittel [57].

Ein Beispiel von nur durch angiografische Maßnahmen sichtbarer Gefäßmissbildungen ist eine arteriovenöse Malformation (AVM). Dieses blutschwammigen Gewebe im Gehirn sind häufig seit Geburt vorhanden und werden tendenziell mit der Zeit größer, weshalb es plötzlich zu Gehirnblutungen kommen kann. Zur Behandlung muss ein interdisziplinäres Team entscheiden, ob die AVM verschlossen werden soll (Embolisation), chirurgisch entfernt (Resektion) oder bestrahlt werden sollte [58]. Abbildung 7 zeigt die Therapie einer solchen AVM an einem Angiographiesystem, welches beispielhaft in Abbildung 6 zu sehen ist.

Diese Maßnahmen der chirurgischen Intervention (Operation) gelten allgemein als schonend dem Patienten gegenüber, weil die Eingriffszeit relativ kurz ist und lediglich eine Punktionsstelle im Lendenbereich als Operationswunde bleibt [59].



ABBILDUNG 6: Ein modernes Angiographiesystem auf dessen Monitor verschiedene Angiographien zu sehen sind.

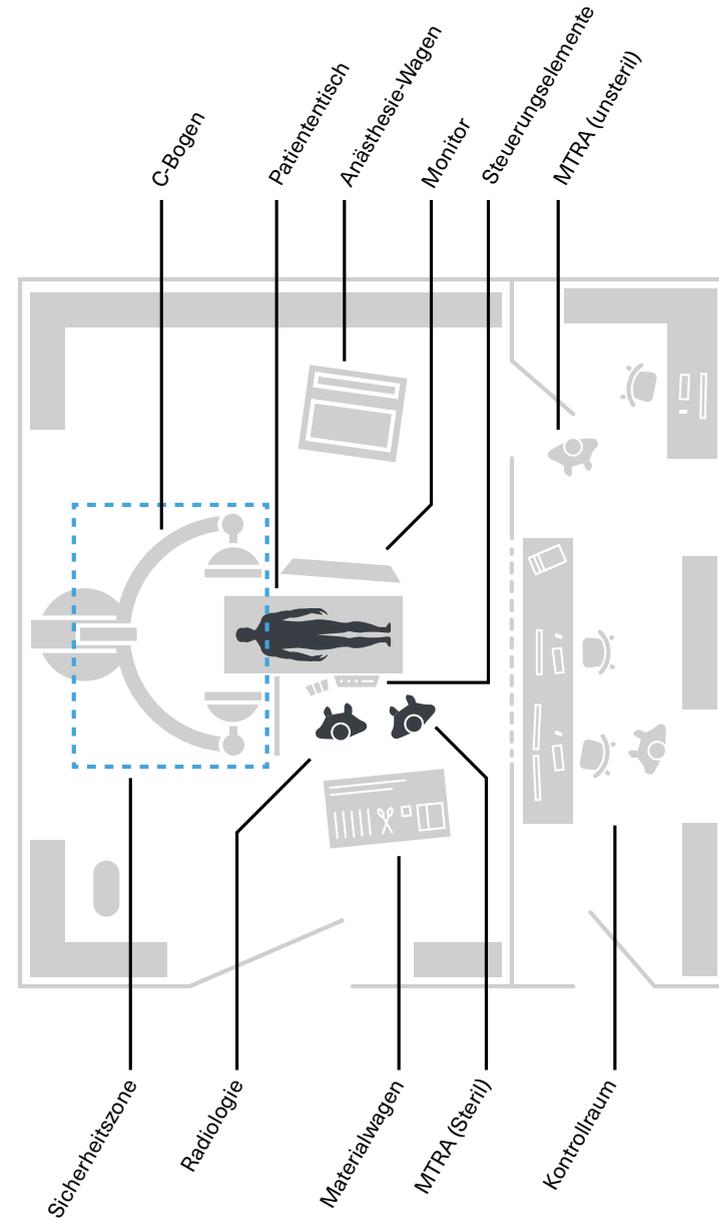


ABBILDUNG 7: Im Verfahren der digitalen Subtraktionsangiographie (DSA) entstandene Darstellung einer Arteriovenösen Malformation (AVM) vor dem Therapiebeginn (links), nach erfolgreicher Teilembolisation (Mitte) und nach vollständiger mikrochirurgischer Resektion (rechts).

Der Operationsaal ist ein Arbeitsplatz mit beschränkten Platzverhältnissen. Speziell das typische Arbeitsumfeld eines Neuroradiologen besitzt einen eingeschränkten Interaktionsraum. In Abbildung 9 wird ein typischer Operationssaal schematisch dargestellt. Wie zu sehen ist, existieren zwei getrennte Räume: Der eigentliche Operationssaal und ein Kontrollraum. Im Operationssaal agieren die Radiologen. Dabei befinden sie sich vorwiegend vor einer Liege, auf der der Patient liegt. Auf Kopfhöhe des Patienten rotiert durch Steuerung der behandelnden Nutzer der Detektor am sogenannten C-Bogen, welcher für die Röntgenaufnahme der Bilder eingesetzt wird. Während der Aufnahme sind Patienten und Ärzte durch Strahlenschutzkleidung sowie einem Strahlenschutzschild geschützt. Aufgrund der Dimensionen dieser Anlagen, sowie dem großen Bewegungsumfang ist ein

Sicherheitsbereich definiert. Der Kontrollraum, oder auch Cockpit oder Workstation genannt, ist durch ein Sichtfenster direkt mit dem OP-Raum verbunden. Zudem sind gewisse Ansichten des Monitors der Angiographieanlage mit denen im Kontrollraum synchronisiert und gewährleistet damit eine begleitende Überwachung der OP. Im Kontrollraum finden zudem die Nachbearbeitung und Archivierung der aufgenommenen Daten statt. So werden dort Strukturen vermessen und Flussinformationen post-operativ von Ärzten begutachtet. Weitere erwähnenswerte Gegenstände der OP sind der hinter den Akteuren befindliche Materialtisch mit sämtlichen sterilen Utensilien (Katheter, Schläuche, Spritzen, Medikamente), sowie der Anästhesiewagen zur Überwachung der Vitalparameter des Patienten.

ABBILDUNG 8: Beispielhafter Aufbau eines interventionellen Operationssaals am neuro-radiologischen Institut des Universitätsklinikums Magdeburg.



3.1.4 MEDIZINISCHE EIN- UND AUSGABEGERÄTE

Zum Verstellen, Steuern und Positionieren der gesamten Angiographieanlage befindet sich am Operationstisch ein Schaltpult. Dieser und alle anderen Geräte mit denen Menschen in Berührung kommen, sind zur Wahrung der Sterilität durch Folien versiegelt. Das Schaltpult ist modular aufgebaut und kann je nach Nutzerpräferenzen und Operationsart konfiguriert werden. Darauf sind verschiedene Bedienelemente, wie Tasten, Touchscreens und Joysticks zu finden, die zusätzlich noch von Fußpedalen ergänzt werden, wie in Abbildung 10 dargestellt. In der oberen Abbildung sind die verschiedenen Module zu sehen. Das Verstellen von Neigung, Drehung und Höhe des Patiententisches erfolgt über den linken Joystick. Daneben befinden sich Module zum Verstellen des C-Bogens, um Arm und Aufnahme Fenster zu positionieren. Es folgt ein Touchscreen-Monitor, der erweiterte Einstellungen und Funktionen bietet und ganz rechts ein Joystick, der ähnlich einer Computermaus fungiert und zur Bildnavigation dient. Die Bildaufnahme geschieht über die Betätigung der beweglichen Fußpedale (unten).

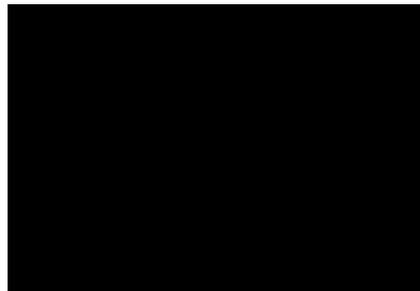
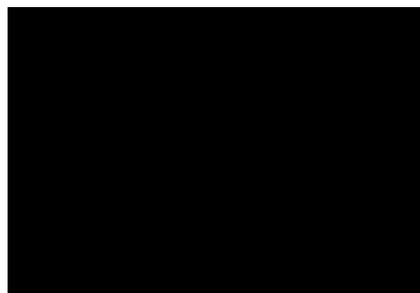


ABBILDUNG 10: Eingabegeräte des Siemens Artis Zee Angiographiesystems. Oben: Modulares Schaltpult, Unten: Kabelloses Fußpedal.



Bildschirme sind wichtige Komponenten des Systems. Dabei ist der Hauptbildschirm, welcher sich direkt im Sichtfeld des Radiologen befindet mit dem Kontrollraum synchronisiert. Es bietet verschiedene Informationen, wie bspw. Echtzeitaufnahmen, Referenzbilder, Patientendaten und genormte Systemparameter, die u.a. Röntgendosis, Bestrahlungsdauer, und Einstellungen ikonografisch darstellen (s. Abbildung 6 und Abbildung 11). Zudem lassen sich verschiedene grafische Anordnungen der verschiedenen Segmente über das Touchinterface einstellen (s. Abbildung 12).



ABBILDUNG 11: In drei Segmente aufgeteiltes Interface eines Angiographiesystems. Im unteren Bereich ist eine grafische Visualisierung des Joysticks sichtbar, welche sich simultan zur realen Eingabe am Schaltpult verhält.

ABBILDUNG 12: Touchinterface zur Auswahl verschiedener Layouts. ▶



Zur Manipulation und Selektion verschiedener Daten wird wie beschrieben ein Joystick verwendet. Es ist dem Nutzer damit möglich Bilder digital zu vergrößern, was jedoch erst durch die Navigation in Untermenüs ermöglicht werden kann (s. Abbildung 13). Dabei wählt der Nutzer das Icon „View 2x“ und dieser vergrößert direkt das angewählte Bild um 150 %. Es sind somit keine Stufen vorhanden und das einzige Feedback dabei, ist die Vergrößerung selbst (s. Abbildung 14). Eine optische Vergrößerung erfolgt durch das Verstellen des Detektors und erneuter Bildaufnahme. Es entstehen Schichtbilder, die als Volumen dargestellt werden können und durch dessen Schichtdicke kann in verschiedenen Ebenen (Bildsequenzen) gescrollt werden.

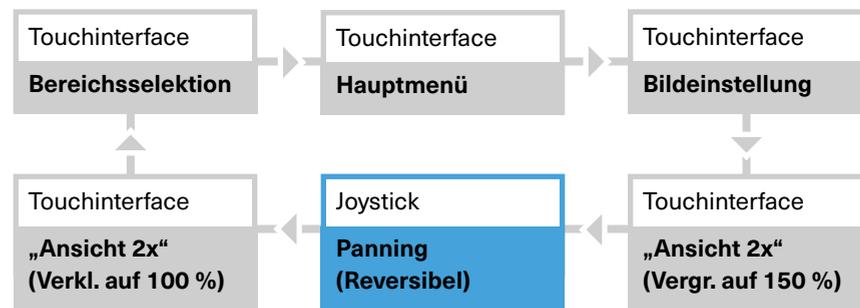


ABBILDUNG 13: Beispielhafte Interaktion zum Durchführen von Bildvergrößerungen in einem Siemens Angiographie System (Artis Q)

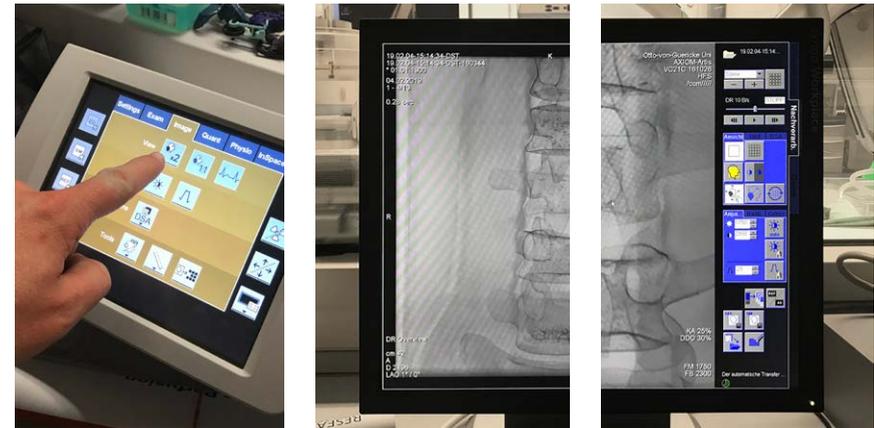


ABBILDUNG 14: Synchronisierter Monitor mit dazugehörigem Schaltpult einer Laboranlage. Eine Bildvergrößerung wird durch das Touchpanel erreicht (links). Das Bild vergrößert sich simultan auf den Monitoren in Kontrollraum (Mitte und rechts) und OP-Saal.

3.2 User Research

Um einen vollständigen Einblick in Handlungen und Abläufe einer radiologischen Intervention zu bekommen, ist es hilfreich Nutzerbedürfnisse zu definieren. Baeck et al. [60] zeigen, dass klassische User Research Methoden im radiologischen Kontext effektiv eingesetzt werden können, wobei die besondere Herausforderung darin liegt, Experten, Designer und Entwickler zusammenzubringen, um innovative und nutzerzentrierte Systeme zu kreieren.

Da aus Kapazitätsgründen in dieser Arbeit keine Expertenworkshops stattgefunden haben, werden auf Grundlage bestehender Recherchen, Analysen und eigenen Beobachtungen sowie Befragungen, Anforderungen einer radiologischen Intervention identifiziert.

Dadurch, dass jede Operation einzigartig ist, jeder Chirurg anders agiert und jeder Patient individuelle Krankheitsbilder aufweist, stellt sich die Definition allgemeingültiger Arbeitsabläufe als schwierig dar. In diesem Abschnitt wird sich vorwiegend auf Probleme des Arbeitsablaufs einer neurologischen Intervention konzentriert. Grundlage bietet die Workflowanalyse von Hübler et al. [4] und wird ergänzt durch eigene Beobachtungen während einer AVM-Embolisation, die grundlegend unter Abschnitt 3.1.2 beschrieben wurde. Der Autor hospitierte dafür am Institut für Neuroradiologie des Universitätsklinikums Magdeburg und sprach mit Personal und Experten der Einrichtung (Radiologen, Physikern und Medizintechnikern). Hierbei sollten praktische Eindrücke aus dem klinischen Alltag von Radiologen gewonnen werden, um mögliche Probleme aufzudecken und Lösungsansätze zu finden.

3.2.1 INTERAKTIONSAUFGABEN

In dieser Arbeit werden wesentlich Grundfunktionalitäten, die denen in aktuellen Angiographiesystemen entsprechen, untersucht. Dafür wurden wesentliche bestehende Interaktionsaufgaben bei der Navigation in Bilddaten anhand der Beobachtungen identifiziert und sind auch durch Hatscher et al. [7] spezifiziert worden:

SELEKTION

Um einen Bereich auf dem Interface (Monitor) anzuwählen, wird ein Cursor mit dem am Schalterpult dafür vorgesehenen Joystick zum Ziel bewegt. Es ergeben sich daraus zwei Teilaufgaben:

- Ansicht auswählen
- Ansicht bestätigen

MANIPULATION

Soll der angewählte Inhalt nun manipuliert werden, stehen je nach Datentyp (Fluoroskopie oder 3D-Volumen) verschiedene Interaktionen zur Verfügung. Auch hier kann per Tastendruck und Bewegung des Joysticks interagiert werden.

2D-DATENSÄTZE

- Scrolling:
Blättern durch einzelne Schichtbilder mittels Vor- und Zurückscrollen
- Panning:
Verschieben eines sichtbaren Bildausschnitts
- Zooming:
Heran- und Herauszoomen eines Schnittbildes

3D-DATENSÄTZE

- Rotate:
Rotieren eines Volumenmodells
- Zooming:
Vergrößern und Verkleinern

3.2.2 **BEOBACHTUNGEN UND WORKFLOWANALYSE**

Sämtliche Geräte mit denen Menschen in Berührung kommen können, sind in Kunststoffolie eingehüllt. Dies ist notwendig, da das Risiko einer Bakterienkontamination besteht [61]. Zudem tragen alle Beteiligten Handschuhe, was gerade die Bedienung des Schaltpults und insbesondere des Touchscreens erschwert.

Die Kommunikation im OP stellte sich als eingeschränkt heraus. Zwar kommunizierte der ausführende Radiologe mit der Assistenz im OP-Saal, wollte er jedoch einen Kollegen um Rat beten, so musste die sterile Assistenz den Saal verlassen und im Kontrollraum Bescheid geben, dass der Kollege Hilfe benötigt. Die Ärzte versuchen nun durch die Verbindungstür, welche oft zwischendurch offenstand, der beiden Räume zu diskutieren und eine Lösung zu finden. Ebenso kam es vor, dass ein Telefon aus dem Kontrollraum reingereicht wurde. Ebenso wurde beobachtet, dass der ausführende Radiologe laut denkt, sprich seine Handlungen stetig beschreibt.

Während der Durchführung des chirurgischen Eingriffs sind die Hände des Radiologen stets am Patienten, um die medizinischen Instrumente zielgerichtet durch die Gefäßstrukturen zu führen. Die Augen richten sich dabei in der Regel auf dem vor ihm befindlichen Bildschirm mit medizinischen Bilddaten, welcher ihm Orientierung und Informationen bietet.

Die pathologischen Strukturen weisen geringe Kontraste auf und sind in Graustufen dargestellt. Zwar lassen sich die Kontraste einstellen, was wiederum mit Handgriffen am Schaltpult zu verbinden ist. Allerdings müssen zur Manipulation dieser Daten nun Arbeitsschritte unterbrochen werden, weil der Arzt zur Bedienung der Steuerungselemente häufig seine Position verändern muss [4] [5]. Es wurde beobachtet, dass die MTRA den Monitor auf Wunsch des Arztes während der Intervention näher rückte. Obwohl dieser nun näher dran war, lehnte sich der Radiologe häufig nach vorne, um Details besser erkennen zu können und einen Weg zur Zielstruktur zu finden. Diese Beobachtung wurde auch bei Hübler et al. getätigt. Dadurch, dass die Steuereinheit relativ groß und nicht zentral gelegen ist, kommt es beim Vorbeugen zur Berührung der Joysticks, was ein unbeabsichtigtes Auslösen verursachen kann.

Während eine Diagnostik in wenigen Minuten getroffen werden kann, können bei einer Intervention Stunden vergehen. Allgemein gilt es die Dauer eines Eingriffs

so kurz wie möglich zu halten. In der dargestellten User Journey Map wurde eine typische Intervention grafisch dargestellt, die auf protokollierten Daten von Hübler et al. beruht. Dabei lässt sich ein beispielhafter Workflow in Teilschritten betrachten, in denen jeweils die Tätigkeiten von Arzt (blau) und MTRA (orange) differenziert wurden. Dabei werden Muster in den Abläufen sichtbar, denn häufig kommt es zu Wechselwirkungen beider Nutzer und die Interaktion muss unterbrochen werden. Es wurden sämtliche Protokolle analysiert und Daten der einzelnen Teilschritte aufsummiert, sodass ein Graph (blau) erstellt werden konnte, der die Häufigkeit an Interaktionen innerhalb der Teilschritte skizziert. Ebenso werden weitere Beispiele typischer Probleme detailliert aufgezeigt und Berührungspunkte (Touchpoints, grün) möglicher Lösungsansätze berührungsloser multimodaler Interaktion eingebracht.

3.2.3 **EXPERTENINTERVIEW**

Um die Eindrücke der Hospitation zu ergänzen, wurde ein schriftliches Experteninterview mit einem Radiologen des Instituts für Neuroradiologie des Universitätsklinikums Magdeburg durchgeführt. Zur Vorbereitung wurde ein unstrukturierter Fragebogen angefertigt (Anhang A1), indem die konkreten Handlungsabläufe, Einschränkungen, Interaktionen mit dem vorhandenen Angiographiesystem abgefragt wurden. Dabei wurden allgemeine Handlungen zur Bildnavigation, Spracheingabe und Bildvergrößerungen abgefragt.

Insgesamt beurteilt der Experte die Bedienung des aktuellen Systems sehr positiv, jedoch werden häufig Einschränkungen in der Bedienung des Touch-/Controllpanels als störend empfunden. Seine Handlungen sind routiniert, die Bedienung des Schaltpults wird als intuitiv empfunden und in der Regel verliert er auch selten die Orientierung beim Sichten der Bilddaten. Besonders interessant ist die Aussage, dass beim Navigieren durch die Bilder kein Blick auf den Joystick gerichtet werden muss, was bedeutet, dass die Handaugenkoordination durch die Routine perfektioniert wurde. Er weist darauf hin, dass individuelle Ansichten je nach Präferenz des Arztes eingesetzt werden. Der befragte Radiologe zum Beispiel bevorzugt eine permanente Vollbildansicht, um Details am Bildschirm besser erkennen zu können.

Auch wird bestätigt, dass eine Bildvergrößerung definitiv notwendig ist, dies jedoch mit vielen Handgriffen verbunden ist und deshalb unpopulär im Kollegium sei. Zoomen findet in seinem Fall in fluroskopischen 2D-, als auch in 3D-Volumen-Daten statt. Diese Aufgabe übernimmt der Befragte in der Regel selbst, weil eine Be-

schreibung des Ziels an die MTRA zu ineffizient wäre. Generell hält er eine berührungslose Zoominteraktion als einen interessanten Aspekt und schätzt diese als relevante Forschungsfrage ein.

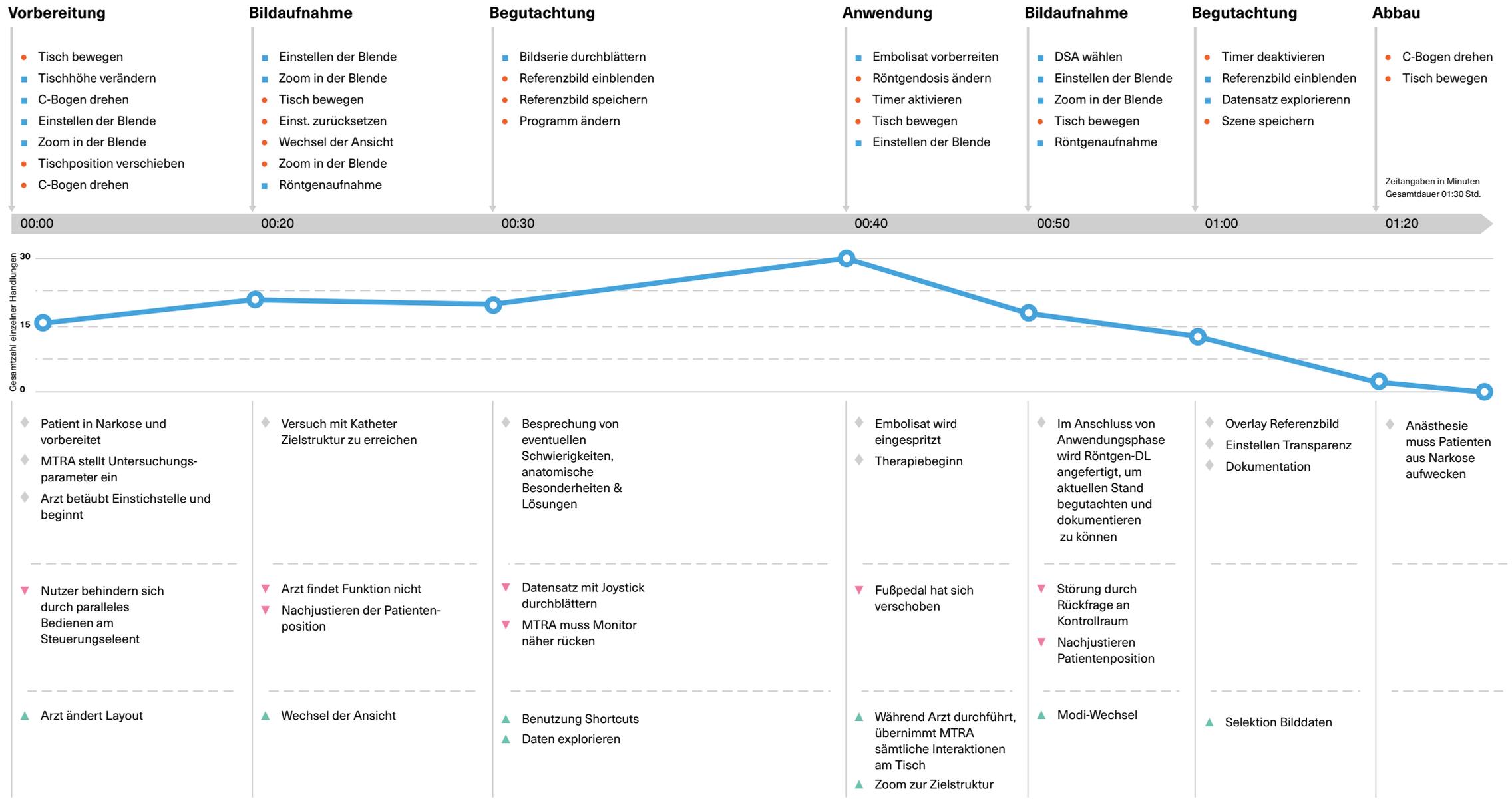
Weiterhin erklärt er auch, dass häufig Einstellungen und Ansichten geändert werden müssen und dies viele Handgriffe erfordert. Dabei wird er nicht immer von der MTRA unterstützt, was mutmaßen lässt, dass er zweifach Schritte vermeiden will. Auf die Frage nach einer denkbaren ergänzenden Modalität, wird sowohl auf einen Blick, als auch auf Spracheingabe hingewiesen. Nach dieser Aussage wären Gesten prinzipiell ausgeschlossen. Da meist der rechte Fuß zum Strahlen genutzt wird, bestände noch „Platz für eine Aufgabe“ mit dem linken Fuß.

Außerdem wurde erfragt, wie hoch die Akzeptanz innovativer Ansätze unter Radiologen ist und ob aus seiner Sicht generationsbedingte Differenzen existieren. Als junger Arzt ist er aufgeschlossen und interessiert an neuen Technologien und den einhergehenden möglichen Verbesserungen. Allgemein schätzt er die Bereitschaft, Dinge zu ändern, generell als gering ein, weil die Radiologen speziell auf diese Systeme ausgebildet sind und demnach Handgriffe routiniert sind, wodurch kein Grund gesehen wird, Neues zu lernen.

User Journey Map

Zusammengefasste Darstellung einer typischen radiologischen Intervention

- Legende:**
- Arzt
 - MTRA
 - ◆ Bemerkung
 - ▼ Probleme
 - ▲ Touch Point



3.3 Kriterienkatalog

In diesem Kapitel wurde bisher aufgezeigt, welche Probleme und Hindernisse während einer radiologischen Intervention auftreten können. Dabei wurde heuristisch versucht Nutzerbedürfnisse zu skizzieren und systematisch Probleme zu identifizieren. Die folgende Zusammenfassung orientiert sich an aktuellen Normen und Kriterien der berührungslosen Interaktion der Arbeit von Manresa-Yee et al. [62].

Beim Entwurf von Benutzerschnittstellen lassen sich nach Norman und Nielsen [63] sieben Grundprinzipien zur Gestaltung von Interaktionen ableiten: Sichtbarkeit, Rückmeldung, Konsistenz, umkehrbare Operation, Auffindbarkeit, Skalierbarkeit und Verlässlichkeit. Schaffer [64] überträgt diese Prinzipien auf gestenbasierte Interaktionen und vermittelt dabei spezifische Eigenschaften, die diese aufweisen sollten: erfassbar, vertrauenswürdig, reaktionsfähig, geeignet, bedeutungsvoll, intelligent, verspielt, angenehm und zielgruppenorientiert. Bei der Entwicklung und Umsetzung von medizinischen Produkten, ist die Anwendung eines nutzerzentrierten Designprozesses gesetzlich vorgeschrieben. So sind nach den Normen DIN EN 60601-1-6 und DIN EN 62366 Hersteller von Medizintechnik dazu verpflichtet, einen detaillierten Usability-Engineering-Prozess zu folgen und diesen zu dokumentieren. Zur Sicherstellung der Benutzerfreundlichkeit gilt es, die Usability-Kriterien zu beachten, die nach ISO-Norm 9241-11 unter „Anforderungen an die Gebrauchstauglichkeit“ festgehalten sind: Effektivität, Effizienz und Nutzerzufriedenheit. Da diese Arbeit nicht das Ziel hat eine reale Software umzusetzen, werden diese Anforderungen an dieser Stelle nicht weiter erläutert, weil sie zur Evaluation unter den Bewertungskriterien 6.2.5 angewendet wurden. Zusammengefasst ergeben sich folgende klinische Anforderungen an das System:

STERILITÄT

Die Eingabemodalitäten dürfen keine zusätzlichen hygienischen Probleme hervorbringen.

DAUER

Der zeitliche Aufwand der Interaktion sollte nicht größer sein als bisher.

PRÄZISION

Die Bildnavigation sollte präzise durchführbar sein und das Gefühl von Kontrolle vermitteln.

VERTRAUTHEIT

Die Interaktion sollte sich auf bekanntes Verhalten beziehen und natürlich sein.

ERGONOMIE

Die physische Beanspruchung gilt es so gering wie möglich zu halten.

AUFMERKSAMKEIT

Der kognitive Aufwand sollte gering gehalten werden.

ARBEITSUMFELD

Der eingeschränkte Arbeitsplatz und die beteiligten Nutzer sollten berücksichtigt werden.

INTERAKTIONSRaum

Die auszuführenden Interaktionen sollten platzsparend sein.

SICHERHEIT

Unbeabsichtigte Auslösungen sollten vermieden werden.

FEHLER

Mögliche Fehleingaben sollten vermieden werden und bei etwaigem Auftreten selbstständig korrigiert werden können.

ERLERNBARKEIT

Die Interaktionstechniken sollten leicht trainierbar sein.

FEEDBACK

Der Nutzer sollte kontinuierliche Rückmeldung über den aktuellen Systemzustand haben.

Konzeption

I < O

KAPITEL 4

Aktiv Mode	Passiv Mode	Gemischt (Skandal)
Kommando > System start > Aktivierung Zoom Start > Selektion > Abbruch	Manipulation > Natürlich > Panning > Zooming	Kombinationen Pattern > Keyword/Sortelbrat + Passive
> Geste > Mieren Kopf > Schütteln Kopf > Augenbraue > Blinzeln > Schütteln > Sprembaechel	> Kopfrotation > Oberkörper Vor/zie > Augenbraue Auf/zu > Schütteln/Krucken > Kopf Nicken	Bsp: Augenbraue runter + Vorlehnen (Aktiv) (Passiv)

Während im vorherigen Kapitel die generellen System- und Nutzeranforderungen klassifiziert wurden, geht es in diesem Abschnitt darum, diese in Konzepte zu übertragen. Die Vorgehensweise zur Konzeption war iterativ, d.h. es wurden zunächst freie Ideen entwickelt, prototypisch umgesetzt und anschließend den Anforderungen gegenübergestellt, woraufhin wieder neue Ideen entstanden sind. All dies geschah mit Hinblick auf eine grundlegende Untersuchung geeigneter Interaktionstechniken und nicht mit dem Ziel einer technischen Umsetzung und Integration in ein bestehendes Angiographiesystem. Ebenso wurden die Konzepte vorwiegend so entwickelt, dass eine Evaluation der entstandenen Hypothesen innerhalb einer Studie untersucht werden konnten. Der gestalterische Schwerpunkt dieser Arbeit liegt in Entwicklung des Userinterfaces. Zu Beginn der Konzeptionsphase wurde ein klassisches Brainstorming eingesetzt, sowie Mindmaps erstellt und ebenso mit Hinblick auf die Studie versucht geeignete Kombinationen verschiedener Eingabemodalitäten ausfindig zu machen.

Anhand der Anforderungsanalyse wurden wesentliche Interaktionen des Radiologen ausfindig gemacht. Auf folgende wesentliche Interaktionsaufgaben wird sich im Konzept beschränkt:

- Trigger/Kommando
- Blinzeln
- Sprache
- Nicken
- Schütteln
- Augenbraue

Sekundär Sate

Workflow (Diskrete Interaktion)

	Start	Select	Exit	Stop
A	Augenbraue runter	Augenbraue runter	Schütteln	Schütteln
B	Nicken	Nicken	Schütteln	Schütteln

	Zoom In	Zoom Out
C	Vorlehnen	Zurücklehnen
D	Augen runter	Augen hoch

Primär Sate

Mit Schütteln

1	CE
2	CE
3	CE
4	DE
5	DE
6	DE
7	CF
8	CF
9	DF
10	DF

- Start
- Selektion
- Spea

- zurück/Abbruch Default
- Zoom Start
- Zoom Direkt
- Zoom Stufen

Gesicht (Augenbraue)

Ks	Augenbraue		
An			
Oz			
Sr			

AKTIVIERUNG

Der Nutzer soll das System und die gewünschte Funktionalität starten und beenden können.

NAVIGATION

Es soll dem Nutzer ermöglicht werden sich frei, d.h. ohne Einschränkungen in den Freiheitsgraden, innerhalb der am Monitor dargestellten Bilddaten bewegen zu können.

SELEKTION

Der Nutzer soll Bereiche auf den Bildschirm selektieren können.

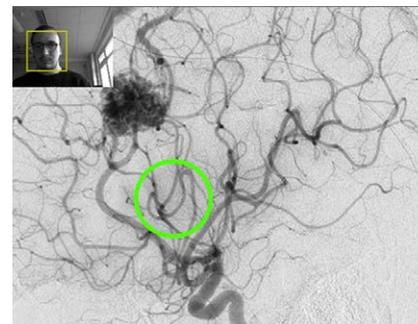
MANIPULATION

Eine Bildvergrößerung und Verkleinerung soll stufenlos gewährleistet sein.

4.1 Vorexperiment

Wie sich u.a. in den Beobachtungen herausstellte, lehnen sich Radiologen während der Intervention häufig nach vorne, um Details auf dem Monitor besser erkennen zu können. Dies wurde zum Anlass genommen zu überprüfen, ob sich das Vor- und Zurücklehnen zum Vergrößern von Bilddaten an einem Monitor im Allgemeinen eignet. Es wurde ein Softwareprototyp mithilfe von Processing umgesetzt und in einem nicht repräsentativen Experiment evaluiert. Processing ist eine stark typisierte objektorientierte Programmiersprache mit interner Entwicklungsumgebung und für die Einsatzbereiche Grafik, Simulation und Animation spezialisiert. Ziel dieser Untersuchung war der Gewinn erster Erkenntnisse des natürlichen Nutzerverhaltens in Bezug auf diese Interaktionstechnik, sowie der Entwicklung geeigneter Prinzipien, Ideen und Technologien.

ABBILDUNG 16: Screenshot die ersten Prototypen. Zusehen ist eine Digitale Subtraktionsangiographie, die mithilfe eines Cursors (grüner Ring) exploriert werden kann



Aufgebaut wurde ein Setup bestehend aus einem Monitor und einer darauf installierten Webcam (s. Abbildung 15). Die Programmierbibliothek OpenCV diente der Erfassung des Gesichts. Ein Cursor wurde eingeblendet, dessen Position relativ zu der durch die Webcam aufgenommenen Gesichtsposition des Betrachters ist (s. Abbildung 16). Nähert sich dieser und überschreitet den definierten Abstand, vergrößert sich das erfasste Gesicht, womit sich das Bild proportional dazu vergrößert. Das Verschieben des Bildausschnitts gelang, wenn das Bild vergrößert wurde und anschließend eine Veränderung der Gesichtsposition stattgefunden hat.

Insgesamt fanden die Probanden die Interaktion intuitiv, bewerteten diese jedoch als sehr anstrengend. Dies ist auf die Limitierung der Bewegung zurückzuführen, was auch als größter Kritikpunkt angesehen wurde

ABBILDUNG 15: Versuchsperson befindet sich in einem Abstand von 1,5m frontal zum Aufbau.



Im Experiment sahen die Probanden eine Weltkarte und wurden nach Ländern und Hauptstädten befragt. Diese sollte mithilfe der implementierten Zoomtechnik lokalisiert werden. Beobachtet wurde das Verhalten von vier Versuchspersonen hinsichtlich der Durchführung der Interaktion mit folgenden Ergebnissen:

- Mit den Augen wurde das Ziel zunächst fixiert und dann versucht den gesamten Körper so auszurichten, dass der Cursor die gewünschte Position erreicht. Die Fixierung des Cursors erwies sich dabei als schwierig.
- Horizontale Bewegungen fielen leichter. Um Regionen oberhalb des Äquators zu erreichen, musste sich der Nutzer strecken und war aufgrund körperlicher Abmessungen limitiert.

- Aufgrund der steifen Haltung veränderte sich der Kopfwinkel nicht. In den meisten Fällen verschränkten die Probanden automatisch konsequent die Arme hinter ihrem Rücken.
- Es kamen keine zusätzlichen Gesten oder auffälligen Sprachausdrücke vor.
- Zwei Probanden kniffen die Augenbrauen zusammen, um Beschriftungen lesen zu können. Niemand hatte eine Sehschwäche.

4.2 Limitation

Wie Oviatt und Cohen [65] bemerken, ist die Wahlfreiheit der Modalitäten ein großer Vorteil der Multimodalität, weil es die Flexibilität bietet, in einem großen Spektrum mit situationsbedingten Bedürfnissen umzugehen. Für die Entwicklung von neuen Ideen kann eine zu große Auswahl auch von Nachteil sein.

Aufgrund der Bewegungseinschränkungen des Nutzers und den Anforderungen an ein multimodales System schränkt sich die Wahl der Modalitäten ein. Eine weitere Einschränkung wurde vom Autor selbst festgelegt: Wie im Kapitel 2.4 Verwandte Arbeiten beschrieben, existieren bereits viele Studien, die im medizinischen Kontext Sprach-, Hand-, Fuß-, Blickinteraktion untersuchen, weshalb sich dafür entschlossen wurde, eigene Ansätze zu entwickeln. Des Weiteren werden die kollaborativen Bedürfnisse und Aktivitäten der verschiedenen Nutzer zur Bedienung des Systems nicht berücksichtigt. Zielgruppe soll in diesem Fall nur der durchführende Radiologe sein, wodurch die Eingaben zunächst nur von einem Nutzer erfolgen sollen.

Für den Radiologen soll es möglich sein, seine Aufmerksamkeit vollkommen auf die Durchführung des Eingriffs zu richten. Da dafür die Hände stetig am Patienten sind, um u.a. den Katheter zu führen, soll es ermöglicht werden, mit dem System, ohne die Benutzung dieser zu interagieren. Die Interaktion mit den Füßen soll ebenso nicht berücksichtigt werden, weil sie bereits eingesetzt wird und als praktisch angesehen wird und somit keinen Innovationsgehalt mit sich bringt. Außerdem kann die Verlagerung des Gewichts und die unnatürliche Stellung negativ auf die Ergonomie wirken.

4.2.1

SENSORIK

Für die Erfassung von berührungslosen Benutzereingaben bieten sich zahlreiche Technologien an, wie in Kapitel 2.3 erläutert wurde. Dabei kann das Anbringen von Sensoren direkt am Nutzer eine genaue Erfassung von Eingaben begünstigen. Dafür müsste der Nutzer, in diesem Fall das medizinische Personal, diese im realen Szenario selbst anbringen und stetig tragen. Es kommt zusätzlich zu einem weiteren Gefahrenpotenzial in Bezug auf die Sterilität, denn die Hardware muss medizinischen Normen entsprechen. Das zusätzliche Anbringen und Tragen einer am Körper angebrachten Sensorik kann körperliche Einschränkungen mit sich bringen und als störend empfunden werden. Das Anbringen von Markern kann ein gezieltes und effizientes Tracking ermöglichen, ist jedoch aufgrund von Einwegkleidung (Jacken, Hosen, Mundschutz, Handschuhe), die in der OP vorgeschrieben ist und nach OP-Ende sofort gereinigt bzw. entsorgt werden sollte, hinderlich und eine weitere Mehrbelastung für die Nutzer.

Kamerabasierte Systeme sind zur handfreien Interaktion eine vielversprechende Technologie. Mittels der Microsoft Kinect wurden umfangreiche Studien durchgeführt, die eine berührungslose Manipulation von medizinischen Bilddaten gewährleisten [66]. Zudem sind Systeme dieser Art besonders in Innenräumen, wie hier im OP-Saal, besonders vielversprechend, weil sie relativ einfach einzurichten sind und eine akkurate Datenübertragung begünstigen, was wiederum eine für eine valide Implementierung von Gesten spricht [29].

4.3 Workflow

Um den Untersuchungsrahmen einzugrenzen, wurden Teilschritte im Workflow zur Bedienung der Benutzerschnittstelle und Grundfunktionalitäten festgelegt, die denen einer realen OP entsprechen. Im ersten Schritt wurden lose Ideen entwickelt und anschließend durch die beschriebenen Limitationen verfeinert und ausgeschlossen. Diese Elemente und deren Zusammenspiel sind in Abbildung 17 visualisiert und werden wie folgt beschrieben:

AUTHENTICATE

In Fragen der Sicherheit spielt die Authentifizierung eine wichtige Rolle. Der Nutzer sollte sich beim System anmelden können und anhand seiner Nutzerdaten identifiziert werden, denn dies würde eine individuelle Systemkonfiguration ermöglichen. Dabei könnte sich die Tischhöhe automatisch auf den jeweiligen Arzt anpassen oder das Layout sich nach seinen Präferenzen verändern. Denkbar wäre, dass der Nutzer anhand seiner Stimmfarbe oder seiner Gesichtszüge erkannt wird. Viele aktuelle Smartphones bieten heute ähnliche Funktionen, um das System zu entsperren. Maschinen-Lerntechniken könnten eingesetzt werden, um die Leistung und Anpassungsfähigkeit der Personalisierung zu optimieren [18].

Zu Prüfen gilt hierbei jedoch der gesetzliche Rahmen zu Datenschutzordnungen. Die Personalisierung und Möglichkeiten verschiedener Nutzer zu differenzieren, können in dieser Arbeit nicht berücksichtigt werden und sind weniger relevant für die Evaluation der Interaktionstechniken. Grundsätzlich ist es aber sinnvoll zu erfassen, ob und wie viele Nutzer sich im Sichtfeld des Systems befinden. Dafür kann erfasst werden an welcher Position im Raum und in welcher Distanz relativ zum System sich Nutzer befinden.

ACTIVATE & DEACTIVATE

Im Konzept soll es möglich sein berührungslos das System zu starten und zu stoppen bzw. Funktionen zu aktivieren und zu deaktivieren. Dabei ist es wichtig, dass etwaige Eingaben, bspw. in Form von Gesten korrekt interpretiert werden damit das System nicht unfreiwillig manipuliert wird [67]. Die Funktionalität, ein System aus dem Ruhemodus zu bringen, wird in der Literatur als „Clutching Mechanism“ bezeichnet [49] [68].

PAN

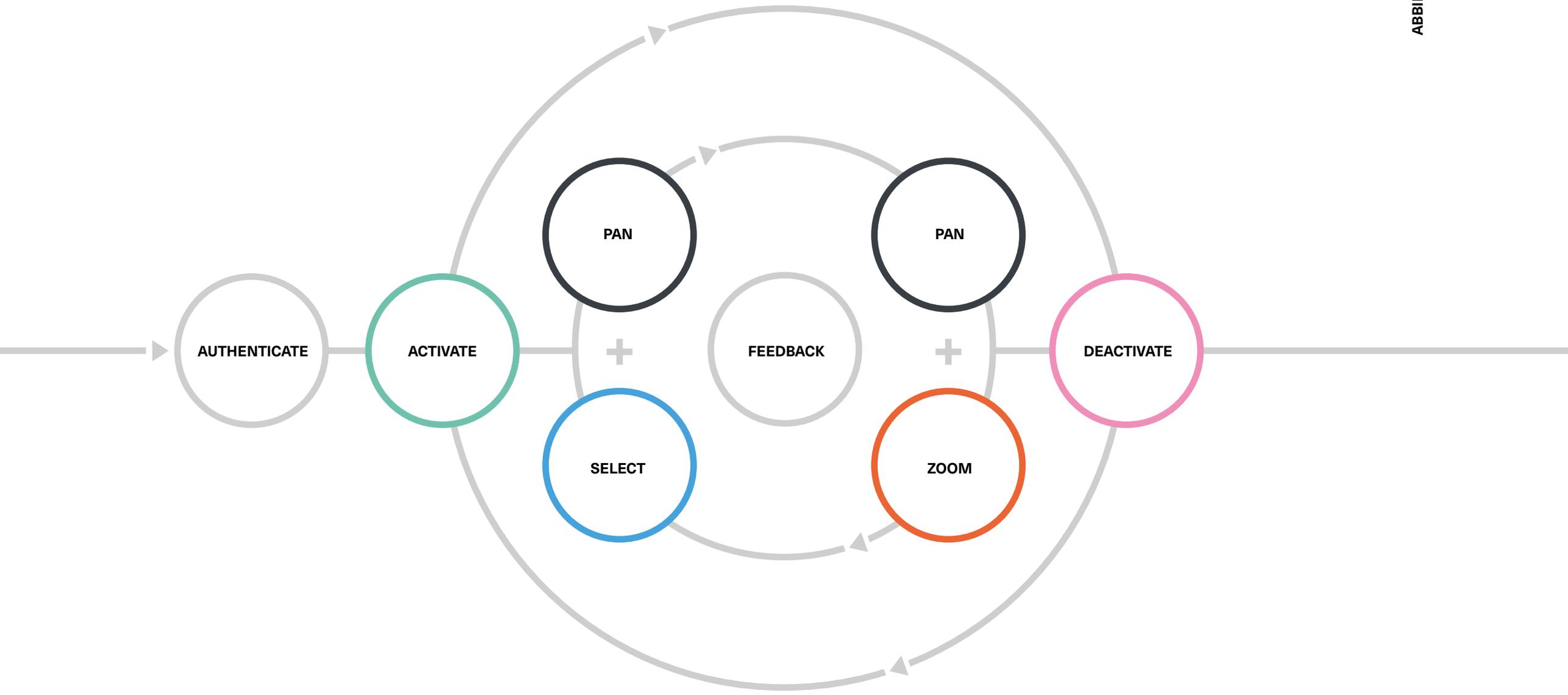
Das Verschieben eines Bildausschnitts, auch Panning (engl. Schwenken) genannt, wird mit Hilfe der Rotation des Kopfes ermöglicht. Dieselbe Interaktionstechnik wird zum Navigieren zu bestimmten Bereichen auf dem Interface genutzt.

SELECT

Ein explizites Kommando soll es ermöglichen, gewünschte Ansichtsfenster zu wählen.

ZOOM

Die Bildvergrößerung soll kontinuierlich, das heißt stufenlos, möglich sein. Während der Bildausschnitt vergrößert ist, dienen Kopfbewegungen als Mittel der Exploration.



4.4 Eingabemodalitäten

Zusammengefasst ergeben sich aus dem vorherigen Abschnitt vorwiegend Eingabemodalitäten oberhalb des Hüft- bzw. Ellenbogenbereichs. Diese sind in Abbildung 18 kenntlich gemacht. Demnach werden Interaktionstechniken präsentiert, die sich als Eingabemodalität eignen.

OBERKÖRPERGESTEN

Weil sich die Hände des ausführenden Arztes während der Operation, die meiste Zeit am Patienten befinden, tritt durch die Armbeugung zugleich eine Einschränkung der Armbewegung auf. Mögliche Interaktionsideen sind hierbei das Vor- und Zurücklehnen, eine Links-Rechts-Bewegung oder das Bewegen der Schultern. Letztere könnten Probleme mit der Handkoordination hervorrufen, weil sie die Arme gleichermaßen mitbewegen. Die bessere Lösung scheint das Vor- und Zurücklehnen zu sein. Da Körpergesten dieser Art nach Preim et al. [30] wenig Präzision erlauben, sind sie besonders als unterstützende Modalitäten interessant, da, wie hier, die Hände nicht benötigt werden. In diesem Fall bietet sich diese Technik zum Zoomen an, wie Schöning et al. bereits untersuchten [48]. Auch aus dem Alltag ist das Verhalten des Vorlehns ganz natürlich. Soll eine nähere Betrachtung stattfinden, ist eine Bewegung in Richtung des Zielobjektes nötig.

Durch Entfernen vom Zielobjekt erhält der Nutzer stattdessen einen Überblick über die Situation. Zur Erfassung des Oberkörpers eignet sich besonders die Kinect [2].

KOPFGESTEN

Eyetracking ist eine vielversprechende und verbreitete Methode multimodaler Interaktion. So realisierten Hatscher et al. [7] eine multimodale Interaktion anhand einer Kombination von Eyetracking und Fußgesten. Der Blick als Zeiger (Pointer) stellte sich demnach als Eingabemodalität als besonders praktikabel heraus, jedoch können leichte Komplikationen auftreten, wenn andere physische Gesten hinzugezogen werden. Der größte Nachteil an dieser Technologie ist, dass wie in diesem Fall der Nutzer ein mobiles System tragen muss, weil auf die relativ große Distanz zwischen Arzt und Display keine exakten Augeninformationen erfasst werden können. Die Idee in dieser Arbeit ist, die Erfassung der Kopfrichtung. Wagner et al. [69] präsentierten dafür Umsetzungsmöglichkeiten verschiedener Kopfgesten, wie u.a. Nicken und Schütteln.

GESICHTSGESTEN

Die Mimik, dazu zählen alle physischen Bewegungen der Gesichtsoberfläche, ist Teil des menschlichen Ausdrucksverhaltens. Wobbrock et al. [70] stellten ein

Gestenalphabet für Mimikinteraktion vor, in dem durch in symbolische Gesten übertragende Augenbewegungen eine Texteingabe am Monitor möglich ist. Im medizinischen Kontext gilt zu prüfen, inwieweit derartige Eingaben möglich sind, denn das Tragen von Schutzhaube und Mundschutz könnte das kontaktlose Tracking erschweren. Dennoch stellt die Augenbewegung eine geeignete Interaktion dar, weil Zustandsänderungen und Bewegungsinformationen abgeleitet werden können.

SPRACHEINGABE

Wie zahlreiche Literatur zeigt, sind Spracheingaben im medizinischen Kontext, als auch historisch betrachtet, eine exzellente Wahl bei der Entwicklung multimodaler Systeme. Nach Turk [18] unterstützen diese besonders bei der Manipulation grafischer Informationen eine verbesserte Effizienz, weisen eine hohe Genauigkeit in der Eingabe auf, sind nicht – im Gegensatz zu Gesten – räumlich beschränkt und können dazu beitragen, eine Überbeanspruchung durch eine weitere Modalität zu vermeiden. Nach Preim et al. [30] sind „Gesten für manipulierende Aufgaben sehr gut geeignet, während Sprache durch ihren beschreibenden Charakter für Kommandos, logische Operationen, Parametereinstellungen, Kontextinformationen u. ä. geeignet ist“. Auch durch die in dieser Arbeit durchgeführten Beobachtungen und Befragungen lässt sich Sprache als geeignete Modalität

feststellen. Es gilt jedoch in einer realen und individuellen Situation zu prüfen, ob dies zutrifft. Denn in Studien von Hatscher et al. äußerten die Versuchspersonen Bedenken, weil ein relativ hoher Geräuschepegel während einer OP herrscht [49]. Neben den Geräuschen die von Lüftungsanlage, Bewegung der Angiographieanlage, menschlichen Bewegungen und weiteren medizinischen Geräten ausgeht, ist auch zu untersuchen welchen Einfluss Konversationen der Nutzer und lautes Denken diese Modalität beeinflussen.

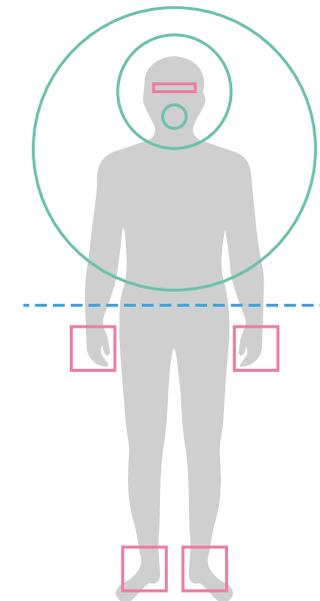


ABBILDUNG 18: Identifizierte (grün) und von der Untersuchung ausgeschlossene (rot) Eingabemodalitäten.

4.4.1 INTERAKTIONSKONZEPTE

Aufgrund der Analyse beschriebener Eingabemodalitäten, wurde sich für folgende Interaktionstechniken entschieden:

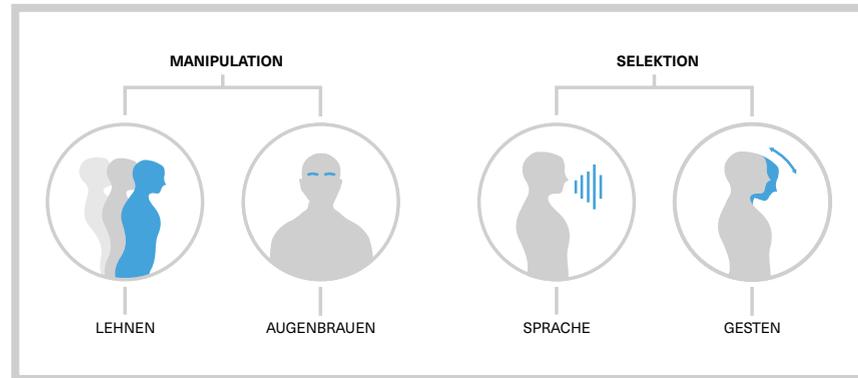


ABBILDUNG 19: Übersicht der Interaktionstechniken

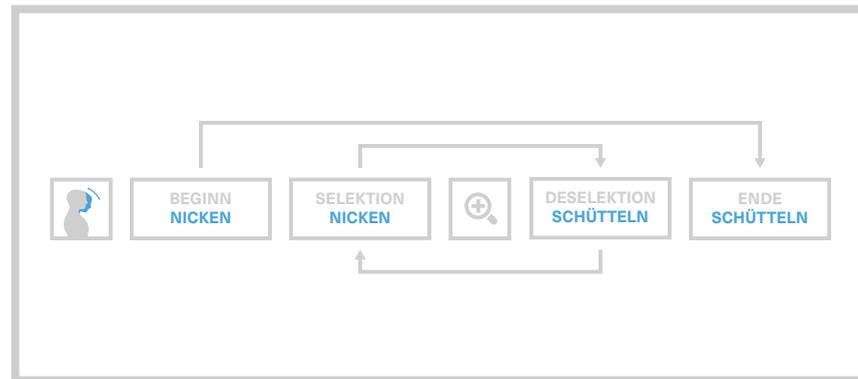


ABBILDUNG 20: Nicken und Schütteln des Kopfes zum (de-)aktivieren und (de-)selektieren.



ABBILDUNG 21: Spracheingabe zum (de-)aktivieren und (de-)selektieren.

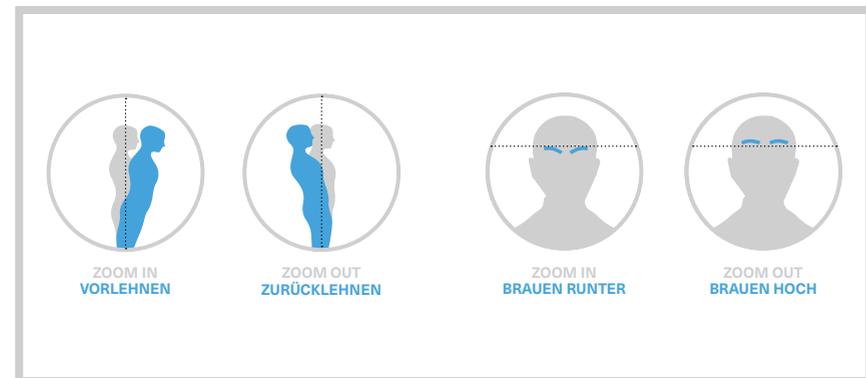


ABBILDUNG 22: Zoominteraktionen bzw. Bildmanipulationstechniken durch Körpergesten (links) und verändern der Mimik (rechts)

Eine angemessene und direkte Rückmeldung über den aktuellen Systemzustand sind für den Nutzer essentiell. Preim et al. [30] stellen heraus, dass speziell bei Gesteninteraktionen nicht nur Auskunft darüber gegeben werden wollte, ob diese generell erkannt wurde, sondern auch, kontinuierliche Auskunft in Echtzeit über die Eingabe bieten. Dies gibt den Nutzer Vertrauen und Sicherheit in der Bedienung des Systems. Für diese Arbeit kommen zwei Ausgabemodalitäten in Frage:

VISUELLES FEEDBACK

Die Benutzeroberfläche sollte visuelle Auskunft über die aktuelle Eingabe geben und wichtige Informationen im sichtbaren Bereich des Nutzers zur Verfügung stellen. Dabei spielen grafische Elemente und deren farbliche Codierung eine Rolle.

AUDITIVES FEEDBACK

Ein weiterer Sinneskanal des Menschen könnte genutzt werden, um das visuelle Feedback zu ergänzen: das Hören. Auf Eingabe des Nutzers könnte eine Sprachausgabe folgen, die ihm mitteilt, welche Eingabe das System erkennt. Ebenso könnten Fehleingaben durch Warntöne oder Bestätigungstöne, ergänzende Auskunft über die Eingabe bieten. Speziell bei der Bildmanipulation ist es außerdem denkbar, dass eine Wiedergabe von zusätzlichen Soundeffekten Hinweise zur Limitation bietet. So könnte ein Tonsignal bei Erreichen der maximalen Vergrößerungsstufe ausgesondert werden. In der Arbeit von Black et al. [71] wurde gezeigt, dass die Kombination von auditivem und visuellem Feedback bei gestenbasierten Interaktionen generell denen einer einzelnen Ausgabemodalität überlegen ist. Jedoch wurde dies in dieser Arbeit vorerst nicht betrachtet, weil auditive Ausgaben Spracheingaben behindern könnten. Ebenso gilt es zu überprüfen, ob Soundausgaben in einer OP angebracht sind, weil zum Beispiel Warntöne Auskunft über einen kritischen Zustand des Patienten geben und mit Signalen des Systems möglicherweise kontrahieren.

4.5 Interface

Um ein grundlegendes Verständnis wichtiger Elemente des Monitorinterfaces zu erlangen, wurde die Oberfläche einer Siemens Angiografieanlage genauer untersucht und anhand dessen erste Entwürfe entwickelt (s. Abbildung auf Folgeseite). Dabei wurde mit verschiedenen Farbakzenten untersucht, welche Farbwahl geeignet erscheint, um unterschiedliche Systeminformationen aufzuzeigen. Die Entwürfe beinhalteten radiologische Aufnahmen, um gerade in Anbetracht von Bildvergrößerungen zu prüfen, welche Skalierung die Bilddaten haben müssen, damit Strukturen gut erkennbar werden. (A) zeigt das Interface ohne geladene Bilddaten mit den spezifischen Systeminformationen. In (B) wird dargestellt, dass als Feedback der Auswahl durch einen farblich abgesetzten Rahmen denkbar ist. Zudem zeigt eine Statusleuchte (orange, links unten), ob das System bereit ist oder ob ein Nutzer erkannt wurde. Unterstützend zum Training könnten verfügbare Sprachbefehle auftauchen und ein weiteres Element im Sichtbereich zeigt an, in welchem Modus sich der Nutzer aktuell befindet (C). Eine weitere Idee war eine Lupenfunktion (D), die punktiert Bereiche vergrößert, wo der Nutzer derzeit hinblickt. (E) zeigt die Vergrößerung eines Bildes innerhalb eines Segments, während (F) eine Funktion beinhaltet, wodurch kleinere Segmente um ein Vielfaches skaliert werden können, um Bilddaten nebeneinander zu stellen.

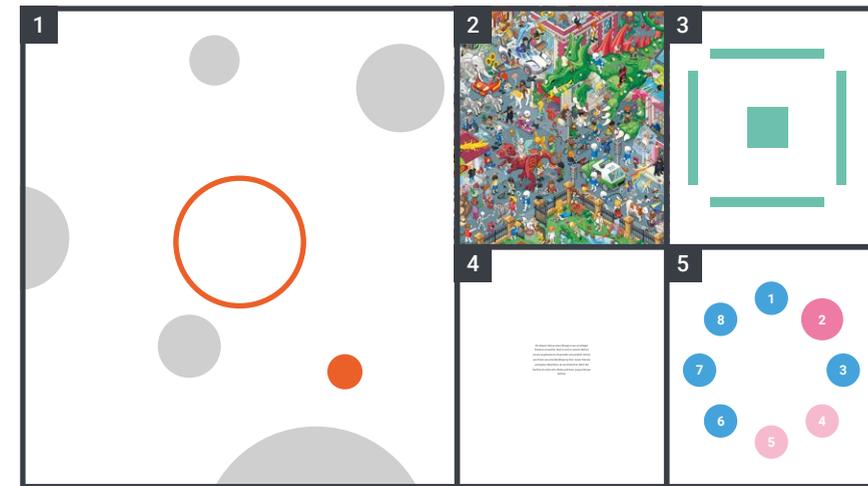
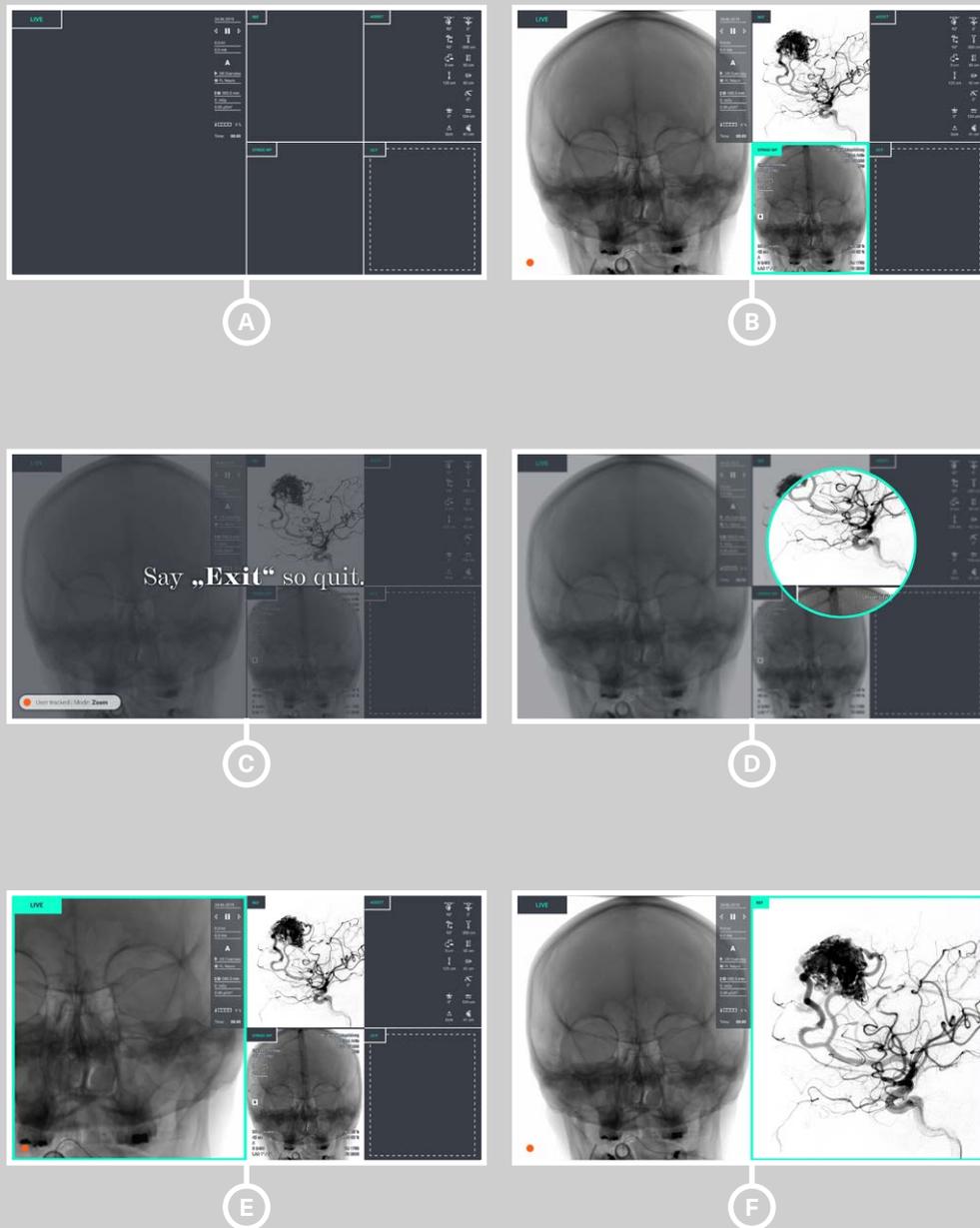


ABBILDUNG 23: Entwürfe unterschiedlicher Aufgaben zur Evaluation verschiedener Interaktionen.

Im nächsten Schritt der Entwicklung wurde sich primär auf die Evaluation der verschiedenen Interaktionstechniken fokussiert (s. Abbildung 23). Der Nutzer sollte darin prinzipiell einen Quadranten mit der Kopfrichtung selektieren und anschließend das Bild manipulieren. Folgende Ideen sind dabei entstanden:

(1) Der Nutzer sieht eine farbliche Kontur (orange) und unterschiedlich große Ellipsen (grau), wobei eine der Farbe der Kontur entspricht (orange). Das Ziel besteht darin, die farbliche Ellipse soweit zu vergrößern, dass sie die Kontur vollständig ausgefüllt wird. Dabei kann die Kontur als Sucher bezeichnet werden (Testen des Skalierungsfaktors).

(2) Ein Wimmelbild ist zu sehen und den Nutzer würde die Aufgabe übertragen,

mit Hilfe einer bestimmten Zoomtechnik das Bild zu vergrößern und ein Objekt darauf zu finden (Untersuchung der kognitiven Belastung).

(3) Ein Rechteck (grün, Zentrum) sollte entlang der umgebenen Texturen mithilfe der Kopfrichtung geführt werden (Prüfung horizontaler und vertikaler Bewegung).

(4) Der dargestellte Text sollte soweit vergrößert werden, dass dieser lesbar wird (Test auf Sehvermögen).

(5) In einer angegebenen Reihenfolge sollten mittels Kopfrotation verschiedene Objekte an- und abgewählt werden (Überprüfung der Präzision von Kopfrichtung).

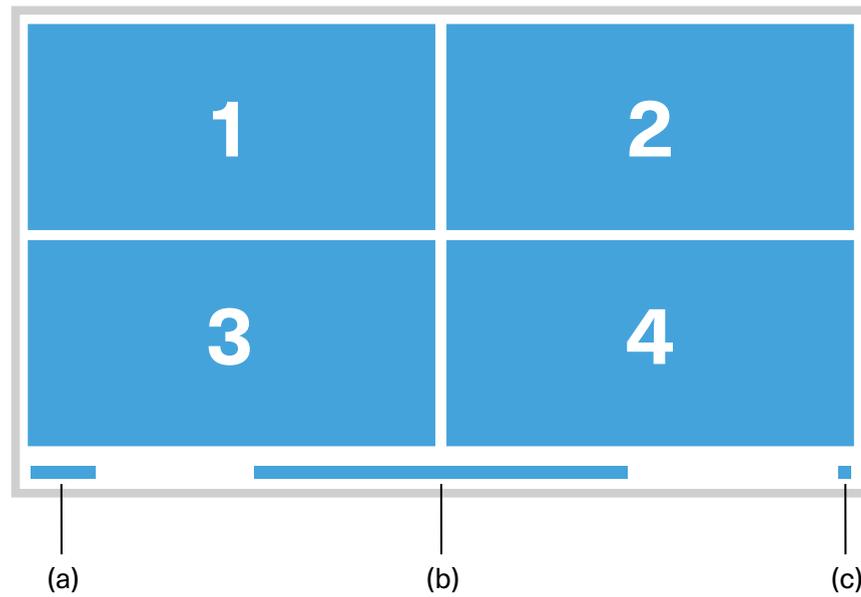


ABBILDUNG 24: Aufbau des Interfaces zur Durchführung der Studie. Im oberen Bereich: 1 – 4 Segmente in den interagiert wird. Eingabemodus (a), Interaktiver Slider (Schieberegler), als Element visuellen Feedbacks bei Durchführung der Primäraufgabe (b), Statusleuchte, welche signalisiert, ob ein Nutzer erfasst wurde (b).

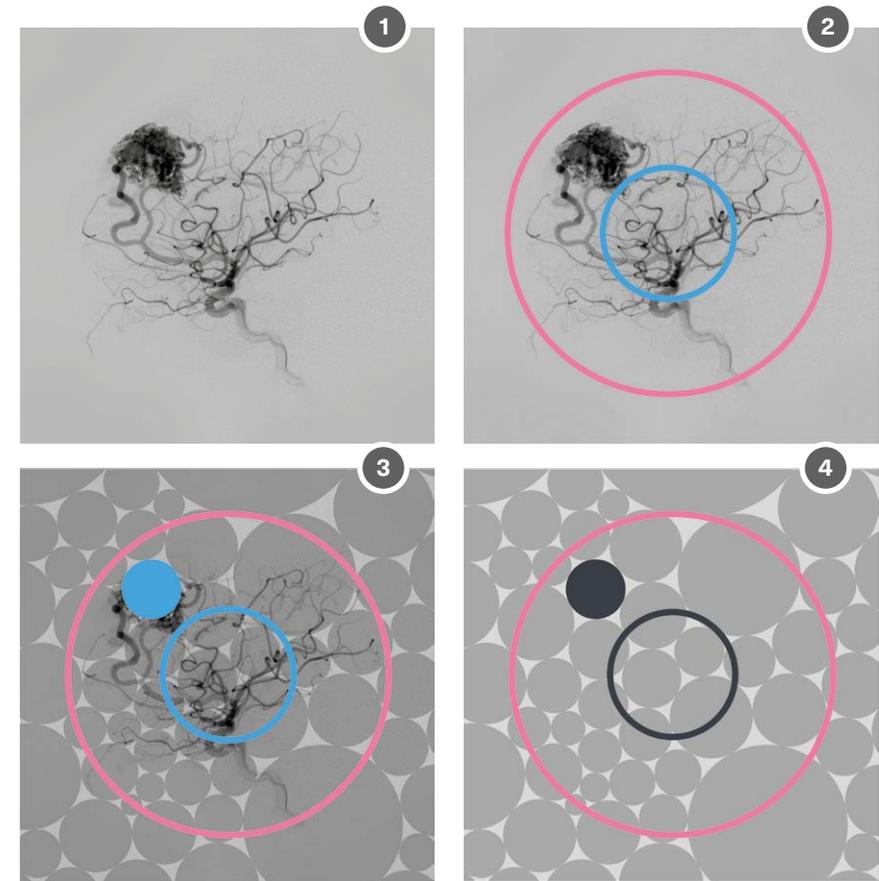


ABBILDUNG 25: Prozess zur Abstrahierung einer AVM Struktur.

Es folgte eine Abstraktion des Bildmaterials einer DSA, worauf eine typische Gefäßstruktur mit einer AVM zu sehen ist (s. Abbildung 25). (1) zeigt die ursprüngliche Aufnahme. In Punkt (2) wurde der Radius der gesamten Struktur eingegrenzt (rot) und ein Sucher platziert (blau). Der wenig vorhandene Kontrast in der Darstellung des Bildmaterials wird überblendet durch ein enges Muster verschiedener Formen in Grautönen (3). Die Position der (möglichen) Zielstruktur (AVM) wird durch eine kontrastreichere Form dargestellt. (4) zeigt das Ergebnis der Abstraktion.

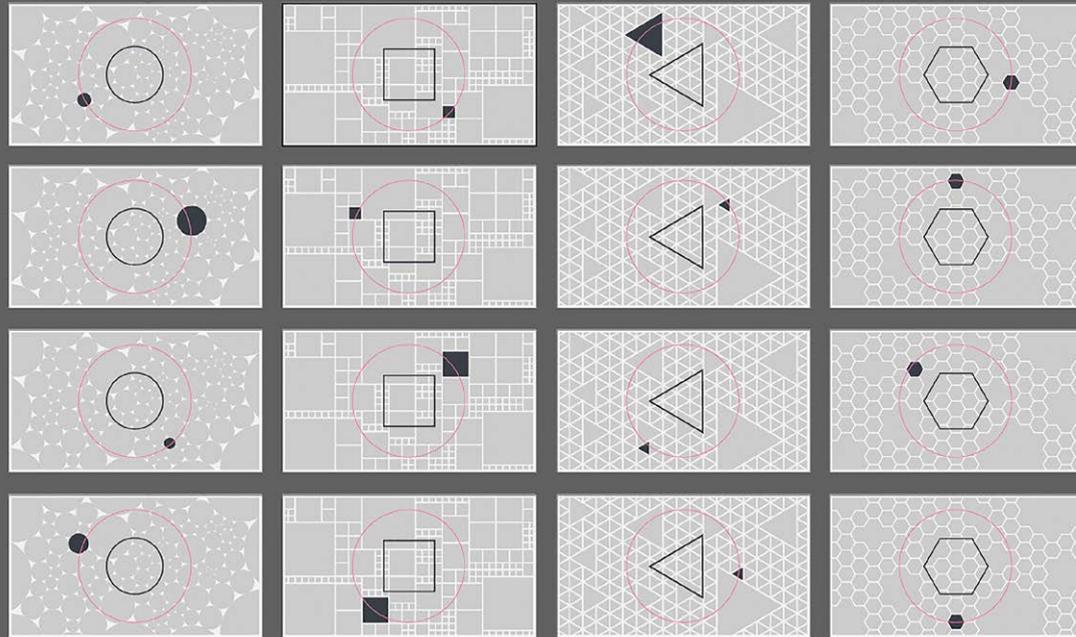


ABBILDUNG 26

Um unterschiedliche Bildaufnahmen zu simulieren, erfolgte die beschriebene Abstraktion mittels unterschiedlicher geometrischer Formen (s. Abbildung 26). Dabei wurden die Zielelemente an vier unterschiedlichen Positionen im beschriebenen Radius platziert, womit eine „Bilddatenbank“ von 16 Bildern entstanden ist.

4.6 Interaktionsaufgaben

4.6.1 PRIMÄRAUFGABE

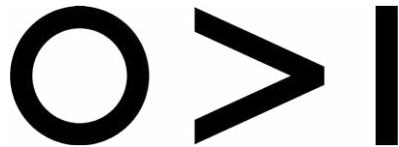
Die in weiteren Kapiteln als Primäraufgabe bezeichnete Aufgabe simuliert dabei die eigentliche Behandlungsaufgabe des Radiologen. Dafür sollen grundsätzlich die Handgriffe des Arztes und die Bewegung des Katheters zur Führung in eine Gefäßstruktur in eine abstrakte Form übertragen werden. Diese Aufgabe stellt die wichtigste Interaktion für den Radiologen dar, denn sein primäres Ziel gilt der Durchführung der Operation und optimalen Versorgung des Patienten. Da in der Realität der Katheter und der darin zu befindlichem Draht sehr fein ist und der Arzt laut Aussagen von Nutzern so gut wie kein haptisches Feedback spürt, wird dieses in der Umsetzung nicht berücksichtigt.

4.6.2 SEKUNDÄRAUFGABE

Die Subaufgabe, wird in dieser Arbeit Sekundäraufgabe genannt und simuliert die Interaktion zur Bildnavigation am Monitor des Angiographiesystems. Damit sollen sekundär Aktionen während der Operation, die die Manipulation und Selektion von Daten umfassen, hinsichtlich den möglichen Eingabemodalitäten an einer grafischen Benutzerschnittstelle untersucht werden. Die Umsetzung dieser Sekundäraufgabe stellt den praktischen Hauptteil der Entwicklung dar.

4.7 Finales Studienkonzept

Im folgenden Abschnitt wird das finale Konzept „OVI“ vorgestellt, welches in Kapitel 6 evaluiert wird. Das Userinterface wurde grafisch aufbereitet und eine konsistente Farbwahl gewählt, welche zu den wichtigsten Faktoren im Design gehört, angewendet. Denn besonders bei medizinischen Anwendungen ist auf die Verwendung einer einheitlichen Farbcodierung zu berücksichtigen, wobei rot, grün, gelb, orange und blau als die an den leichtesten differenzierbaren Farben sind [73], weshalb dieses Farbschema angewendet wurde.



Die Grafik auf den folgenden Seiten zeigt Screenshots der finalen Oberfläche und stellt dabei den Ablauf der Sekundäraufgabe zur Durchführung der Studie in Kapitel 6 dar.

1. Im Ausgangspunkt sieht der Nutzer vier gleichgroße in einem Raster angeordneten Segmente, die ein Muster aus verschiedenen geometrischen Formen und einer gestrichelten Kontur der jeweiligen Form in deren Zentrum beherbergen. Eine dunkle, halbtransparente Fläche liegt dabei über den Segmenten und signalisiert einen Standby-Modus. Im unteren Bereich signalisiert die Statusleuchte, dass kein Nutzer erfasst wurde (Orange). Der Eingabemodus steht auf „Mode“ und erwarte Befehle, genauso wie der Slider auch noch nicht aktiv ist.

2. Ein Nutzer wurde erfasst, was durch die grüne Statusleuchte ersichtlich wird. Ein Cursor (Kreuz) wird eingeblendet und der Nutzer kann diesen durch Veränderung seiner Kopfrichtung bewegen.

3. Nach Eingabe eines initialen Startbefehls (Geste oder Sprache), klart die Oberfläche auf, der Eingabemodus zeigt den aktuellen Status an (Start), der Slider wird aktiviert und kann bewegt werden. Zudem werden einzelne Formen aus dem Muster des Hintergrunds anhand einer zufälligen Auswahl durch das System dunkel (vgl. Abbildung 26). Dem Nutzer ist es nun möglich durch einen Selektionsbefehl (Geste oder Sprache) ein Segment anzuwählen, auf welches er mit dem Cursor zeigt.

4. Bei erfasster Befehlseingabe wird die Selektion des Bereichs mit einem blauen Rahmen hervorgehoben und das Wort „Select“ erscheint im Bereich des Eingabemodus. Der Nutzer befindet sich damit direkt im Zoom-Modus und kann das Bild manipulieren. Gleichzeitig verschwindet der Cursor, weil dieser gleichzusetzen ist mit der geometrischen Kontur.

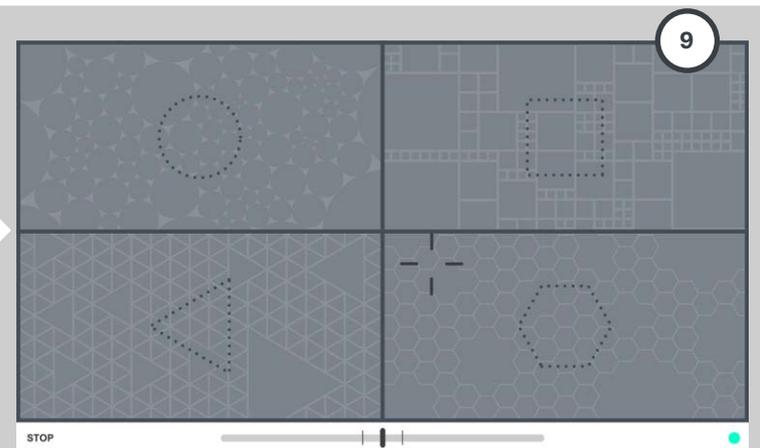
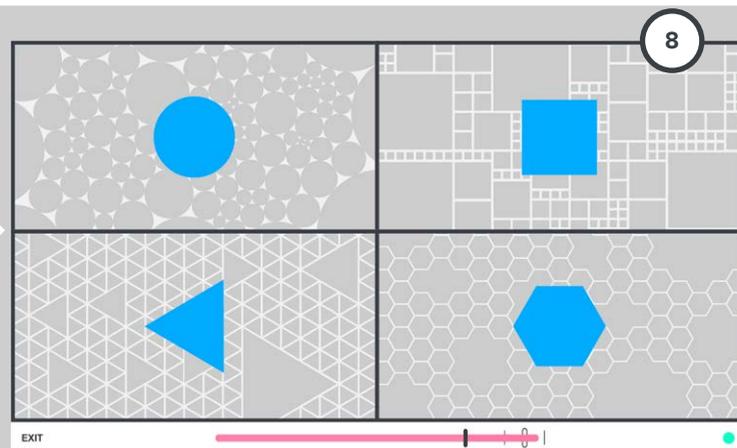
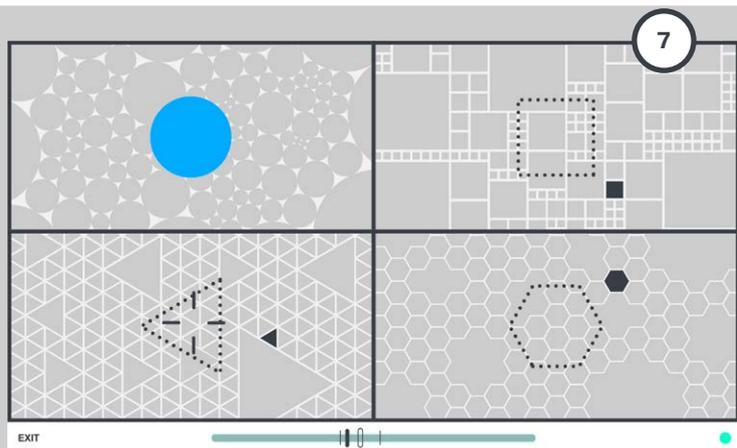
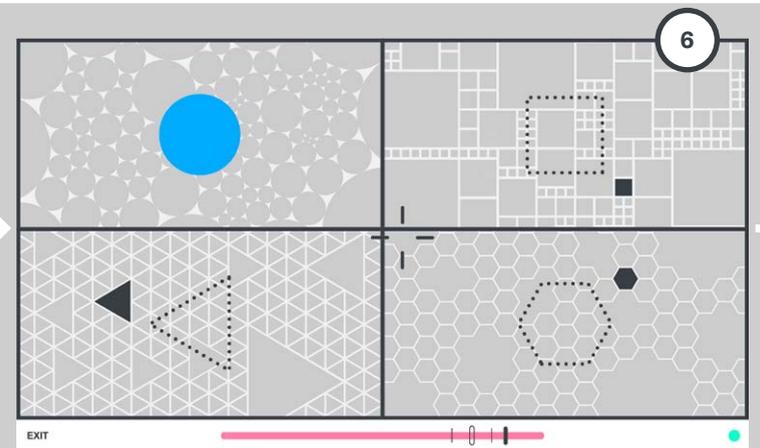
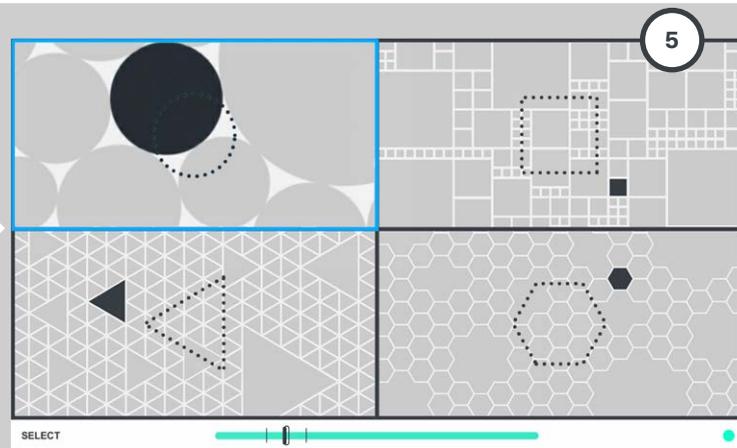
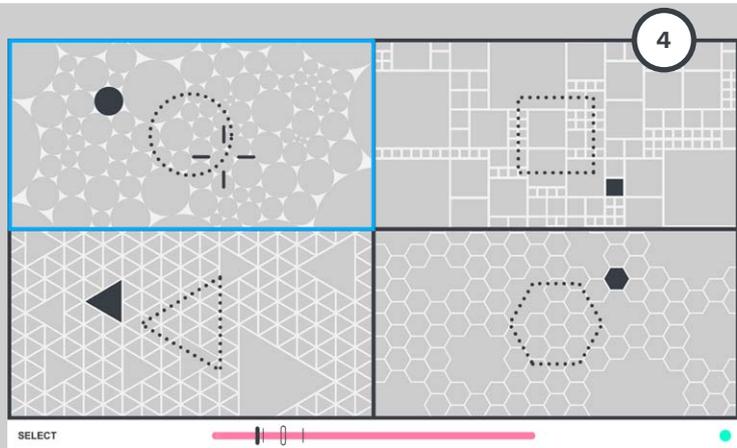
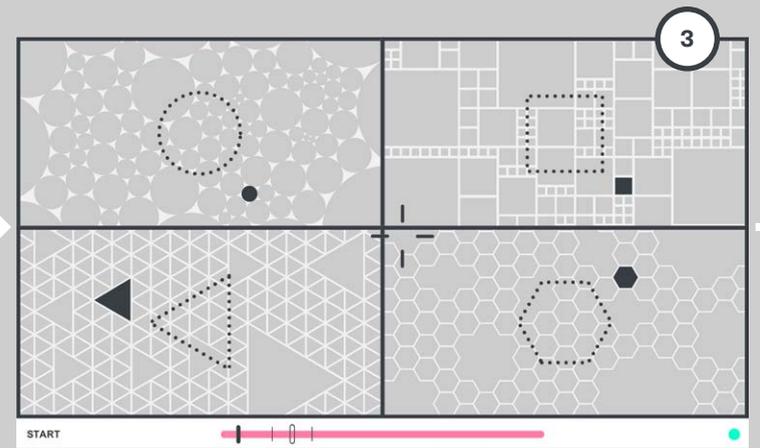
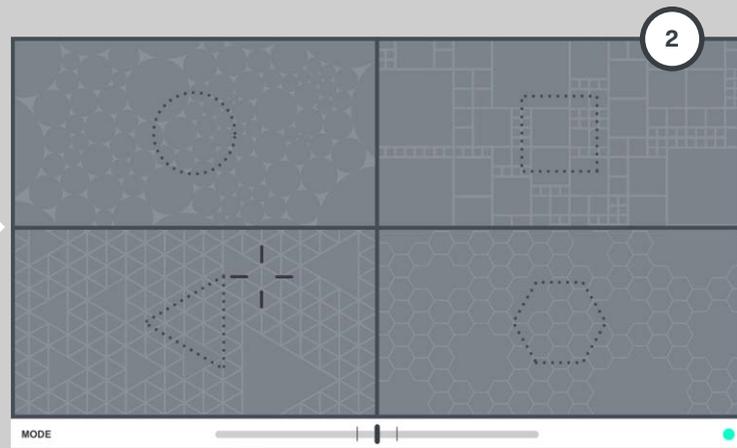
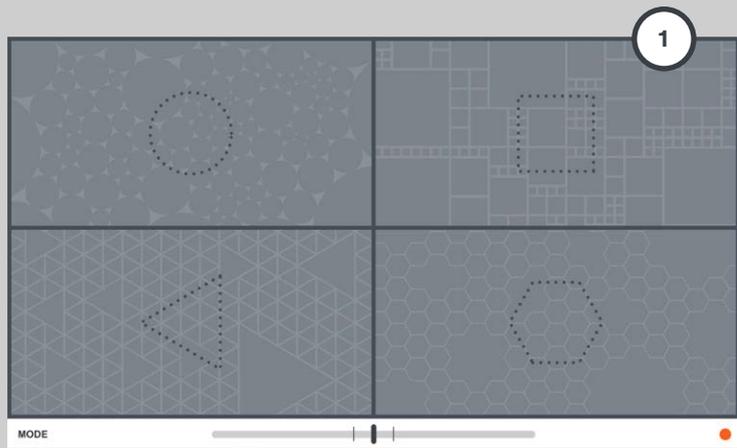
5. Durch die jeweilige Manipulationstechnik (Vorlehnen oder Augenbrauen) vergrößert der Nutzer die sich im jeweiligen Segment befindende Textur, mit dem Ziel die jeweilige dunkle geometrische Form auf Größe und Position des Suchers (gestrichelte Kontur) zu bringen. Parallel bewegt der Nutzer den Slider, welcher durch einen Verlaufsfarbwechsel zwischen Grün und Rot die Korrektheit der Eingabe visualisiert.

6. Sobald die Position und Größe übereinstimmen, verschwindet die Kontur und das dunkle Element zeigt durch eine blaue Farbgebung, dass die Aufgabe abgeschlossen ist. Der blaue Rahmen bleibt solange bestehen, bis die Eingabe des Deselektionskommandos „Exit“ erfolgt, welches ebenso im Eingabemodus dargestellt wird. Nach dieser Eingabe erscheint der Cursor wieder.

7. Im Selektionsmodus kann der Nutzer in beliebiger Reihenfolge das nächste Segment anwählen und darin Zoomen.

8. Sobald alle Elemente in den vier Segmenten vergrößert wurden und sich der Nutzer im Auswahlmodus befindet, kann dieser durch einen abschließenden Befehl (Sprache oder Geste) das System bzw. die Aufgabe beenden.

9. Der Eingabemodus zeigt das Wort „Stop“, der obere Bereich wird erneut überblendet und der Slider inaktiv. Solange ein Nutzer erfasst wurde, bleibt der Cursor bestehen. Dies dient dem später erläuterten Studiendesign, weil mehrere Durchläufe gemessen werden sollen.



KAPITEL 5

I<O

Prototyping

Um die entstandenen Entwürfe zur berührungslosen Interaktion zur Steuerung der konzipierten Benutzeroberflächen des vorangegangenen Kapitels zu testen, wurde ein Hard- und Softwareprototyp entwickelt, dessen Umsetzung in diesen Abschnitt erläutert wird. Da grundlegende Interaktionen entwickelt wurden und dabei wesentliche Rückschlüsse auf das Zusammenspiel der verschiedenen Modalitäten gezogen werden sollen, dienen die Prototypen speziell der Evaluation in der im späteren Verlauf dieser Arbeit durchgeführten Nutzerstudie. Es wurde, wie bereits erwähnt, nicht versucht eine reale Schnittstelle in ein Angiographiesystem zu integrieren, sondern exemplarisch für Forschungszwecke ein Aufbau errichtet, welches wesentliche Interaktionen und Gegebenheiten in einer Intervention in abstrakter Form und prototypisch simulieren soll. Zunächst wird der Entwicklungsprozess erläutert und die verwendeten Entwicklungswerkzeuge für einen Soft- und Hardwareprototypen präsentiert. Es folgt eine Beschreibung der Implementierung der entworfenen Mechanismen, sowie die Vorstellung der finalen Prototypen und deren Einsatz in der darauffolgenden Nutzerstudie.

5.1 Softwareprototyp

Der Schwerpunkt des praktischen Teils dieser Arbeit liegt in der prototypischen Umsetzung der berührungslosen und multimodaler Interaktionen zum Zweck der beschriebenen Sekundäraufgabe. Die entwickelte Benutzeroberfläche wird folgend als Softwareprototyp bezeichnet.

Aufgrund des großen Funktionsumfangs erfolgte die Implementierung des Softwareprototypen in Unity 3D 2018.2. Bei der Unity Game Engine handelt es sich um eine umfangreiche Entwicklungs- und Laufzeitumgebung, die vordergründig bei Entwicklung von Computerspielen und interaktiver 3D-Grafikanwendungen zum Einsatz kommt [74]. Multiple Ressourcen, wie Audio, Video, 3D-Modelle und Grafiken können eingebunden werden, um, wie auch in dieser Arbeit, komplexe Benutzeroberflächen umzusetzen. Die Verknüpfung dieser Ressourcen und Erstellung interaktiver Anwendungen erfolgt durch die Einbindung von Skripten. Die Umsetzung des Prototyps erfolgte in der Programmiersprache C# mithilfe der Entwicklungsumgebung Microsoft Visual Studio 2017 [75]. Die Sprach- und Gestensteuerung wurde über das mitgelieferte SDK (Software Development Kit) der Microsoft Kinect v2 realisiert.

Im Zusammenspiel mit der Konzeptphase, wurde auch der Prototyp iterativ entwickelt. Im ersten Schritt wurden verschiedene Bibliotheken getestet und grundlegende Interaktionen implementiert. Die Abbildung 29 zeigt den ersten Entwicklungsschritt, bei dem zuerst die Manipulation durch das Vor- und Zurücklehnen des Oberkörpers integriert wurde, weil vermutet wurde, dass dessen Interaktion am zuverlässigsten funktioniert. Reale radiologische 2D- und 3D-Bilddaten fanden Verwendung, um einen Eindruck im Einsatz in einer realen Software zu gewinnen. Verschiedene Bildanordnungen, sowie die Zuverlässigkeit der Spracheingabe wurden

erprobt und auch im finalen Prototyp verwendete Sprachbefehle, wie „Scale“, was zum Vergrößern eines einzelnen Quadranten diente, eingesetzt. Auf die Darstellung der Zoomstufe wurde im finalen Prototyp ebenso verzichtet, weil Vorexperimente ergaben, dass dies keinen Vorteil bei der Exploration von Bild-daten bietet. Die Einbettung des finalen Prototyps ist in Abbildung 28 zu sehen. Die in der Konzeption gestalteten Grafiken wurde als Texturen und sogenannten UI-Elemente in einem 3D-Raum übereinander angeordnet. Je nach Eingabe werden diese ein oder ausgeblendet bzw. in ihrer Größe und Position verändert.

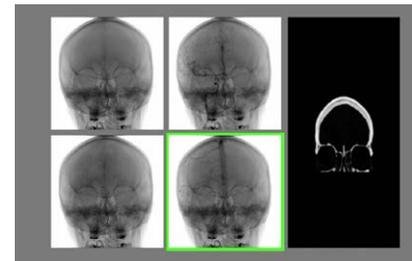


ABBILDUNG 27: Screenshot der Testversion mit medizinischen Bilddaten.

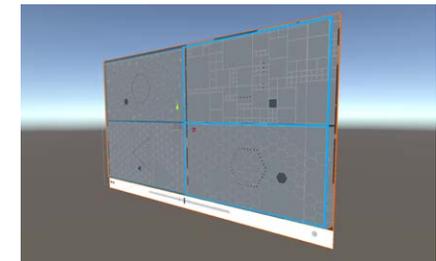


ABBILDUNG 28: Screenshot aus der Unity-Entwicklungsumgebung.

KOPF- UND BLICKRICHTUNG

Aus technischer Sicht gestaltet sich das Erfassen des Blicks als schwierig, da die Fixation eines Objekts in Realität sehr unruhig ist. Sogenannte Sakkaden, also schnelle, ruckartige Augenbewegungen, lassen den Blick regelrecht zittern. Des Weiteren gestaltet sich das Eyetracking aufgrund der relativ großen Entfernung zwischen Monitor und Betrachter als schwierig, zumindest wenn keine am Körper angebrachte Hardware getragen werden soll. Aus diesem Grund wurde sich dafür entschieden die Kopfrichtung zu erfassen. Ein weiterer Vorteil ist zudem, dass der Nutzer seinen Blick während der Interaktion auf andere Elemente, wie Systeminformationen oder das Ausführen einer Tätigkeit mit den Händen konzentrieren kann. Erfasst wurde zunächst das Gesicht des Betrachters. Der Infrarotsensor der Kinect liefert dafür Kamerakoordinaten, um 3D-Punkte im Raum zu finden. Darüber hinaus bietet die Kinect neben Farb- und Tiefenkoordinaten, ein integriertes Skelett-Tracking zur Erfassung des menschlichen Körpers durch den Zugriff auf 25 Joints (Gelenkverbindungen). Durch Orientierungskoordinaten lässt sich dabei zusätzlich bestimmen, in welchem Winkel die Joints ausgerichtet sind. Zur Interpretation der Mimik kann mittels der verwendeten Bibliothek auf insgesamt 35 Joints im Gesicht zugegriffen werden, welche wie später beschrieben zu Umsetzung der Interaktion mit den Augenbrauen ihren Einsatz fanden. Die Umsetzung zur Erfassung der Kopfrichtung wird in Abbildung 29 schematisch dargestellt.

Durch Rotation des Gesichts kann der Nutzer einen Cursor bewegen und diese Funktion auch zum Verschieben (sog. Panning) in Bildern nutzen. Die empfangenen Werte zum Tracking der aktuellen Kopfposition seitens der Kinect weisen geringe Toleranzen auf. Damit eine ruhige Bewegung zu sehen ist, wurden die einzelnen Messungen mit dem Verfahren der einfachen exponentiellen Glättung korrigiert. Nach diesem Prinzip wird eine gewisse Anzahl an zurückliegenden Werten mit dem aktuellen Wert gewichtet und aufaddiert, mit dem Ziel eine Vorhersage über den folgenden Wert zu treffen und damit ein Rauschen aus den Daten herauszufiltern. Eine kleine Verzögerung, die den Bewegungen einer Computermaus angelehnt ist, wurde implementiert, um der Nachvollziehbarkeit der erfassten Kopfbewegung ein natürliches Verhalten zu geben. Der Mittelpunkt des Cursors auf dem Bildschirm konnte individuell durch Verschiebung der Orientierungskoordinaten verschoben werden, sodass Rücksicht auf die Körpergröße des Nutzers genommen werden konnte und damit eine entspannte Haltung gewährleistet ist.

KOPFGESTEN

Ein Gestenbaukasten ist durch die SDK gegeben, konnte jedoch in dieser Arbeit nicht verwendet werden, weil nur rudimentäre Interaktionen mit den Extremitäten gegeben sind. Zur Implementierung der Kopfgesten wurden anhand der Orientierungskoordinaten Richtungsvektoren erstellt. Üblicherweise wird gemessen wie lange eine Geste in einem bestimmten Intervall ausgeführt wird. Trotz der Trägheit des Cursors hat dies den Nachteil, dass die Selektion nicht präzise ausgeführt werden kann, weil durch die gemeinsame Verwendung der Koordinaten der Auswahlbereich ständig verschoben werden würde. Deshalb wurde ermittelt, wie schnell sich der Vektor in einem Intervall ändert. Dafür wurde in verschiedenen Status (bspw. Kopf hoch – runter – hoch – runter – hoch – runter) überprüft, ob es eine Veränderung in der Position gab, womit jeweils eine Geste für Kopfnicken und -schütteln realisiert werden konnte. Somit wurde die physische Belastung möglichst geringgehalten, weil selbst minimale Bewegungen des Kopfes erfasst werden konnten.

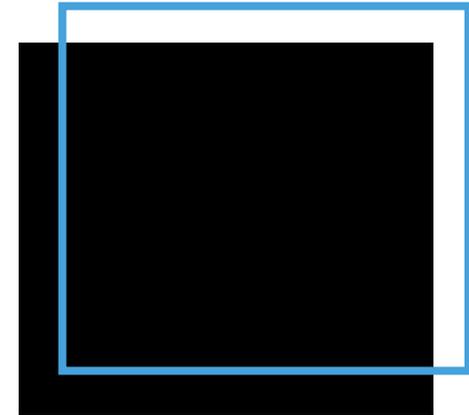


ABBILDUNG 29: Diese Abbildung zeigt die verschiedenen Kamera- bzw. Orientierungskoordinaten der Kinect (x , y , z), woraus die Rotationswinkel ($r1$ und $r2$) zur Bestimmung der Gesten errechnet werden konnte. Die z -Achse kann als Zeiger zur Interpretation der Blickrichtung verstanden werden.

5.1.5 LOG-FUNKTIONEN & INDIVIDUELLE PARAMETER

In Softwareschnittstellen wird in der Regel das Mitschreiben von Anwendungs- und Anwenderdaten eingesetzt, um Funktionen zu verbessern, Probleme in der Nutzung ausfindig zu machen oder Erweiterungen zu entwickeln. Auch für diese Anwendung ist die Datenanalyse essenziell, um Aussagen über das Nutzerverhalten zu treffen. Somit wurde ein Datenlogger integriert, der die Darstellung von Befehlseingaben und Zeitausgaben ermöglicht. In einer Log-Datei werden die Werte gespeichert, um sie zur späteren Evaluation in Tabellen zu überführen. Wichtigster Aspekt ist dabei die Zeiterfassung: Zu dem Zeitpunkt, wenn der Nutzer mit der Bearbeitung der Sekundäraufgabe durch initiale Ausführung der jeweiligen Interaktion beginnt, zählt eine Stoppuhr, bis der Nutzer durch den jeweiligen Befehl die Aufgabe beendet. Des Weiteren wurde die Fehleingabe des Nutzers erfasst, also die Zeit wie lange die Eingabe außerhalb des Toleranzbereichs war (s. Abbildung 32). Damit das System auf die Probanden und die jeweilige zu testende Interaktion eingestellt werden konnte, wurde die Einstellung verschiedener Parameter ermöglicht (s. Abbildung 31).

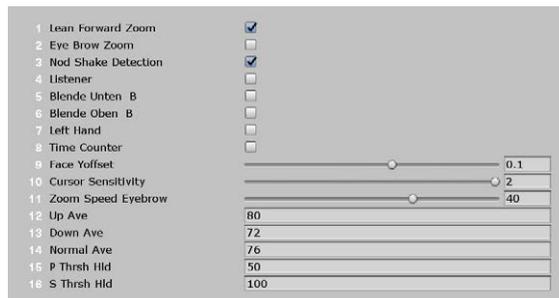


ABBILDUNG 31: Einstellbare Parameter der individuellen Einrichtung des Softwareprototypen. Auswahl der Interaktionstechniken (1 – 4), Ausblenden des oberen und unteren Sichtfelds zur separaten Evaluation von Primär- und Sekundäraufgabe (5 & 6), Linkshänder Modus (7), Stoppuhr (8), Verschiebung des Cursors und dessen Sensibilität (9 & 10), Geschwindigkeit des Zooms mit den Augenbrauen (11), gemessene Werte der Kalibrierung (12 -14), Schwellenwerte zum Erreichen der Zielposition und Größe (15 & 16).

5.2 Hardwareprototyp

5.2.1 UMSETZUNG

Um eine den Anforderungen optimale und zugleich in der Benutzung simple Hardwarechnittstelle zu kreieren, erfolgte die Umsetzung auf Basis der Physical-Computing-Plattform Arduino und dessen dazugehöriger Entwicklungsumgebung. Dabei kam der Mikrocontroller Arduino Uno R3 und ein Time-of-Flight Sensor zum Einsatz. Der Sensor sendet Lichtimpulse im Infrarotbereich und ermittelt die Distanz über die Dauer der Reflektion zum nächsten Objekt. Es werden Werte zwischen 30 – 2000 mm mit Auflösung von 1 mm vom Arduino interpretiert und via USB an Unity übermittelt. Da geringe hardwarebedingte Abweichungen mit einem Toleranzbereich von 3% gegeben sind, werden die empfangenen Daten über das Verfahren der exponentiellen Glättung, also dem selben Algorithmus, der zum Glätten der Werte der Kinect verwendet wurde, korrigiert. Der Sensor sitzt dabei in einem Rohr und misst die Distanz zum vom Nutzer geführten Stab (s. Abbildung 33). Im unteren Bereich des Interfaces befindet sich ein Schieber, auf dem die Nutzereingabe 1:1 übertragen wird. Der Nutzer bewegt vertikale Linie in horizontaler Richtung dessen maximale Skala der der Eingabe entspricht. Ziel ist es, diese Linie in einem definierten Toleranzbereich zu halten. So lange sich die Linie im Bereich befindet, erscheint die Anzeige grün. Wird dieser verlassen, so färbt sie sich rot (s. Abbildung 32). Dabei bewegt sich dieses Element in kleinen Schritten in eine zufällige Richtung. In einem Zeitintervall von 10 Sekunden verändert sich der Bereich dabei vollständig und ebenso randomisiert seine Position über die Gesamtbreite. Alle Komponenten wurden in ein standfestes Gehäuse mithilfe einer 3D-Druck-Konstruktion zusammengeführt. Abbildung 34 zeigt das Ergebnis.



ABBILDUNG 32: Schieberegler zur Übertragung und Rückmeldung der Nutzereingabe. Der dunkle fette Strich gibt die Richtung wieder. Befindet sich dieser innerhalb des Toleranzbereichs (dünne Linien) ist die Anzeige grün gefärbt (unten). Außerhalb dessen färbt diese sich schrittweise rot.

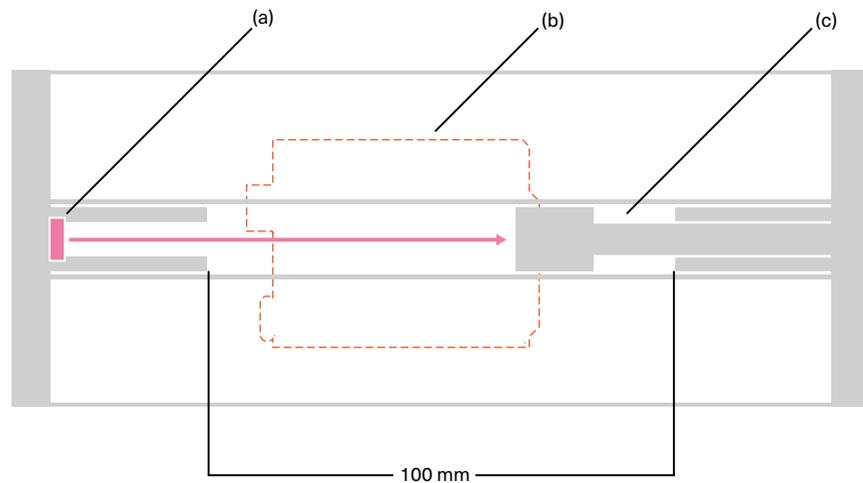


ABBILDUNG 33: Darstellung zeigt den Aufbau des Hardware-Prototypen. Links in einer Röhre befindet sich der Distanzsensor (a). Dieser misst die Distanz bis zum Ende des vom Nutzer geführten Stabs (c). Für den Nutzer nicht sichtbar ist am Boden ist der Mikrocontroller installiert (b).

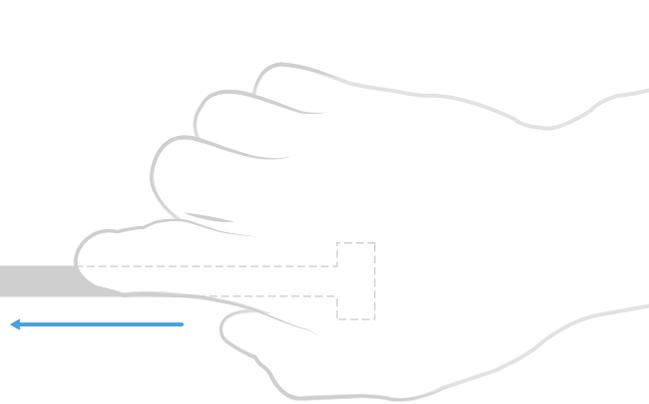
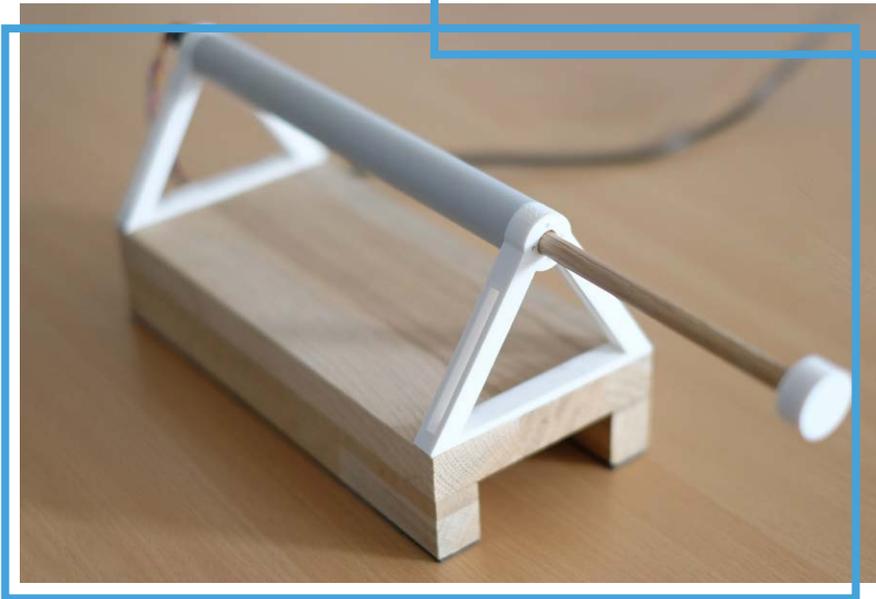


ABBILDUNG 34: Konstruktion aus 3D-Druck-Elementen, Holz und Kunststoffrohren



5.3 Setup

Im letzten Schritt der Entwicklung folgte das Zusammenbringen des Hard- und Software-Prototypen. Vor Beginn der Evaluation wurden verschiedene Situationen und räumlichen Gegebenheiten im Labor getestet (s. Abbildung 35). Störelemente wie Licht oder Geräusche galt es zu identifizieren und die Einrichtung danach auszurichten. Die Installation ist zunächst nur darauf ausgerichtet in normalen räumlichen Laborbedingungen bedient zu werden. Es besteht weiterhin die Möglichkeit diese an einem Angiographiesystem zu implementieren, doch aus Zeitgründen und da es elementare Interaktionen zu testen galt, wurde darauf verzichtet. Eine schematische Darstellung des prototypischen Aufbaus ist in Abbildung 36 gegeben bei vollständig und ebenso randomisiert seine Position über die Gesamtbreite. Alle Komponenten wurden in ein standfestes Gehäuse mithilfe einer 3D-Druck-Konstruktion zusammengeführt. Abbildung 34 zeigt das Ergebnis.



ABBILDUNG 35: Installation der Prototypen in der Testphase.

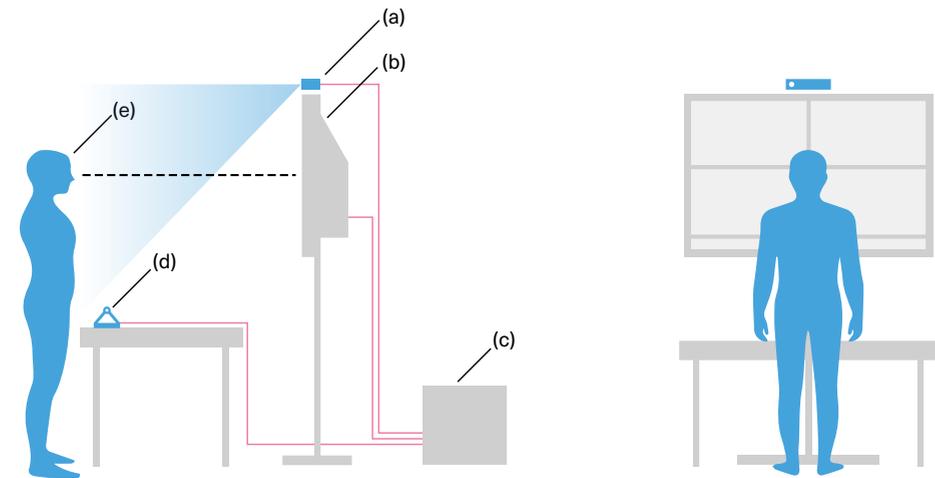


ABBILDUNG 36: Finaler Aufbau zur Durchführung der Laborstudie. Im Winkel verstellbare Microsoft Kinect (a), höhenverstellbarer Monitor (b), Computer (c), Hardwareprototyp, dessen Höhe angepasst werden kann (d), Versuchsperson (e).

Evaluation

KAPITEL 6

100

Um die entstandenen Interaktionskonzepte zu testen, wird das entwickelte prototypische System im Rahmen einer Nutzerstudie unter Laborbedingungen evaluiert. Die Beurteilung der gesammelten Ergebnisse bildet den wichtigsten Abschnitt dieser Arbeit. Dabei sollen Stärken und Schwächen der implementierten Modalitäten und deren Zusammenspiel genau identifiziert, sowie Optimierungsmöglichkeiten dargestellt werden.

6.1 Fragestellungen

Die Ergebnisse der Nutzerstudie sollen dabei helfen, Aussagen zu formulieren, inwieweit die entwickelten Interaktionstechniken zur Selektion und Manipulation von medizinischen Bilddaten während einer Intervention anwendbar sind. Anhand der erstellten Anforderungen in Kapitel 3 ergeben sich folgende Fragestellungen:

KLINISCHE ANFORDERUNGEN

Können die Konzepte den gegebenen klinischen Anforderungen gerecht werden?

KONTROLLE

Gewährleistet die berührungslose Systembedienung hinreichend Sicherheit in der Durchführung der Interaktionstechniken?

PRÄZISION

Ist die Durchführung von Selektion und Manipulation präzise genug?

BELASTUNG

Wie hoch ist die kognitive und physische Belastung?

DOPPELAUFGABEN

Welchen Einfluss hat das parallele Ausführen der Primäraufgabe auf die Sekundäraufgabe und umgekehrt?

BENUTZERFREUNDLICHKEIT

Sind die Interaktionstechniken intuitiv, vertraut und verständlich?

LEISTUNG

Wie viel praktische Übung ist notwendig, um die Interaktionen zu erlernen?

UMGEBUNGSEINFLÜSSE

Beeinflusst, und wenn ja wie stark, die Umgebung das System und die Interaktionen?

FEEDBACK

Ist genügend Feedback seitens des Systems vorhanden oder benötigt es Erweiterungen?

6.2 Planung

6.2.1 _____ VERSUCHSPERSON

Da die Aufgaben abstrakt gehalten sind und aus der Kombination der verschiedenen Modalitäten universelle Aussagen formuliert werden sollen, sind keine Fachkenntnisse nötig. Zu diesem Zweck wurden Probanden ohne medizinische Vorkenntnisse ausgewählt.

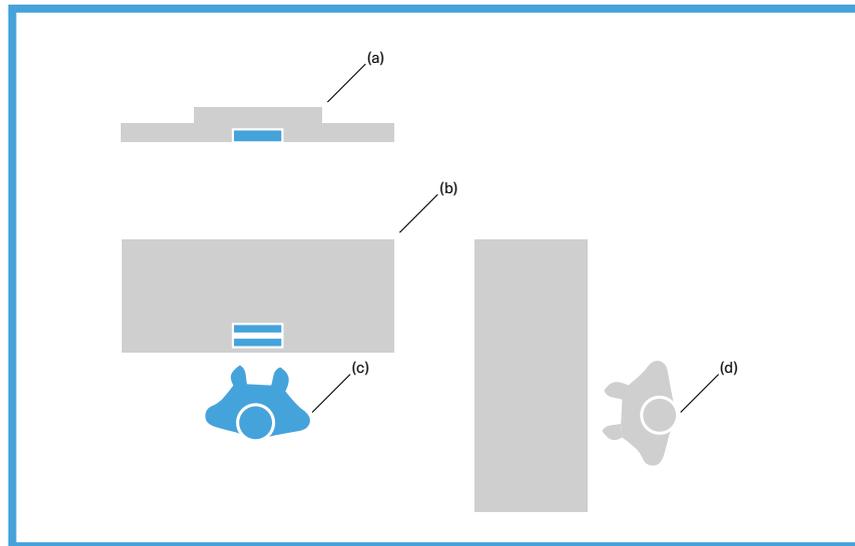


ABBILDUNG 37: Darstellung der Versuchsumgebung. (a) Monitor inkl. Kinect, (b) Tisch auf dem sich der Hardwareprototyp befindet, (c) Versuchsperson, (d) Versuchsleiter

6.2.2 _____ VERSUCHSUMGEBUNG

Die Durchführung des Experiments erfolgte in einem abgelegenen Raum. In diesem Raum befanden sich ausschließlich je eine Versuchsperson und der Versuchsleiter. Die Hardware umfasste einen Computer, einen 75" Monitor, einen Tisch, auf dem sich der Hardwareprototyp zur Durchführung Primäraufgabe befand, sowie eine oberhalb des Monitors installierte Microsoft Kinect. Die Versuchspersonen hatten exakt 180 cm Abstand, gemessen von Tischkante zum Display. Tisch und Monitor waren höhenverstellbar, sodass es pro Versuchsperson einen individuellen Aufbau gab. Schematisch ist dieser Aufbau in Abbildung 37 dargestellt. Einen Eindruck vom realen Aufbau lässt sich in Abbildung 38 gewinnen.

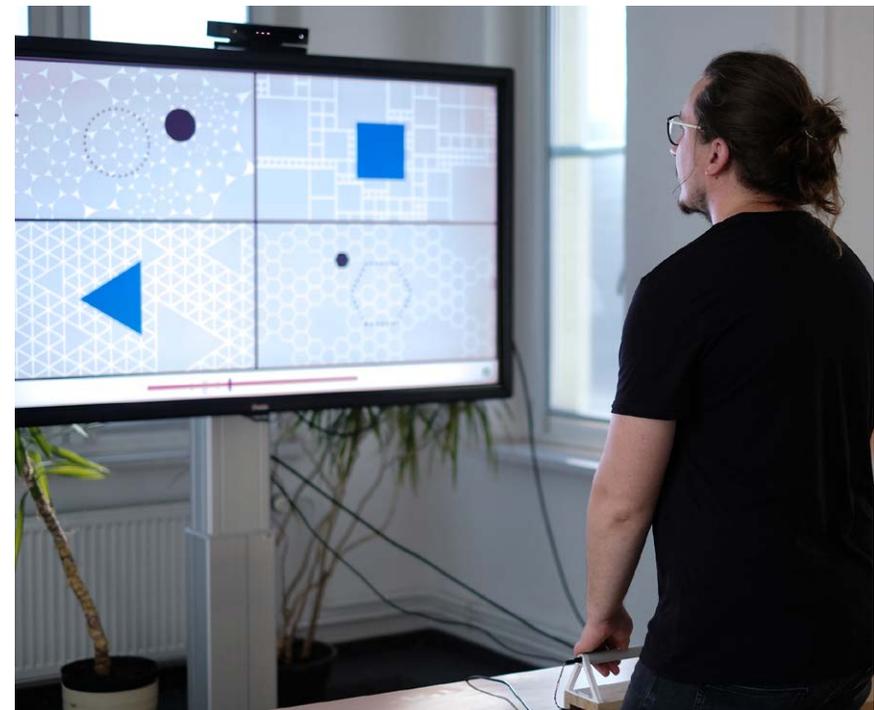


ABBILDUNG 38: Fotografie vom Aufbau während der Nutzerstudie.

6.2.3 STUDIENDESIGN

Ein ausgewogenes Zusammenspiel von Sekundär- und Primäraufgabe zur Durchführung der Operation ist essenziell, daher gilt es herauszufinden, welchen Einfluss die jeweilige Modalitätskombination auf den Workflow hat. Aus diesem Grund wurden die Probanden in zwei Gruppen unterteilt (s. Abbildung 39): Versuchsgruppe A testete das System mit Primäraufgabe → Versuchsgruppe B ohne, weshalb von einem between-groups-Design gesprochen wird. Jedoch wird auch unabhängig der Versuchsgruppen analysiert, welche Interaktionstechniken und Kombinationen sich generell eignen.

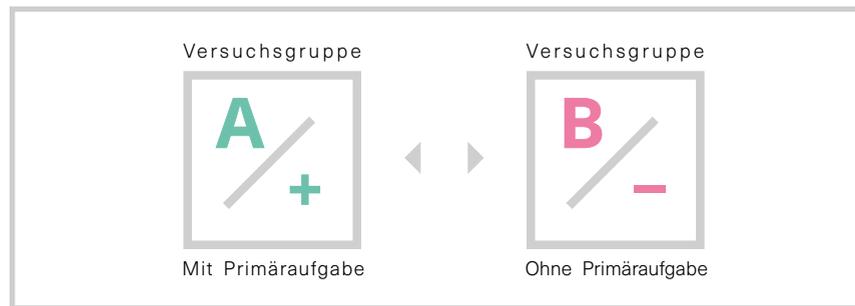


ABBILDUNG 39: Darstellung des Vergleichs beider Versuchsgruppen hinsichtlich des Einflusses der Primäraufgabe.



ABBILDUNG 40: Abkürzungen und Farbuordnung der Kombinationen. (v.l.n.r.) Sprache und Lehnen, Gesten und Lehnen, Sprache und Augenbrauen, Gesten und Augenbrauen.

Jede Versuchsperson durchläuft dabei jedoch alle vier möglichen Modalitätskombinationen zur Erfüllung der Sekundäraufgabe. Diese Kombinationen werden fortan als Abkürzung und dazugehöriger farblicher Kennzeichnung geführt, welche in Abbildung 40 aufgeschlüsselt sind. Die Zuordnung der Kombinationen und Reihenfolge in der Durchführung erfolgte randomisiert (s. Abbildung 41).

	A	B	Kombination			
Probanden in Versuchsreihe	1	2	S _L	S _E	G _L	G _E
	3	4	S _E	S _L	G _E	G _L
	5	6	G _L	G _E	S _L	S _E
	7	8	G _E	G _L	S _E	S _L

ABBILDUNG 41: Zeigt die Zuordnung der Kombinationen unter allen Probanden.

Nach Begrüßung des Probanden erfolgte die persönliche Vorstellung des Untersuchungsleiters und die Übergabe eines Dokuments, in dem allgemeine Informationen zum generellen Ablauf, Ziel der Studie, Forschungsvorhaben, Einwilligung und Rücktritt sowie und Datenschutz erläutert wurden (Anhang A3). Während die randomisierte Zuweisung und Vorbereitung der Versuchsreihe erstellt wurde, erfolgte anschließend die Erfassung demographischer Daten (Anhang A4).

Bestanden keine Fragen seitens des Probanden, so begann die Vorstellung des Prototyps und die Instruktion der Aufgaben. Es wurde erläutert, dass die Situation einer radiologischen Intervention simuliert wurde, indem kurz über die Tätigkeiten des Radiologen informiert wurde. Dann erfolgte die Erläuterung der Benutzeroberfläche, sowie die verschiedenen Modalitäten zur Erfüllung der Sekundäraufgabe, dessen Erfüllung so schnell wie möglich erfolgen sollte. Befand sich die Versuchsperson in Versuchsgruppe A (mit Primäraufgabe), so wurde die Primäraufgabe am Hardwareprototypen als adaptierte Interaktion der Katheterführung erklärt, dessen Ausführung so genau wie möglich durchgeführt werden soll. Versuchsgruppe B wurde die Erklärung vorenthalten. Vor Beginn der Durchführung wurde das System auf die Körpergröße der Person angepasst: Monitorhöhe, Richtung (links/rechts) und Höhe Hardwareprototyps, Kalibrierung der Cursorhöhe sowie Messung der Augenbrauenposition.

Aufgrund der implementierten Bewegungen des Toleranzbereichs, um eventuelle Abweichungen bestimmen zu können und um ein Gefühl für die Interaktion mit dem Hardwareprototypen zu bekommen, wurde bei Versuchsgruppe A ermittelt, wie lange sie sich im Toleranzbereich halten können. Der obere Bereich des Interfaces wurde dafür ausgeblendet, womit nur der Schieberegler sichtbar war. Die Messung der Abweichung (Baseline) erfolgte in einer Zeit von 1:30 Minute, da aus dem Non-Experiment hervorging, dass dies die durchschnittliche Zeit zur Erledigung der Aufgaben ist. Der Schieberegler wurde bei Versuchsgruppe B vollständig ausgeblendet, um keine zusätzliche Ablenkung zu schaffen.

Im ersten Schritt vor der Durchführung wurde die Kontrolle des Cursors mithilfe der Kopfrichtung geübt. Die zugewiesenen Kombinationen wurden in Form von Kärtchen zur Orientierung im Sichtfeld des Probanden platziert (s. Abbildung 42). Anschließend erfolgten 2 – 3 Trainingsdurchläufe pro Kombination/Interaktionstechnik. Fühle sich der Proband auf Nachfrage des Untersuchungsleiters bereit,

so startete der erste Durchlauf und es wurde die Zeit gemessen. Nach exakt drei Messungen wurden die Daten (Zeit & Auszeit) protokolliert und die Versuchsperson erhielt ein Formular zur subjektiven Selbsteinschätzung. Nach einer kurzen Pause und situationsbedingter Neukalibrierung des Systems wurden die fehlenden drei Kombinationen nach gleichem Ablauf durchgeführt. Abschließend fand eine unstrukturierte Nachbefragung statt, in der das allgemeine Empfinden während der Durchführung des Versuchs abgefragt wurde und die Möglichkeit bestand, Feedback zu den jeweiligen Techniken zu geben. Das Vorgehen in Form eines Ablaufplans ist dem Anhang A2 zu entnehmen.

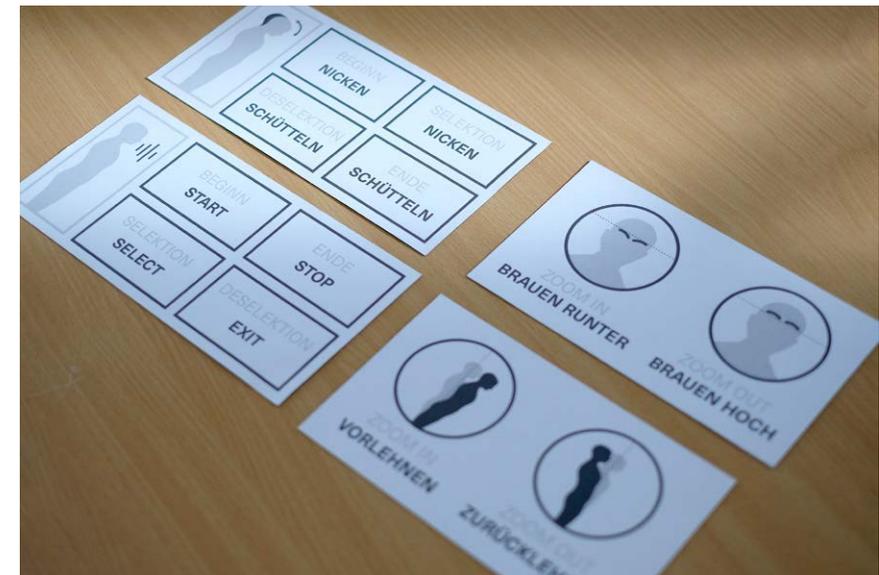


ABBILDUNG 42: Informationskarten im Sichtfeld der Versuchsperson mit aktuell zur Verfügung stehender Selektions- und Manipulationstechniken.

Anhand des in Kapitel 4 beschriebenen Studienkonzepts und den Anforderungen des Systems aus Kapitel 3 werden folgend klare Kriterien definiert, die eine konkrete Bewertung des Prototyps ermöglichen. Dabei steht die Messung der Usability im Vordergrund. Sie ist nach ISO 9241-11 definiert als „Das Verhältnis zur Genauigkeit und Vollständigkeit eingesetzte Aufwand, mit dem Benutzer ein bestimmtes Ziel erreichen“. Dies bedeutet, dass der Nutzer anhand der Funktionalität, als Attribut der Interaktion, die im Kontext gestellte Aufgabe effizient, effektiv und zufriedenstellend lösen kann. Eine objektive Datenerhebung sowie das Messen der subjektiven Beanspruchung gewährleisten dabei, eine gezielte Analyse und Bewertung der Usability.

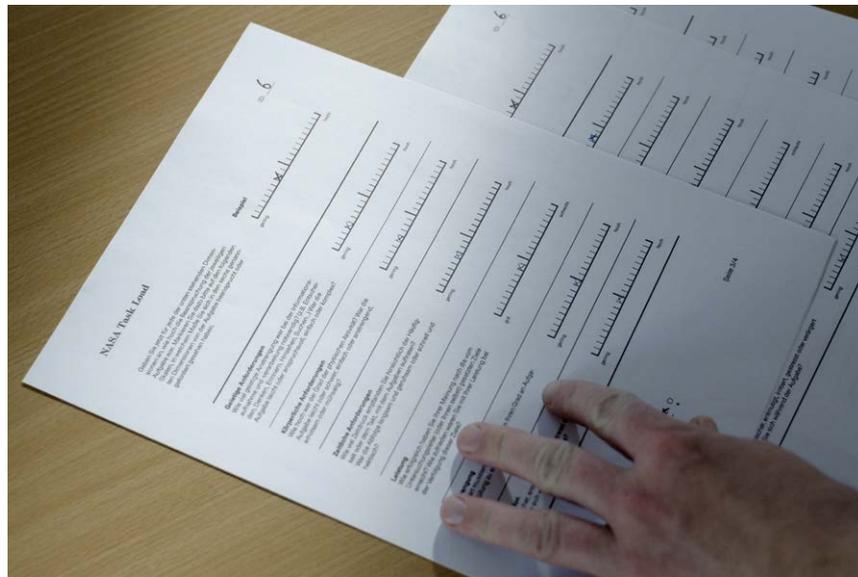


ABBILDUNG 43: Zu sehen sind die vier NasaTLX-Fragebogen, die je Proband Verwendung fanden. Unter einer kurzen Erklärung und einem Beispiel befindet sich die Dimensionen und der Erklärung: Geistige Anforderungen, Körperliche Anforderungen, Zeitliche Anforderungen, Leistung, Anstrengung und Frustrationsniveau. Die Skalen umfassen 20 Stufen, die mit jeweils 5 Punkten bewertet werden können.

GENAUIGKEIT

Die Effektivität ist, laut oben beschriebener ISO-Norm, die Genauigkeit und Vollständigkeit, mit dem der Nutzer ein bestimmtes Ziel erreicht. Die Vollständigkeit wird in der Studie gemessen, in dem die Sekundäraufgabe vollständig abgeschlossen ist. Dafür muss, wie in Kapitel Fehler! Verweisquelle konnte nicht gefunden werden. beschrieben, jede Zoominteraktion innerhalb der Quadranten abgeschlossen sein. In Betrachtung der Genauigkeit wird der Primäraufgabe der größten Relevanz zugeschrieben. Denn auch in der echten Intervention sollte das Ziel sein, so genau wie möglich zu arbeiten und damit dem Patienten am bestmöglichen zu versorgen. Aus diesem Grund wird die im Studienablauf beschriebene Abweichung gemessen, in der sich ein Nutzer außerhalb des Toleranzbereichs befand.

BEARBEITUNGSZEIT

Die Effizienz steht im Verhältnis der Effektivität und kann bestimmt werden, indem die Zeit gemessen wird, die ein Nutzer braucht, um eine Aufgabe zu erledigen. In der Studie wird die Dauer ermittelt, wie lange der Nutzer benötigt, um die Sekundäraufgabe zu erfüllen. Dabei wird ebenso verglichen, wieviel Zeit die jeweilige Interaktionstechnik und die Kombination der verschiedenen Modalitäten in Anspruch nimmt.

BEANSPRUCHUNG

Um die subjektive Schwierigkeitsempfinden zu messen, wurde die mentale, physiologische und zeitliche Beanspruchung während der Bearbeitung der Aufgaben durch den mehrdimensionalen NASA-TLX (Task Load Index) Fragebogen erfasst [76] (Anhang A5). Die Erhebung nach NASA-TLX ist standardisiert, hoch etabliert und die erfassten Daten gelten als valide. Sie eignen sich besonders im Usability Test in der Kontextanalyse und des Testphasen [77]. Neben der Erfassung der kognitiven und physischen Beanspruchung lassen sich mit diesem Fragebogen (s. Abbildung 43). Ein weiterer Vorteil des Fragebogens ist zudem, dass dieser zwischen 0 und 100 liegt und eine prozentuale Skala ermöglicht. Jedoch wird aus Gründen der besseren Visualisierung von Differenzen die Dimension 0 bis 20 dargestellt. Im Abschnitt Diskussion werden, die erfassten Daten zudem abgeglichen mit den Aussagen der Nutzer in der Nachbefragung.

6.2.6 NON-EXPERIMENT

In einem Non-Experiment wurde der technische Aufbau, die beschriebenen Interaktionstechniken und der Untersuchungsablauf mit einer Kontrollgruppe, die denen der späteren Versuchsgruppe ähnelt, vorgetestet. Dies war wichtig zur Ermittlung der gesamten Studiendauer, der Zeit, die zu Durchführung einer Aufgabe benötigt wird (u.a. Bestimmung der Baseline) und dem Identifizieren von Fehlerquellen. Ebenso wurden durchschnittliche Parameter zur individuellen Kalibrierung erfasst, Feinheiten in den Interaktionen abgestimmt und das System dadurch optimiert. Aufgrund der durchgeführten Systemanpassungen sind diese Daten nicht aussagekräftig und fließen somit nicht in die Gesamtbeurteilung ein. Die Rohdaten dieses Vorexperiments sind im Anhang A8 zu entnehmen.

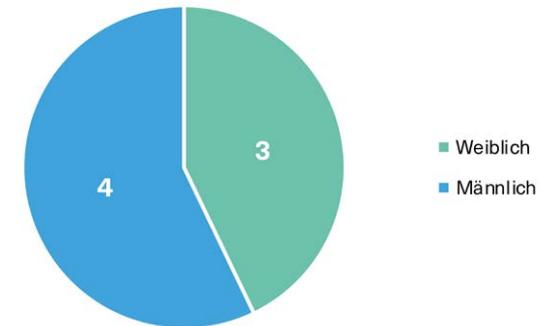


ABBILDUNG 44: Zusammensetzung aller Probanden nach Geschlecht.

6.3 Durchführung

6.3.1 ZUSAMMENSETZUNG DER PROBANDEN

Insgesamt umfasste die Nutzerstudie acht Versuchspersonen, welche zu jeweils vier Personen in zwei Versuchsgruppen aufgeteilt wurden. Versuchsgruppe A testet alle Bedingungen des Experiments mit Primäraufgabe, während Versuchsgruppe B dieselben Bedingungen ohne Primäraufgabe durchlief.

Die demographischen Angaben umfassten Alter, Geschlecht und Tätigkeitsstatus. Das Durchschnittsalter aller Probanden betrug 28,1 Jahre. Die Versuchsgruppe A setzte sich aus einer Frau und drei Männern zusammen. In Versuchsgruppe B war das Verhältnis 2:2. Das Spektrum der angegebenen Expertisen, bezogen auf dem Tätigkeitsstatus, reichte von Interaction und Industrie Design, über Informatik und Medizintechnik. Die spezifischen Daten werden in Abbildung 44 und Abbildung 45 dargestellt.

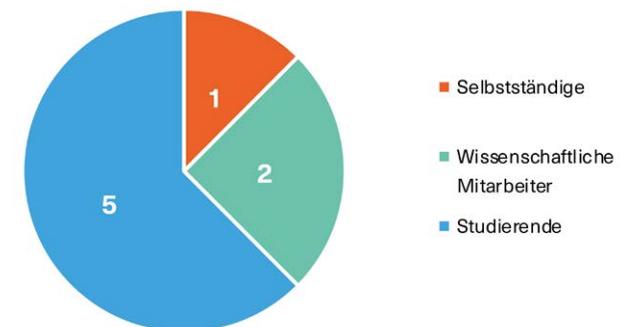


ABBILDUNG 45: Zusammensetzung der Probanden nach Tätigkeitsstatus.

Ausnahmslos gaben die Probanden die rechte Hand als dominant an. Diese Angabe war notwendig, weil das System ggf. darauf eingestellt und das Instrument der Richtung entsprechend ausgerichtet werden musste. Ebenso wurde erhoben, ob eine Sprachstörung vorliegt, um eventuelle Komplikationen mit der Spracheingabe zu prüfen, wobei alle Probanden diese Angabe verneinten.

Darüber hinaus wurde nach einer Einschränkung des Sehvermögens gefragt. Deswegen Aufteilung wird in Abbildung 46 dargestellt. Eingeladen wurden nur Personen, dessen Fehlsichtigkeit nicht zu stark ausgeprägt ist, weil ein Brillengestell teilweise die Interaktion mit den Augenbrauen behindern kann. Durch Non-Experiment und Selbsttest konnte im Vorfeld gesagt werden, dass die Bedienung des Systems bei einer geringen Kurzsichtigkeit uneingeschränkt möglich ist.

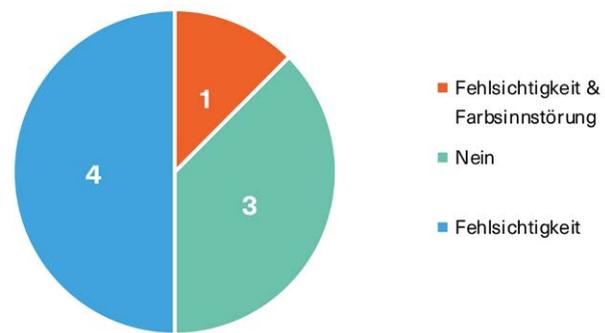


ABBILDUNG 46: Angaben der Versuchspersonen bei Einschränkung des Sehvermögens. Nein bedeutet, dass keine Fehlsichtigkeit oder Farbsinnstörung vorhanden ist.

Weiteren Angaben grenzten die Beherrschung der englischen Sprache, sowie Kenntnisse in den Bereichen Mensch-Computer-Interaktion, Gestensteuerung bzw. Berührungslose Interaktion, Gesichts-, Emotions- oder Eyetracking, sowie Sprachsteuerung ein. Hierbei sollte geprüft werden, ob und wenn ja, welchen Einfluss die jeweiligen Fähigkeiten und Kenntnisse auf das System im Allgemeinen haben. Die Befragung wurde mittels einer Likert-Skala von gar nicht (1) bis sehr gut (5) – bezogen auf Englisch – und keine (1) bis sehr erfahren (5) – bezogen auf die Erfahrungen in den jeweiligen Bereichen – durchgeführt (s. Abbildung 47).

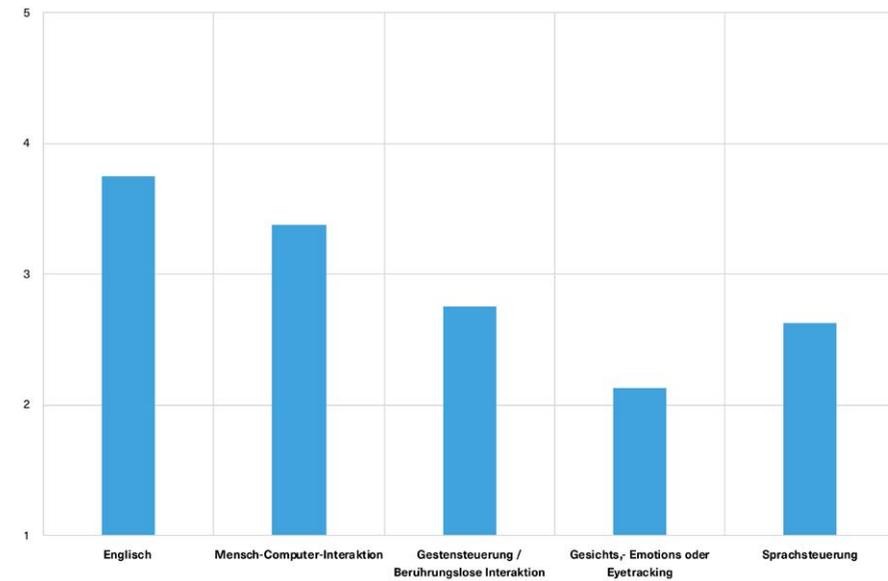


ABBILDUNG 47: Einschätzung der Versuchspersonen von Fähigkeiten der englischen Sprache und Kenntnisse in spezifischen technologischen Bereichen.

Bezogen auf die angegebenen Daten lässt sich kein signifikanter Einfluss von Farb- und oder Fehlsichtigkeit zur Bedienung des Systems erkennen. Ebenso hat das Niveau der englischen Sprache und dementsprechend die Aussprache der Sprachbefehle keinen Einfluss auf die Benutzung des Systems.

Trotz der niedrig eingestellten Spracherkennung wurden teilweise Befehle nicht sofort erkannt, wodurch der Nutzer diese wiederholen musste. Trotzdem die Versuchsumgebung abgelegen war, waren leichte Umgebungsgeräusche, wie entfernte Gespräche und Straßenlärm wahrnehmbar. Ebenso beeinträchtigten kurze Konversationen mit der Versuchsperson teilweise die Eingabe. Beim ersten Kontakt in Durchläufen mit Sprachbefehlen kam es außerdem, trotz vorherigen Übungen und sichtbaren Merktzettel, zu Verwechslungen. Hierbei sei besonders hervorzuheben, dass die Befehle „Stop“ und „Exit“ durcheinandergebracht wurden. Eine Fehlinterpretation der Eingaben seitens des Systems kam aufgrund der wenigen und von der Aussprache bewusst unterschiedlich gewählten Wörter nicht vor.

Im Mittel betrug die Dauer eines Durchlaufs zur Erledigung der Sekundäraufgabe 1:14 Minute. Die durchschnittliche Standardabweichung der zeitlichen Messung in allen Kombinationen betrug in Gruppe A acht Sekunden und in Gruppe B sechs Sekunden (vgl. Abbildung 50). Versuchsgruppe B (ohne Primäraufgabe) wurde pro Durchlauf stetig um 10 % schneller, wohingegen Versuchsgruppe A im letzten Durchlauf exakt 10 % langsamer wurde, wie in Abbildung 48 zu sehen ist. Der Abbildung 49 ist zu entnehmen, dass beide Versuchsgruppen Anfangs Zeit benötigten, um sich an das System zu gewöhnen, jedoch nur bei Versuchsgruppe B ein Lerneffekt ersichtlich wird.

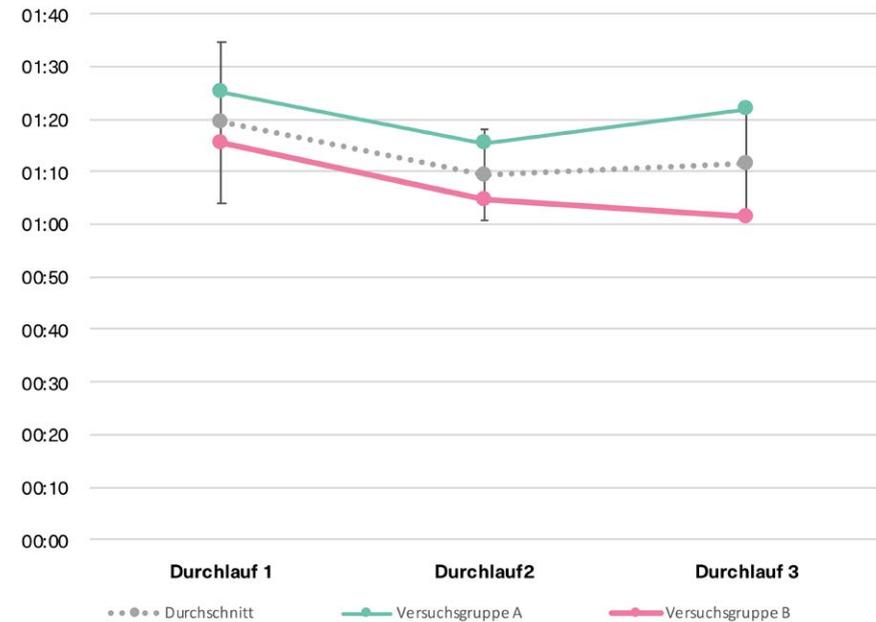


ABBILDUNG 48: Entwicklung der Bearbeitungszeit der drei gemessenen Durchläufe anhand der durchschnittlichen Dauer (Zeit in Min. unter allen Probanden in beiden Versuchsgruppen).

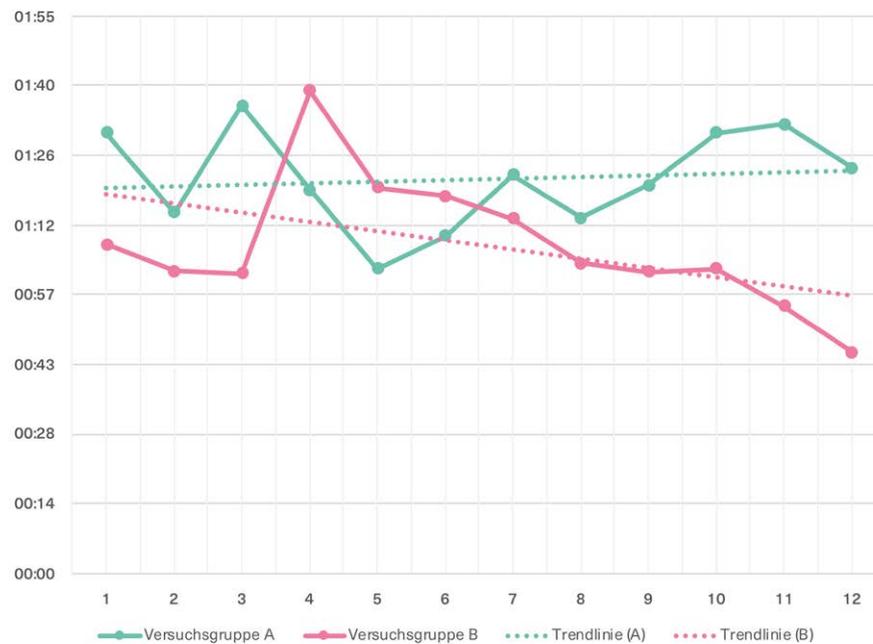


ABBILDUNG 49: Darstellung der erfassten Dauer unter allen Probanden in allen Durchläufen und mit allen Interaktionskombinationen. Y-Achse = Zeit in Minuten. X-Achse = drei Durchläufe mal vier Kombinationen.

Die durchschnittlich gemessene Zeit zur Erledigung der Aufgaben betrug summiert 14:55 Minuten. Der Mittelwert der Versuchsgruppe A lag bei 16:17 Minuten, womit diese 2:44 Minuten langsamer als Versuchsgruppe B mit 13:33 Minuten war. Unterschiede sind ebenso signifikant beim Vergleich beider Versuchsgruppen in Bezug auf die jeweilige Interaktionstechnik, wie in Abbildung 50 zu sehen ist. Es zeigt sich auf beiden Seiten: Unter den Selektionstechniken sind Gesten- und Spracheingaben gleichwertig. In Kollation der Techniken zur Bildmanipulation, ist das Zoomen durch Vorlehnen vorteilhafter, als der Gebrauch der Augenbrauen. Die Kombination GL ist, gemessen an der Zeit, deutlich die effizienteste Interaktionstechnik, wenn keine chirurgische Tätigkeit (Primäraufgabe) ausgeführt wird. Die Interaktion mit den Augenbrauen ist in beiden Gruppen schlecht zu bewerten, wobei die Kombination GE besonders ineffizient erscheint.

Das Hinzuziehen der subjektiven Beurteilung anhand des Fragebogens NASA-TLX zeigt ein ähnliches Bild. Wie in Abbildung 51 zu entnehmen, sind Gesteninteraktionen mit einer höheren Beanspruchung als Sprachbefehle verbunden. Und auch hier schneidet Vorlehnen als Technik zur Bildmanipulation besser ab, als die Interaktion mit den Augenbrauen. Es fällt jedoch auf, dass es nur einen geringen Unterschied zwischen den Kombinationen GL und SL gibt. Insgesamt sind alle Kombinationen nahezu gleichwertig bewertet, wobei auch hier GE eine besonders hohe Beanspruchung zugeschrieben wird.

Die durchschnittliche Standardabweichung in Versuchsgruppe A betrug 1,6 und in Versuchsgruppe B 1,9. Betrachtet man die einzelnen Dimensionen des Fragebogens in Abbildung 52, so wird die erbrachte Leistung insgesamt als positiv eingeschätzt, wobei auch hier die Augenbraueninteraktion negativ auffällt. Signifikante Differenzen innerhalb der Versuchsgruppen sind vor allem in der geistigen Beanspruchung und dem Grad der Frustration erkennbar. So differenziert sich mentale Anforderung bei der Versuchsgruppe mit Primäraufgabe besonders zu Versuchsgruppe B.

Der Einfluss der Sekundäraufgabe auf die Primär wird in Abbildung 53 erkenntlich gemacht. Die gemessene Baseline beträgt 32 Sekunden. Im Schnitt waren die Probanden 31 Sekunden außerhalb des Bereichs. SE beeinträchtigt am meisten die Genauigkeit in der Durchführung einer motorischen Tätigkeit. Das Lehnen (SL + GL) hat den geringsten Einfluss auf die Primäraufgabe. Die erfassten Rohdaten der beiden Versuchsgruppen sind im Anhang A6 und A7 dokumentiert.

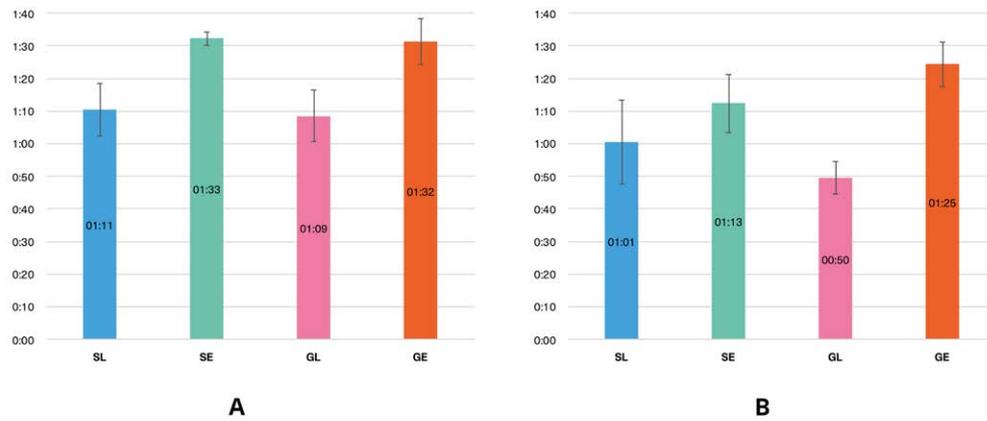


ABBILDUNG 50: Vergleich der durchschnittlich benötigten Zeit pro Kombination. Links Versuchsgruppe mit Primäraufgabe (A), rechts ohne Primäraufgabe (B).

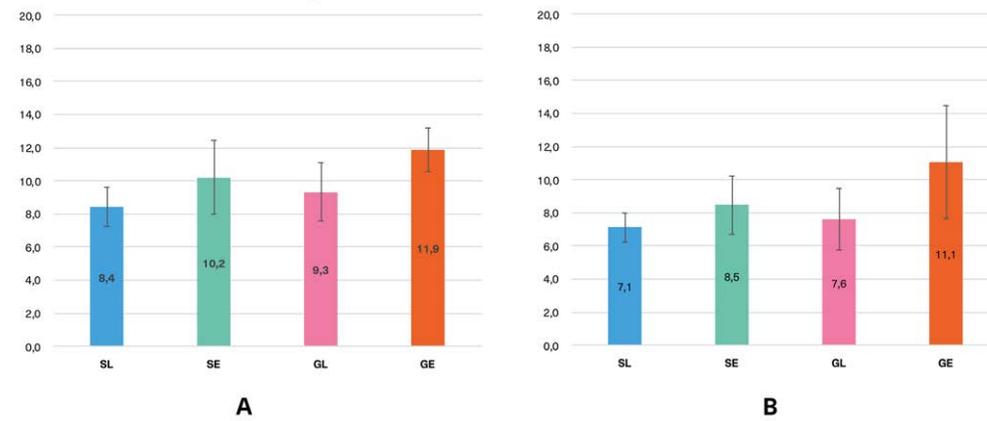


ABBILDUNG 51: Bewertung der Interaktionstechniken nach NasaTLX. Gering (0) Hoch (20)

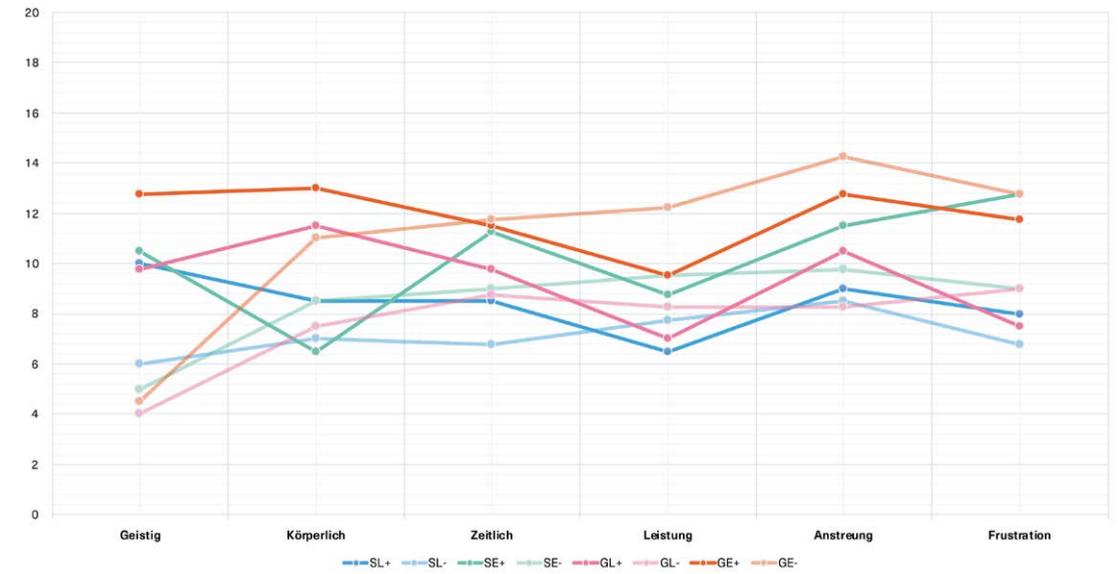


ABBILDUNG 52: Dimensionen nach NasaTLX. Der volle Farbton stellt Versuchsgruppe A (+) und der helle Ton Versuchsgruppe B (-) dar.

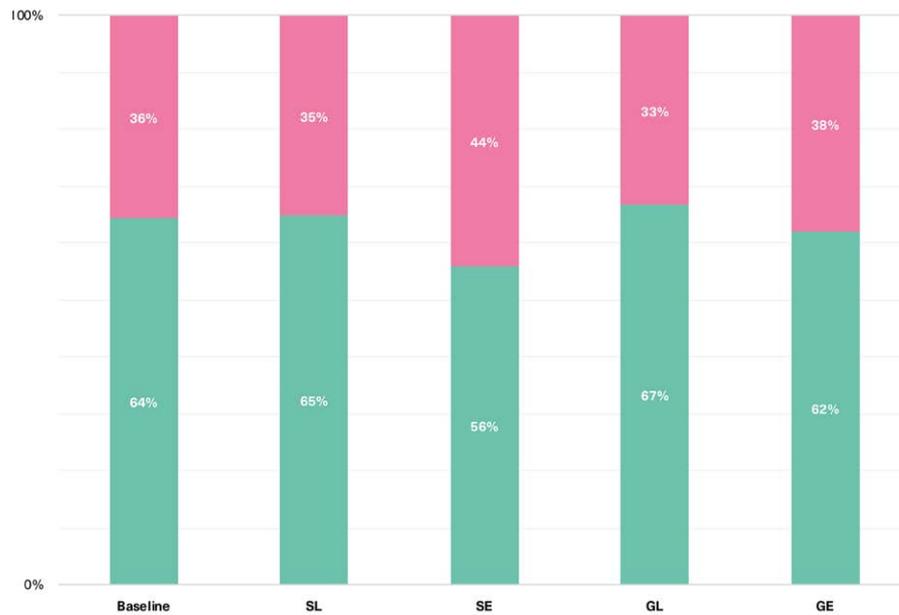


ABBILDUNG 53: Einfluss der Primäraufgabe bezogen auf Zeit in Prozent

6.3.3 BEOBACHTUNG UND NACHBEFRAGUNG

Insgesamt lässt sich festhalten, dass die Probanden, besonders während der Ausübung der Interaktionstechniken, spielerisch agierten. „Es hat sich wie ein Spiel angefühlt“ bemerkte einer der Probanden, wobei andere auch mitteilten, dass der Ehrgeiz sie gepackt habe pro Durchlauf in der Sekundäraufgabe schneller zu werden und sie besonders motiviert waren die Primäraufgabe exakt auszuführen. Dabei war zu beobachten, dass die Probanden einen stetigen Augenkontakt zur Primäraufgabe pflegten und häufig ihre Konzentration zuerst dieser Aufgabe widmeten und dann nach erfolgter Selektion in das Bild zoomten. Gerade in den ersten Durchgängen wurde teilweise die Primäraufgabe außer Acht gelassen und kurzzeitig ver-

gessen. Auch wurden zu Beginn die Interaktionstechniken durcheinandergebracht. Nach einigen Wiederholungen fielen die Interaktion leichter und gerieten nicht mehr in Vergessenheit. Alle Probanden begannen bei der Selektion beim oberen linken Quadranten. Teilweise wurde vor Start des nächsten Durchlaufs der Cursor schon vorsorglich auf den intendierten Quadranten bewegt. Das Feedback der Selektion (blauer Rahmen) schien teilweise nicht eindeutig, was einige Probanden bemerkten und beobachtet werden konnte, dass sich die Nutzer nicht sicher waren, ob das System ihre Eingabe erkannt hat.

Der Gebrauch der Sprache hat sich laut den Versuchspersonen insgesamt am besten angefühlt und war „intuitiv und leichtgängig“. Die Gesten waren negativ behaftet, weil sich laut den Probanden das Kopfnicken wesentlich angenehmer gestaltet als das Schütteln. In der Manipulation bemerkten fast alle, dass die Interaktion mit der Mimik als unangenehm und es war gar von befürchtigtem Muskelkater die Rede. Auch die Bewegung mithilfe der Kopfrichtung stieß auf Demut, weil der Kopf gerade an den Rändern zu sehr gedreht werden musste. Selektion und Manipulation fiel in den unteren Feldern leichter, weil der Kopf nicht zu sehr gehoben werden musste. Negativ beurteilt wurde zudem, dass die Zoomstufe und das Verschieben des Bildes begrenzt war. Nach der Selektion zoomten die Probanden zunächst hinein und explorierten selbständig das zu findende Objekt. Die Textur im Hintergrund bietet dabei laut den Aussagen genug Orientierung und ebenso ausreichend Feedback über die aktuelle Zoomstufe.

Die Interaktion mit den Augenbrauen fiel teilweise schwer, weil die Eingabe trotz Kalibrierung nicht immer exakt erkannt wurde. Vor allem im oberen Bereich des Interface und wenn das Gesicht frontal zur Kamera gerichtet ist, erfolgte ein besseres Tracking, da die Kamera das Gesicht besser erfassen konnte. Des Weiteren bemängelten die Probanden an dieser Technik, dass sie das Gefühl hatten, nicht präzise genug interagieren zu können.

6.4 Diskussion

Die zu Beginn der Evaluation erstellten Fragestellungen, anhand der Anforderungen multimodaler Systeme, gilt es im Folgenden zu diskutieren.

Aus den Daten ist zu entnehmen, dass die Dauer der Durchführung unter Anwendung der Primäraufgabe zunimmt. Die marginale Zunahme der Dauer innerhalb der drei Durchläufe in Abbildung 48 lässt sich durch diese Belastung und den Verlust der Aufmerksamkeit begründen. Betrachtet man alle Durchläufe hinsichtlich der Kombinationen, so wird, wenn man die Reihenfolge außer Acht lässt, ersichtlich, dass sich ein Lerneffekt bei Versuchsgruppe Beinstellt, wohingegen es bei Versuchsgruppe A mit Primäraufgabe keine bzw. eine leicht negative Tendenz zeigen lässt. (vgl. Abbildung 49) Das bedeutet, dass die Primäraufgabe beim Erlernen der Sekundäraufgabe hinderlich ist. Unter Ärzten, die ihr Handwerk beherrschen und somit keine neue Primäraufgabe lernen müssen, könnte die Tendenz anders sein. Die starken Schwankungen und die Tatsache, dass sich kein Plateau eingespielt hat, ist zurückzuführen auf die vergleichsweise geringe Anzahl der Probanden und zeigt, dass zu wenige Durchläufe gemessen wurden.

Es stellte sich heraus, dass neben einer

subjektiven Beurteilung seitens der Nutzer, auch stets eine objektive Datenerhebung herangezogen werden sollte, weil Übereinstimmungen, aber auch Differenzen sichtbar gemacht werden konnten.

Abbildung 54 zeigt eine Übersicht der Ergebnisse, mit dessen Hilfe Empfehlungen ausgesprochen werden können. Zusammengefasst sind die Interaktionstechniken, die ohne Primäraufgabe durchgeführt wurden, besser zu bewerten. Dies hängt mit der beschriebenen sensorischen und motorischen Zusatzbelastung zusammen. Klarer Favorit ist dabei die Gestenselektion in Kombination mit dem Lehnen, wenn keine weitere Aufgabe getätigt wird. Besteht für den Nutzer also die Möglichkeit seine Tätigkeit zu unterbrechen, so ist diese Technik am effizientesten. Kann er diese nicht unterbrechen, fällt die Wahl auf eine Kombination aus Spracheingabe und Lehnen. Das Vor- bzw. Zurücklehnen ist in beiden Dimensionen die beste Interaktionstechnik. Es ist besonders vorteilhaft, weil es sich um eine natürliche Form der Interaktion handelt und aus ergonomischer Sicht am wenigsten physische Beanspruchung mit sich bringt. Hinsichtlich der zeitlichen Messung (Dauer) gibt es in den Selektionstechniken (Sprache o. Gesten) keine klaren Sieger, wohingegen in der subjektiven Bewertung die Spracheingabe vorne liegt. Im Gegensatz zu den Ges-

ten, die häufig wiederholt werden mussten, weil das System die Eingabe nicht erkannte, war die Reaktion der Sprachbefehle direkter und die Nutzer mussten keine zusätzliche physische Interaktion ausführen. Ebenso nahmen sie auch etwas weniger Zeit in Anspruch und der Nutzer konnte sich auf das Zielen mit der Kopfrichtung fokussieren. Die Laborumgebung begünstigte jedoch eine vergleichsweise optimale Erkennung der Sprachbefehle, was in einem OP Szenario nicht so zu übertragen ist.

Die Kombination aus Gesten- und Augenbraueninteraktion stellt sich im Ganzen besonders schlecht heraus und ist demnach nicht weiterzuempfehlen. Begründen lässt sich dies, da die Kopfge- sten von Seiten der Versuchspersonen als besonders anstrengend empfunden wurden. Das könnte zum einen an der Art der Implementierung liegen, denn eine Optimierung der Empfindlichkeit könnte dies abschwächen, aber auch, dass hier zwei konträre Bewegungen ausgeführt wurden: Das Schütteln erforderte mehr Anstrengung als das Nicken. Die Verwendung der Mimik als Mittel der Interaktion ist in diesem speziellen Fall der Augenbrauenteknik negativ behaftet. Selten nutzen Menschen die Mimik als bewusstes Werkzeug, denn vielmehr ist es ein unbewusstes Mittel der Kommunikation. Dadurch dass die Muskulatur selten genutzt wird, ist diese Technik, gerade in Anbetracht einer Intervention bzw. der hier realisierten

Umsetzung, als Interaktionstechnik gescheitert. In der Gesamtheit ist anzumerken, dass keine der Interaktionstechniken die Ausübung der Primäraufgabe signifikant einschränkt. Gerade die Steuerung mit der Kopfrichtung ist vorteilhaft, weil der Nutzer seinen Blick auf die kurzfristige Ausübung seiner Primäraufgabe konzentrieren kann.

		OHNE PA	MIT PA
DAUER	F1	+	-
	F2	S = G	S = G
	F3	L > E	L > E
	IX	GL	/
NTLX	F1	+	-
	F2	S > G	S > G
	F3	L > E	L > E
	IX	GE besonders schlecht	
PA	F1	Leistung = Baseline	
	F2	S = G	S = G
	F3	L = E	L = E
	IX	/	

ABBILDUNG 54: Tabellarische Übersicht zur Zusammenfassung der Ergebnisse. FN = Faktor (Interaktionstechniken), F1 = Auswirkung Primäraufgabe (PA), F2 = Selektionstechniken, F3 = Manipulationstechniken, IX = Interaktion zw. Faktoren, > / < = besser / schlechter, + / - = positive / negative Bewertung hinsichtlich der Interaktionstechniken

6.4.1 OPTIMIERUNGSMÖGLICHKEITEN

Anhand der Auswertung und auch durch die Beobachtungen während der Laborstudie, können Verbesserungsmöglichkeiten im Studiendesign, aber auch bei den Prototypen selbst in Aussicht gestellt werden, wodurch Hilfe bei etwaigen Folgeuntersuchen geleistet werden kann.

PRIMÄRAUFGABE

Es fiel den Nutzern besonders leicht im unteren Bereich des Interfaces zu interagieren, was an der Positionierung der Skala der Primäraufgabe gelegen hat. Dieses Element sollte im Sichtfeld des Nutzers angezeigt werden, also direkt im jeweiligen Quadranten. Dafür spricht, dass auch in einem realen Angiographiesystem das Instrument (bspw. die Katheterspitze) direkt im Bild zu sehen ist. Weitere Hospitationen sowie der Austausch mit Radiologen können dabei helfen, detailliertere Ableitungen der Tätigkeiten zu schaffen und somit eine bessere Simulation der Primäraufgabe zu gestalten.

KOPFGESTEN

Wie beschrieben, bescherte das Kopfschütteln den meisten Probanden Probleme, weshalb die Bewertung der Gesten gelitten hat. Das Schütteln könnte durch eine andere Kopfgeste ersetzt werden oder es werden gezielt Mög-

lichkeiten gesucht, verschiedene Arten des Nickens umzusetzen.

MIMIKGESTEN

Zur Manipulation erwies sich die Interaktion mit den Augenbrauen als unvoreilhaft. Denkbar ist eine Kombination zu schaffen, mit der eine Selektion möglich wäre, die gleichzeitig weniger zum Einsatz kommen soll. So könnten die Augenbrauen als Sicherheitsmechanismus eingesetzt werden, um bestimmte Funktionen zum Starten, Abbrechen bzw. Beenden dienen.

SPRACHE

Auch bei Sprachbefehlen ist die Verwendung von Mechanismen in Form von speziellen Schlüsselwörtern nützlich. Denn hier kann sichergestellt werden, dass der Nutzer zuerst ein Feedback erhält, ob die Eingabe (richtig) erkannt wurde. Um ein Beispiel zu nennen: der Nutzer initialisiert das System mit dem Schlüsselwort „Ovi“ und fügt anschließend den Befehl hinzu „Ovi ... Start“. Zwischen den Signalwort und Befehl verarbeitet das System bereits die Information und gibt ein visuelles Feedback, dass eine Eingabe erkannt wurde. Zudem müsste ein spezielles Vokabular entwickelt werden, mit denen das System trainiert werden kann, womit die Erkennungsquote ständig besser wird. Zusätzliche könnte eine bessere Mikro-

fonierung des Raums eine genauere Erkennung gewährleisten. Denkbar wäre ein Richtmikrofon, welches der Kopfbewegung des Nutzers folgt und die Stimmfarbe und Aussprache mit seiner Person in Verbindung bringen kann.

KOPFBEWEGUNGEN

Durch die Installation am oberen Monitorrand war die Sicht der Kamera teilweise beschränkt, was dazu führte, dass die Gesichtserfassung gelegentlich verloren ging, sobald ein Nutzer zu weit nach unten schaute. Auch beim Blick nach links und rechts stieß die Erkennung an seine Grenzen. Dieses Problem könnte gelöst werden, indem mehrere Kameras am Monitor bzw. im Raum installiert werden, denn hiermit kann die Körper- und Gesichtsposition mehrerer Nutzer exakt bestimmt werden.

INTERFACE

Das Layout des Prototyps wurde bewusst in vier gleichmäßige Segmente aufgeteilt. Diese Vereinheitlichung war im Zusammenhang der Studie nützlich, spiegelt allerdings nicht die reale Situation wieder, weil am echten System die Möglichkeit der Individualisierung besteht. Aus diesem Grund sollten verschiedene und modulare Aufbauten getestet werden. Es sind auch nicht überall die gleichen Interaktionen nötig. So sollte integriert werden, dass sich spezifische Interaktionen nur in bestimm-

ten Bereichen oder mit besonderen Bedienelementen durchführen lassen, während in anderen Bereichen Informationen dargestellt werden können.

So könnten funktionale Elemente, mit dem Blick angewählt und mit einer Geste gestartet werden. ▶

UMGEBUNG

Die Laborbedingungen waren bekanntermaßen steril. Die Lichtbedingungen und Geräuschkulisse war optimal und gleicht nicht dem einer realen OP Situation, weshalb die Robustheit des Systems hinsichtlich Umgebungseinflüssen optimiert werden muss.

TRACKING

Um eine bessere Analyse, insbesondere mit Hinblick der Überführung in eine echte Software zu ermöglichen, ist es sinnvoll die Blickbewegungen zu erfassen. Ähnlich wie beim Eyetracking kann die Kopfposition in einem Wärmebild dargestellt werden und ermöglicht so weitere Optimierungen des Interfaces.

FEEDBACK

Das Feedback beschränkte sich nur auf visueller Ebene, was gerade in Anbetracht der Komplexität der Bedienung mit den Interaktionstechniken, nicht immer ausgereicht hat. Es könnten andere aber auch zugleich weitere Sinneskanäle angesprochen werden. Dabei ist es zum Beispiel denkbar, dass das System auditive Rückmeldung gibt, wenn eine Eingabe erkannt wurde oder nicht eindeutig zugewiesen werden konnte. Dies könnte gerade beim Zoom nützlich sein, wo durch Sounds Rückmeldung zur maximalen Zoomstufe oder Begrenzung des Bewegungsraums angedeutet werden können. Vibrotaktiler Feedback direkt am Instrument könnte während der Ausführung der Primäraufgabe von Nutzen sein.

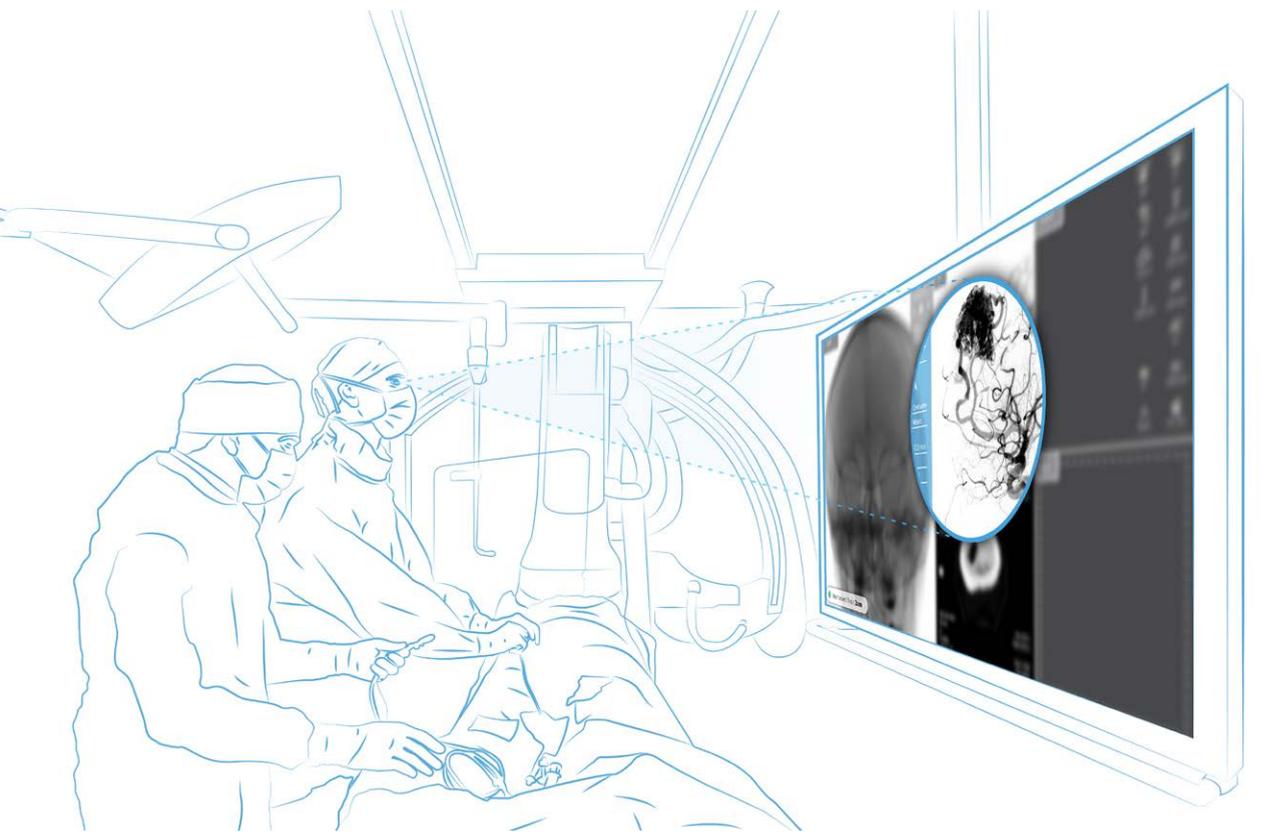
TRAINING

Die Einweisung der Probanden in das System war in beiden Versuchsgruppen gleich. Trotz des vorherigen Trainings konnte kein Plateau dargestellt werden, weil zu wenig Durchläufe erfasst wurden. Die Sekundäraufgabe sollte mehrere Trainingsabläufe beinhalten, bevor sie mit der Primäraufgabe getestet wird. Zusätzlich sollte die doppelte Anzahl an Versuchspersonen herangezogen werden.

Zusammenfassung

I < O

KAPITEL 7



In dieser Arbeit wurden multimodale Mensch-Computer-Schnittstellen im medizinischen Kontext anhand eines nutzerzentrierten Designprozess erforscht. Auf Grundlage einer theoretischen Aufarbeitung multimodaler Systeme wurde explorativ der Arbeitsbereich von Radiologen untersucht. Durch Beobachtungen, Interviews und einer systematischen Analyse im Nutzungskontext einer radiologischen Intervention konnten spezifische Probleme im Workflow identifiziert und An-

forderungen für eine berührungslose Benutzerschnittstelle klassifiziert werden.

Innerhalb eines iterativen Prozesses wurden durch gegebene Nutzungsanforderungen neuartige Interaktionstechniken am Beispiel einer radiologischen Intervention erforscht. Der gestalterische Entwurf umfasste dabei die Abstraktion der bestehenden Schnittstelle, in eine speziell zur Evaluation der konzipierten Techniken grafischen Benut-

zeroberfläche. Ergebnis ist das Konzept „OVI“, mit dem untersucht wurde, welche multimodalen Kombinationen handfreier Interaktionstechniken zur Navigation in Bilddaten geeignet sind. Die entstandenen Eingabemodalitäten dienten der Selektion und Manipulation und umfassten Ganzkörper-, Kopf-, und Augengesten, sowie Spracheingabe. Dabei wurde konkret überprüft, welche manipulative Modalität sich am besten für Zoominteraktionen eignet. Die spezifischen Interaktionsaufgaben des Radiologen wurden in eine Primär- und Sekundäraufgabe unterteilt. Dafür wurde ein prototypisches System umgesetzt, welche ein Soft- und Hardwareinterface bietet, wodurch die Tätigkeiten des Nutzers simuliert wurden.

Abschließend erfolgte die Evaluation des entwickelten Konzepts innerhalb einer Nutzerstudie unter Laborbedingungen. Die Studie umfasste acht Teilnehmer, die in zwei Versuchsgruppen unterteilt wurden und diente der Beurteilung der verschiedenen Modalitäten hinsichtlich der Benutzerfreundlichkeit, sowie des Einflusses der Primäraufgabe auf den Workflow. Es konnten Stärken und Schwächen sowohl der einzelnen Interaktionstechniken, als auch Kombination verschiedener Modalitäten, herausgestellt werden. Eine wichtige Erkenntnis des Studiendesigns ist, dass zu objektiven Messdaten eine subjektive Beurteilung herangezogen werden sollte, weil nur durch dieses Zu-

sammenspiel differenzierte Ergebnisse ermittelt werden konnten.

Kopfgesten zur Selektion, in Kombination des Vorbeugens zur Bildvergrößerung, stellte sich in dieser Studie als Favorit heraus, wenn keine Doppelaufgabe ausgeführt wird und der Nutzer die Tätigkeit kurzweilig unterbrechen kann. Ein kontinuierlicher Workflow kann gewährleistet werden, wenn die Selektionen mittels Sprachbefehlen ersetzt werden. Das Vor- und Zurücklehnen ist unter den hier evaluierten Techniken geeignet zur Bildvergrößerung. Eine Kombination aus Kopf- und Augengesten kann nicht empfohlen werden, weil die Dauer, Genauigkeit und Beanspruchung besonders negativ auffielen.

Zusammenfassend ist zu sagen, dass jegliche zukünftige Untersuchung dieser Art Rücksicht auf das Zusammenspiel von verschiedenen Modalitäten hinsichtlich der Durchführungen unterschiedlicher Aufgaben nehmen sollte. Vor allem die kognitiv belastende und motorisch präzise Durchführung einer chirurgischen Tätigkeit, die in dieser Studie annähernd nachempfunden werden sollte, zeigt, dass es Wechselwirkungen angesichts der Interaktionstechniken mit dem System gab. Abschließend gilt festzuhalten, dass multimodale Ansätze wie diese, einen ergänzenden Mehrwert der Mensch-Computer-Interaktion in bestehenden Systemen bieten können.

Der menschenzentrierte Designprozess, der in dieser Arbeit verfolgt wurde, beinhaltet laut ISO-Norm weitere Iterationen in Hinblick der Umsetzung eines Produkts nach menschlichen Bedürfnissen. Die entwickelten multimodalen Konzepte bedürfen weiterer Tests mit einhergehenden Optimierungen in der Zielumgebung. Die Benutzeroberfläche sollte weiterentwickelt, mit medizinischen Bilddaten ersetzt und im Idealfall an einer Angiografieanlage getestet werden. Dabei sollte eine Studie mit realen Nutzern, sprich angewandten Radiologen, durchgeführt werden, um konkrete Bedürfnisse zu identifizieren und die Interaktionstechniken dahingehend zu optimieren. Als erster Schritt sollte über die Ergebnisse dieser Arbeit mit Ärzten diskutiert und der Prototyp praktisch getestet werden. Besonders interessant ist die Frage, ob die Primäraufgabe den realen Bedingungen gerecht wird und wie in Absprache mit den Radiologen weitere motorische Tätigkeiten und Anforderungen entwickelt werden können. Die Kommunikation mit dem Computer über die Sprache wird stetig weiterentwickelt und langfristig eine vielversprechende Lösung bieten, weshalb ein Grundvokabular mit Medizinern zu entwickeln ist.

Es konnten nicht alle Ideen untersucht und jegliche Eventualitäten multimodaler Systeme in Betracht gezogen werden. Die Themen Authentifizierung und Sicherheit wurden lediglich theoretisch betrachtet und es Bedarf der Implementierung solcher Mechanismen. Ebenso wurde das Multi-User-Szenario außen vorgelassen. Dabei ist es interessant herauszufinden, ob und wie mehrere Nutzer berührungslos interagieren können und wie sich dies im Workflow integrieren lässt.

Als Feedback wurde in dieser Arbeit auf visuelle Systemrückmeldungen gesetzt, wobei auch Töne, als zusätzliche Modalität, eine unterstützende Wirkung erzielen könnten. (Mikro-)Animationen können zusätzliches Feedback und eine noch natürlichere Wirkung der Eingabebefehle bieten.

Die Manipulationstechniken wurden für zweidimensionalen Ansprüche optimiert, jedoch findet die Exploration von 3D-Volumen genauso oft Verwendung, weshalb neue Überlegungen zur Übertragung der entwickelten Techniken getroffen werden müssen.

LITERATURVERZEICHNIS

- [1] Bolt, „Put-that-there: Voice and gesture at the graphics interface.“ MIT Press, Bd. Computer Graphics, Nr. 14 (3), pp. 262-270, 1980.
- [2] O'Hara, Kenton, et al., „Touchless interaction in surgery.“, Commun. ACM, 57.1, pp. 70-77, 2014.
- [3] Odisio and Wallace, „Image-guided interventions in oncology.“, Surgical Oncology Clinics, Bd. 23, Nr. 4, pp. 937-955, 2014.
- [4] Hübler, Hansen, Beuing, Skalej and Preim, „Workflow Analysis for Interventional Neuroradiology using Frequent Pattern Mining“, CURAC, pp. 165-168, 2014.
- [5] Mentis, O'Hara, Gonzalez, Sellen, Corish, Criminisi and Theodore, „Voice or gesture in the operating room.“, In Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems, pp. 773-780, 2015.
- [6] Meyer, Levine, Egan, Cohen, Spitz, Garcia and Sandberg, „A computerized perioperative data integration and display system.“, International Journal of Computer Assisted Radiology and Surgery, Bde. (3-4), pp. 191-202, 2007.
- [7] Hatscher, „GazeTap: towards hands-free interaction in the operating room.“, Proceedings of the 19th ACM international conference on multimodal interaction. ACM, 2017.
- [8] Rein, „Einführung in die Physiologie des Menschen“, Berlin: Springer, 1964.
- [9] Hedicke, „Mensch-Maschine- Systemtechnik, 2. Auflage.“, In Multimodalität in Mensch-Maschine-Schnittstellen, Düsseldorf, Symposion Verlag, pp. 203-232, 2000.
- [10] Engesser, „Duden Informatik 2. Auflage“, Mannheim: Duden Verlag, 1993.
- [11] Gallace, „In Touch with the Future: The Sense of Touch From Cognitive Neuroscience to Virtual Reality“, Oxford University Press, 2014.
- [12] Park, „Designing Across Senses: A Multimodal Approach to Product Design“, O'Reilly Media, 2018.
- [13] Beuter, „Gestenbasierte Positionsreferenzierung für die multimodale Interaktion mit einem anthropomorphen Robotersystem“, Diplomarbeit, Technischen Universität Bielefeld, 2007.
- [14] Wickens, „Multiple resources and performance prediction“, Theoretical Issues in Ergonomics Science, Bd. 3(2), pp. 159-177, 2000.
- [15] Kiefer, „Multitasking-Heuristiken in dynamischer Mensch-Technik-Interaktion“, DGLR, Bonn, 2006.
- [16] Bunt, Harry, Beun and Borghuis, „Multimodal human-computer communication: systems, techniques, and experiments.“, Vol. 1374. Springer Science & Business Media, 1998.
- [17] Neuss, Dissertation, „Usability Engineering als Ansatz zum Multimodalen Mensch- Maschine-Dialog“, Technische Universität München, 2000.
- [18] Turk, „Multimodal interaction: A review“, in Pattern Recognition Letters, pp. 189-195, 2014.
- [19] Charwat, „Lexikon der Mensch-Maschine-Kommunikation (2. Auflage)“, München: Oldenbourg Verlag, 1994.
- [20] Boles, „Begleitbuch zur Vorlesung Multimedia- Systeme“, Oldenburg: Carl von Ossietzky Universität Oldenburg, 1998.
- [21] Foley, „Computer Graphics: Principles and Practice, Second Edition“, Addison-Wesley Publishing Company Inc., 1992.
- [22] Lee, „Search of a Natural Gesture.“, Bd. Crossroads The ACM Magazine for Students 16, Nr. 4, pp. 11-14, 2010.
- [23] Blake, „nui.joshland.org“, 6 April 2010. [Online]. Available: <http://nui.joshland.org/2010/04/nui-reuse-existing-skills.html>. [Zugriff am 2.7.2019].
- [24] Bollhoefer, Meyer and Witzsche, „Microsoft Surface und das Natural User Interface (NUI)“, 2009.
- [25] Saffer, „Designing Gestural Interfaces“, O'Reilly Media Inc., 2009.
- [26] Oviatt, „Ten myths of multimodal interaction“, Communications of the ACM, Bd. 42, Nr. 11, pp. 74-81, 1999.
- [27] Oviatt, „The human-computer interaction handbook: fundamentals, evolving technologies and emerging applications“, ACM, pp. 414 – 428, 2013.
- [28] Rötting, „Multimodale Mensch-Maschine Interaktion“, In Beiträge zur Mensch-Maschine-Systemtechnik aus Forschung und Praxis, Band 1, Düsseldorf, Symposion Publishing GmbH, pp. 283-300, 2005.
- [29] Liu, Hongyi and Wang, „Gesture recognition for human-robot collaboration: A review“, International Journal of Industrial Ergonomics , Bd. 68, pp. 355-367, 2018.
- [30] Preim, „Interaktive Systeme, User Interface Engineering, 3D-Interaktion, Natural User Interfaces Band 2.“, Heidelberg: Springer-Verlag Berlin, 2015.
- [31] Ultrahaptic, „leapmotion“, 2019. [Online]. Available: <https://www.leapmotion.com/>. [Zugriff am 14.6.2019].
- [32] Stereolabs, „stereolabs.com“, 7.2019. [Online]. Available: <https://www.stereolabs.com/zed/>. [Zugriff am 1.7.2019].
- [33] Terabee, „lidarmag.com“, 16.4.2019. [Online]. Available: <https://lidarmag.com/2019/04/16/terabee-introduces-latest-3d-time-of-flight-camera/>. [Zugriff am 2.7.2019].
- [34] Microsoft, „developer.microsoft.com“, 2019. [Online]. Available: <https://developer.microsoft.com/de-de/windows/kinect>. [Zugriff am 2.7.2019].
- [35] Microsoft, „microsoft.com“, 2019. [Online]. Available: <https://www.microsoft.com/de-de/hololens/buy>. [Zugriff am 28.6.2019].

- [36] Nickel, „www.golem.de“, 24.2.2019. [Online]. Available: <https://www.golem.de/news/azure-kinect-dk-xbox-kinect-ist-jetzt-ein-iot-sensor-1902-139600.html>. [Zugriff am 2.6.2019].
- [37] Dent, „engadget.com“, 20.2.2019. [Online]. Available: https://www.engadget.com/2019/02/20/lg-g8-thinq-gesture-sensor-module/?guccounter=1&guce_referrer=aHR0cHM6Ly93d3cuZ29vZ2x1LmNvbS8&guce_referrer_sig=AQAAABFS_GY7FTR-pGbHZWkDDX2bSO8Ye5wF-0dLrgt6d0R0pwOnzzWNgSKbMCDYNoGENxwrnzaWHsekfpuWZW5vFBOeUzQemsCLBcTt6iz. [Zugriff am 2.7.2019].
- [38] Totilo, „https://kotaku.com/the-switch-joy-con-s-infrared-sensor-is-cooler-than-we-1822669059“, 2.2.2018. [Online]. Available: <https://kotaku.com/the-switch-joy-con-s-infrared-sensor-is-cooler-than-we-1822669059>. [Zugriff am 7.7.2019].
- [39] Magic Leap Inc., „magic Leap.com“, 2019. [Online]. Available: <https://www.magic Leap.com/>. [Zugriff am 1.7.2019].
- [40] Meta Vision, „www.metavision.com“, 2019. [Online]. Available: <https://www.metavision.com/>. [Zugriff am 1.7.2019].
- [41] Manus VR, „manus vr“, 2019. [Online]. Available: <https://manus-vr.com/>. [Zugriff am 3.7.2019].
- [42] Ireviews, „ireviews.com“, 2019. [Online]. Available: <https://www.ireviews.com/review/myo-armband>. [Zugriff am 6.7.2019].
- [43] Google, „atap.google.com“, 2019. [Online]. Available: <https://atap.google.com/soli/>. [Zugriff am 7.7.2019].
- [44] Nigay, Laurence and Coutaz, „A design space for multimodal systems - concurrent processing and data fusion“, INTERCHI '93, Bd. Proceedings of the Conference on Human Factors and Computing Systems, pp. 172-178, 1993.
- [45] Billinghurst, „medium.com“, 21.4.2019. [Online]. Available: <https://medium.com/@marknb00/multi-modal-mixed-reality-b4ab8ddc304b>. [Zugriff am 1.7.2019].
- [46] León, „medium.com“, 7.12.2016. [Online]. Available: <https://medium.com/@germanleono/life-after-mobile-vr-and-ar-will-be-powered-by-multimodal-interactions-5338847db203>. [Zugriff am 2.7.2019].
- [47] Wenge et al., „DMove: Directional Motion-based Interaction for Augmented Reality Head-Mounted Displays.“, Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. ACM, 2019.
- [48] Schöning, Daiber and Krüger, „Whole body interaction with geospatial data.“, International Symposium on Smart Graphics. Springer, Berlin, Heidelberg, 2009.
- [49] Hatscher and Hansen, „Hand, Foot or Voice: Alternative Input Modalities for Touchless Interaction in the Medical Domain.“, In Proceedings of the 2018 on International Conference on Multimodal Interaction, ACM, 2018.
- [50] Hatscher, Luz and Hansen, „Foot Interaction Concepts to Support Radiological Interventions.“, In Mensch und Computer Tagungsband, 2017.
- [51] Wagner et al., „Konzepte mit unterschiedlichen Platzanforderungen zur Navigation in medizinischen Bilddaten mittels eines Sensorfußbodens Problemstellung Material und Methoden.“, Hannover, CURAC - Computer- und Roboterassistierte Chirurgie, 2017.
- [52] Universitätsklinikum Magdeburg, Institut für Neuroradiologie, „http://www.kneu.ovgu.de“, 10.6.2019. [Online]. Available: http://www.kneu.ovgu.de/Patienten/Spezialsprechstunden+_+Poliklinik/Neuroradiologie.html. [Zugriff am 17.5.2019].
- [53] Hasselmann, „www.operation-karriere.de“, 16.7.2018. [Online]. Available: <https://www.operation-karriere.de/karriereweg/assistentenarzt/arzt-mit-durchblick-facharzt-fuer-radiologie.html>. [Zugriff am 4.5.2019].
- [54] Deutsche Gesellschaft für Neuroradiologie e.V., „www.dgnr.org“, 15.6.2019. [Online]. Available: <https://www.dgnr.org/de-DE/74/was-ist-neuroradiologie>. [Zugriff am 3.6.2019].
- [55] AUBI-plus GmbH, „www.aubi-plus.de“, 2019. [Online]. Available: <https://www.aubi-plus.de/berufe/medizinisch-technischer-radiologieassistent-269/>. [Zugriff am 18.6.2019].
- [56] Radiologische Allianz GbR, „radiologische-allianz.de“, 2019. [Online]. Available: <https://www.radio-logische-allianz.de/methoden/diagnose/angiographie/>. [Zugriff am 20.5.2019].
- [57] Wikipedia, „wikipedia.org“, 2019. [Online]. Available: https://de.wikipedia.org/wiki/Digitale_Subtraktionsangiographie. [Zugriff am 15.5.2019].
- [58] Neurochirurgie Uniklinik Tübingen, „neurochirurgie-tuebingen.de“, 2019. [Online]. Available: <https://www.neurochirurgie-tuebingen.de/de/spezialgebiete/vasculaere-neurochirurgie/intrakranielle-gefassmissbildungen/arteriovenose-malformationen/>. [Zugriff am 20.4.2019].
- [59] Deutsche Gesellschaft für Neuroradiologie e.V., „dgnr.org“, 2019. [Online]. Available: <https://www.dgnr.org/de-DE/74/was-ist-neuroradiologie>. [Zugriff am 3.5.2019].
- [60] Baeck, Johannes, Thäle and Weichert, „Nutzerzentrierte Entwicklung eines KI-Assistenten für Radiologen“, In Hess, S. & Fischer, H. (Hrsg.), Mensch und Computer 2018 - Usability Professionals, Bde. Mensch und Computer 2018 - Usability Professionals, Gesellschaft für Informatik e.V. Und German UPA e.V., pp. 337-352, 2018.
- [61] A. Mewes, P. Saalfeld, O. Riabikin, M. Skalej und C. Hansen, „A gesture-controlled projection display for CT-guided interventions.“, International journal of computer assisted radiology and surgery, Bd. 11, Nr. 1, pp. 157-164, 2016.
- [62] Manresa-Yee, Cristina and Asensio, „Usability of vision-based interfaces“, Actas del XIV Congreso Internacional Interacción Persona Ordenador, 2013.
- [63] Norman and Nielsen, „Estural interfaces: a step backward in usability“, interactions 17.5, pp. 46-49, 2010.
- [64] Saffer, „Designing Gestural Interfaces.“, O'Reilly Media, Inc., 2009.
- [65] Oviatt, „The Paradigm Shift to Multimodality in Contemporary Computer Interface“, Morgan & Claypool, Williston, VT, 2015.

- [66] Tan and Justin et al. „Informatics in radiology: Developing a touchless user interface for intra-operative image control during interventional radiology procedures.“, Radiographics 33.2, pp. 61-70, 2013.
- [67] Mauser, Stanislas and Burgert, „Touch-free, gesture-based control of medical devices and software based on the leap motion controller“, MMVR, pp. 265–270, 2014.
- [68] Seán and Gavin, „Touchless computer interfaces in hospitals: A review“, Health informatics journal, 2018.
- [69] Wagner, Petra, Malisz and Kopp, „Gesture and speech in interaction: An overview.“, pp. 209-232, 2014.
- [70] Wobbrock, Rubinstein, Sawyer and Duchowski, „Longitudinal Evaluation of Discrete Consecutive Gaze Gestures for Text Entry.“, In Proc. of Eye Tracking Research & Applications, pp. 11-18, 2008.
- [71] Black, Ganze, Hettig and Hansen, „Auditory Display for Improving Free-hand Gesture Interaction“, In Mensch und Computer 2017-Tagungsband, pp. 1-11, 2017.
- [72] Meinel, „Spezifische Effekte visueller und kognitiver Ablenkung bei der Kraftfahrzeugführung“, Dissertationschrift, 2012.
- [73] Weinger, Wiklund and Gardner-Bonneau., „Handbook of human factors in medical device design“, CRC Press, 2010.
- [74] Unity Technologies, „unity3d.com“, 2019. [Online]. Available: <https://unity3d.com/de/unity>. [Zugriff am 30.5.2019].
- [75] Microsoft Corporation, „visualstudio.microsoft.com“, 2019. [Online]. Available: <https://visualstudio.microsoft.com/de/>. [Zugriff am 9.6.2019].
- [76] Hart, Sandra and Staveland, „Development of NASA-TLX (Task Load Index):Results of empirical and theoretical research.“, Advances in psychology. Vol. 52. North-Holland, pp. 139-183, 1988.
- [77] Hörold et al., „Der „Faktor Mensch“ im Fokus der Entwicklung von 3D-Technologien - Methodensammlung“, Chemnitz: Technische Universität Chemnitz Fakultät für Maschinenbau, 2015.
- [78] Stein, Barry and Meredith, „The merging of the senses.“, The MIT Press, 1993.
- [79] Mitra, Sushmita and Acharya, „Gesture recognition: A survey.“, IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) 37.3, pp. 311-324, 2007.
- [80] Mewes et al., „Touchless interaction with software in interventional radiology and surgery: a systematic literature review.“ International journal of computer assisted radiology and surgery, Bd. 2, Nr. 2, pp. 291-305, 2017.
- [81] Grange et al., „M/ORIS: a medical/operating room interaction system.“ In Proceedings of the 6th international conference on Multimodal interfaces, pp. 159-166, 2004.

ABBILDUNGSVERZEICHNIS

SEITE

- 13 „Kapitel 1“ – jc-gellidon @ unsplash.com
 33 „Kaipitel 3“ – <https://www.pexels.com/photo/two-person-doing-surgery-inside-room-1250655/>
 55 „Kapitel 4“ – you-x-ventures @ unsplash.com
 125 „Kapitel 7“ Anonyme Aufnahme, Institut für Neuroradiologie am Universitätsklinikum Magdeburg

NR

- 4 Rechts: <https://brekel.com/wp-content/uploads/2018/05/Dcr7lwpV4AAJnkH.jpg>
 Links: <https://www.microsoft.com/en-us/p/azure-kinect-dk/8pp5vxdm9nhq>
 5 Aus Bolt, Richard A. "Put-that-there": Voice and gesture at the graphics interface. Vol. 14. No. 3. ACM, 1980.
 6 https://www.i-med.ac.at/pr/pressebilder/images/angio_breit.jpg, Foto: MUI/C. Lackner)
 7 https://www.neurochirurgie-tuebingen.de/index.php?rex_media_type=content&rex_media_file=avm1.jpg, https://www.neurochirurgie-tuebingen.de/index.php?rex_media_type=content&ex_media_file=avm1.jpg
 8 Institut für Neuroradiologie am Universitätsklinikum Magdeburg
 10 <https://www.siemens-healthineers.com/de/angio/options-and-upgrades/components-and-options>
 28 & 28 kyryll-ushakov- @ unsplash.com

Fragebogen

1. **Welche Bilddaten betrachten Sie während der Intervention?**
Das hängt natürlich von der jeweiligen Intervention und Umgebungssituation ab. Bei einem Schlaganfall muss alles sehr schnell gehen und man macht deutlich weniger Bilder als bei einer geplanten Aneurysma- oder AVM-Embolisation. Prinzipiell natürlich unsere Durchleuchtungsbilder, also Flouroskopie sowie nach speziellen Rotationsangiographien auch 3D-Volumendatensätze.
2. **Wie ist dabei die Anordnung dieser Daten auf dem Bildschirm?**
Die ist ebenfalls immer anders und auch stark vom Interventionalisten abhängig. Ich benutze zum Beispiel ein anderes Layout als OA Beuing, dem du zugeschaut hast. Prinzipiell ist es so, dass es darauf ankommt, ob beispielsweise beide Ebenen gebraucht werden. Falls nicht, kann man die benutzte Ebene als Vollbild anzeigen lassen.
3. **Wie und wie oft wechseln sie währenddessen zwischen Daten?**
Lässt sich auch schwer pauschalisieren, der Idealgedanke ist ja, dass man erst Flouroskopie benötigt, um die Kathetersysteme an die Schädelbasis zu bringen. Dann wird eine Rotationsangiographie durchgeführt und der 3D-Datensatz angeschaut, um ein räumliches Verständnis der Anatomie zu bekommen. Nach entsprechender Einstellung der Ebenen reicht dann idealerweise wieder Flouroskopie. Aber natürlich kann es sein, dass bei komplizierten Anatomien mehrfach der Volumendatensatz angeschaut werden muss.
4. **Welche Einschränkungen bestehen dabei aktuell?**
Dazu muss man ständig auf dem Touchpad das Layout ändern und dann mit dem Joystick den Datensatz durchblättern oder rotieren.
5. **Werden Sie bei diesem Vorgehen unterstützt?**
Manchmal durch die MTRA (abhängig vom Interventionalisten und der MTRA).
6. **Müssen Sie währenddessen andere Geräte benutzen?**
Nein.
7. **Sind Sie zufrieden mit der Position der Bildschirm(e)?**
Prinzipiell schon, da eine andere Anordnung in Relation zur Patientenposition und den verfügbaren Katheterlängen mir nicht plausibel erscheint.
8. **Wird der Monitor oft verstellt?**
Für jeden Patienten.
9. **Welche (intuitive) Handlungsabläufe bestehen bereits?**
Je nach dem welcher Interventionalist mit wieviel Erfahrung am Tisch steht, kann der natürlich unmittelbar nach Generieren einer neuen Bildserie diese durchblättern ohne dabei den Joystick oder so suchen zu müssen. Das läuft intuitiv bzw no-look ab.
10. **Ist ein optischer Zoom während der Intervention möglich/notwendig?**
Ist definitiv notwendig- deine Intervention war ja anscheinend ein gutes Beispiel dafür. Ich versuche mir dann immer die Ebene als Vollbild anzeigen zu lassen. Prinzipiell ist ein zoomen möglich, aber mit sehr vielen Handgriffen verbunden und daher nicht populär unter uns.
11. **In welchen Situationen ist ein radiologischer Zoom notwendig?**
Immer wenn es um filigrane Fragestellungen, meist in peripheren Gefäßen geht. Beispielweise, ob irgendwo noch ein kleiner Restverschluss eines Gefäßes vorhanden ist (Schlaganfall) oder umgekehrt, ob ein pathologisches Gefäß weiterhin durchblutet ist (AVM)
12. **Haben Sie häufig Probleme bei der Orientierung?**
Das kommt extrem selten vor, nur bei sehr komplexen AVMs, die aus mehreren Stromgebieten gespeist werden, braucht man mal ne Weile bis man den Blutfluss verstanden hat.
13. **Was wäre hilfreich ergänzend zu sehen?**
Ich finde die Idee des optischen Zooms schon sehr gelungen. Es ist prinzipiell sehr schwer an etwas weiteres zu denken, wenn man bereits an der besten Angio auf diesem Planeten arbeiten kann
14. **In welchen Datensätzen zoomen Sie und wie oft kommt es vor?**
Sowohl Flouroskopie als auch 3D-Volumen.
15. **Wobei liegt dann der Fokus / welche Bildbereiche wollen Sie dabei vergrößert haben?**
Bei Flouroskopie siehe Frage 12. Bei 3D geht es dann am meisten um ein Aneurysma, welches so „freigedreht“ werden muss, dann man verstehen kann, woher das Blut kommt und wohin es nach dem Aneurysma fließt, um eine geeignete Therapie dafür zu finden.
16. **Ist überhaupt eine weitere Interaktion möglich?**
Allenfalls optisch, durch Sprache oder durch einen Fuß. Ich benutze immer meinen rechten Fuss zum Strahlen, der linke hätte quasi noch Platz für eine Aufgabe. Ansonsten ist Sprache sicher eine gute Sache.
17. **Kann dies auch von anderen Personen ausgeführt werden?**
Das würde in manchen Situationen zu lange dauern. Wenn ich der MTRA erst erklären will, was ich jetzt vergrößert haben möchte.
18. **Wäre eine berührungslose Steuerung denkbar?**
Vielleicht via Eyetracking oder eben Sprache.
19. **Wäre eine Sprachsteuerung denkbar?**
Ja.
20. **Wie schätzen Sie die Lernbereitschaft und Offenheit gegenüber neuen Technologien und Interaktionstechniken im Kollegium ein? Gibt es generationsbedingte Differenzen?**
Da bin ich mir sicher, dass es die gibt. Generell gibt es wohl nur eine geringe Bereitschaft Dinge zu Knern, in denen man sonst schon sehr gut ist. Für die geplanten Interventionen sind aber eh immer maximal zwei Leute pro halbes Jahr in der DSA eingeteilt, sodass ich das quasi erstmal alles testen könnte.
21. **Welche Erfahrung bestehen aktuell im Gebiet der berührungslosen Interaktion?**
Wir haben schon ein paar Sachen mit Hatscher, Mewes und Julian gemacht.
22. **Wie schätzen Sie die Relevanz der Thematik, sowie die technischen Möglichkeiten in Bezug auf berührungslose Interaktion ein?**
Definitiv ein spannender Punkt und einer der wenigen Punkte, an denen man sich vorstellen könnte, die Prozeduren noch zu verbessern.

Ablaufplan

Vor Start: Checkliste Materialien

- Informationen & Einwilligungen
 - Demografische Angaben
 - 4x Nasa TLX
 - Protokoll
1. Begrüßung, Vorstellung, Kontrolle Teilnehmerliste
 2. Information zur Studie, Forschungsvorhaben und Ziel
 3. Einwilligungserklärung, Rücktritt, Datenschutz, Fragen des Probanden
 4. Erfassung Demografische Angaben
 - a. Währendessen Zuweisung Versuchsreihe
 5. Instruktion, Vorstellung Prototypen und Aufgaben
 - a. Simulation eines OP-Szenarios
 - b. Sekundäraufgabe: Bedienung, Bildmanipulation, Zooming > so schnell wie möglich
 - c. Primäraufgabe: Katheterführung > so genau wie möglich
 6. Ergonomische Einrichtung des Systems und Kalibrierung individueller Parameter
 - a. Monitorhöhe
 - b. Tischhöhe
 - c. Instrument links/rechts
 - d. Kalibrierung Augenbrauen
 - e. Kalibrierung Cursorhöhe
 7. Messung Baseline bei Primäraufgabe
 8. Übung Kopfrichtungs- und Spracheingabe, Gesten
-
9. Vergabe 1 von 4 Kombination
 10. 2-3 Trainingsdurchläufe pro Kombination
 11. Frage ob sich Proband bereit fühlt
 12. Messung von drei Versuchen und Erfassung der Daten (Zeit, Auszeit) im Protokoll
 13. Fragebogen NasaTLX für jeweilige Kombination
 14. Kurze Pause zur eventuellen Einrichtung/Neukalibrierung
 15. Wiederholung der Punkte 9 bis 14 bis alle Durchläufe abgeschlossen
-
16. Nachbefragung: Empfindung, Stimmung, Schwierigkeiten, Interesse, Wirkung
 - a. Wie ging es dir im Versuch? Was ist dir aufgefallen? Was war gut? Was war schlecht?
 17. Verabschiedung

Informationen zur Studie

Sehr geehrter Teilnehmer, sehr geehrte Teilnehmerin,
wir freuen uns, dass Sie bereit sind, an unserem Forschungsvorhaben teilzunehmen. Ich werde Ihnen jetzt einige Informationen zur Studie geben und Sie können dann anschließend entscheiden, ob Sie an der Studie teilnehmen möchten oder nicht. Bitte zögern Sie nicht, nachzufragen, wenn Sie etwas nicht verstehen oder wenn Sie etwas zusätzlich wissen möchten.

Ziel und Auftraggeber der Studie

Dieses Forschungsvorhaben wird im Zuge einer Masterarbeit im Bereich Interaction Design an der Hochschule Magdeburg-Stendal sowie für eine danach geplante wissenschaftliche Veröffentlichung der Arbeitsgruppe Computer-Assisted-Surgery der Fakultät für Informatik der Otto-von-Guericke-Universität durchgeführt. Ziel dieser Arbeit ist, die Exploration freihändiger Interaktionsmuster, zur Bedienung von Benutzeroberflächen an einem Angiographiesystem. Die vorliegende Studie untersucht dabei die Anwendbarkeit, Intuitivität und mögliche Benutzertauglichkeit des erstellten Prototypen, um diesen bewerten und verbessern zu können. Die Erkenntnisse dienen der Weiterentwicklung dieses Systems, mit dem Ziel des zukünftigen Einsatzes für eine radiologische Intervention.

Ablauf der Studie

Nach Ihrer Einwilligung zur Studie werden zu Beginn Ihre demografischen Daten erfasst, woraufhin Ihnen der Prototyp erläutert wird. Sie werden ein multimodales System bedienen, welches, je nach Versuchsreihe, Primär- und Sekundäraufgaben beinhaltet. Die Primäraufgabe umfasst die Ausübung einer Tätigkeit, welche die Situation während einer Katheterintervention simulieren soll. Während der Sekundäraufgabe werden vier verschiedene Modalitäten zur Benutzung des Systems zur Verfügung gestellt, woraufhin jeweils eine Kombination aus zwei verschiedenen Interaktionstechniken evaluiert wird. Der Untersuchungsleiter weist Sie in die verschiedenen Aufgaben ein, kalibriert das System auf Ihre Parameter und lässt Sie zunächst die verschiedenen Interaktionstechniken trainieren. Anschließend werden drei Durchgänge zur Datenerhebung pro Interaktionstechnikombination mit oder ohne Primäraufgabe protokolliert. Nach jedem Durchgang folgt ein Fragebogen in Form des NASA Task Load Index. Abschließend erfolgt eine kurze informelle Nachbefragung. Insgesamt dauert die Studie maximal 50 min.

Haben Sie noch Fragen?

Sollten Sie noch weitere Fragen zum Studienablauf haben, so zögern Sie nicht, diese zu stellen. Wir sind Ihnen dankbar, wenn Sie sich dazu bereit erklären, an dieser Studie teilzunehmen. Für weitere Fragen stehen wir Ihnen jederzeit gerne zur Verfügung: Danny Schott, Mail: d.schott90@gmail.com, Telefon: 015209262072

Rücktritt und Datenschutz

Die Teilnahme an dieser Studie ist freiwillig. Sie können jederzeit ohne Angabe von Gründen die Teilnahme beenden. Grundsätzlich werden alle während der Studie erstellten Aufzeichnungen und Daten gemäß den geltenden Datenschutzrichtlinien vertraulich behandelt. Das heißt, Ihre persönlichen Daten und Studienergebnisse werden Zugriffssicher aufbewahrt und ausschließlich anonymisiert und damit ohne Bezug zu Ihrer Person verantwortungsvoll ausgewertet. Sollten Sie von der Studie zurücktreten, werden Ihre bereits erhobenen Daten anonymisiert weiter genutzt, jedoch keine weiteren erhoben.

Einwilligung zur Studie

Ich bestätige hiermit, dass ich durch den Untersuchungsleiter mündlich über Wesen, Bedeutung, Risiken und Tragweite der beabsichtigten Studie aufgeklärt wurde und für meine Entscheidung genügend Bedenkzeit hatte. Mir wurden die „Informationen zur Studie“ erklärt, ich fühle mich ausreichend informiert und habe verstanden, worum es geht. Der Untersuchungsleiter hat mir ausreichend Gelegenheit gegeben, Fragen zu stellen, die alle für mich ausreichend beantwortet wurden. Ich hatte genügend Zeit mich zu entscheiden.

Datenschutz

Ich habe verstanden, dass bei wissenschaftlichen Studien persönliche Daten erhoben werden. Die Weitergabe, Speicherung und Auswertung dieser studienbezogenen Daten erfolgt nach gesetzlichen Bestimmungen und setzt vor Teilnahme an der Studie meine freiwillige Einwilligung voraus: Ich erkläre mich damit einverstanden, dass im Rahmen dieser Studie erhobene Daten auf Fragebögen und elektronische Datenträger aufgezeichnet und ohne Namensnennung weitergegeben werden an die Universität Magdeburg, Universitätsplatz 2, 39106 Magdeburg und im Rahmen einer wissenschaftlichen Veröffentlichung publiziert werden.

Freiwilligkeit der Teilnahme

Meine Einwilligung, an diesem Forschungsvorhaben als Proband/in teilzunehmen, erfolgt ganz und gar freiwillig. Ich wurde darauf hingewiesen, dass ich meine Einwilligung jederzeit ohne Angaben von Gründen widerrufen kann, ohne dass mir dadurch irgendwelche Nachteile entstehen. Ich erkläre mich einverstanden an der oben erläuterten Studie teilzunehmen.

_____	den _____	_____	Danny Schott
Ort	Datum	Name Proband/in (Druckschrift)	Name Untersuchungsleiter
_____	_____	_____	_____
Unterschrift Proband/in	_____	_____	Unterschrift Untersuchungsleiter

Demografische Angaben

Datum: _____ Uhrzeit: _____ ID: _____

- Alter:** _____ Jahre
- Geschlecht:** weiblich männlich divers
- Tätigkeitsstatus:** Schüler
 Student, im Studienfach: _____
 erwerbs- bzw. berufstätig als: _____
 Sonstiges: _____
- Dominante Hand:** links rechts

Ist ihr Sehvermögen beeinträchtigt?

- nein
 Fehlsichtigkeit (Allgemeine Sehschwäche)
 Farbsinnstörung (Bspw. Rot-Grün-Sehschwäche)
 Sonstiges: _____

Liegt eine Sprachstörung vor?

- nein
 leicht
 schwer

Wie gut beherrschen Sie die englische Sprache?

Gar nicht Sehr gut Muttersprache

Welche Erfahrungen haben Sie in folgenden Bereichen?

Fachwissen Mensch-Computer-Interaktion
 Keine Sehr erfahren

Gestensteuerung / Berührungslose Interaktion
 Keine Sehr erfahren

Gesichts-, Emotions oder Eyetracking
 Keine Sehr erfahren

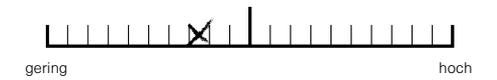
Sprachsteuerung
 Keine Sehr erfahren

NASA Task Load

ID: _____

Geben Sie jetzt für jede der unten stehenden Dimensionen an, wie hoch die Beanspruchung der jeweiligen Aufgabe war. Markieren Sie dazu bitte auf den folgenden Skalen, in welchem Maße Sie sich in den sechs genannten Dimensionen von der Aufgabe beansprucht oder gefordert gesehen haben.

Beispiel



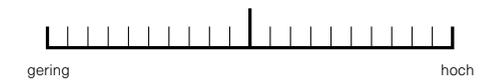
Geistige Anforderungen

Wie viel geistige Anstrengung war bei der Informationsaufnahme und -verarbeitung notwendig? (z.B. Entscheiden, Denken, Erinnern, Hinsehen, Suchen...) War die Aufgabe leicht oder anspruchsvoll, einfach oder komplex?



Körperliche Anforderungen

Wie hoch war der Grad der physischen Aktivität? War die Aufgabe leicht oder schwer, einfach oder anstrengend, erholsam oder mühselig?



Zeitliche Anforderungen

Wie viel Zeitdruck empfanden Sie hinsichtlich der Häufigkeit oder dem Takt, mit dem Aufgaben auftraten? War die Abfolge langsam und geruhsam oder schnell und hektisch?



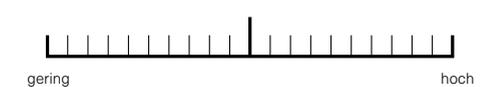
Leistung

Wie erfolgreich haben Sie Ihrer Meinung nach die vom Untersuchungsleiter (oder Ihnen selbst) gesetzten Ziele erreicht? Wie zufrieden waren Sie mit Ihrer Leistung bei der Verfolgung dieser Ziele?



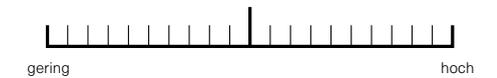
Anstrengung

Wie hart mussten Sie arbeiten, um Ihren Grad an Aufgabenerfüllung zu erreichen?



Frustration

Wie unsicher, entmutigt, irritiert, gestresst oder verärgert fühlten Sie sich während der Aufgabe?



Versuchsgruppe A – mit Primäraufgabe

ID	Reihe	D1	D1 Out	D2	D2 Out	D3	D3 Out	Baseline	G	K	Z	L	A	F
ID: 1	SL	01:27	00:50	01:13	00:28	00:56	00:23	00:36	13	8	9	7	11	7
	SA	00:58	00:29	01:00	00:21	01:14	00:44		10	15	10	5	15	10
	GL	01:16	00:40	01:03	00:40	00:55	00:19		11	15	11	10	18	10
	GA	01:32	00:40	01:26	00:34	01:33	00:35		14	17	10	10	16	11
ID: 3	SA	01:51	00:30	01:46	00:22	02:30	01:09	00:09	15	2	14	15	15	18
	SL	01:36	00:25	01:20	00:14	01:23	00:31		9	9	7	9	7	8
	GA	02:01	00:43	01:59	00:36	02:07	00:48		8	12	13	10	12	15
	GL	01:11	00:18	01:17	00:15	01:35	00:34		5	10	8	8	7	8
ID: 5	GL	01:26	00:25	00:56	00:14	01:35	00:32	00:47	13	8	10	5	8	3
	GA	01:28	00:54	01:08	00:26	01:27	00:18		11	5	8	5	11	9
	SL	01:20	00:23	00:55	00:19	01:16	00:34		6	4	5	3	6	2
	SA	02:20	01:22	02:40	01:21	01:28	00:55		11	5	13	9	9	11
ID: 7	GA	01:20	00:34	01:03	00:29	01:26	00:32	00:38	18	18	15	13	12	12
	GL	01:15	00:34	00:44	00:13	00:35	00:19		10	13	10	5	9	9
	SA	00:52	00:17	00:57	00:19	01:03	00:32		6	4	8	6	7	12
	SL	01:01	00:21	00:48	00:20	00:59	00:21		12	13	13	7	12	15
ID: 1	Alter	Geschlecht	Tätigkeit	Hand	Sehvermögen	Sprachstörung	Englisch	MCI	Gesten	Tracking	Sprachst.			
ID: 3	38	w	Wiss. Mit.	rechts	nein	nein	4	4	3	2	2			
ID: 5	29	m	St.	rechts	Farb- & Fehlsicht	nein	4	5	3	4	3			
ID: 7	23	m	St. DE	rechts	nein	nein	5	2	2	1	2			
ID: 7	25	m	St. ID	rechts	Fehlsicht	nein	3	4	4	3	4			

■ Erfasste Zeiten pro Durchlauf in Minuten, Out = Fehler

■ NASA TLX (Dimensionen als Abkürzungen)

■ Demografische Angaben, Fähigkeiten und Kenntnisse in Likert-Skala

Versuchsgruppe B – ohne Primäraufgabe

ID	Reihe	D1	D2	D3	D3	G	K	Z	L	A	F			
ID: 2	SL	01:34	01:06	01:03	01:03	4	17	13	11	13	8			
	SA	02:36	02:21	01:21	01:21	4	19	18	16	15	15			
	GL	01:09	00:58	00:57	00:57	3	12	12	11	10	10			
	GA	01:36	01:11	01:06	01:06	3	18	17	15	17	14			
ID: 4	SA	01:10	01:11	01:12	01:12	1	5	3	3	9	5			
	SL	01:12	00:48	01:12	01:12	4	5	4	10	11	10			
	GA	01:29	01:55	01:17	01:17	4	7	6	16	14	16			
	GL	00:48	00:49	00:40	00:40	3	6	5	10	7	12			
ID: 6	GL	00:44	00:59	00:37	00:37	5	6	7	9	10	10			
	GA	02:02	01:15	01:56	01:56	4	8	10	11	11	11			
	SL	01:30	00:40	00:49	00:49	8	3	7	5	7	6			
	SA	00:57	00:45	00:37	00:37	6	4	9	6	5	3			
ID: 8	GA	01:04	00:54	01:16	01:16	7	11	14	7	15	10			
	GL	00:49	00:55	00:43	00:43	5	6	11	3	6	4			
	SA	00:45	00:43	01:06	01:06	9	6	6	13	10	13			
	SL	00:51	00:56	00:40	00:40	8	3	3	5	3	3			
ID: 2	Alter	Geschlecht	Tätigkeit	Hand	Sehvermögen	Sprachstörung	Englisch	MCI	Gesten	Tracking	Sprachst.			
ID: 4	35	w	Wiss. Mit.	rechts	nein	nein	5	5	3	3	3			
ID: 6	22	w	St. MT	rechts	Fehlsichtigkeit	nein	3	2	2	2	1			
ID: 8	29	m	Selbstst.	rechts	Fehlsichtigkeit	nein	3	3	3	3	2			
ID: 8	24	m	St. ID	rechts	Fehlsichtigkeit	nein	3	2	2	2	1			

■ Erfasste Zeiten pro Durchlauf in Minuten, Out = Fehler

■ NASA TLX (Dimensionen als Abkürzungen)

■ Demografische Angaben, Fähigkeiten und Kenntnisse in Likert-Skala

ID: A & B mit Primäraufgabe, ID: C & D nur Sekundär

ID: A	Reihe	D1	D1 Out	D2	D2 Out	D3	D3 Out	Baseline	G	K	Z	L	A	F
ID: A	SL	01:10	00:27	00:54	00:16	01:17	00:22	00:14	6	8	11	6	12	6
	SA	02:48	01:27	01:15	00:28	01:01	00:26		7	7	7	15	17	17
	GL	01:08	00:19	01:14	00:25	00:57	00:26		7	9	7	7	11	8
	GA	01:51	01:08	01:29	00:32	00:38	00:15		7	9	7	4	9	5
ID: B	SA	01:13	00:12	01:26	00:32	01:23	00:31	00:11	12	10	11	11	7	11
	SL	02:02	01:06	01:24	00:37	01:16	00:32		14	13	14	11	9	3
	GA	01:21	00:39	01:33	00:55	01:04	00:22		17	13	8	16	13	15
	GL	01:48	00:42	00:57	00:13	01:24	00:51		17	10	10	6	10	11
ID: A	Alter	Geschlecht	Tätigkeit	Hand	Schvermögen	Sprachstörung	Englisch	MCI	Gesten	Tracking	Sprachst.			
	21	m	St.	links	nein	nein	4	3	3	3	1			
ID: B	21	w	St.	rechts	nein	nein	4	4	4	2	4			

ID: C	Reihe	D1	D2	D3	G	K	Z	L	A	F				
											D1	D2	D3	G
ID: C	SL	01:30	01:03	01:38	15	13	13	4	11	15				
	SA	01:26	00:55	01:43	19	17	14	8	18	18				
	GL	01:57	00:42	00:30	12	13	14	17	19	19				
	GA	00:38	00:44	00:33	7	9	9	5	10	6				
ID: D	GA	01:46	01:42	00:45	15	17	15	15	17	17				
	GL	01:03	00:52	00:36	7	10	10	5	8	6				
	SA	00:52	00:39	00:37	14	16	14	9	12	10				
	SL	00:55	00:36	00:43	9	10	8	9	9	11				
ID: C	Alter	Geschlecht	Tätigkeit	Hand	Schvermögen	Sprachstörung	Englisch	MCI	Gesten	Tracking	Sprachst.			
	23	m	St.	links	nein	nein	4	4	3	2	3			
ID: D	29	m	WiMi	rechts	fehl, fehl	nein	4	4	4	2	2			

■ Erfasste Zeiten pro Durchlauf in Minuten, Out = Fehler

■ NASA TLX (Dimensionen als Abkürzungen)

■ Demografische Angaben, Fähigkeiten und Kenntnisse in Likert-Skala