

# ARGUS VISION

---

Design und Evaluation eines Tracking Tools  
für Ausstellungsgestaltende

Bachelorarbeit

# ARGUS VISION

---

Design und Evaluation eines Tracking Tools  
für Ausstellungsgestaltende

Bachelorarbeit

Vorgelegt von:

An der:

Sektion:

Fachbereich:

Erstgutachter:

Zweitgutachter:

Moritz Skowronski

Universität Konstanz

Mathematisch- Naturwissenschaftliche Sektion

Informatik & Informationswissenschaft

Prof. Dr. Harald Reiterer

Jun.-Prof. Dr. Bela Gipp

Konstanz, Juni 2016

### **Zusammenfassung**

Zeitgenössische Ausstellungen werden zunehmend mittels komplexer, meist interaktiver Medien inszeniert. Um derartige Ausstellungen konzipieren und umsetzen zu können, setzen sich Ausstellungsgestaltungsbüros aus Vertreter\*innen einer Vielzahl unterschiedlicher gestalterischer und technischer Disziplinen zusammen. Die Unterstützung von transdisziplinär, also disziplinübergreifend arbeitenden Ausstellungsgestaltenden im Design-Prozess ist ein Ansatz in der Forschung zur Mensch-Computer-Interaktion. Dieser beinhaltet die Bereitstellung von Do It Yourself Tools, welche die Entwicklung interaktiver Medien für sämtliche beteiligten Disziplinen ermöglichen. In dieser Arbeit wird das Design eines solchen Do It Yourself Tools – Argus Vision – beschrieben und das Tool evaluiert. Mit Argus Vision lässt sich Kamera-Tracking mit Tiefenkameras ohne Programmierung umsetzen. Dazu werden sämtliche Funktionen des Tracking Tools in einer Bedienoberfläche gebündelt. In dieser lassen sich zusätzlich virtuelle Bereiche im Raum, sogenannte Triggerzones, definieren, die bei Berührung durch Personen aktiviert werden. Damit lassen sich schnell einfache Interaktionskonzepte testen und umsetzen. Um zu prüfen, inwiefern Argus Vision sich in interaktive Medien einbinden lässt, wurde das Tool in einer Installation im Rahmen einer Ausstellung genutzt sowie an ein bereits bestehendes Tool, welches in Ausstellungen angewendet wird, angebunden. In einem weiteren Schritt wurden Interviews mit Ausstellungsgestaltenden unterschiedlicher Disziplinen geführt. Diese bewerteten die Bedienoberfläche, deren Verständlichkeit sowie das Triggerzone-Prinzip positiv und konnten sich vorstellen, dass kleinere Ausstellungsgestaltungsbüros Argus Vision einsetzen würden. Auf Basis der Interviews konnten gleichzeitig zahlreiche Weiterentwicklungsmöglichkeiten für das Tracking Tool ausgearbeitet werden.

### **Abstract**

Contemporary exhibitions are increasingly staged using complex and often interactive media. In order to create such exhibitions exhibition design firms employ professionals from a wide range of different design and engineering disciplines. The support of transdisciplinary exhibition designers in their design process is one goal in recent human-computer interaction research. This includes the deployment of Do It Yourself Tools that enable all disciplines involved to design and create interactive media themselves. In this work the design and evaluation of Argus Vision is presented. Argus Vision is a Do It Yourself Tool which allows exhibition designers the use of camera-tracking using multiple depth cameras. All functions of Argus Vision are bundled in a user interface therefore eliminating the need for writing code. Additionally, the user can test and create simple interaction concepts by creating virtual areas, so called Triggerzones, which get activated when visitors come in contact with them. To prove that Argus Vision can be used in interactive media, the Tracking Tool was both used in an installation at an exhibition and to control another Do It Yourself Tool for exhibitions. Furthermore, interviews with exhibition designers from various disciplines were conducted. Especially the user interface, its comprehensibility and the Triggerzone-concept were rated positively by the interviewees. The interviews are a source from which it was possible to deduce a large amount of ideas for future work on Argus Vision.

## Inhaltsverzeichnis

Zusammenfassung	I
Abstract	II
Inhaltsverzeichnis	III
Abkürzungsverzeichnis	V
Abbildungsverzeichnis	VI
Tabellenverzeichnis	VIII
<b>1 EINLEITUNG</b>	<b>1</b>
<b>2 THEORIEBASIS</b>	<b>3</b>
2.1 Ausstellungen und Ausstellungsgestaltung .....	3
2.2 Mensch-Computer-Interaktion und Ausstellungsgestaltung.....	5
2.3 Tracking-Technologien in Ausstellungen.....	6
2.3.1 Einsatz von IKT mit Tracking-Technologien am realen Objekt.....	8
2.3.2 Einsatz von IKT mit Tracking-Technologien als eigenständiges Exponat .....	9
2.3.3 Einsatz von IKT mit Tracking-Technologien als Mittel zur Narration und Navigation .....	11
<b>3 ANFORDERUNGSANALYSE</b>	<b>12</b>
3.1 Erhobene Anforderungen.....	12
3.2 Tracking Tools.....	15
3.3 Tabellarischer Vergleich: Tracking Tools und Anforderungen.....	17
<b>4 ARGUS VISION</b>	<b>18</b>
4.1 Architektur von Argus Vision .....	19
4.2 Verwendung und Oberfläche von Argus Vision.....	20
4.2.1 Argus Kinect.....	21
4.2.2 Argus Control .....	22
4.3 Implementierung von Argus Vision .....	30
4.3.1 Tracking in Argus Vision .....	31
4.3.2 Kommunikation in Argus Vision.....	35
4.4 Zusammenfassung .....	37
<b>5 EVALUATION VON ARGUS VISION</b>	<b>39</b>
5.1 Fallstudie – Installation „Bruch“ .....	39
5.1.1 Die Installation „Bruch“ .....	40
5.1.2 Umsetzung der Installation .....	41
5.1.3 Ergebnisse aus der Anwendung von Argus Vision.....	44

---

5.1.4	Verbesserungsmöglichkeiten für Argus Vision .....	44
5.2	Fallstudie – Anbindung von Argus Vision an smartPerform.....	45
5.2.1	Die Präsentation in smartPerform.....	45
5.2.2	Integration von Argus Vision .....	46
5.2.3	Ergebnisse aus der Anwendung von Argus Vision.....	47
5.2.4	Verbesserungsmöglichkeiten für Argus Vision .....	48
5.3	Interviews mit Ausstellungsgestaltenden.....	50
5.3.1	Durchführung der Interviews.....	50
5.3.2	Analyse der Interviews .....	51
5.3.3	Ergebnisse der Interviews.....	51
5.4	Zusammenfassung der Ergebnisse.....	54
<b>6</b>	<b>FAZIT UND AUSBLICK</b>	<b>55</b>
	<b>LITERATURVERZEICHNIS</b>	<b>57</b>
	<b>ANHANG</b>	<b>61</b>
	Interview-Leitfaden	62
	Interview mit Dominik Hegemann	63
	Interview mit Sebastian Oschatz	69
	Interview mit Prof. Eberhard Schlag	80
	Interview mit Christoph Diederichs	88
	Interview mit Prof. Thomas Hundt	94
	USB-Stick	103

---

## Abkürzungsverzeichnis

2D .....	Zweidimensional
3D .....	Dreidimensional
CHP.....	Cultural Heritage Professional
DIY .....	Do It Yourself
IKT.....	Informations- und Kommunikationstechnologien
MCI.....	Mensch-Computer-Interaktion

## Abbildungsverzeichnis

Abbildung 1: Organigramm von Ralph Appelbaum Associates (nach Hughes 2010).....	4
Abbildung 2: Die fünf proxemischen Dimensionen (Quelle: Greenberg et al. 2011).....	6
Abbildung 3: a) Eine Skeletterkennung ermöglicht die Interaktion an der Medienstation (Quelle: <a href="https://artcom.de/wp-content/uploads/2015/02/2014_micropia_07-1189x765.jpg">https://artcom.de/wp-content/uploads/2015/02/2014_micropia_07-1189x765.jpg</a> , Abgerufen am 10.06.16). b) In der Installation ICE wird nur die Kontur der Interagierenden abstrakt repräsentiert (Quelle: Bullivant 2005).....	9
Abbildung 4: a) Die Position der Besuchenden beeinflusst die Bodenprojektion (Quelle: Snibbe & Raffle 2009). b) Die Beleuchtung der LED-Platten wird über die Position der Besuchenden gesteuert (Quelle: Gestalten & Hanschke 2011). ....	10
Abbildung 5: Die Bewegung und Position der Besuchenden wird durch eine Bodenprojektion dargestellt (Quelle: avedition 2015). ....	11
Abbildung 6: Die Kommunikations-Struktur von Argus Vision. ....	20
Abbildung 7: Die Bedienoberfläche von Argus Kinect. ....	21
Abbildung 8: Die Bedienoberfläche von Argus Control. ....	22
Abbildung 9: Ausschnitt des Bereichs I aus Abbildung 8: Die Kommunikations-Einstellungen in Argus Control. ....	23
Abbildung 10: Ausschnitt des Bereichs II aus Abbildung 8: Der Kinect-Bereich von Argus Control. ....	24
Abbildung 11: Ausschnitt des Bereichs III aus Abbildung 8: Der Stream-View von Argus Control. ....	25
Abbildung 12: Der Tiefen- (a) und Infrarot-Stream (b) in Argus Control. Der Tiefen-Stream stellt die Tiefe des Raums in Graustufen dar. Der Infrarot-Stream liefert ein realitätsgetreues Bild des Raums.....	26
Abbildung 13: Der Tracking-Stream zeigt die Konturen erkannter Objekte im Raum (b). Ist das Tracking deaktiviert, so zeigt er Änderungen gegenüber dem leeren Raum in Weiß (a). ....	26
Abbildung 14: Im Point Cloud-Stream kann mit einer virtuellen Kamera der Raum dreidimensional erkundet werden. In a) ist eine Hand von vorne, in b) ist dieselbe Szene von hinten gezeigt.....	27
Abbildung 15: Die Triggerzone-Einstellungen in Argus Control. ....	28
Abbildung 16: Drei Triggerzones sind auf einer Tischfläche platziert. ....	29
Abbildung 17: Die Arbeitsschritte im Trackingzyklus pro Kinect-Bild. ....	31



---

Abbildung 18: Der Trackingprozess in Argus Vision. In a) wird ein nicht kalibriertes, in b) ein kalibriertes Bild mit erfolgreicher Background Subtraction und in c) eine vollständig erkannte Kontur gezeigt. ....	32
Abbildung 19: Ein Beispiel für eine Distanzmatrix, welche aus drei neuen und zwei alten Personen errechnet wird. Farblich hervorgehoben sind die zueinander zugeordneten Personen. Person_Neu <sub>B</sub> ist eine neue Person. ....	34
Abbildung 20: Das Format der von Argus Control übermittelten Nachrichten. ....	36
Abbildung 21: Eine Besucherin steht vor einer der drei Stoffseiten der Stoffvitrine – das projizierte Glas ist zerbrochen. Im Vordergrund ist die reale Vitrine zu sehen. ....	40
Abbildung 22: Die technische Planung der Installation. ....	41
Abbildung 23: Eine Person bewegt sich auf eine Gaze zu und bringt dadurch das Glas zunehmend zum Zerschlagen. ....	42
Abbildung 24: Der Tiefen- und Tracking-Stream einer Kinect in der Installation. Hier schlägt eine Person auf die Gaze ein, um das Glas zerspringen zu lassen. ....	43
Abbildung 25: Risse bilden sich an der x-Position der Centroids der zwei Besuchenden. ....	43
Abbildung 26: Verarbeitung der Argus Vision-Daten in der Animation. ....	44
Abbildung 27: a) Ein Mann steht in einem, mit einem Kreuz markierten Bereich und betrachtet die präsentierten Bilder. Eine Frau bewegt sich in Richtung eines Kreuzes. b) Durch das Betreten der Fläche um das Kreuz werden nun auch an ihrer Position Bilder gezeigt. ....	46
Abbildung 28: Skizzenhafter Aufbau der Fallstudie. ....	46
Abbildung 29: a) Eine Frau zeigt auf ein präsentiertes Bild. b) Dieselbe Szene im Trackingview von Argus Control. ....	47
Abbildung 30: Die Aufgaben der einzelnen Tools in der Fallstudie. ....	48
Abbildung 31: Schematischer Ablauf des Interviewprozesses. ....	51

---

## **Tabellenverzeichnis**

Tabelle 1: Die Anforderungen an ein Tracking Tool für Ausstellungsgestaltende.....	14
Tabelle 2: Die bereits existierenden Tracking Tools und inwiefern diese die Anforderungen an ein Tracking Tool für Ausstellungsgestaltende erfüllen.....	17
Tabelle 3: Argus Vision im Vergleich mit den in Abschnitt 3.2 vorgestellten Tracking Tools. ....	38

# 1 Einleitung

Ausstellungsgestaltung ist heute ein interdisziplinäres Feld. Architekt\*innen und Kurator\*innen arbeiten eng mit Medien-, Produkt- und Lichtgestalter\*innen, Grafiker\*innen und Informatiker\*innen zusammen, um moderne Ausstellungen zu konzipieren und umzusetzen (Hughes 2010). Dabei ist der Begriff Ausstellung längst nicht mehr an die Institution Museum gekoppelt (Teufel 2001). Zu den Aufgabenbereichen eines Ausstellungsgestalters\*iner Ausstellungsgestalterin zählt heute zusätzlich zur klassischen musealen Dauerausstellung die Gestaltung von Wechsel- und Sonderausstellungen, Messe-Events, Einkaufszentren und Besucher\*innenzentren (Lorenc et al. 2007)<sup>1</sup>.

Insbesondere das Aufkommen der Informations- und Kommunikationstechnologien (IKT) hatte im zwanzigsten Jahrhundert starken Einfluss auf kontemporäre Ausstellungen (Barthelmes & den Oudsten 2011). In Science-Centern oder „hands-on“-Ausstellungen präsentierten Ausstellungsgestaltende Ende der 1960er Jahre erstmalig Exponate, mit welchen der\*die Besucher\*in interagieren konnte (Oppenheimer 1976). Besucher\*innen wurden so von reinen Betrachtenden zu Agierenden, die selbst Teile der Ausstellung steuern und verändern konnten (Hughes 2010; Snibbe & Raffle 2009). Dieser Entwicklung folgten weitere Ausstellungsformen in den 1980er Jahren. Mit der Einsicht, dass auch ein weiß gestrichener, ansonsten leerer Raum Ausstellungsstücke in einen spezifischen Kontext stellt, wurde die Szenografie – die bewusste, kontextualisierende Inszenierung des Raums durch den Einsatz von Medien wie Licht, Grafik und Ton – in der Ausstellungsgestaltung zunehmend eingesetzt (Barthelmes & den Oudsten 2011).

Die Durchdringung des Alltags durch die IKT und deren Verwendung in Ausstellungen stellt Ausstellungsgestalter\*innen heutzutage vor neue Herausforderungen. Um die Attraktivität von Ausstellungen für eine neue Generation von potentiellen Besuchenden aufrecht zu erhalten, werden diesen eine zunehmende Zahl interaktiver Exponate, aufwendiger Medieninstallationen und personalisierter Inhalte präsentiert (Dernie 2006; Hughes 2010; Simon 2010).

Ein enges Zusammenarbeiten der im Ausstellungsgestaltungsprozess involvierten Disziplinen ist deshalb unumgänglich, um eine in sich stimmige, konsistente Ausstellung zu entwickeln. Dies wird auch in der Fachliteratur oft gefordert (Borchers 2000; Hughes 2010; Teufel 2001). Dennoch werden Spezialist\*innen, wie beispielsweise Medienproduzierende, bis heute häufig erst in der Realisierungsphase eingebunden (John 2001; Schwarz 2001). Insbesondere in kleineren und mittelgroßen Ausstellungen, deren Budget und Personal begrenzt ist, ist deshalb ein transdisziplinärer Ansatz gefragt. Ein Ansatz also, bei dem die unterschiedlichen Disziplinen nicht nur im konstanten Austausch stehen, sondern Ausstellungen disziplinübergreifend konzipiert und realisiert werden (Baur 2001; John 2001; Schaeffer & Lindell 2016).

---

<sup>1</sup> Aus diesem Grund wird in der vorliegenden Arbeit der Begriff Ausstellung stellvertretend für alle aufgezählten Ausprägungen verwendet. Ausgenommen davon sind lediglich Kunstmuseen und –ausstellungen, da diese in der Museumswelt einen gesonderten Platz einnehmen und Aussagen zur Inszenierung nicht durchgängig auf diese übertragbar sind (Barthelmes & den Oudsten 2011).

Eine Herausforderung ist dabei die Kommunikation zwischen den einzelnen beteiligten Disziplinen. Judith Oexle, frühere Chefin des Landesamts für Archäologie Sachsen, bringt die Problemstellung wie folgt auf den Punkt: „Kurator, Designer und Architekt sprechen Griechisch, Lateinisch und Hebräisch“ (Oexle 1999, o.S.). Für eine Kommunikationsbasis muss also eine möglichst große Wissens- und Fähigkeiten-Schnittmenge geschaffen werden (McDermott et al. 2013). Dies lässt sich nicht nur durch eine gemeinsame Sprache, sondern auch durch die Bereitstellung von Tools erreichen, die alle Disziplinen gleichermaßen bedienen und benutzen können. Insbesondere Do It Yourself (DIY) Tools wie Arduino<sup>2</sup> und Gadgeteer<sup>3</sup>, die speziell für Anwender\*innen mit wenig bis keinen Programmierkenntnissen entwickelt wurden, ermöglichen den Bau einfacher, interaktiver Prototypen und werden daher auch in der Ausstellungswelt angewendet (Maye et al. 2014).

Mit dem Design und der Evaluation des Tracking Tools *Argus Vision* setzt auch die vorliegende Arbeit an diesem Punkt an: *Argus Vision* ist ein Tool, welches die einfache und schnelle Verwendung von Kamera-Tracking in Ausstellungen ermöglicht. Dabei kann es zur Erstellung interaktiver Räume und Installationen sowie als Rapid Prototyping Tool verwendet werden. Die Zielgruppe sind Ausstellungsgestaltende – so bietet *Argus Vision* eine leicht zu bedienende grafische Oberfläche und vereinfacht die sonst aufwendige und programmierintensive Verwendung von Tracking in Ausstellungen.

In dieser Arbeit werden zunächst die interdisziplinäre Arbeitsweise von Ausstellungsgestaltenden sowie die Techniken beschrieben, mit welchen diese Ausstellungen inszenieren. In den Fokus rücken daraufhin einige Forschungsarbeiten der Mensch-Computer-Interaktion (MCI), in welchen das Ziel verfolgt wurde, Ausstellungsgestalter\*innen im Design-Prozess – unter anderem durch den Einsatz von DIY Tools – zu unterstützen. Um zu verstehen, wie ein DIY Tool zur Entwicklung von interaktiven Medien mit Tracking-Technologien eingesetzt werden kann, werden die möglichen Einsatzbereiche von diesen in Ausstellungen sowie für die Arbeit wichtige Konzepte der MCI präsentiert. In Kapitel 3 werden die Anforderungen, welche für die Entwicklung eines Tracking Tools erhoben wurden, vorgestellt und bereits bestehende Tracking Tools auf deren Erfüllung dieser untersucht. Darauf folgend werden die Architektur, die Bedienoberfläche sowie die Funktionen von *Argus Vision* ausführlich beschrieben. In einem weiteren Kapitel erfolgt die Evaluation des Tracking Tools anhand zweier Fallstudien sowie Expert\*inneninterviews mit Ausstellungsgestalter\*innen. Abschließend werden die Ergebnisse dieser Arbeit zusammenfassend präsentiert und ein Ausblick auf zukünftige Arbeit gegeben.

---

<sup>2</sup> <https://www.arduino.cc/>, Abgerufen am: 22.04.16

<sup>3</sup> <http://www.netmf.com/gadgeteer/>, Abgerufen am: 22.04.16

## 2 Theoriebasis

In diesem Kapitel werden zwei Ziele verfolgt: Es wird zum einen begründet, warum die Entwicklung eines Tracking Tools für Ausstellungsgestaltende sinnvoll ist, zum anderen werden für diese Arbeit relevante Konzepte der MCI vorgestellt. Dazu wird beginnend die Arbeit von Ausstellungsgestaltenden und wie diese moderne Ausstellungen inszenieren beschrieben. Darauf folgend wird der Ansatz, ein DIY Tool für den Bereich des Trackings in Ausstellungen zu entwickeln, in bereits bestehende Forschungsansätze in der MCI eingebettet. Abschließend wird dargelegt, wie Tracking-Technologien in Ausstellungen verwendet werden können und welche Messdaten dabei benötigt werden.

### 2.1 Ausstellungen und Ausstellungsgestaltung

“We conceive of exhibitions as ‘narrative environments’: ones in which the designer’s task is to translate stories and academic objectives that are often complex into a thought-provoking, spatial narrative.” (Kossmann & de Jong 2010, 7)

Diese Definition von Kossmann & de Jong, den Kreativdirektoren des Ausstellungsgestaltungsbüros Kossmann.dejong<sup>4</sup>, beschreibt den heutigen Anspruch an Ausstellungen. Längst lassen sich diese nicht mehr auf das bloße Ausstellen von Exponaten reduzieren. Vielmehr ist die Aufgabe einer modernen Ausstellung das „zugänglich Machen und Vermitteln“ (Bertson et al. 2012, 8) eines bestimmten Themas. Ob nun historische Artefakte oder Produkte eines Unternehmens präsentiert werden – Ziel ist die Einordnung dieser Objekte in ein für die Besuchenden verständliches, emotional involvierendes Narrativ. So generieren Ausstellungen ihre Bedeutung nicht mehr über ihre Exponate, sondern über deren Beziehung zueinander (Kossmann & de Jong 2010). Hiernach besteht die Aufgabe von Ausstellungsgestaltenden<sup>5</sup> mitunter darin, diese Beziehung für Besucher\*innen verständlich zu gestalten. Dabei liegt in Ausstellungen der Fokus nicht darauf, die Erwartungshaltung des breiten Publikums zu erfüllen, sondern Vermittlungsziele zu erreichen und die Betrachtenden emotional zu involvieren (Kilger 2011).

Unter diesen Gesichtspunkten lässt sich der seit den 1980er Jahren vollziehende Wandel von klassischen, kaum inszenierten Ausstellungen, von „bildungsbürgerlichen Tempeln“ (Derks 2000, 199) zu modernen Ausstellungen begreifen, welche mit einer Vielzahl unterschiedlicher Gestaltungsmittel inszeniert werden. Zu diesen zählen Mittel der Architektur, Innenarchitektur, Grafik, Mediengestaltung, Kunstgeschichte, Dramaturgie, Regie, des Bühnenbilds, Kommunikationsdesigns und Licht- und Produktdesigns (Barthelmes & Den Oudsten 2011). Diese Art der Raumgestaltung wird als *Szenografie* bezeichnet (ebda.). Zusätzlich werden den Besuchenden vermehrt interaktive Medien präsentiert (ebda.). Aufgrund dieser Vielzahl

---

<sup>4</sup> <http://www.kossmanndejong.nl>, Abgerufen am 23.04.16

<sup>5</sup> Ausstellungsgestaltung ist keine geschützte Berufsbezeichnung. In der Fachliteratur werden daher auch oft die Termini Museologie oder Szenografie benutzt. In dieser Arbeit wird sich auf den Begriff Ausstellungsgestaltung beschränkt.

von unterschiedlichen Inszenierungsmöglichkeiten sind Ausstellungsgestaltungsteams heute interdisziplinär aufgestellt (Hughes 2010). Die Zusammenarbeit aller beteiligten Disziplinen wird dabei von Beginn der Konzeptionsphase bis zum Abschluss der Realisation gefordert (Teufel 2001).

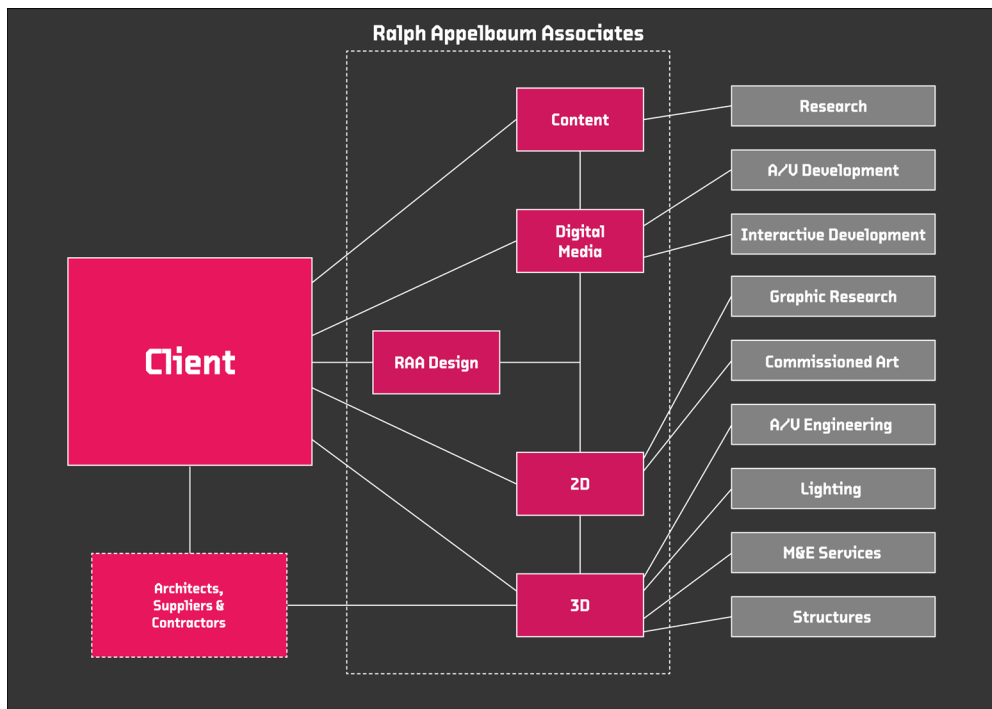


Abbildung 1: Organigramm von Ralph Appelbaum Associates (nach Hughes 2010)

Das Organigramm des Ausstellungsgestaltungsbüros Ralph Appelbaum Associates<sup>6</sup> (siehe Abbildung 1) zeigt beispielhaft die Menge an unterschiedlichen Planungsaufgaben eines Büros. Insbesondere im Kontext von Digital Media – wozu auch die Entwicklung interaktiver Medien zählt – ist zu erwähnen, dass in den meisten Fällen zwar die Planung dieser Medien in den Aufgabenbereich der Ausstellungsgestaltenden fällt, die Realisation dann aber zumeist an Medienproduktionsunternehmen weitergegeben wird (Schwarz 2001).

<sup>6</sup> <http://www.raany.com>, Abgerufen am 26.04.16

## 2.2 Mensch-Computer-Interaktion und Ausstellungsgestaltung

Interaktive Medien zeichnen sich insbesondere dadurch aus, dass mit ihnen spielerisch Wissen vermittelt und das Besuchererlebnis von Besuchenden verbessert werden kann (Wang 2015). In ihnen lassen sich ästhetische und spielerische Aspekte mit dem Ziel der Informationsvermittlung verknüpfen (Hassenzahl et al. 2003). Dementsprechend beschäftigt sich ein Großteil der Forschung in der MCI mit der Usability von interaktiven Medien in Ausstellungen und der User Experience der Besuchenden<sup>7</sup>.

Weitere Forschungsarbeiten beschäftigen sich mit der Unterstützung der Arbeit von Ausstellungsgestaltenden im Design-Prozess. Zwei dieser Arbeiten werden im Folgenden beschrieben. Dabei verfolgt die erste dieser Arbeiten einen interdisziplinären Ansatz, während die letztere einen transdisziplinären Ansatz vorstellt, welcher Argus Vision zu Grunde liegt.

Jan Borchers, Chef der Media Computing Group der RWTH Aachen, welche unter anderem zu interaktiven Exponaten forscht, formulierte die Notwendigkeit Pattern Languages (Mustersprachen) für die Arbeit in interdisziplinären Teams zu verwenden (Borchers 2000) und damit das von Judith Oexle formulierte Kommunikationsproblem (Oexle 1999) zu lösen. Jede Disziplin verwendet ihre eigene Mustersprache – diese beschreibt in einer Sammlung von Termini mögliche Lösungen für verschiedene Aufgabenstellungen dieser Disziplin (Borchers 2000). Die Kommunikation innerhalb eines interdisziplinären Teams wird so zum einen durch eine klar definierte Terminologie, zum anderen durch ein größeres Verständnis des Arbeitsprozesses der beteiligten Personen verbessert. Mittel der Kommunikation ist dabei die Sprache – die Aufgabenbereiche sind aber weiterhin klar getrennt (ebda.).

In dem Projekt Material EncounterS with digital Cultural Heritage (meSch)<sup>8</sup> hingegen wird von Kurator\*innen und anderen sogenannten Cultural Heritage Professionals (CHPs) gefordert, nicht nur bei der Konzeption, sondern auch bei der Umsetzung von Ausstellungen und deren Exponaten aktiv teilzunehmen (Petrelli et al. 2013). Das Ziel des Projekts ist das Design von berührbaren (tangible) Exponaten für Ausstellungen (ebda.). In co-design Workshops werden neben Focus Groups und Interviews auch Hardwaresketching und Low Fidelity Prototyping in interdisziplinären Teams durchgeführt (McDermott et al. 2014). Das in den Workshops erlernte technische Know-how soll CHPs nicht nur ermöglichen, ihre eigenen Ideen besser zu kommunizieren, sondern interaktive Ausstellungen vollständig selbst zu gestalten und umzusetzen (Maye et al. 2014; Petrelli et al. 2013). Dabei reicht das aus Workshops gewonnene Wissen aber nicht aus, um ein interaktives Exponat von Grund auf zu entwickeln (Wolf et al. 2015). Daher müssen einfach zu verwendende DIY Tools bereitgestellt werden, mit denen eine solche Eigenentwicklung ermöglicht wird (ebda.). Mit dem Arduino oder Microsoft .Net Gadeteer sind bereits eine Reihe kostengünstiger Hardwaretools erhältlich. Im Rahmen des meSch-Projekts wurden zusätzlich einige Softwaretools mit diesem Ziel entwickelt (ebda.).

---

<sup>7</sup> Beispiele hierfür finden sich unter anderem in Hesperhol et al. 2013, Hakvoort 2013 oder Snibbe & Raffle 2009

<sup>8</sup> <http://mesch-project.eu>, Abgerufen am: 29.04.16

## 2.3 Tracking-Technologien in Ausstellungen

In Abschnitt 2.2 wurden zwei Ansätze aufgezeigt, welche das Ziel verfolgen, die Arbeit von Ausstellungsgestaltenden im Design-Prozess zu unterstützen. Einer dieser Ansätze beinhaltet das Bereitstellen von DIY Softwaretools, um die eigenständige Entwicklung interaktiver Medien zu ermöglichen.

Die Wichtigkeit, ein DIY Softwaretool, insbesondere für den Bereich des Trackings in Ausstellungen zu entwickeln, lässt sich aus dem aktuellen Trend sogenannter immersiver Ausstellungen & Räume ableiten (Kilger 2011). In diesen können Besuchende durch Interaktion im Raum nicht nur einzelne interaktive Medien, sondern auch das szenografische Erscheinungs-, also das Raumbild selbst beeinflussen – der\*die Besucher\*in soll in die Ausstellung eintauchen (Buchanan 2008; Hakvoort 2013; Snibbe & Raffle 2009; Monaci et al. 2011).

Das zunehmende Interesse an derartigen Ausstellungskonzepten lässt sich auch durch die Verbreitung von ubiquitous computing (ubicomp) erklären. Der Begriff wurde 1991 von Mark Weiser geprägt und beschreibt seine damalige Vision eines Alltags, in dem Technologie sich nahtlos in diesen integriert (Weiser 1991). Mit Geräten wie Tablets und Smartphones, die viele Menschen täglich bei sich tragen, ist ein Teil dieser Vision tatsächlich Realität geworden. Zusätzlich beschreibt Weiser auch die Kommunikation all dieser Geräte untereinander und die automatische Anpassung dieser auf Nutzungssituationen (ebda.). So stellt sich ein Smartphone beispielsweise automatisch stumm, wenn der\*die Besitzer\*in in einem Konzert oder Kino ist.

Die Vision Mark Weisers lässt sich auf immersive Ausstellungen übertragen. Ein Exponat könnte beispielsweise nur dann zusätzliche Informationen zeigen, wenn ein\*e Besucher\*in sich diesem zuwendet oder eine gewisse Zeit daran verweilt; eine Raumbelichtung könnte sich je nach Anzahl der im Raum befindlichen Personen verändern.

Daten dieser Form, welche mit Tracking-Technologien gemessen werden können, werden als *proxemische* Daten bezeichnet (Ballendat et al. 2010). Der Begriff *Proxemics* (*Proxemik*), ursprünglich ein Begriff aus der Sozialforschung, beschreibt den Einfluss der räumlichen Beziehung von zwei Objekten auf ihre Kommunikation miteinander (Hall 1966). Diese proxemischen Daten lassen sich in fünf Kategorien unterteilen, welche in Abbildung 2 dargestellt sind (Greenberg et al. 2011).

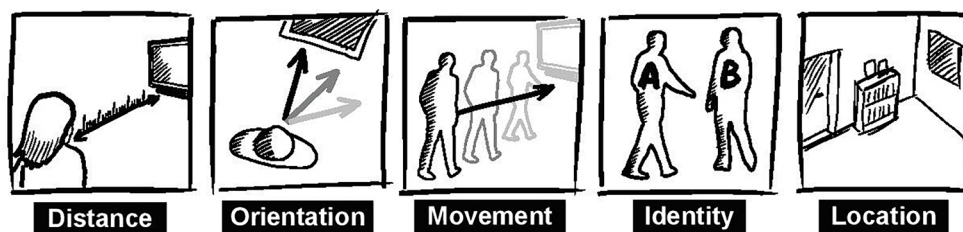


Abbildung 2: Die fünf proxemischen Dimensionen (Quelle: Greenberg et al. 2011)



Die *Distanz (Distance)* beschreibt den Abstand zwischen zwei Objekten, die *Orientierung (Orientation)*, in welchem Winkel diese zueinander stehen. Aus der Kombination der Distanz und Orientierung über eine bestimmte Zeitspanne ergibt sich die *Bewegung (Movement)*. Mit ihr lässt sich beispielsweise die Bewegungsgeschwindigkeit von Personen messen oder, ob diese sich von einem anderen Objekt weg- oder zu diesem hindrehen. Das eindeutige Unterscheiden von Objekten, also beispielsweise die Zuschreibung einer eindeutigen ID zu jedem Objekt, wird durch die Kategorie *Identität (Identity)* gekennzeichnet. Direkten Einfluss auf die vier zuvor genannten Kategorien hat der *Ort (Location)*. Mit dieser Kategorie wird der umgebende Raum beschrieben, in dem sich die Objekte befinden. Der Begriff Raum ist dabei mehr als geografischer Raum zu verstehen. Mit der Kategorie Ort kann also sowohl ein Zimmer als auch eine größere, öffentliche Fläche gemeint sein oder auch nur bestimmte virtuelle Bereiche in diesen. Die Messung der Distanz kann also, je nach Ort, anders erfolgen (Greenberg et al. 2011).

Um die fünf proxemischen Dimension messen zu können werden in Ausstellungen unterschiedliche Tracking-Technologien eingesetzt. Dazu zählen Tiefenkameras (avedition 2015), RFID (Cafaro et al. 2013), Bluetooth (Oschatz, S. 70), Infrarotsensoren (Monaci et al. 2011), eigens hergestellte Lösungen (Kuflik et al. 2011) oder eine Kombination aus mehreren Technologien (Cafaro et al. 2013). Der Einsatz unterschiedlicher Technologien ist dadurch bedingt, dass diese jeweils Vor- und Nachteile in spezifischen Situationen besitzen. So wird RFID häufig verwendet, um Besuchende über eine gesamte Ausstellung zu verfolgen und diesen personalisierte Inhalte zu bieten (Cafaro et al. 2013). Infrarotsensoren hingegen lassen sich für die Erkennung der Identität nicht einsetzen, da diese lediglich bei Bewegung ein elektrisches Signal senden (eine genauere Analyse der möglichen einsetzbaren Technologien findet sich in der dieser Arbeit vorausgegangenen Seminararbeit<sup>9</sup>).

Die proxemischen Daten, die eine Tracking-Technologie liefern muss, sind also insbesondere von deren Einsatzbereich abhängig. Diese Einsatzbereiche lassen sich Daniel Klinkhammer zufolge für die IKT, also auch für Tracking-Technologien, in vier Bereiche aufteilen – *Einsatz von IKT am realen Objekt*, *Einsatz von IKT als eigenständiges Exponat*, *Einsatz zur Kontextbildung zwischen Objekten* und *Einsatz zur Orientierung und Navigation* (Klinkhammer 2009). Da ein Pfad, also eine Navigation durch die Ausstellung oftmals dazu dient, Kontexte zwischen Objekten zu bilden, werden die letzten beiden Kategorien zum *Einsatz von IKT als Mittel zur Narration und Navigation* zusammengefasst. Tracking-Technologien funktionieren dabei allerdings nie als eigenständige IKT, da sie lediglich Daten über die Position oder andere Eigenschaften einer Person ausgeben – die Verarbeitung dieser Daten muss zusätzlich mithilfe anderer IKT erfolgen. Die Einsatzbereiche für Tracking-Technologien lassen sich also wie folgt kategorisieren:

- Einsatz von IKT mit Tracking-Technologien am realen Objekt
- Einsatz von IKT mit Tracking-Technologien als eigenständiges Exponat
- Einsatz von IKT mit Tracking-Technologien als Mittel zur Narration und Navigation

---

<sup>9</sup> Diese lässt sich dem der Arbeit beiliegenden USB-Stick entnehmen.

Diese Kategorien sind nicht trennscharf – eine Kontextbildung zwischen zwei Objekten kann auch an einem der beiden Objekte stattfinden und ließe sich demnach auch in die erste Kategorie einordnen.

Im Folgenden werden die Kategorien anhand von Anwendungsbeispielen aus aktuellen und früheren Ausstellungen beschrieben. Dabei wird herausgestellt, welche proxemischen Daten in den Anwendungsbeispielen verwendet wurden. Es werden nur solche Technologien als Tracking-Technologien bezeichnet, die Besuchende bereits ohne deren explizite Interaktion mit dem System erkennen. Eine Medienstation beispielsweise, an der sich Besuchende durch das Auflegen eines RFID-Chips anmelden können, wird dementsprechend nicht als IKT mit Tracking-Technologien bezeichnet.

### **2.3.1 Einsatz von IKT mit Tracking-Technologien am realen Objekt**

Unter dem Einsatz von IKT am realen Objekt versteht Klinkhammer die Anwendung dieser, um Besuchenden zusätzliche Informationen zu einem Exponat zu vermitteln (Klinkhammer 2009). In diesem Kontext werden Tracking-Technologien hauptsächlich als Lichtschranken verwendet, um beim Herantreten eines\*einer Besuchenden an ein Objekt zusätzliche Informationen visuell oder auditiv präsentieren.

So wurden beispielsweise in der Ausstellung *Mariko Mori – Oneness* auf den Boden vor bestimmten Kunstwerken Kreise gezeichnet. Betrat ein\*e Besucher\*in einen solchen Kreis, so wurde dies über Infrarotsensoren erkannt. Aus gerichteten Lautsprechern konnten die Besuchenden sodann zusätzliche Informationen zum vor ihnen befindlichen Kunstwerk hören (Kortbek & Grønbæk 2008).

Da in solchen Anwendungen lediglich die Präsenz eines\*einer Besuchenden oder dessen\*deren Distanz zum Objekt festgestellt werden ist hier nur die proxemische Dimension der Distanz von Relevanz. Komplexere Anwendungen können jedoch zusätzlich die Kategorien der Orientierung und Bewegung nutzen, um akkurater auf Besuchende zu reagieren, wenn beispielsweise einer dem Objekt weggedrehten Person keine Informationen gezeigt werden sollen.

### 2.3.2 Einsatz von IKT mit Tracking-Technologien als eigenständiges Exponat

Am häufigsten werden IKT mit Tracking-Technologien als eigenständige Exponate eingesetzt. Hier lassen sich sowohl klassische Medienstationen, interaktive Installationen als auch immersive Räume und Ausstellungen finden.

In einer Medienstation kombinieren Cafaro et al. eine Kinect und RFID, um mehreren Besucher\*innen gleichzeitig personalisierte Inhalte zu präsentieren (Cafaro et al. 2013). RFID wird dabei verwendet, um die Identität der Besuchenden festzustellen, die Kinect, um deren Distanz zur Medienstation zu bestimmen (ebda.).

Kamera-Tracking kann auch eingesetzt werden, um die Kontur von Personen zu erkennen oder eine Gestensteuerung zu ermöglichen. An einer Medienstation in der Ausstellung *Micropia* beispielsweise können Besuchende durch Gestensteuerung ihren eigenen Körper erkunden (vgl. Abbildung 3a). Dieser wird, mit zusätzlichen Informationen über verschiedene Körperteile versehen, an der Medienstation repräsentiert<sup>10</sup>. Oftmals wird jedoch nur die Kontur von Personen abstrakt repräsentiert, ohne eine komplexere Skeletterkennung zu verwenden. Hierbei wird, wie beispielsweise in der Installation *ICE*, zumeist nur ein ästhetisches, spielerisches Ziel verfolgt (vgl. Abbildung 3b) (Bullivant 2005).

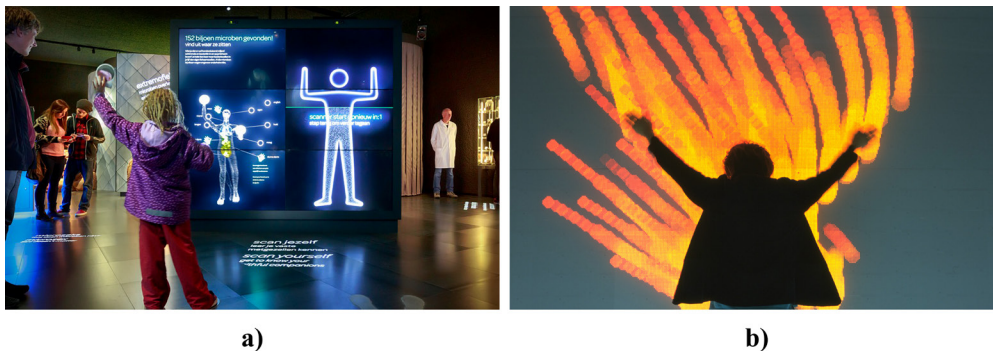


Abbildung 3: a) Eine Skeletterkennung ermöglicht die Interaktion an der Medienstation (Quelle: [https://artcom.de/wp-content/uploads/2015/02/2014\\_micropia\\_07-1189x765.jpg](https://artcom.de/wp-content/uploads/2015/02/2014_micropia_07-1189x765.jpg), Abgerufen am 10.06.16). b) In der Installation *ICE* wird nur die Kontur der Interagierenden abstrakt repräsentiert (Quelle: Bullivant 2005).

Derartige Exponate nutzen zwar die Distanz der Personen zur Installation, um die Interaktion ab einer bestimmten Nähe zu ermöglichen, die Skelett- oder Konturerkennung ist jedoch nicht durch die proxemischen Dimensionen abgedeckt.

<sup>10</sup> vgl. <https://artcom.de/project/micropia/>, Abgerufen am 10.06.16

In interaktiven Installationen wird häufig die Position der Besuchenden innerhalb eines virtuellen Interaktionsraums verwendet oder geprüft, ob diese einen bestimmten Bereich innerhalb des Interaktionsraums betreten haben.

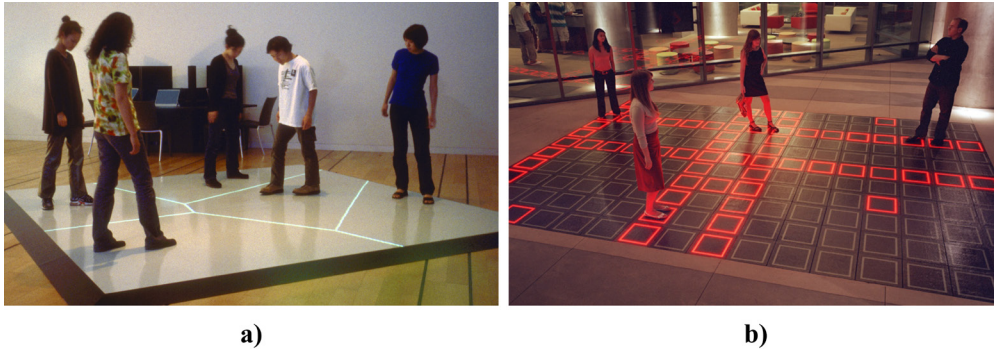


Abbildung 4: a) Die Position der Besuchenden beeinflusst die Bodenprojektion (Quelle: Snibbe & Raffle 2009).  
b) Die Beleuchtung der LED-Platten wird über die Position der Besuchenden gesteuert (Quelle: Gestalten & Hanschke 2011).

In der Installation *Boundary Functions* werden Voronoi-Diagramme um die Besuchenden, welche in einem Interaktions-bereich sind, gezeichnet (Snibbe & Raffle 2009). Die Position der Besuchenden bestimmt dabei die Konstruktion der Diagramme (vgl. Abbildung 4a). In *Interactive* hingegen wird die Beleuchtung der Flächen eines mit LED-Platten bestückten Bodens über die Position der Besuchenden gesteuert (vgl. Abbildung 4b) (Gestalten & Hanschke 2011). Während in *Boundary Functions* die Distanz der Besuchenden zu einer Kamera genutzt wird, um deren Position zu bestimmen<sup>11</sup>, werden in *Interactive* Gewichtssensoren verwendet, um festzustellen, ob ein Mensch auf einer Platte steht<sup>12</sup>. Die Messung erfolgt hier also binär – nach Greenberg et al. kann jede der LED-Patten als eigener virtueller Bereich angesehen werden (Greenberg et al. 2011).

<sup>11</sup> vgl. <http://www.snibbe.com/projects/interactive/boundaryfunctions/>, Abgerufen am 10.06.16

<sup>12</sup> vgl. <https://segd.org/interactive-11th-and-flower>, Abgerufen am 10.06.16



Abbildung 5: Die Bewegung und Position der Besuchenden wird durch eine Bodenprojektion dargestellt (Quelle: avedition 2015).

Immersive Räume und Ausstellungen nutzen zumeist die Position und Bewegung der Besuchenden im Raum, um mit diesen das Raumbild zu verändern. In *Time Machine* beispielsweise werden in einem Raum sowohl die Wand als auch der Boden durch eine Projektion inszeniert. Um Besucher\*innen das Gefühl zu geben, Teil der in der Projektion erzählten Geschichte zu sein, werden deren Position und Bewegungen abstrakt auf dem Boden widergespiegelt (avedition 2015).

### 2.3.3 Einsatz von IKT mit Tracking-Technologien als Mittel zur Narration und Navigation

Unter der Kontextbildung zwischen Objekten versteht Klinkhammer die, in Abschnitt 2.1 erläuterte, narrative Verknüpfung der Exponate einer Ausstellung (Klinkhammer 2009). Diese Verknüpfung kann Besuchenden bereits bei der Orientierung in der Ausstellung und der Navigation durch diese helfen. Umgekehrt kann auch eine klare Navigation durch eine Ausstellung Besuchende darin unterstützen, eine thematische Verknüpfung zwischen Objekten herzustellen.

In diesen Bereichen lassen sich kaum Beispiele für IKT mit Tracking-Technologien finden. Dies lässt sich durch die konzeptuelle sowie technische Schwierigkeit erklären, die Objekte thematisch miteinander zu verknüpfen, da die Reihenfolge, Anzahl und Auswahl der betrachteten Objekte zwischen den Besuchenden variiert. Diese Informationen müssen demnach für sämtliche Besuchenden gespeichert werden, um diesen beispielsweise personalisierte Inhalte in Abhängigkeit ihrer bisher betrachteten Objekte zu präsentieren.

## 3 Anforderungsanalyse

Aus der Theoriebasis wurden in der Seminar- und Projektarbeit<sup>13</sup> Anforderungen an ein Tracking Tool für Ausstellungsgestaltende abgeleitet. Diese werden in diesem Kapitel erneut aufgezeigt. Im Anschluss daran werden bereits bestehende Tracking Tools vorgestellt und geprüft, inwiefern diese die Anforderungen erfüllen.

### 3.1 Erhobene Anforderungen

In Abschnitt 2.2 wurden verschiedene Ansätze zur Unterstützung von Ausstellungsgestaltenden im Design-Prozess vorgestellt. Teil eines dieser Ansätze ist die Bereitstellung von DIY Tools, mit welchen die Umsetzung interaktiver Medien für transdisziplinär arbeitende Ausstellungsgestalter\*innen ermöglicht wird. Um ein solches DIY Tool für die Anwendung von Tracking in Ausstellungen zu schaffen, wurden in Abschnitt 2.3 Einsatzbereiche von Tracking anhand einiger Fallbeispiele aus existierenden Ausstellungen beschrieben und aufgezeigt, welche Daten der jeweiligen Tracking-Technologien in diesen verwendet werden.

Vor diesem Hintergrund werden nachfolgend die final erhobenen Anforderungen an ein Tracking Tool für Ausstellungsgestaltende vorgestellt und aufgezeigt, inwiefern diese aus der Theoriebasis abgeleitet wurden.

#### **A1. Auf Ausstellungsgestaltende ausgerichtete Bedienoberfläche**

Das Tool richtet sich an transdisziplinär arbeitende Ausstellungsgestaltende. Um allen Disziplinen, also auch Ausstellungsgestaltenden mit wenig oder keinen Programmierkenntnissen die Arbeit mit dem Tracking Tool zu ermöglichen, bietet dieses eine übersichtliche Bedienoberfläche mit begrenzten Einstellungsmöglichkeiten. Dies dient dazu, die Komplexität der Bedienung zu reduzieren und das Tool leicht verständlich zu gestalten. Der\*die Benutzer\*in wird zusätzlich die Auswirkungen jeder Interaktion mit der Bedienoberfläche auf dieser repräsentiert sehen.

#### **A2. Erkennen von Personen aus beliebigem Winkel**

Das Erkennen von Personen ist sicherlich die wichtigste technische Funktion, die ein Tracking Tool bieten muss. Da die meisten in Abschnitt 2.3 vorgestellten interaktiven Medien die Position der Besuchenden im Raum oder deren Kontur benötigten, wurde sich bereits in der Seminararbeit für die Verwendung von Tiefenkameras entschieden. Proprietäre Lösungen für das Erkennen von Personen – wie beispielsweise die Skeletterkennung der Microsoft Kinect – funktionieren dabei jedoch nur, wenn die Ausrichtung der Kinect parallel zum Boden ist, sodass Personen frontal gefilmt werden können, wenn sie im Bereich vor der Kamera stehen (Kamerabereich). Dass die Erkennung in einem Tracking Tool für Ausstellungsgestaltende aus jedem beliebigem Winkel funktioniert, ist in erster Linie den ästhetischen Ansprüchen in Ausstellungen geschuldet: Eingesetzte Hardware soll zumeist nicht sichtbar sein. Es

---

<sup>13</sup> Diese lässt sich dem der Arbeit beiliegenden USB-Stick entnehmen.

muss also möglich sein, die Kameras beispielsweise an der Decke anzubringen oder in einer Verkleidung zu verstecken.

### **A3. Erstellen von Triggerzones**

Die erfolgreiche Erkennung von Personen erfordert von Ausstellungsgestaltenden weiterhin, dreidimensionale Daten (3D-Daten) zu verstehen und mit diesen zu arbeiten. Diese Weiterverarbeitung der Daten kann erneut einen großen Programmieraufwand bedeuten. Mit dem Einsatz von sogenannten *Triggerzones* kann dieser für einfache interaktive Anwendungen jedoch minimiert werden. *Triggerzones* sind virtuelle Bereiche, welche sich im Kamerabereich erstellen und frei platzieren lassen und sich wie komplexe Lichtschranken verhalten. Betritt eine Person eine Zone, so erhält der\*die Benutzer\*in des Tracking Tools Zugriff auf verschiedene Daten, etwa, welche Person die Zone betreten hat.

Anwendungen, wie beispielsweise die Installation *Enteractive* aus Abschnitt 2.3.2 (vgl. Abbildung 4b), lassen sich so durch die Verwendung mehrerer *Triggerzones* leicht umsetzen. Zusätzlich können *Triggerzones* während der Prototyping-Phase eingesetzt werden, um in Prototypen einfach und effizient die Interaktionskonzepte zu testen, ohne komplexere Abläufe programmieren zu müssen.

### **A4. Steuerung von mehreren Tiefenkameras über ein zentrales Tool**

In immersiven Räumen oder größeren Installationen (vgl. *Time Machine* aus Abschnitt 2.3.2) werden mehrere Kameras benötigt, um den kompletten Interaktionsraum abzudecken. Diese Kameras werden über eine zentrale Software konfiguriert und überwacht, die gleichzeitig auch die Daten der einzelnen Kameras bündelt und an andere IKT weiterleitet.

### **A5. Automatisierter Start**

Das Tracking Tool kann in Ausstellungen über einen beliebigen Zeitraum ohne Wartung betrieben werden. Dies beinhaltet, dass das Tool so konfiguriert wird, dass es beispielsweise täglich startet und herunterfährt. Gleichzeitig werden einmal vorgenommene Einstellungen beim Neustart wieder geladen.

### **A6. Geeignetes Protokoll zur Übertragung der Daten**

In Abschnitt 2.3 wurde bereits aufgezeigt, dass Tracking-Technologien nie autonom, sondern immer in Kombination mit anderen IKT, die die Daten der Tracking-Technologien verwenden, arbeiten. Es wird also ein Protokoll zur Übertragung der Daten gewählt, das zum einen von möglichst vielen IKT unterstützt wird und zum anderen für den\*die Benutzer\*in leicht zu verstehen ist.

### **A7. Tracking von Personen**

Aus der in Abschnitt 2.3 erfolgten Analyse ging hervor, dass nur wenige IKT (insbesondere solche, die ein Kameratracking verwenden) tatsächlich ein Tracking von Personen vornehmen. Der Begriff Tracking lässt sich von dem der Erkennung von Personen über die Komponente Zeit trennen. Letzteres beschreibt dabei das Erkennen einer oder mehrerer Silhouetten im Kamerabereich zu einem bestimmten Zeitpunkt. Das Tracking von Personen beschreibt hingegen die Zuweisung einer eindeutigen ID zu derselben Person, solange diese im Kamerabereich ist.

In Tabelle 1 werden sämtliche Anforderungen an das Tracking Tool zusammenfassend kurz beschrieben.

<b>Anforderung</b>	<b>Beschreibung</b>
<b>A1. Auf Ausstellungsgestaltende ausgerichtete Bedienoberfläche</b>	Sämtliche Funktionen sind über eine intuitive grafische Bedienoberfläche nutzbar.
<b>A2. Erkennen von Personen aus beliebigem Winkel</b>	Personen vor der Kamera werden zu jedem Zeitpunkt erkannt. Dabei ist die Position der Kamera irrelevant.
<b>A3. Erstellen von Triggerzones</b>	Der*die Benutzer*in kann virtuelle Zonen im Bereich vor der Kamera erstellen, die bei Berührung, wie bei Lichtschranken, ein Signal auslösen.
<b>A4. Steuerung von mehreren Tiefenkameras über ein zentrales Tool</b>	Es können beliebig viele Tiefenkameras verwendet werden, welche alle über eine zentrale Steuerung gewartet und konfiguriert werden.
<b>A5. Automatisierter Start</b>	Das Tool kann so konfiguriert werden, dass es im Ausstellungsbetrieb in bestimmten Zeitintervallen startet und sich dabei selbstständig konfiguriert.
<b>A6. Geeignetes Protokoll zur Übertragung der Trackingdaten</b>	Das Tool lässt sich an eine Vielzahl verschiedener IKT anbinden. Dabei sind die übertragenen Daten leicht zu verwenden.
<b>A7. Tracking von Personen</b>	Personen wird beim Betreten des Kamerabereichs eine eindeutige ID zugewiesen. Diese wird über den gesamten Zeitraum, in der die Person im Kamerabereich ist beibehalten.

*Tabelle 1: Anforderungen an ein Tracking Tool für Ausstellungsgestaltende*



## 3.2 Tracking Tools

Es lassen sich nur wenige Tools finden, welche die Verwendung von Tiefenkameras vereinfachen. Diese Tracking Tools werden im Folgenden vorgestellt und auf die Erfüllung der in Abschnitt 3.1 erhobenen Anforderungen untersucht. Teile dieser Analyse wurden bereits in dem Projektbericht durchgeführt.

### **Proximity Toolkit**

Das Proximity Toolkit ist ein Tool für das Prototyping von Applikationen, die Kategorien der proxemischen Interaktion nach Greenberg et al. nutzen (siehe Abschnitt 2.3) (Marquardt et al. 2011). Obwohl das Toolkit die Anbindung von Kinects (V1) unterstützt, lässt es sich in seinem vollen Umfang nur in Kombination mit Motion Capturing-Systemen, wie beispielsweise dem VICON Tracking-System<sup>14</sup>, nutzen. Mit diesen ist eine präzise Erkennung von Personen und deren Skelett im Raum möglich. Werden nur Kinects verwendet, so beschränkt sich die Funktion des Tools auf eine grafische Darstellung der von der Kinect gelieferten Skelettdaten und der Verwendung dieser. Die Erkennung und das Tracking von Personen funktioniert demzufolge über das proprietäre Tracking der Kinect. Dies impliziert auch, dass das Proximity Toolkit mit Kinects nur dann verwendbar ist, wenn die Ausrichtung dieser parallel zum Boden des Raums ist, da ansonsten das Skelett-Tracking nicht funktioniert. Die Menge der Kinects ist dabei durch die Anzahl der USB Ports am verwendeten Gerät oder der Leistung desselben beschränkt. Über eine grafische Bedienoberfläche lassen sich der von den Kameras abgetastete Raum und alle darin befindlichen Personen, Geräte und sogenannte Fixed Feature Entities darstellen und überwachen. Die Bedienoberfläche ist komplex und richtet sich offensichtlich an Menschen mit Programmiererfahrung. Das Tool bietet zusätzlich einige nützliche Funktionen für das Prototyping, wie beispielsweise die Möglichkeit der Aufnahme von Interaktionsabläufen, um diese nicht für jeden Test neu auszuführen. Die Fixed Feature Entities entsprechen den in Abschnitt 3.1 beschriebenen Triggerzonen. Die Tracking-Daten, die das Proximity Toolkit liefert, können über eine event-basierte C# API abgegriffen werden.

### **XDKinect**

XDKinect ist ein Framework für die Entwicklung von Cross-Device-Bedienoberflächen mithilfe einer Kinect (V1) (Nebeling et al. 2014). Der Begriff Cross-Device Interaction ist in die ubicomp-Vision Mark Weisers einzuordnen und beschreibt sowohl die in Abschnitt 2.3 beschriebene Kommunikation zwischen verschiedenen Geräten als auch die gleichzeitige Interaktion eines Nutzers\* einer Nutzerin über mehrere dieser Geräte hinweg. Das Framework vereinfacht dabei die proprietären Skelett-Tracking-Daten der Kinect und liefert diese über das Websocket-Protokoll im JSON-Format an beliebig viele Clients weiter. XDKinect besitzt keine Bedienoberfläche.

---

<sup>14</sup> <http://www.vicon.com/products/camera-systems>, Abgerufen am 12.05.16

### SoD-Toolkit

Ebenfalls für die Unterstützung von Multi- und Cross-Device-Applikationen wurde das SoD-Toolkit entwickelt (Seyed et al. 2015). Das Toolkit kann Sensordaten von mehreren verschiedenen Geräten, wie der Kinect (V1 und V2), von Smartphones und Tablets, des iBeacons und der Leap Motion verarbeiten. Die Kinects können dabei jeweils an ein eigenes Endgerät angeschlossen werden, welches deren Tracking-Daten über LAN weiterleitet. Diese Daten werden in einer zentralen Applikation gebündelt und an Clients über das Websocket-Protokoll im JSON-Format gesendet. Eine grafische Bedienoberfläche zeigt unter anderem die verfügbaren Geräte, die Personen, die über die Geräte erkannt werden und deren Position. Auch wenn ein Ziel des Tools das Senken der Hürde für weniger erfahrene Programmierer\*innen ist, richtet sich das Tool weiterhin an Entwickler\*innen und Forscher\*innen im Forschungsfeld der MCI (ebda.). Auch im SoD-Toolkit wird lediglich das proprietäre Skelett-Tracking der Kinect verwendet.

### TSPS

Das Toolkit for Sensing People in Spaces (TSPS)<sup>15</sup> ist das einzige Toolkit, welches unter anderem für den Einsatz in künstlerischen Installationen entwickelt wurde. Dementsprechend lässt sich die Applikation über eine simple grafische Bedienoberfläche verwenden. Es kann eine beliebige Kamera, also auch eine gewöhnliche, die nur ein zweidimensionales Bild (2D-Bild) liefert, genutzt werden. Dabei nutzt TSPS Computer Vision Algorithmen, um sich bewegende Objekte zu erkennen. Da diese sowohl für die Kinect als auch für 2D-Kameras identisch funktionieren, nutzt TSPS die zusätzlichen Tiefeninformationen nicht, die eine genauere Erkennung möglich machen. Die Trackingdaten werden über verschiedene Protokolle, wie OSC<sup>16</sup> und TUIO<sup>17</sup>, welche häufig in interaktiven Installationen und generativer Kunst<sup>18</sup> angewendet werden, an einen Client weitergeleitet. Zusätzlich existieren Bibliotheken für Javascript sowie für die Programmiersprache und -umgebung Processing, welche insbesondere die Erstellung von generativer Kunst und Rapid Prototyping unterstützt.

---

<sup>15</sup> <http://www.tsps.cc/>, Abgerufen am 12.05.16

<sup>16</sup> OSC ist ein einfaches Protokoll, welches auf dem UDP Netzwerkprotokoll basiert und ursprünglich für die Steuerung von Musiksoftware verwendet wurde.

<sup>17</sup> <http://www.tuio.org/>, Abgerufen am 12.05.16

<sup>18</sup> Der Begriff Generative Kunst beschreibt Kunst, deren Entstehungsprozess durch das Abarbeiten eines Regelsatz' gekennzeichnet ist. Dazu zählen somit insbesondere programmierte Abläufe, die, in Abhängigkeit von Variablen, bei jedem Durchlauf ein anderes Ergebnis erzielen.

### 3.3 Tabellarischer Vergleich: Tracking Tools und Anforderungen

In Tabelle 2 werden die einzelnen bereits existierenden Tracking Tools aufgeführt und es wird vermerkt, ob diese die in Abschnitt 3.1 erhobenen Anforderungen erfüllen. Zusammenfassend lässt sich feststellen, dass keines der analysierten Tools diese vollständig erfüllt. Dies ist unter anderem dadurch bedingt, dass diese nicht für die Zielgruppe Ausstellungsgestaltende entwickelt wurden. Das Proximity Toolkit, XDKinect und das SoD-Toolkit richten sich klar an Programmier\*innen und Forscher\*innen. Gleichzeitig ist deren Hauptziel die Entwicklung von Cross-Device-Applikationen. Dementsprechend bieten sie auch eine Vielzahl von Funktionen, die nicht oder nur selten in Ausstellungen verwendet werden und damit die Tools selbst unnötig komplex machen.

TSPS ist das einzige Tool mit einer klaren Ausrichtung auf die Entwicklung von interaktiven Installationen. Dementsprechend unterstützt es als einziges Tool ein Tracking, welches unabhängig vom Winkel und der Position der Kamera ist. Nachteile von TSPS sind – im Vergleich zu den anderen Tools – das ungenaue Tracking sowie, dass es nicht möglich ist, mehrere Kameras zu verwenden.

Tracking Tool	A1	A2	A3	A4	A5	A6	A7
Proximity Toolkit	■	■	■	■	■	■	■
XDKinect	■	■	■	■	■	■	■
SoD-Toolkit	■	■	■	■	■	■	■
TSPS	■	■	■	■	■	■	■

Tabelle 2: Die bereits existierenden Tracking Tools und inwiefern diese die Anforderungen an ein Tracking Tool für Ausstellungsgestaltende erfüllen.

■ = Erfüllt die Anforderungen, ■ = Erfüllt die Anforderungen teilweise, ■ = Erfüllt die Anforderungen nicht.

## 4 Argus Vision

In Abschnitt 3.3 wurde gezeigt, dass keines der bereits existierenden Tracking Tools die Anforderungen an ein Tracking Tool für Ausstellungsgestaltende erfüllt. Zusätzlich besitzt keines eine Softwarearchitektur, welche die einfache Erweiterung dieser um die benötigten Funktionen ermöglicht. Aus diesem Grund wurde das Tracking Tool Argus Vision vollständig neu entwickelt.

Das Tool wurde auf GitHub veröffentlicht<sup>19</sup> und ist dementsprechend für Ausstellungsgestaltende und andere Nutzer\*innen frei zugänglich. Argus Vision wurde in Processing (siehe Abschnitt 3.2 TSPS) programmiert, einer Programmiersprache also, die von manchen Ausstellungsgestaltenden beherrscht wird. So können diese bei Bedarf das Tool um weitere Funktionen erweitern.

Die Aufgabe von Argus Vision lässt sich wie folgt beschreiben:

---

Argus Vision vereinfacht die Anwendung von Tracking mit mehreren Kinects mithilfe einer Plug and Play-Struktur und einer leicht zu bedienenden grafischen Oberfläche. Die Zielgruppe sind dabei Ausstellungsgestaltende mit wenig Programmierkenntnissen, welchen mit dem Tool die Möglichkeit gegeben wird, interaktive Installationen und Räume eigenständig zu entwickeln. Mithilfe von sogenannten Triggerzones, frei im Raum positionierbaren Quadern, welche wie Lichtschranken funktionieren, lassen sich zusätzlich einfache Interaktionsarten schnell testen und umsetzen.

---

Dabei ist Argus Vision als Middleware zu verstehen. Es vereinfacht die Anwendung von Tracking in einer IKT; die Verarbeitung und Verwendung dieser Daten erfolgt jedoch weiterhin in anderen Applikationen. Beispielhaft zu nennen sind hier Mediensteuerungen wie e:cue<sup>20</sup> und resolume<sup>21</sup> oder Programmierumgebungen für generative Gestaltung wie vvvv<sup>22</sup> und Touchdesigner<sup>23</sup>.

In diesem Kapitel wird zuerst die Architektur von Argus Vision vorgestellt. Die Arbeitsweise mit dem Tool und die Funktionsweise dessen werden daran anknüpfend erläutert. Abschließend werden einige Implementierungsdetails gesondert hervorgehoben. Designentscheidungen, welche durch die in Kapitel 3 angestellten Überlegungen motiviert sind, stehen dabei durchgängig im Fokus.

---

<sup>19</sup> <https://github.com/MoritzSkowronski/ArgusControl> und <https://github.com/MoritzSkowronski/ArgusKinect>, Abgerufen am 25.05.16

<sup>20</sup> <http://www.ecue.de/>, Abgerufen am 19.05.16

<sup>21</sup> <https://resolume.com/>, Abgerufen am 19.05.16

<sup>22</sup> <https://vvvv.org/>, Abgerufen am 19.05.16

<sup>23</sup> <http://www.derivative.ca/>, Abgerufen am 19.05.16

## 4.1 Architektur von Argus Vision

Argus Vision besteht aus zwei Programmteilen, der zentralen Steuereinheit und Middleware *Argus Control* und beliebig vielen *Argus Kinect*-Instanzen.

Wichtig ist dabei, dass eine Kinect V2<sup>24</sup>, welche in Argus Vision Verwendung findet, jeweils einen eigenen PC benötigt. Es existieren also genau so viele Argus Kinect-Instanzen, wie Kinects eingesetzt werden. Zwar existieren Lösungen, um mehrere Kinects an einen PC zu verbinden<sup>25</sup>, womit Kosten eingespart werden können, die Vorteile einer Multi-PC Lösung überwiegen jedoch. Zum einen ist die Verarbeitung der Datenmenge bei bereits zwei Kinects nur mit sehr leistungsstarken Rechnern möglich. Nutzt man für jede Kinect einen eigenen, baugleichen Rechner, so lässt sich davon ausgehen, dass die Daten jeder Kinect mit einer gleichwertigen Geschwindigkeit verarbeitet werden. Dies ist insofern wichtig, als ein funktionierendes Tracking von der Anzahl der verarbeiteten Bilder pro Sekunde abhängt. Bricht die Leistung an einer oder aller Kinects ein, so beeinflusst dies wiederum die Besucher\*innenerfahrung am IKT selbst. Zum anderen lassen sich die Kinects auf einen beliebig großen Raum oder gar mehrere Räume verteilen. Da die Kinect ihre Daten über den USB 3.0-Standard überträgt, für welchen eine Kabellänge von drei Metern empfohlen wird, ist dies mit einer Ein-PC Lösung nicht möglich.

Argus Control nimmt die wichtigste Rolle in der Architektur von Argus Vision ein. Über das Programm lassen sich sämtliche damit verbundene Argus Kinect-Instanzen steuern und überwachen. Gleichzeitig empfängt es die Trackingdaten der einzelnen Argus Kinects, bündelt diese und sendet sie an einen Client weiter. Dieser kann ein beliebiges Programm sein, welches in der Lage ist, die Daten zu verarbeiten.

Die Aufgabe von Argus Kinect ist lediglich die Verarbeitung der Daten der Kinect und die Weiterleitung dieser an Argus Control. Diese Kommunikations-Struktur wird in Abbildung 6 verdeutlicht.

---

<sup>24</sup> Im Weiteren ist, sofern nicht anders angegeben, mit dem Begriff Kinect immer die Version 2 der Kinect gemeint.

<sup>25</sup> <https://github.com/hanasaan/ofxMultiKinectV2>, Abgerufen am 15.05.16

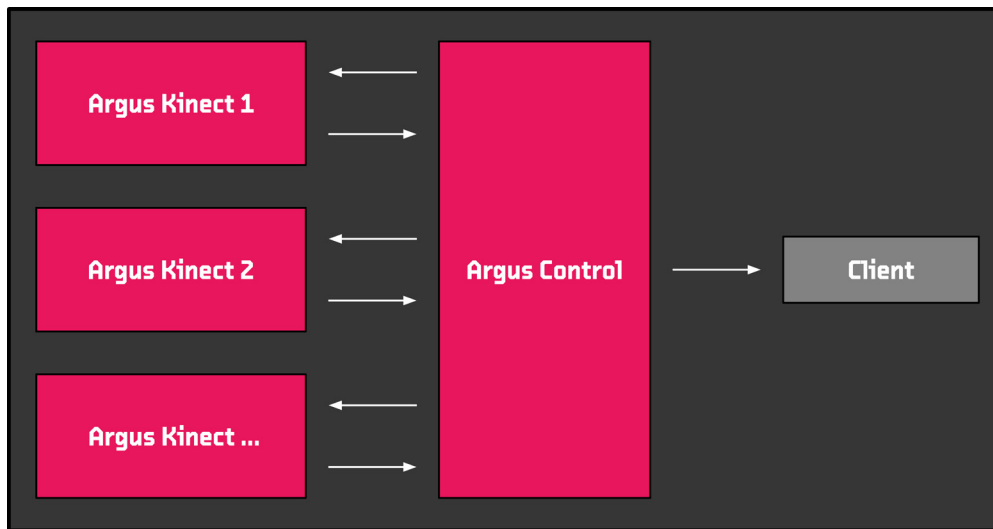


Abbildung 6: Die Kommunikations-Struktur von Argus Vision.

## 4.2 Verwendung und Oberfläche von Argus Vision

In diesem Abschnitt wird eine optimale Arbeitsweise mit Argus Vision vorgestellt. Anhand dieser lassen sich die Bedienoberfläche und Funktionen des Tracking Tools aufzeigen.

Argus Vision ist dabei so konzipiert, dass das Tool beim Beenden sämtliche vorgenommene Einstellungen speichert und diese beim erneuten Start wiederherstellt. Im Folgenden wird darum beispielhaft eine erstmalige Konfiguration beschrieben. Ob zuerst die Argus Kinect-Instanzen gestartet werden oder Argus Control, ist in der Nutzung irrelevant. Um das Verständnis für die Bedienoberfläche von Argus Control zu erhöhen, wird hier jedoch mit der Beschreibung von Argus Kinect begonnen.

Auf die Entwicklung der Bedienoberflächen von Argus Control und Argus Kinect wurde ein besonderes Augenmerk gelegt. Diese müssen zum einen ansprechend, zum anderen leicht verständlich sein, um von möglichst vielen Ausstellungsgestaltenden benutzt werden zu können. Daher wurde daher die Anzahl der Einstellungsmöglichkeiten in der jeweiligen Bedienoberfläche auf ein Minimum reduziert.

### 4.2.1 Argus Kinect

Die Aufgabe von Argus Kinect ist der Empfang von Einstellungen von Argus Control, die Verarbeitung der Rohdaten der Kinect sowie das Senden dieser an Argus Control. Mit dem Start der Anwendung wird die Kinect gestartet und gespeicherte Einstellungen werden, sofern vorhanden, geladen. Da sämtliche Einstellungen in Argus Control getätigt werden, beschränkt sich die Bedienoberfläche von Argus Kinect auf die Eingabe der zur Kommunikation mit Argus Control benötigten Daten.

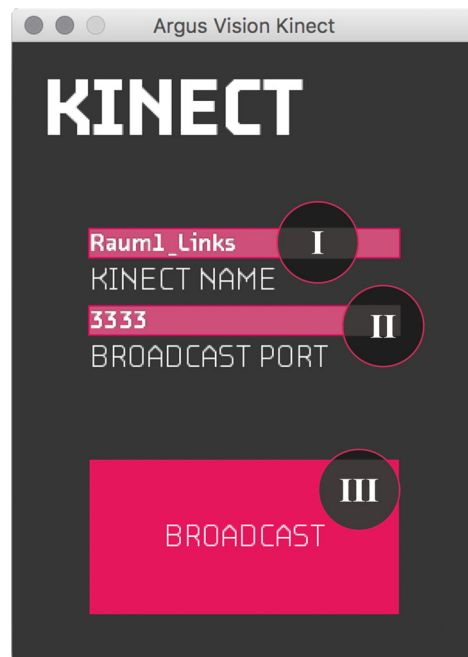


Abbildung 7: Die Bedienoberfläche von Argus Kinect.

Diese sind ein eindeutiger Name für die Kinect (Abbildung 7 I) und ein Port (Abbildung 7 II), über welchen die Kommunikation zwischen Argus Control und Argus Kinect stattfindet. Mit dem Aktivieren des *Broadcast*-Buttons (Abbildung 7 III) werden die eingegebenen Daten als Einstellungen gespeichert und der sogenannte *Broadcast*-Prozess beginnt. Während dieses Prozesses versucht Argus Kinect sich mit einer Argus Control-Instanz zu verbinden

### 4.2.2 Argus Control

Die Bedienoberfläche des Herzstücks von Argus Vision, Argus Control, gliedert sich in vier Bereiche. Dabei ist die Oberfläche so aufgebaut, dass sie gegen den Uhrzeigersinn von links nach rechts zu bearbeiten ist.

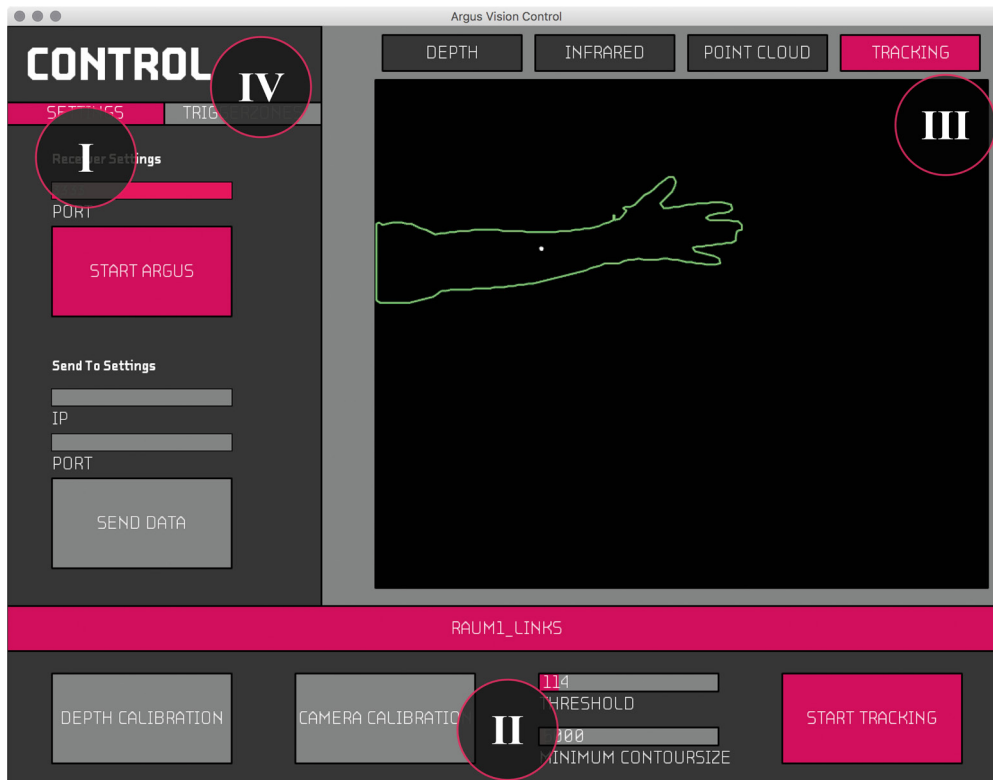


Abbildung 8: Die Bedienoberfläche von Argus Control.

Dies wird in Abbildung 8 dargestellt. Die Bearbeitung beginnt folglich im Bereich I, den *Kommunikations-Einstellungen*. In diesem Bereich wird das Design der Oberfläche von Argus Kinect wiedergespiegelt, um zusätzlich darauf hinzuweisen, dass hier die gleichen Einstellungen vorgenommen werden müssen. In Abbildung 9 wird dieser Bereich noch einmal getrennt dargestellt.



In den *Send To Settings* (Abbildung 9 II) wird die unidirektionale Kommunikation mit einem Client festgelegt. Darum muss zusätzlich zu einem Port auch die IP-Adresse des Clients angegeben werden. Mit dem Betätigen des *Send Data*-Buttons werden die in Argus Control kumulierten Tracking-Daten der einzelnen Argus Kinect-Instanzen an den Client weitergeleitet. Eine genauere Beschreibung der Daten findet sich in Abschnitt 4.3.1, zwei Beispiele für die Verwendung dieser in Kapitel 5.

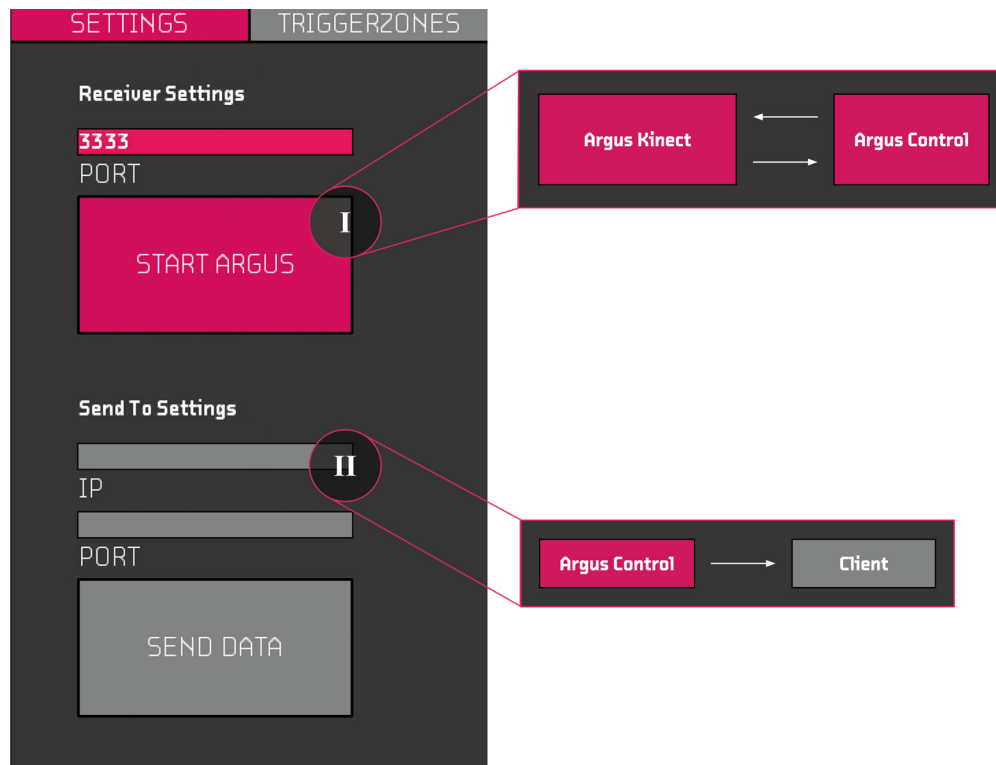


Abbildung 9: Ausschnitt des Bereichs I aus Abbildung 8: Die Kommunikations-Einstellungen in Argus Control.

In den *Receiver Settings* hingegen wird die Kommunikation mit den Argus Kinect-Instanzen festgelegt (Abbildung 9 I). Dementsprechend muss hier nur derselbe Port wie in Argus Kinect eingetragen werden. Ist der *Start Argus*-Button betätigt, so wartet Argus Control fortan auf diesem Port auf Nachrichten von Argus Kinect-Instanzen.

Erkennt Argus Control eine dieser an dem Port, so wird eine Verbindung zwischen Argus Control und dieser Argus Kinect-Instanz hergestellt. Diese lässt sich nun mit Namen in der *Kinect-Leiste* des *Kinect-Bereichs* auswählen, wie in Abbildung 10 I dargestellt.

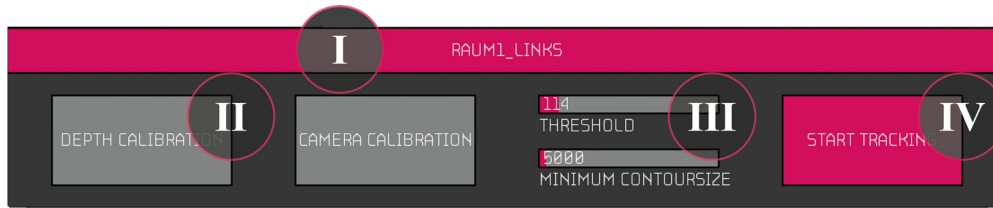


Abbildung 10: Ausschnitt des Bereichs II aus Abbildung 8: Der Kinect-Bereich von Argus Control.

Verbinden sich weitere Argus Kinect-Instanzen mit Argus Control, so werden diese ebenfalls mit Namen der Kinect-Leiste hinzugefügt. Ist eine dieser Instanzen ausgewählt, so kann diese individuell konfiguriert werden. Die Konfiguration besteht aus zwei Schritten – der Aufnahme des leeren Raums und dadurch Kalibrierung des Tracking-Bereichs über den *Depth Calibration*-Button (Abbildung 10 II) und des Festlegens zweier Schwellenwerte für das Tracking, dem *Threshold* und der *Minimum Contoursize* (Abbildung 10 III). Dieser Vorgang wird in Abschnitt 4.3.1 detailliert erklärt. Abschließend kann das Tracking über den Button *Start Tracking* (Abbildung 10 IV) gestartet werden. Der Button *Camera Calibration* verfügt in der in dieser Arbeit vorgestellten Version von Argus Vision noch über keine Funktionalität. Mit diesem sollen mehrere Kinects mit sich überschneidendem Raumbild zu einer virtuellen Kamera verbunden werden können, um ein raumspannendes Tracking zu ermöglichen (vgl. die Möglichkeiten zur Weiterentwicklung von Argus Vision in Abschnitt 5.3.3).

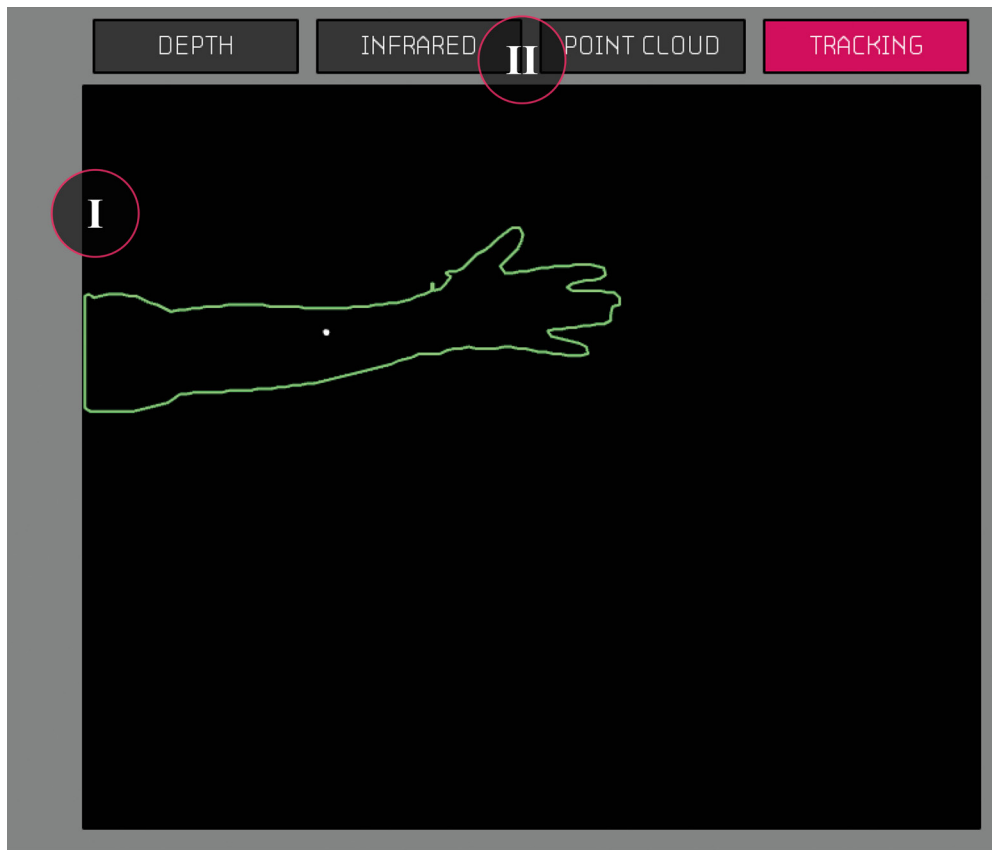


Abbildung 11: Ausschnitt des Bereichs III aus Abbildung 8: Der Stream-View von Argus Control.

Ebenfalls erst nach dem Anwählen einer Argus Kinect-Instanz in der Kinect-Leiste nutzbar ist der Stream-View von Argus Control, welcher in Abbildung 11 dargestellt ist. In diesem lassen sich verschiedene Darstellungen der Kinect-Daten abrufen (Abbildung 11 I) – Tiefen- (Depth), Infrarot- (Infrared), Point Cloud- und Tracking-Stream (Abbildung 11 II). Die Funktion der jeweiligen Streams wird im Folgenden beschrieben.

Mit den Tiefen- und Infrarot-Streams lässt sich der reale Raum am besten abbilden. Im *Tiefen-Stream* werden die Tiefenwerte der Kinect, welche in einem Intervall von null bis 8.000 Millimetern liegen, auf Grauwerte in einem Intervall von null bis 255 abgebildet (Abbildung 12a). Kleine Werte, also der Kamera nahe Objekte, werden dunkler abgebildet, große Werte, also der Kamera ferne Objekte, heller. Der *Infrarot-Stream* zeigt das Infrarotbild der Kinect und damit die realitätsgetreueste Aufnahme der vier Streams (Abbildung 12b). Der Vorteil eines Infrarotbilds gegenüber einem Farbbild ist, dass dieses von der Helligkeit des Raums unabhängig ist. Auch in einem vollständig dunklen Raum lassen sich also noch sämtliche Objekte vor der Kamera erkennen.



Abbildung 12: Der Tiefen- (a) und Infrarot-Stream (b) in Argus Control. Der Tiefen-Stream stellt die Tiefe des Raums in Graustufen dar. Der Infrarot-Stream liefert ein realitätsgetreues Bild des Raums.

Der *Tracking-Stream* besitzt zwei Modi. Ist das Tracking nicht aktiviert, so wird ein schwarzweißes Bild angezeigt. Schwarz gezeichnet sind hierbei Stellen im Raum, bei denen keine Änderung gegenüber dem leeren Raum zu erkennen ist, weiß hingegen diejenigen Stellen, bei denen eine Änderung angenommen wird. In Abbildung 13a wird ein solches Bild dargestellt. Hier wird Arm und Hand einer Person weiß dargestellt, der Rest des Bildes ist schwarz. Auffällig dabei ist, dass noch einige Störungen im Bild, also weiße Punkte, die keinem Objekt zuzuordnen sind, zu sehen sind. Diese entstehen beispielsweise durch Oberflächen, die Infrarotlicht nicht oder nur teilweise reflektieren. Ist das Tracking aktiviert, so werden die Ergebnisse des Trackings – die Konturen der erkannten Objekte – angezeigt. Dies wird in Abbildung 13b dargestellt. Zusätzlich zur Kontur wird das gewichtete Mittel der Kontur als weißer Punkt gekennzeichnet. Die noch in Abbildung 13a vorhandenen Störungen sind nun entfernt, da nur Objekte einer bestimmten Minimalgröße erkannt werden.

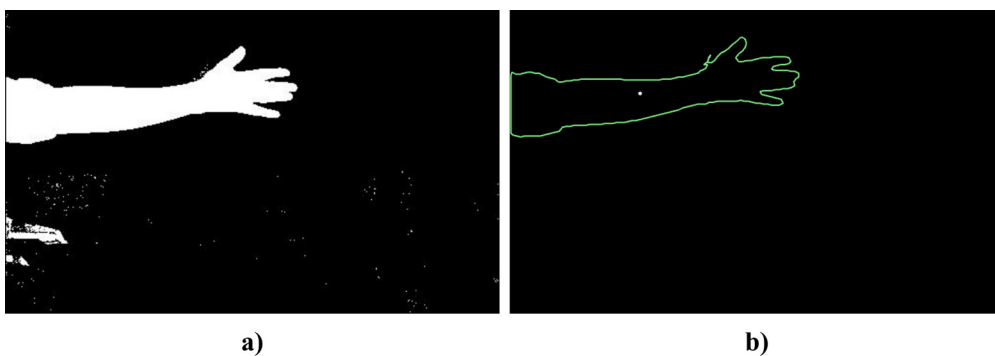


Abbildung 13: Der Tracking-Stream zeigt die Konturen erkannter Objekte im Raum (b). Ist das Tracking deaktiviert, so zeigt er Änderungen gegenüber dem leeren Raum in Weiß (a).

Die während der Konfiguration wichtigste Visualisierungsart ist der *Point Cloud-Stream*. Eine Point Cloud zeigt dreidimensionale Punkte in einem virtuellen dreidimensionalen Raum. Da die Kinect zusätzlich zu den normalen zweidimensionalen x,y-Werten einer Kamera noch einen Tiefenwert (z-Wert) liefert, lassen sich diese x,y,z-Punkte in einen virtuellen dreidimensionalen Raum zeichnen. Durch diesen Raum kann dann mit einer virtuellen Kamera navigiert werden. So lässt sich eine Szene aus sämtlichen Perspektiven betrachten. In Abbildung 14 ist die gleiche Szene wie in Abbildung 12 als Point Cloud abgebildet<sup>26</sup>. Auch hier ist ein Arm zu erkennen. Dieser lässt sich nun aber sowohl von vorne als auch von hinten betrachten.

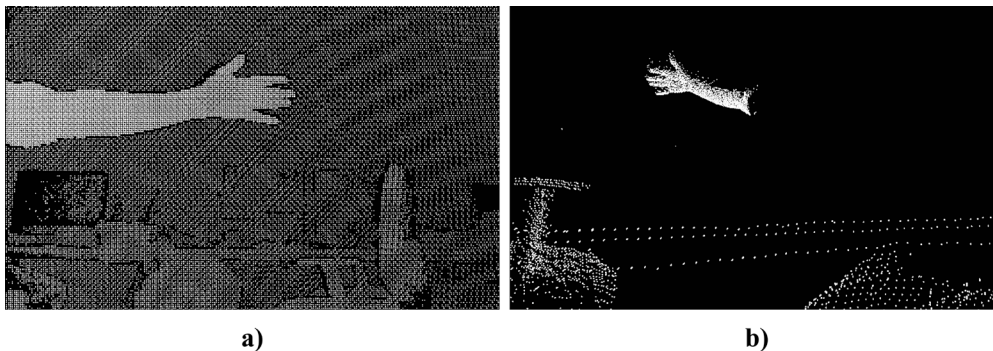


Abbildung 14: Im Point Cloud-Stream kann mit einer virtuellen Kamera der Raum dreidimensional erkundet werden. In a) ist eine Hand von vorne, in b) ist dieselbe Szene von hinten gezeigt.

Der Vorteil des Point Cloud-Streams ist, dass der\*die Benutzer\*in nicht mehr auf seine\*ihre perspektivische Vorstellungskraft angewiesen ist. Die Darstellung eines Objekts auf einem zweidimensionalen Bild ist immer von der Person des\*der Betrachtenden beziehungsweise der Kamera abhängig. Die tatsächliche Größe und Position eines Objekts einzuschätzen ist demzufolge schwierig. Durch das virtuelle Navigieren durch den Raum kann diese Hürde überwunden werden. Wird Argus Control auf einem Laptop betrieben, so kann der\*die Benutzende sich sogar selbst in die Point Cloud begeben, um ein besseres Gespür für die Tiefe des Raums zu entwickeln.

---

<sup>26</sup> Um sämtliche Aufnahmen der Point im Druck besser abbilden zu können, wurden diese nachträglich bearbeitet. Die Originale finden sich auf dem dieser Arbeit beiliegenden USB-Stick.

Dies erweist sich insbesondere bei der Erstellung und Platzierung von Triggerzones als nützlich. Dieses Feature ermöglicht es, dreidimensionale Zonen im Raum zu erstellen, die beim Betreten dieser eine Aktion auslösen (dieses Konzept ist in Abschnitt 3.1 detailliert beschrieben). Bei angewählter Argus Kinect-Instanz können die Triggerzones über die *Triggerzone-Einstellungen* (Abbildung 8 IV) erstellt (Abbildung 15 II), positioniert (Abbildung 15 IV) und gelöscht (Abbildung 15 III) werden.

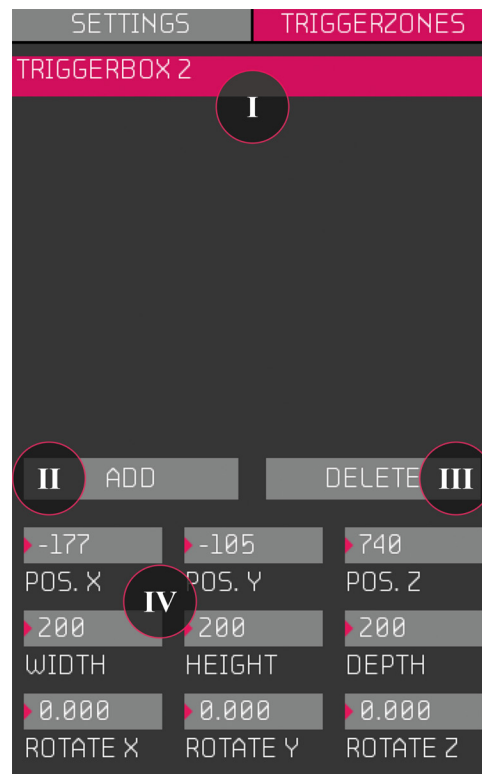


Abbildung 15: Die Triggerzone-Einstellungen in Argus Control.

Die vorhandenen Triggerzones werden sowohl in der Liste in den Triggerzone-Einstellungen (Abbildung 15 I) dargestellt als auch in jedem Stream. Im Point Cloud-Stream können die Triggerzones zusätzlich per Mausklick auf die Visualisierung positioniert werden. Aus diesem Grund wird hier die Darstellung der Triggerzones in der Point Cloud gezeigt.

Wie Abbildung 16 zu entnehmen ist, sind die Triggerzones Quader, deren Position, Größe und Rotation frei bestimmbar sind. Die Rotation ist insbesondere dann sinnvoll, wenn die Kinect geneigt auf eine horizontale Oberfläche zeigt. Durch die Rotation ist der\*die Benutzer\*in in der Lage, eine Triggerzone auf dieser Fläche zu platzieren, ohne, dass sie diese schneidet.

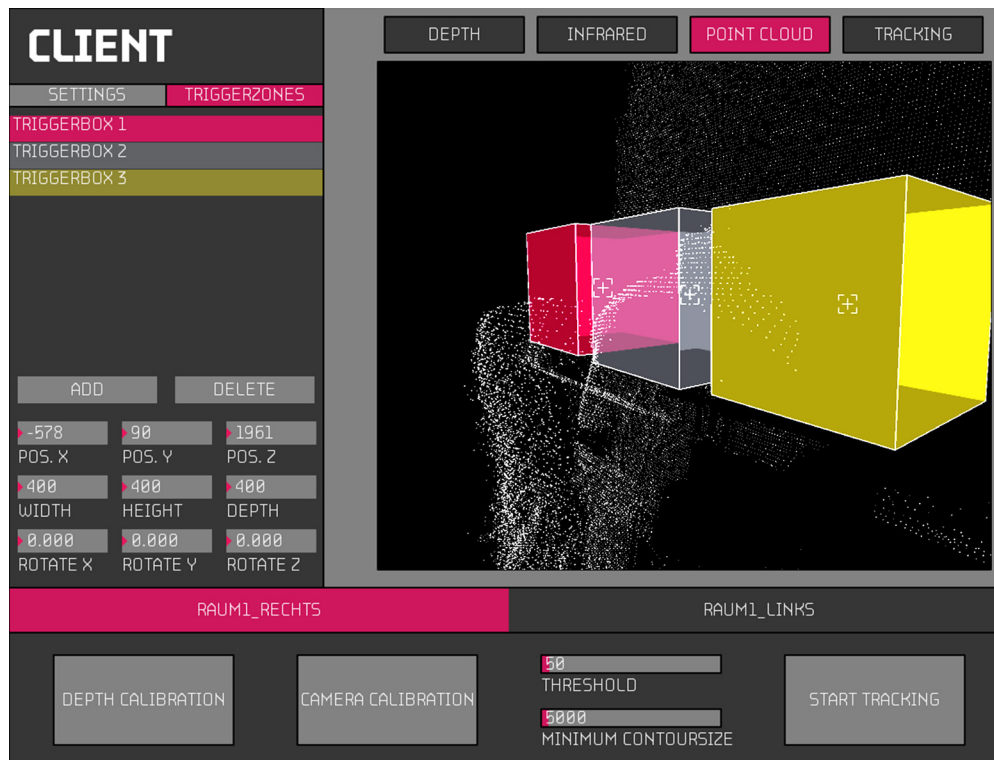


Abbildung 16: Drei Triggerzones sind auf einer Tischfläche platziert.

In Abbildung 16 sind drei Triggerzones auf einer Tischfläche platziert. Die aktuell ausgewählte Zone ist in der Auszeichnungsfarbe des Tools – Magenta – hervorgehoben. Die Auswahl erfolgt dabei entweder über die Liste in den Triggerzone-Einstellungen oder direkt in der Point Cloud durch das Hovern mit dem Mauszeiger über den *Grabber* der jeweiligen Zone. Dieser ist als umrahmtes Kreuz dargestellt. Ein Linksklick auf diesen ermöglicht die freie Rotation des Quaders um seinen Mittelpunkt. Mit einem Rechtsklick kann die Triggerzone im Raum verschoben werden. Die Größe kann nur über die Triggerzone-Einstellungen geändert werden.

Die Zonen können auch in den jeweiligen 2D-Streams – dann aber nur über die Triggerzone-Einstellungen – platziert werden. Die frei bewegliche 3D-Ansicht ermöglicht allerdings eine deutliche exaktere Positionierung. Der\*die Benutzer\*in kann zusätzlich bei aktiviertem Tracking testen, ob die Zonen richtig platziert wurden.

Er\*Sie tritt dazu in den Raum, in dem er\*sie die Triggerzone vermutet. Ist diese korrekt gesetzt, so leuchtet sie auf. Dabei leuchtet sie umso heller, je mehr Fläche der Zone durch die getrackte Person belegt wird.

### 4.3 Implementierung von Argus Vision

Die Implementierung von Argus Vision wurde in der dieser Arbeit vorausgegangen Projektarbeit bereits umfassend beschrieben. Aus diesem Grund werden in der vorliegenden Arbeit nur die wichtigsten Teile von Argus Vision beschrieben – das Tracking und die Kommunikation. Letzteres umfasst dabei sowohl die Übermittlung der Tracking-Daten an den Client als auch die Kommunikation zwischen Argus Control und Argus Kinect.

Dabei existieren zu bestimmten in dieser Arbeit verwendeten Algorithmen bereits fortschrittlichere Ansätze. Für Object Detection und Tracking beispielsweise – das Erkennen und Verfolgen von Objekten in einem Kamerabild – finden sich zahlreiche komplexe Algorithmen für unterschiedliche Einsatzzwecke<sup>27</sup>.

Der Fokus bei der Entwicklung von Argus Vision liegt jedoch auf der Entwicklung eines umfassenden grafischen Tools mit besonderen Eigenschaften für Ausstellungsgestaltende. Um diese in dem für die Projektarbeit anberaumten Zeitrahmen umzusetzen, wurden dementsprechend insbesondere beim Tracking von Personen und dem Streaming der Point Cloud – für den Anwendungsfall ausreichende – eigene Lösungen implementiert.

---

<sup>27</sup> Ein umfassender Überblick zu den wichtigsten Algorithmen und deren Verwendung findet sich in Yilmaz et al. 2006.



### 4.3.1 Tracking in Argus Vision

Das Tracking findet in Argus Vision vollständig in Argus Kinect statt. Dafür werden ausschließlich die Tiefendaten der Kinect – Ganzzahlwerte von null bis 8.000 (Millimeterwerte) – verwendet. Das Tracking basiert auf einem einfachen Zyklus, welcher pro aufgenommenem Bild der Kinect vollständig durchlaufen werden muss. Dieser ist in Abbildung 17 dargestellt und beginnt mit dem Empfang der Tiefendaten der Kinect.

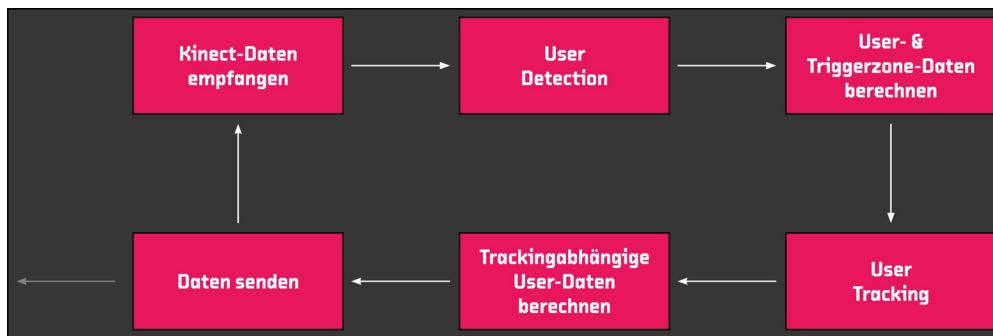


Abbildung 17: Die Arbeitsschritte im Trackingzyklus pro Kinect-Bild.

Wie aus dem Zyklus ersichtlich ist, bauen alle Arbeitsschritte aufeinander auf – in diesem Abschnitt wird der Prozess des Trackings demnach am Beispiel der Verarbeitung eines Bilds erklärt.

Da der Arbeitsschritt *Kinect-Daten empfangen* nur den Abruf dieser von der Kinect beinhaltet, wird dieser Vorgang hier nicht weiter beschrieben.

#### User Detection

Der Begriff *User Detection* beschreibt das Erkennen von Personen (hier: Users) aus einem beliebigen Winkel (siehe Abschnitt 3.1 Erkennen von Personen aus einem beliebigen Winkel).

Um dies zu ermöglichen wird eine modifizierte Form der Background Subtraction (Hintergrundsubtraktion) angewandt. Der Begriff beschreibt das Speichern eines Bildes als Hintergrund. Darauf folgende Bilder werden auf Unterschiede zu diesem geprüft. Die Pixel, welche sich von dem Hintergrund unterscheiden, werden daraufhin weiterverarbeitet. In den meisten Fällen wird dieser Prozess bei 2D-Bildern angewandt (so beispielsweise in Wolf et al. 2015 und Yilmaz et al. 2006).

Durch die Verwendung der zusätzlichen Tiefendaten kann der Prozess zum einen vereinfacht, zum anderen verbessert werden. Während in 2D-Bildern die Farbwerte der Pixel verglichen werden müssen, wird in Argus Kinect der Tiefenwert der jeweiligen Pixel auf Abweichungen gegenüber der Tiefenwerte des Hintergrundbilds untersucht.

Ist die Differenz in der Tiefe eines Pixels größer als ein Schwellenwert wird dieser weiß, andernfalls schwarz, eingefärbt. Dieser Schwellenwert kann von dem\*der Benutzer\*in selbst in der Oberfläche von Argus Control über den Slider Threshold eingestellt werden (Abbildung 10 III).

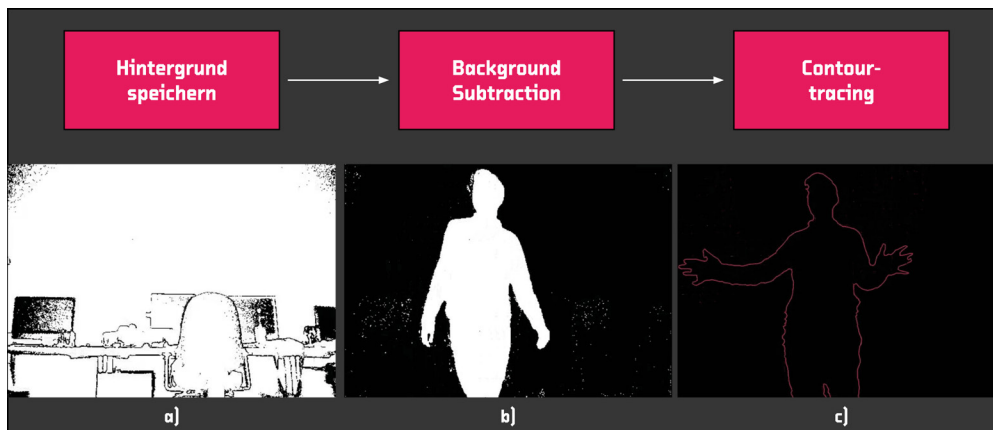


Abbildung 18: Der Trackingprozess in Argus Vision. In a) wird ein nicht kalibriertes, in b) ein kalibriertes Bild mit erfolgreicher Background Subtraction und in c) eine vollständig erkannte Kontur gezeigt.

In Abbildung 18 werden zwei aus dieser Regel entstandene Bilder gezeigt. Abbildung 18a ist dabei aus der Anwendung des Background Subtraction-Algorithmus auf einen leeren Hintergrund entstanden, einen Hintergrund also, bei dem der Tiefenwert für jeden Pixel null beträgt. Abbildung 18b zeigt eine erfolgreiche Background Subtraction. Für den Hintergrund wurde dabei ein Bild des Raums von Abbildung 18a gewählt. In Argus Control lässt sich dieses über das Drücken des Buttons Depth Calibration erstellen (Abbildung 10 II). Dadurch wird das aktuelle Bild der Kinect als Hintergrundbild gespeichert.

Für das menschliche Auge ist in Abbildung 18b bereits eine Person zu erkennen, tatsächlich werden aber nur schwarze und weiße Pixel dargestellt – Daten also, welche unverarbeitet keine Auskunft über eine Person, deren Kontur und andere Eigenschaften geben. Um aus diesem Schwarzweißbild die Konturen von Personen zu extrahieren, wird ein Contourtracing-Algorithmus von OpenCV<sup>28</sup> verwendet. Das Ergebnis dieses Algorithmus liefert eine endliche Menge an Konturen, welche wiederum aus einer endlichen Menge von Pixeln bestehen. Da auch die Störfaktoren im Bild als Kontur erkannt werden, können Konturen ausgeschlossen werden, welche kleiner als ein bestimmter Schwellenwert sind. Dies kann in der Oberfläche von Argus Control über den Slider Minimum Contoursize (Abbildung 10 III) eingestellt werden. Abschließend werden die Pixel wieder mit den Tiefenwerten der Kinect versehen. Als Ergebnis dieses Arbeitsschrittes erhält man also die aus dreidimensionalen Punkten, sogenannten Voxeln, bestehende Kontur jeder Person im Raum.

<sup>28</sup> <http://opencv.org/>, Abgerufen am 19.05.16

Durch das Reduzieren der Tiefenwerte auf ein 2D-Bild entsteht im letzten Schritt des skizzierten Algorithmus ein Nachteil. Der Contourtracing-Algorithmus ist nicht in der Lage zwischen zwei Personen zu unterscheiden, wenn deren Konturen sich schneiden. Diese werden dann als eine Person erkannt.

#### **Trackingunabhängige und -abhängige User- und Triggerzone-Daten berechnen**

Aus dem in Abbildung 17 dargestellten Trackingzyklus ist zu entnehmen, dass das Berechnen der Daten einer Person in zwei Schritte aufgeteilt ist (*User- und Triggerzone-Daten berechnen* sowie *Trackingabhängige User-Daten berechnen*). Die Daten setzen sich für jede Person wie folgt zusammen:

- Eindeutige ID
- Alter (Aufenthaltsdauer, gemessen als Anzahl der Bilder der Kinect)
- Kontur
- Centroid (Arithmetisches Mittel über x, y und z-Werte der Kontur)
- Center (Mittelpunkt des kleinstmöglichen Quaders um die Kontur)
- Geschwindigkeit
- Beschleunigung

Zusätzlich werden mit den Personendaten auch die Daten der jeweiligen Triggerzones errechnet. Diese bestehen nur aus drei Attributen:

- Eindeutige ID
- Anzahl der belegten Voxel
- Anzahl der belegten Voxel pro Person

Wie die Berechnung dieser Daten erfolgt, lässt sich dem Projektbericht entnehmen. Da sowohl das Alter als auch die Geschwindigkeit und Beschleunigung einer Person sich nur in Abhängigkeit der Zeit ermitteln lassen, sind diese von einem erfolgreichen Tracking abhängig und werden demnach erst nach dem User Tracking berechnet.

### User Tracking

Für das Tracking wird die Position aller Personen im aktuellen Bild mit der Position aller Personen aus dem vorherigen verglichen. Dadurch lässt sich eine Distanzmatrix aus den Abständen der einzelnen Centroids errechnen.

Alte Person Neue Person	Person_Alt <sub>A</sub>	Person_Alt <sub>B</sub>
Person_Neu <sub>A</sub>	385,2337	3,36744
Person_Neu <sub>B</sub>	365,9685	471,5642
Person_Neu <sub>C</sub>	0	380,5642

Abbildung 19: Ein Beispiel für eine Distanzmatrix, welche aus drei neuen und zwei alten Personen errechnet wird. Farblich hervorgehoben sind die zueinander zugordneten Personen. Person\_Neu<sub>B</sub> ist eine neue Person.

In Abbildung 19 wird eine solche gezeigt. Der Trackingprozess ergibt sich nun wie folgt:

Die Distanzmatrix wird auf ihr Minimum geprüft. Ist dieses gefunden (hier in einem ersten Schritt die Distanz zwischen Person\_Alt<sub>A</sub> und Person\_Neu<sub>C</sub>), so wird der neuen Person die ID der bereits erfassten Person zugeordnet. Beide Personen werden daraufhin aus der Distanzmatrix entfernt. Dieser Vorgang wird so lange wiederholt, bis entweder nur noch neue oder alte Personen vorhanden sind.

Ersteres tritt genau dann ein, wenn eine oder mehrere Personen den Kamerabereich betreten haben, die Anzahl der Personen insgesamt sich also erhöht. Diese werden mit einer neuen ID den bereits getrackten Personen hinzugefügt.

Umgekehrt bedeutet das, dass letzterer Fall genau dann eintritt, wenn eine oder mehrere Personen den Kamerabereich verlassen haben. Diese werden aus der Liste der getrackten Personen entfernt.

### User- & Triggerzone-Daten senden

Abschließend werden die Daten zu getrackten Personen und Triggerzones an Argus Control gesendet. Dabei werden genau diese Daten übermittelt, die bereits im Arbeitsschritt *Trackingunabhängige und -abhängige User- und Triggerzone-Daten berechnen* beschrieben wurden. Wie diese Daten versendet werden, wird im folgenden Abschnitt beschrieben.

### 4.3.2 Kommunikation in Argus Vision

Argus Vision übermittelt Daten über drei verschiedene Protokolle – TCP, UDP und OSC.

Dabei wird TCP für die verlustfreie Übermittlung der Einstellungen von Argus Control an die jeweilige Argus Kinect-Instanz genutzt.

UDP hingegen ist ein verlustbehaftetes Protokoll – es wird sowohl für das Übertragen der verschiedenen Streams als auch für die initiale Kopplungsphase zwischen Argus Kinect und Argus Control verwendet. Letztere funktioniert über ein Broadcasting-System. Argus Kinect sendet eine Kommunikationsaufforderung an jedes Gerät in sämtlichen verfügbaren Subnetzen. Empfängt Argus Control diese Nachricht, so bestätigt es den Empfang. Daraufhin wird eine TCP-Verbindung hergestellt und ein zusätzlicher Port für die Kommunikation festgelegt.

Über diesen erfolgt fortan die Übertragung der Tracking-Daten mittels OSC. Dieses basiert auf dem UDP-Standard und ist demnach ebenso verlustbehaftet. Der entscheidende Vorteil von OSC gegenüber anderen Protokollen ist, dass es von vielen Anwendungen, die im Ausstellungsbetrieb verwendet werden, unterstützt wird. Gleichzeitig ist die Arbeit mit dem Protokoll durch die simple Struktur der Nachrichten sehr einfach.

In diesem Abschnitt wird zuerst auf die Übertragung des Point Cloud-Streams eingegangen, bevor das Kapitel mit einer kurzen Erläuterung der verwendeten OSC-Syntax beendet wird.

#### Übertragung des Point Cloud-Streams

Die Übertragung von Point Cloud-Daten gestaltet sich durch die Menge der zu übertragenden Daten schwierig. Da in Argus Kinect nur Tiefenwerte für die Konstruktion der Point Cloud verwendet werden, reduziert sich diese zwar, es bleiben aber immer noch 217.088 Ganzzahlwerte zu senden. Zur Übertragung von Point Clouds existieren – genau wie für das Object Tracking – einige aufwendige Lösungen, wie beispielsweise die Komprimierung der Daten in einer Octree-Struktur<sup>29</sup>.

Ein einfacher Ansatz, welcher in Argus Vision verwendet wird, ist die Reduzierung der übertragenen Werte. Nach Abschluss der Projektphase wurde nur jeder vierte Tiefenwert übertragen, was für ein hinreichend komplexes Bild genügt.

Um die Datenmenge zusätzlich zu verringern, werden die Tiefenwerte auf RGB-Werte abgebildet. Daraus kann dann ein 2D-Farbbild erstellt werden, welches in einem weiteren Schritt mit gewöhnlichen Bildkompressionsverfahren bearbeitet werden kann. Dieser Schritt folgt dem Vorgehen von Pece et al., welche Tiefenwerte so auf RGB-Werte abbildeten, dass diese selbst nach einem verlustbehafteten Kompressionsverfahren wieder in ausreichendem Umfang hergestellt werden konnten (Pece et al. 2011).

---

<sup>29</sup> <http://pointclouds.org/documentation/tutorials/compression.php>, Abgerufen am 19.05.16

In Argus Vision wurde die Basis für eine Kompression nach diesem Vorbild zwar implementiert, es wurde jedoch ein verlustfreies Komprimierungsverfahren – die Komprimierung in das Grafikformat PNG – verwendet. Dies vereinfacht die Komplexität der Abbildung von Tiefenwerten auf Farbwerte.

Das komprimierte Bild wird anschließend über UDP von Argus Kinect an Argus Control übermittelt, dort dekomprimiert und die einzelnen Farbwerte werden wieder in Tiefenwerte übersetzt. Aus diesen wird schließlich die Point Cloud-Visualisierung generiert.

### **Kommunikation zwischen Argus Control und einem Client**

Die Daten, die Argus Control an einen Client weiterleitet, werden über das OSC-Protokoll übertragen. Die Kommunikation der Trackingdaten erfolgt also analog zu der Übertragung dieser von einer Argus Kinect-Instanz zu Argus Control. Die OSC-Nachrichten, welche Argus Control versendet, orientieren sich an dem TUIO-Standard, erweitern jedoch die Größe der Nachricht. Das Format der Nachrichten ist in Abbildung 20 dargestellt.

**/argusControl/Eventname/Attribut<sub>1</sub>/Variable<sub>1</sub>/.../Attribut<sub>n</sub>/Variable<sub>n</sub>/fseq/Frameanzahl**

Abbildung 20: Das Format der von Argus Control übermittelten Nachrichten.

Dabei sind solche Teile der Nachricht in Magenta hervorgehoben, die sich je nach Event unterscheiden. Verlässt beispielsweise eine Person den Kamerabereich, so wird dies mit dem Eventnamen *personLeft* signalisiert. Darauf folgt lediglich das Attribut *set* und die *ID der Person*. Geschlossen wird jede Nachricht mit einer Schlusssequenz (*fseq*), die die Laufzeit von Argus Control in Bildern angibt. So können die Nachrichten bei verlustbehafteter Kommunikation in die richtige Reihenfolge gebracht werden.

Für das Tracking von Personen existieren folgende Events:

- personEntered (Eine Person betrat den Kamerabereich)
- personMoved (Eine Person bewegte sich vor der Kamera)
- personLeft (Eine Person verließ den Kamerabereich)

Für die Triggerzones hingegen wurde eine solche Event-Struktur nicht aufgebaut. Ein Update einer Triggerzone wird nur durch den Eventnamen *triggerzone* gekennzeichnet.

## 4.4 Zusammenfassung

In diesem Kapitel wurde das Tracking Tool Argus Vision mit sämtlichen Funktionen vorgestellt. Gleichzeitig wurden wichtige Teile der Programmierung des Tools erörtert. Zusammenfassend wird aufgezeigt, inwiefern das Tool die in der Anforderungsanalyse erhobenen Anforderungen (siehe Abschnitt 3.1) erfüllt. Diese Überlegungen werden dabei nur auf Basis des Tools zum Zeitpunkt der Fertigstellung der Projektarbeit angestellt. Tiefergehender wird Argus Vision in Kapitel 5 evaluiert.

Um transdisziplinär arbeitenden Ausstellungsgestaltenden die Arbeit mit einem DIY Tool zu ermöglichen, muss dieses über eine übersichtliche und intuitive Bedienoberfläche verfügen (A1). In Argus Vision wurden hierzu die Einstellungsmöglichkeiten sowie die benötigten Schritte zur ersten Inbetriebnahme reduziert – es muss in Argus Kinect und Argus Control nur derselbe Port angegeben werden, um eine Verbindung zwischen den beiden Programmteilen herzustellen. Zusätzlich sieht der\*die Benutzer\*in die Auswirkung seiner\*ihrer Interaktion jederzeit auf der Bedienoberfläche.

Für das Erkennen von Personen aus beliebigem Winkel (A2) wurde ein eigener Tracking-Algorithmus entwickelt. Dieser erfüllt die Anforderung in ausreichendem Maße. Personen werden, sofern sie getrennt voneinander stehen, zuverlässig erkannt. Das Tracking funktioniert dabei aus jedem Winkel. Ein Nachteil ist, dass zwei oder mehrere Personen als eine erkannt werden, wenn deren Konturen sich schneiden.

Über die grafische Oberfläche lassen sich Triggerzones beliebiger Größe an einer ebenfalls beliebigen Position im Raum erstellen (A3). Die Erstellung und Positionierung dieser Triggerzones gestaltet sich dabei einfach und wird durch die Point Cloud-Visualisierung unterstützt. Die Triggerzones haben dabei die Form von Quadern; weitere Auswahlmöglichkeiten, insbesondere Ellipsen, sollten zusätzlich zur Verfügung gestellt werden.

Die Steuerung aller Kinects erfolgt über ein zentrales Tool – Argus Control (A4). Sowohl die Steuerung als auch die Überwachung werden von dort aus übernommen.

Die Einstellungen für sämtliche jemals verbundene Argus Kinect-Instanzen werden in Argus Control gespeichert und, sobald sich eine dieser Instanzen erneut verbindet, automatisch übermittelt. Argus Control sowie Argus Kinect speichern zusätzlich ihre Kommunikations-Einstellungen, um beim Start die Verbindung zueinander herzustellen. Argus Control und Argus Kinect bieten allerdings keine Funktion für das automatische Einschreiben in den Autostart des jeweiligen Betriebssystems – dies muss also manuell erfolgen. Die Anforderung Automatisierter Start (A5) ist demnach nur teilweise erfüllt.

Als Protokoll zur Übertragung der Tracking-Daten (A6) wurde OSC gewählt. OSC ist verbindungslos – es können also Daten einfach von Argus Control an den Client gesendet werden, ohne vorher eine Verbindung aufbauen zu müssen. Gleichzeitig wird es von einer Vielzahl von Programmen, die in Ausstellungen verwendet werden, unterstützt. Zusätzlich ist die Struktur der Nachrichten sehr einfach und verständlich. Das Tracking Tool TSPS (vorgestellt in Abschnitt 3.2) bietet jedoch mehr Konfigurationsmöglichkeiten. In diesem lassen sich

verschiedene Protokolle wählen. Attribute, die nicht benötigt werden, können aus den Daten ausgeschlossen werden, um die Datenlast zu verringern.

Das Tracking von Personen (A7) basiert in Argus Vision auf der Annahme, dass eine Person nur eine minimale Distanz in einem Bruchteil einer Sekunde zurücklegen kann. Es wird also die Position aller Personen im aktuellen Bild mit der Position aller Personen im vorherigen verglichen. Dies ist ein rudimentärer Tracking-Algorithmus – er funktioniert so lange gut, wie Personen auseinander stehen. Sind zwei Personen nah beieinander, kann es zu Fehlern im Tracking kommen. Fehlerhaft ist das Tracking in jedem Fall, wenn zwei Silhouetten sich schneiden, da hier das Erkennen von Personen aus beliebigem Winkel die Anzahl der Personen von zwei auf eine reduziert.

Abschließend wird nun Argus Vision den in Abschnitt 3.2 untersuchten Tools gegenüber gestellt.

Tracking Tool	A1	A2	A3	A4	A5	A6	A7
Proximity Toolkit	■	■	■	■	■	■	■
XDKinect	■	■	■	■	■	■	■
SoD-Toolkit	■	■	■	■	■	■	■
TSPS	■	■	■	■	■	■	■
Argus Vision	■	■	■	■	■	■	■

Tabelle 3: Argus Vision im Vergleich mit den in Abschnitt 3.2 vorgestellten Tracking Tools.  
 ■ = Erfüllt die Anforderungen, ■ = Erfüllt die Anforderungen teilweise, ■ = Erfüllt die Anforderungen nicht.



## 5 Evaluation von Argus Vision

Die Evaluation von Argus Vision wird in dieser Arbeit in zwei Schritten vorgenommen.

Um zu prüfen, inwiefern sich das Tracking Tool für die Umsetzung interaktiver Installationen einsetzen lässt, wurden in einem ersten Schritt zwei Fallstudien durchgeführt. Gleichzeitig konnten in der Anwendung Fehler entdeckt und Verbesserungsmöglichkeiten herausgearbeitet werden. Die Fallstudien unterscheiden sich nicht nur thematisch, sondern auch in der Verwendung von Argus Vision. In der Installation „Bruch“ wurden die Centroids der erkannten Personen verwendet. In einer zweiten Fallstudie wurde hingegen das Triggerzone-Prinzip getestet. Zusätzlich konnte in dieser Fallstudie geprüft werden, inwiefern sich Argus Vision an ein anderes DIY Tool für den Ausstellungsbereich anbinden lässt.

In einem zweiten Schritt wurden Interviews mit Ausstellungsgestaltenden gestalterischer und technischer Disziplinen geführt, welche das Tool in verschiedenen Kategorien bewerten sollten.

### 5.1 Fallstudie – Installation „Bruch“

Im Rahmen der Veranstaltungsreihe Blended Museum<sup>30</sup>, in welcher Studierende aus den Bereichen Architektur, Geschichte, Kommunikationsdesign und Mensch-Computer-Interaktion an neuartigen Ausstellungskonzepten arbeiten, entstand die Ausstellung „Tell Genderes – 20 Meter Menschheitsgeschichte“<sup>31</sup>. In dieser wurden Ergebnisse einer Ausgrabung auf dem Siedlungshügel Tell Genderes im Nordwesten Syriens auf vier Geschossen des BildungstURMs Konstanz<sup>32</sup> präsentiert. In dem vierten Geschoss wurde auf die aktuelle politische Situation in Syrien und den damit verbundenen Verlust von Kulturgütern durch Kunsthandel, Kunstraub und Zerstörung aufmerksam gemacht. Dafür wurde eine Installation mit Argus Vision umgesetzt. Die Installation selbst sowie deren Umsetzung und die Integration von Argus Vision in diese werden im Folgenden erläutert.

---

<sup>30</sup> <http://hci.uni-konstanz.de/index.php?a=research&b=projects&c=8609065>, Abgerufen am 25.05.16

<sup>31</sup> <http://tell-genderes.de/>, Abgerufen am 25.05.16

<sup>32</sup> <http://www.konstanz.de/wirtschaft/00915/00918/index.html>, Abgerufen am 25.05.16

### 5.1.1 Die Installation „Bruch“

Im Zentrum des dritten Stockwerks steht eine leere, schwach beleuchtete, zerstörte Vitrine metaphorisch für die zerstörten und verschwundenen Kulturgüter. In einigem Abstand dazu hängen Gazestoffe von der Decke herab, welche die Seiten der Vitrine imitieren – es wird also eine „Stoffvitrine“ um die eigentliche gespannt, auf welche virtuelles Glas projiziert wird.

Besuchende, die auf die tatsächliche Vitrine zugehen und dabei „in“ die Stoffvitrine treten, brechen damit unabsichtlich oder absichtlich das virtuelle Glas an oder zerbrechen dieses vollständig. In letzterem Fall wird in dem daraus entstandenen Guckloch Bild- und Videomaterial von Kulturgüterstörungen gezeigt (siehe Abbildung 21).

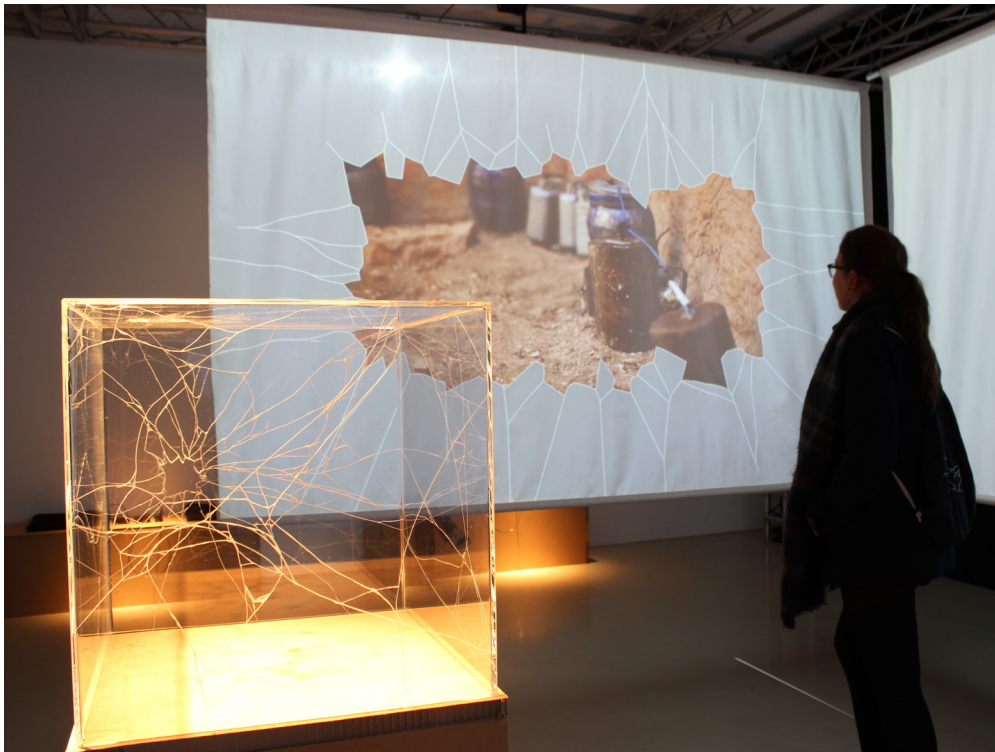


Abbildung 21: Eine Besucherin steht vor einer der drei Stoffseiten der Stoffvitrine – das projizierte Glas ist zerbrochen. Im Vordergrund ist die reale Vitrine zu sehen.

### 5.1.2 Umsetzung der Installation

Die Umsetzung der Installation gliederte sich in drei Teile: Die Planung der Technik und Architektur, die Implementierung der Animation sowie die Integration von Argus Vision in diese. Im Folgenden wird der finale Aufbau der Technik und der Architektur kurz aufgezeigt. Anschließend wird kurz die Funktionsweise des Algorithmus skizziert, um die Integration von Argus Vision beschreiben zu können.

#### Umsetzung der Technik und Architektur

Die komplexe Planung der Architektur und Medientechnik erfolgte gemeinsam mit Architekturstudierenden der HTWG Konstanz und dem Veranstaltungstechnikbüro J&C<sup>33</sup>.

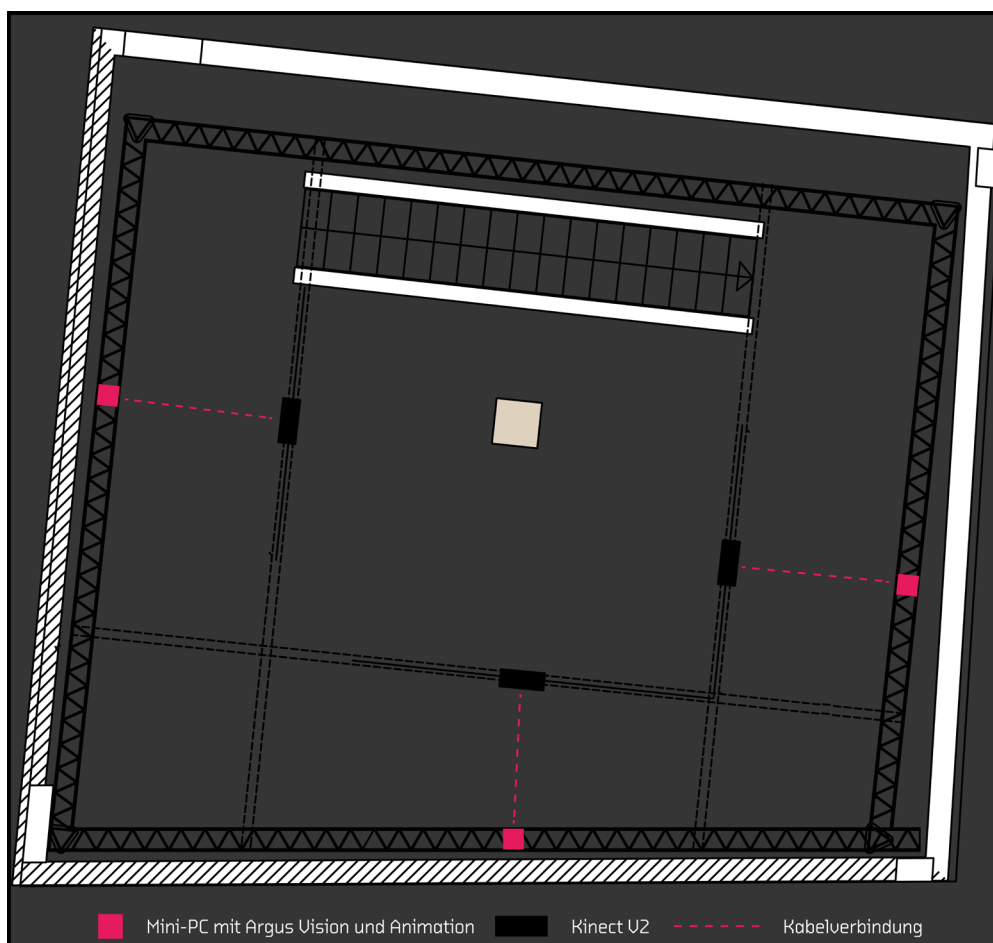


Abbildung 22: Die technische Planung der Installation.

<sup>33</sup> <http://www.jundc.com/>, Abgerufen am 26.05.16

Um die Gaze in der Mitte des Raums aufhängen zu können und eine Konstruktion zu verwenden, die drei Projektoren und zusätzliche Geräte tragen kann, wurde ein raumspannendes Traversensystem aufgebaut. Die Gaze selbst wurden an Stangen, die zwischen dem Traversensystem angebracht wurden, aufgehängt (schwarz gestrichelte Linien in Abbildung 22).

Auf jedem der – auf den Traversen angebrachten Mini-PCs – wurde sowohl Argus Kinect und Argus Control betrieben als auch die Animation abgespielt. In dieser Installation wurde also nicht von der Multi-Kinect-Fähigkeit von Argus Vision Gebrauch gemacht. Die Kinects wurden direkt über den Gaze angebracht und auf den Raum vor diesen ausgerichtet.

### Integration von Argus Vision

In der Animation in der Installation „Bruch“ kann der\*die Besucher\*in bewusst oder unbewusst Risse in dem Glas einer virtuellen Vitrine erzeugen oder dieses gänzlich zerstören. Die Risse im Glas werden dabei anhand der Position der Besuchenden im Raum berechnet. Diese entstehen genau dort, wo die Person vor der Gaze steht. Die Distanz der Person zur Gaze bestimmt die Größe des Risses. Dies wird in Abbildung 23 skizziert.

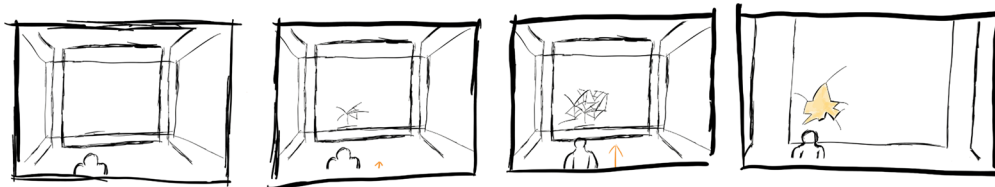


Abbildung 23: Eine Person bewegt sich auf eine Gaze zu und bringt dadurch das Glas zunehmend zum Zerschlagen.

Überschreitet eine Person eine bestimmte Entfernungsgrenze zur Gaze bricht das Glas vollständig ein und gibt ein Bild oder Video frei (siehe Abbildung 23). Nach einer Zeitspanne verblasst die Darstellung und der Algorithmus beginnt erneut.

Die Präsenz von Besuchenden wurde in der Prototyping-Phase fingiert. Mit der Position des Mauszeigers wurde die Position eines Besuchenden vor der Projektion simuliert. Über die Tastatur konnte die Nähe des Besuchenden zur Gaze verändert werden. Die Integration von Argus Vision erfolgte erst vor Ort und nach dem Aufbau der physischen Installation.

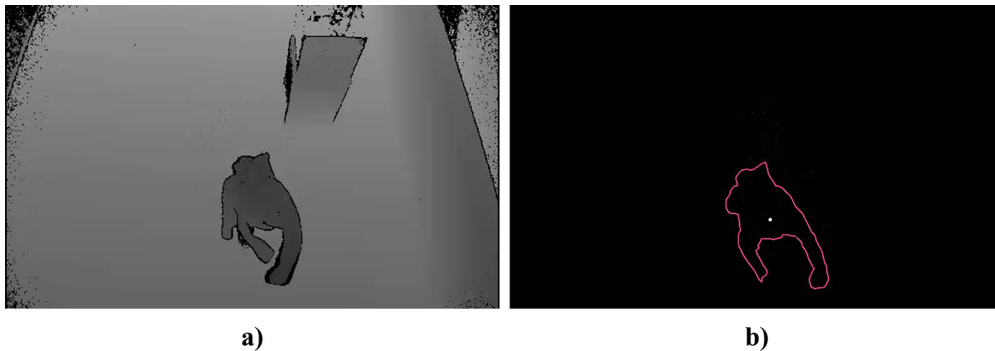


Abbildung 24: Der Tiefen- und Tracking-Stream einer Kinect in der Installation. Hier schlägt eine Person auf die Gaze ein, um das Glas zerspringen zu lassen.

Diese teilte sich in zwei Aufgaben: Zuerst wurde der Empfang und die Verarbeitung der Argus Vision-Daten in der Animation programmiert. Danach musste die Maus- und Tastatursteuerung, die zuvor beschrieben wurde, durch diese Daten ersetzt werden.

Maus- und Tastatursteuerung wurden durch die Centroids der Besuchenden ersetzt. Die x-Position eines Centroids bestimmte, wo auf der Projektion Risse entstanden – der y-Wert, wie groß der Radius war, in dem diese sich ausbreiteten.

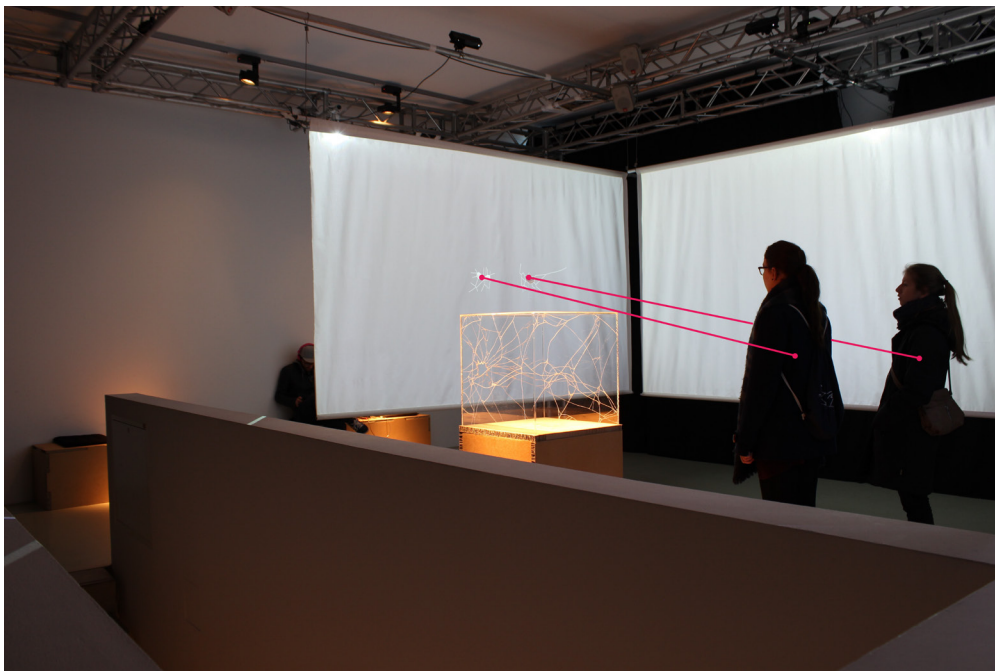


Abbildung 25: Risse bilden sich an der x-Position der Centroids der zwei Besuchenden.

### 5.1.3 Ergebnisse aus der Anwendung von Argus Vision

Die Integration von Argus Vision in die Installation war erfolgreich. Diese wurde über einen Zeitraum von 44 Tagen ausgestellt, während dessen Argus Vision fehlerfrei funktionierte.

Allerdings fiel nach dem Anbringen der Kinects an dem Traversensystem auf, dass der Point Cloud-Stream nicht mehr übertragen wurde, da die Paketgröße der übertragenen Daten das Paketlimit überschritt. Dies hatte für die Fallstudie keine Auswirkung, da der Point Cloud-Stream hauptsächlich für die Platzierung von Triggerzones eingesetzt werden soll.

Für den Empfang der Daten von Argus Vision in der Animation musste zunächst ein OSC-Plug-In verwendet werden, um anschließend die Syntax der empfangenen Daten zu analysieren. So konnten beispielsweise Triggerzone-Daten oder nicht relevante Attribute von Personen direkt verworfen werden. Um diese Daten abzuspeichern wurde die Klasse Person, in der bereits in Argus Vision die Personendaten gespeichert werden, verwendet.

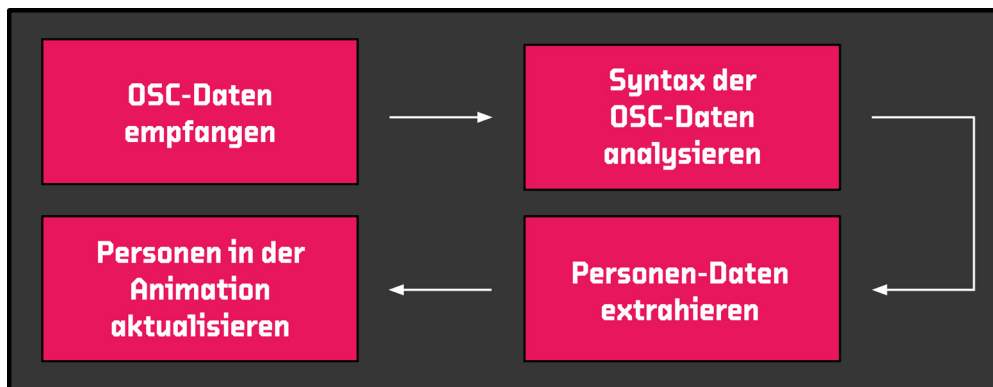


Abbildung 26: Verarbeitung der Argus Vision-Daten in der Animation.

Dieser Prozess, dargestellt in Abbildung 26, findet in jeder Applikation statt, welche Argus Vision nutzen will und sollte demnach standardisiert Ausstellungsgestaltenden zur Verfügung gestellt werden (vgl. 5.1.4).

### 5.1.4 Verbesserungsmöglichkeiten für Argus Vision

Während der Integration von Argus Vision in die Glas-Animation zeigte sich in der Programmierung von Argus Vision ein Fehler sowie eine Möglichkeit zur besseren Anbindung von Argus Control an einen Client.

Letzteres lässt sich daraus ableiten, dass der in Abbildung 26 dargestellte Prozess in dieser Form in jeder Anwendung Verwendung finden muss. Ziel fortführender Arbeiten muss also sein, diesen Prozess in Plug-Ins für eine Vielzahl verschiedener Programmiersprachen und Tools zu kapseln. So würden Ausstellungsgestaltende einen leichten Zugriff auf Personen- und Triggerzone-Daten erhalten, ohne vorher eine Syntaxanalyse durchführen zu müssen oder eine geeignete Datenstruktur zu schaffen.

Für das Problem der Übertragung der Point Cloud-Daten existieren zwei Lösungsansätze. Ersterer ist, wie bereits in Abschnitt 4.3.2 beschrieben, die Verwendung einer komplexeren Methode zur Point Cloud-Komprimierung. Für die Verwendung in der weiteren Fallstudie wurde jedoch lediglich die Menge der übertragenen Tiefendaten reduziert. Dadurch sinkt der Detailgrad der Point Cloud, sie konnte aber in Fallstudie 5.2 in ausreichender Form verwendet werden.

## 5.2 Fallstudie – Anbindung von Argus Vision an smartPerform

Wie bereits in Abschnitt 3.1 beschrieben, ist es wichtig, dass Argus Vision leicht an andere Tools, welche im Ausstellungsbereich Anwendung finden, anzubinden ist. Zusätzlich sollte Argus Vision während einer Prototyping-Phase einsetzbar sein. Die dafür benötigte Flexibilität sollte insbesondere durch die Verwendung von Triggerzones bereitgestellt werden. Diese wurden dementsprechend in einer weiteren Fallstudie verwendet, um Argus Vision an ein bereits bestehendes, kommerzielles Tool anzubinden, welches unter anderem im Ausstellungsbereich eingesetzt wird.

Das Tool, an welches Argus Vision angebinden wurde, ist smartPerform<sup>34</sup>. Mit diesem lassen sich Präsentationen für interaktive Tische und Displays per Drag and Drop umsetzen. Um Programmabläufe festlegen zu können, bietet smartPerform einen sogenannten Logik-Designer. In diesem sind Funktionen in Bausteinen gekapselt. Diese können miteinander verbunden werden, um komplexere Programmabläufe zu definieren. Der Logik-Designer wurde verwendet, um Argus Vision anzubinden.

### 5.2.1 Die Präsentation in smartPerform

Die Präsentation<sup>35</sup> zeigte Bilder und Videos der in Abschnitt 5.1 vorgestellten Ausstellung. Dabei wurde die Präsentationsfläche in vier Bereiche unterteilt, welche die einzelnen Geschosse der Ausstellung repräsentierten. Die Präsentation selbst wurde an eine Wand projiziert, um ein Public Display zu simulieren, an welchem Passanten vorbeigehen.

Während des Standby-Modus der Präsentation befinden sich alle Medien in der Mitte der Präsentationsfläche. Betritt eine Person einen bestimmten – einem Geschoss zugeordneten – Bereich, so bewegen sich die passenden Bilder an die Position dieses Bereichs und vergrößern sich gleichzeitig. Verlässt die Person den Bereich, transformieren sich die Bilder wieder an ihren Ausgangspunkt zurück. Die Bereiche wurden in dem Testaufbau mit einem Kreuz auf dem Boden gekennzeichnet. Die Interaktion mit der Präsentation wird in Abbildung 27 gezeigt.

---

<sup>34</sup> <http://www.smartperform.de/de>, Abgerufen am 28.05.16

<sup>35</sup> Die Präsentation kann dem der Arbeit beiliegenden USB-Stick entnommen werden.



Abbildung 27: a) Ein Mann steht in einem, mit einem Kreuz markierten Bereich und betrachtet die präsentierten Bilder. Eine Frau bewegt sich in Richtung eines Kreuzes. b) Durch das Betreten der Fläche um das Kreuz werden nun auch an ihrer Position Bilder gezeigt.

### 5.2.2 Integration von Argus Vision

Um den Bereich vor der Projektion in Argus Vision abzudecken wurden zwei Kinect-Kameras – jeweils eine an den Wänden orthogonal zur Projektion – angebracht, sodass jede Kinect zwei von vier Interaktionsbereiche überwachte (vgl. Abbildung 28). Um die Daten kumuliert an smartPerform weiterzuleiten, wurden die zwei Argus Kinect-Instanzen mit einer Argus Control-Instanz, die auf einem anderen PC betrieben wurde, über WLAN verbunden.

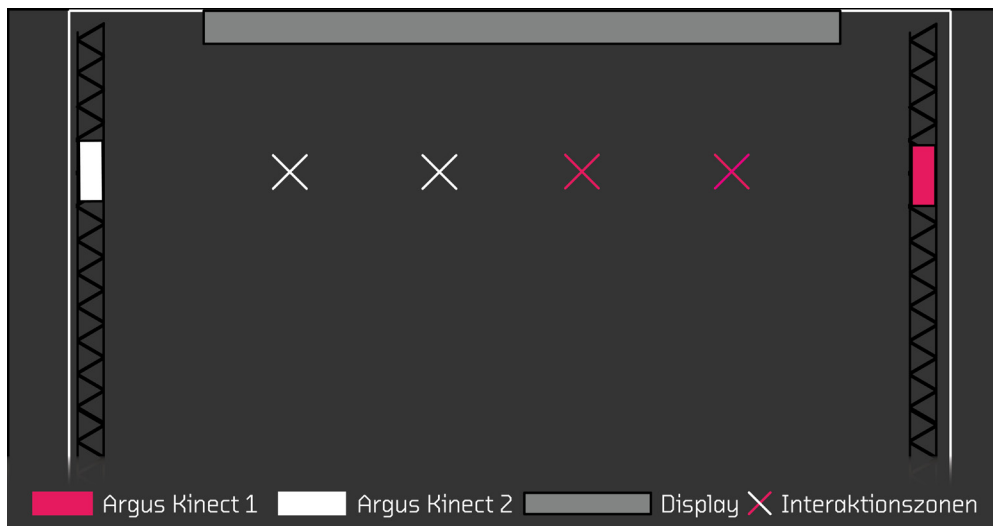


Abbildung 28: Skizzenhafter Aufbau der Fallstudie



Das Konzept zur Integration von Argus Vision sah vor, dass Triggerzonen an den Stellen eingezeichnet werden, welche mit einem Kreuz gekennzeichnet sind. Damit dies möglich ist, wurde davor die übertragene Punktzahl der Point Cloud reduziert, sodass die Übertragung fehlerfrei funktionierte (vgl. Abschnitt 5.1.4).

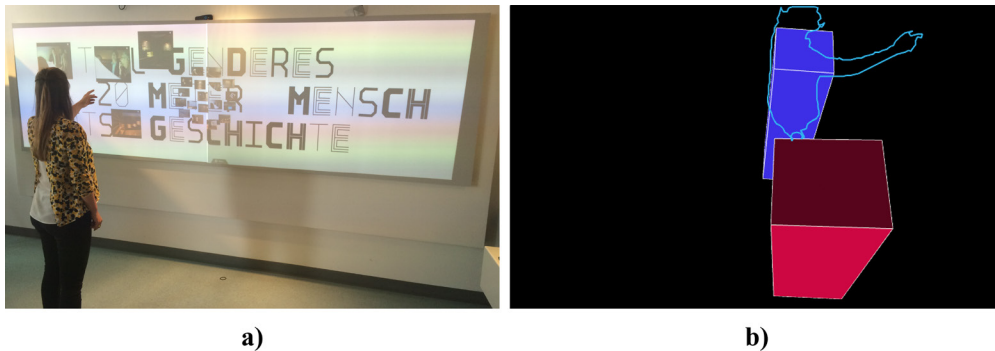


Abbildung 29: a) Eine Frau zeigt auf ein präsentiertes Bild. b) Dieselbe Szene im Trackingview von Argus Control.

Für die Positionierung der Triggerzonen wurde ein großes Objekt an die Position der Interaktionsbereiche gestellt. Darauffolgend mussten die Triggerzonen einfach in der Point Cloud Ansicht über die Objekte gelegt werden.

Für die Kommunikation über ein Netzwerk bietet smartPerform einen UDP Knoten, mit dem die Triggerzone-Daten von Argus Control empfangen werden sollten.

### 5.2.3 Ergebnisse aus der Anwendung von Argus Vision

Auch in dieser Fallstudie ließ sich Argus Vision erfolgreich anbinden. Hierfür waren einige Modifikationen erforderlich. Zusätzlich wurde ein Übertragungsproblem festgestellt, welches sich im Verlauf der Fallstudie nicht lösen ließ.

Wie bereits in Abschnitt 5.2.2 erwähnt, sollten die Daten von Argus Control an smartPerform über eine UDP-Schnittstelle übertragen werden. Da OSC – das Netzwerkprotokoll zur Übertragung von Tracking-Daten von Argus Vision – auf dem UDP-Standard basiert, war smartPerform zwar in der Lage die Nachrichten zu empfangen, nicht jedoch diese auszulesen, da die Zeichenformatierung, welche OSC verwendet, nicht bekannt war. Für den Empfang der OSC-Nachrichten musste demnach ein weiteres Tool verwendet werden, welches sodann die verarbeiteten Daten über UDP an smartPerform weiterleitete. Dafür wurde vvvv verwendet, ein Tool welches für die Gestaltung großer Medienlandschaften entwickelt wurde. vvvv findet demnach häufig Anwendung in Ausstellungen und es können – ähnlich zu smartPerform – Programmabläufe durch das Verbinden von Bausteinen erstellt werden<sup>36</sup>.

<sup>36</sup> Der verwendete vvvv Patch kann ebenfalls dem beiliegenden USB Stick entnommen werden.

Zusätzlich zur Weiterleitung der Daten wurde vvvv außerdem verwendet, um die von Argus Vision gesendeten Daten zu verarbeiten. Argus Vision sendet bei jeder Veränderung der Belegung einer Triggerzone den aktuellen Belegungsgrad in numerischer Form. Dieser kann mit smartPerform mit regulären Ausdrücken ausgelesen oder aber in einer einfacheren Form durch eine bedingte Anweisung (beispielsweise eine if-Abfrage) in vvvv überprüft werden – da die Anbindung möglichst einfach funktionieren sollte, wurde hier die letztere Variante verwendet.

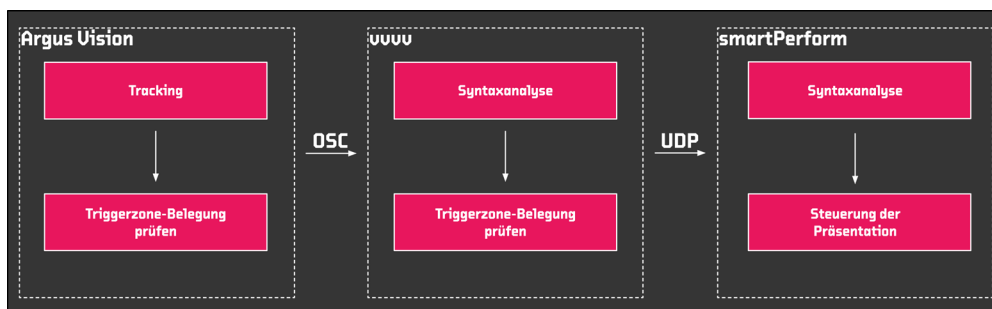


Abbildung 30: Die Aufgaben der einzelnen Tools in der Fallstudie.

Bei der Übertragung über WLAN kam es mitunter zu Verbindungseinbrüchen und dem Verlust von Daten, sobald mehrere Personen erkannt wurden und dadurch die Menge der zu übertragenen Daten zunahm. Letzteres war insbesondere dann problematisch, wenn nicht übermittelt wurde, dass Personen Triggerzonen verlassen hatten. Dies hatte zur Folge, dass die geöffneten Medien sich nicht mehr in ihre Ausgangsposition zurückbewegten. In früheren Tests mit bis zu vier Argus Kinect-Instanzen war dies nicht eingetreten. Diese Begrenzung der Kinect-Zahl lässt sich auf die Verwendung von WLAN gegenüber kabelgebundenem LAN zurückführen.

#### 5.2.4 Verbesserungsmöglichkeiten für Argus Vision

In der Fallstudie ließen sich Probleme ausschließlich in der Kommunikation zwischen Argus Vision und smartPerform feststellen. Diese wiederum lassen sich in zwei Bereiche einteilen – Übertragungs- und Syntaxprobleme.

Die Übertragungsprobleme traten bei der Verwendung von zwei Kinects, bzw. der Übertragung derer kumulierten Daten an vvvv über WLAN auf. Zusätzlich ließen sich mit smartPerform die von Argus Control über OSC gesendeten Daten nicht auslesen.

Wie in Abschnitt 4.4 bereits aufgeführt, ließe sich ein Teil der Übertragungsprobleme dadurch lösen, dass mehrere Netzwerkprotokolle, wie TCP, UDP und TUIO, in Argus Control auswählbar wären. In dem Beispiel smartPerform wäre es für Benutzer\*innen dann möglich gewesen, die Daten auszulesen.

Um die Daten daraufhin über einfache reguläre Ausdrücke zu filtern, sollten Teile der Syntax frei konfigurierbar sein. So sollte der\*die Benutzer\*in die Möglichkeit erhalten, nur bestimmte Daten zu übertragen, bzw. nicht benötigte Daten auszuschließen. Für die Verwendung von Triggerzones sollte zusätzlich in Argus Control die Möglichkeit existieren, ein Belegungslimit einzustellen, ab welchem die Triggerzones als betreten angesehen werden. In einer Applikation wie smartPerform müsste sodann nur noch ausgelesen werden, ob eine Triggerzone aktiviert ist.

Die Behebung dieser Syntaxprobleme kann sich zusätzlich positiv auf die Übertragungsfehler auswirken. Bei Datenverlust in der Übertragung können die Personen- und Triggerzone-Daten einfach reduziert werden, indem bestimmte Attribute von der Übertragung ausgeschlossen werden. Falls sämtliche dieser Daten übertragen werden müssen, ist TCP – ein verlustfreies Netzwerkprotokoll – zu wählen. Hierfür müssten in einer fortführenden Arbeit Performance-Tests durchgeführt werden, um zu prüfen, ob die hohe Datenmenge zu einer starken Verzögerung in der Übertragung führt.

## 5.3 Interviews mit Ausstellungsgestaltenden

Um Argus Vision nicht ausschließlich über die eigene Verwendung im Rahmen zweier Fallstudien zu prüfen, wurden zusätzlich semi-strukturierte Interviews mit Ausstellungsgestaltenden, der Zielgruppe des Tools, geführt. Dazu wurden fünfzehn Ausstellungsgestaltungsbüros kontaktiert und um einen Interviewtermin gebeten. Bei der Auswahl dieser wurde darauf geachtet, Büros zu wählen, welche sich in der Gewichtung der dort vertretenen Disziplinen unterscheiden.

Die Interviewten sind:

- Dominik Hegemann, Assoziierter, Atelier Brückner<sup>37</sup> in Stuttgart
- Sebastian Oschatz, Geschäftsführer, MESO Digital Interiors<sup>38</sup> in Frankfurt am Main
- Prof. Eberhard Schlag, Partner, Atelier Brückner in Stuttgart
- Christoph Diederichs, Interaktionsgestalter, Atelier Markgraph<sup>39</sup> in Frankfurt am Main
- Prof. Thomas Hundt, Gründer und Geschäftsführer, jangled nerves<sup>40</sup> in Stuttgart

### 5.3.1 Durchführung der Interviews

Für die Durchführung der Interviews wurde ein Leitfaden mit Fragen erstellt<sup>41</sup>. Dieser dient zum einen der inhaltlichen Strukturierung des Interviews, zum anderen als Grundlage für die Analyse der Interviews (Kaiser 2014). Der Interviewleitfaden war aufgeteilt in eine Aufwärmphase, Fragen zu Argus Vision, zu den Triggerzonen, zu Verwendungsmöglichkeiten sowie zu Weiterentwicklungsmöglichkeiten und fehlenden Funktionen des Tools.

In semi-strukturierten Interviews dient der Leitfaden dabei nur als grobe Richtlinie. Der\*die Interviewende kann also jederzeit Fragen stellen, welche nicht im Leitfaden vorgesehen sind, auf Fragen aus dem Leitfaden verzichten oder diese zu einem anderen Zeitpunkt stellen (Lazar et al. 2010).

In einer Aufwärmphase wurden den Interviewten jeweils Fragen zu ihrer Arbeitsweise, zu durchgeführten Projekten und zu aktuellen Trends in der Ausstellungsgestaltung gestellt. Neben der Informationsgewinnung ist diese Phase wichtig, um die Befragten sich in der Interviewsituation wohl fühlen zu lassen (ebda.).

Daraufhin wurden dem\*der Interviewpartner\*in Argus Vision in einer kurzen Präsentation vorgestellt<sup>42</sup>. Dies war nötig, da die Ausstellungsgestaltenden nicht die Zeit hatten, das Tool selbst anzuwenden und ausführlich zu testen. Dementsprechend bezogen sich die Fragen,

---

<sup>37</sup> <http://www.atelier-brueckner.de/de>, Abgerufen am 07.06.16

<sup>38</sup> <http://www.meso.net>, Abgerufen am 07.06.16

<sup>39</sup> <https://www.markgraph.de>, Abgerufen am 07.06.16

<sup>40</sup> <http://www.janglednerves.com>, Abgerufen am 07.06.16

<sup>41</sup> Dieser findet sich im Anhang auf Seite 56.

<sup>42</sup> Diese lässt sich dem der Arbeit beiliegenden USB-Stick entnehmen.

welche zu Argus Vision gestellt wurden, auf mögliche Verwendungszwecke, User Experience Aspekte der Oberfläche sowie auf die wahrgenommene Bedienbarkeit des Tools.

Die Fragen, die zu Argus Vision gestellt wurden, sind sämtlich dem generischen qualitativen Fragebogen von Nestler et al. entnommen (Nestler et al. 2011). Dieser ermöglicht die Generierung quantitativer Ergebnisse auf Basis qualitativer Leitfadeninterviews durch die Zuordnung von Fragen zu gewichteten Kategorien (ebda.). Diese Vorgehensweise wurde in dieser Arbeit jedoch nicht angewendet, da die Ausstellungsgestaltenden – wie bereits erwähnt – das Tool lediglich auf Grundlage der Präsentation während des Interviews bewerten konnten.

### 5.3.2 Analyse der Interviews

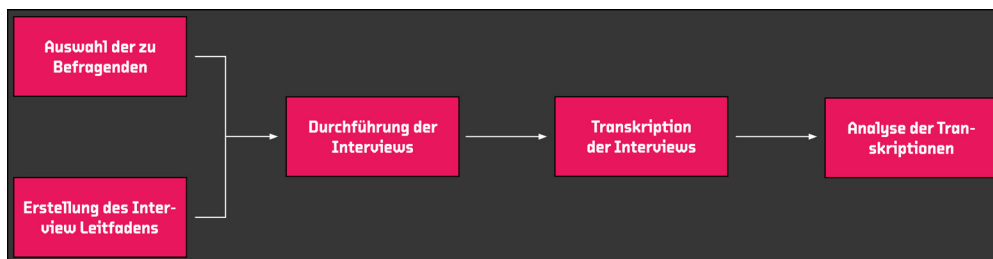


Abbildung 31: Schematischer Ablauf des Interviewprozesses

Nach der Durchführung der Interviews wurden diese in einem ersten Schritt vollständig transkribiert. Für Expert\*inneninterviews ist eine einfache Verschriftlichung des Gesprochenen ausreichend – zusätzliche Informationen, wie Pausen, Laute, Mimik oder Gestik müssen nicht berücksichtigt werden (Burger 2011).

Den Ausstellungsgestaltenden wurde anschließend das Transkript übermittelt. Diese konnten darin Änderungen vornehmen, sofern die Antworten inhaltlich nicht verfälscht wurden. Zusätzlich war dies erforderlich, da Teile der Antworten der Ausstellungsgestaltenden Informationen zu aktuellen Projekten enthielten, die nicht der Öffentlichkeit zugänglich gemacht werden durften. Da ein Großteil der Interviews aus Berichten von eigenen Projekten der Ausstellungsgestaltenden besteht, wurden die Interviews nicht anonymisiert. Die Interviewten stimmten dem zu. Die finalen Transkripte, die vollständig dem Anhang entnommen werden können, waren Basis der Analyse.

Darauffolgend wurden die Transkripte analysiert und einzelne Passagen grob Kategorien zugeordnet. Die Kategorien basierten auf den im Leitfaden formulierten Fragen.

### 5.3.3 Ergebnisse der Interviews

Die Oberfläche von Argus Vision wurde von allen Befragten positiv bewertet. Das Design wurde unter anderem als „slick“ (i.e. glatt, poliert) und „passend“ bezeichnet (Diederichs, S. 90) – es erwecke den Eindruck, „als hätte da auch ein Gestalter mitgewirkt“ (Schlag, S. 85).

Dabei wurde die Übersichtlichkeit und Verständlichkeit der Bedienoberfläche (Hegemann, S. 66; Oschatz, S.77; Schlag S. 85; Diederichs, S. 91; Hundt, S. 100) sowie der Programmierlogik (Oschatz, S. 75; Hundt, S. 101) von mehreren Befragten positiv herausgestellt. Hundt beispielsweise bezeichnete die Oberfläche als „aufgeräumt“ und führte weiter aus, dass diese den Eindruck erwecke, dass sowohl er als auch seine Mitarbeiter sofort „damit loslegen können“ (Hundt, S. 100). Um Fernwartungen durchführen zu können und die Überwachung von Argus Vision von jedem Gerät aus zu ermöglichen, sollte das Tool zusätzlich zentral über einen Browser gesteuert werden können (Oschatz, S. 74). Erstmaligen Benutzer\*innen soll die Anwendung von Argus Vision durch die Verwendung eines „Tooltipmodus“ erleichtert werden (Diederichs, S. 91). Darüber hinaus, so Diederichs, könnte Argus Control in zwei Fenster aufgeteilt werden – einen „Edit Modus“, in dem man „Triggerpunkte setzen kann“ sowie einen „Setup Modus“ für das Vornehmen der restlichen Einstellungen und für die Überwachung der Kinects (Diederichs, S. 91).

Da sämtliche Befragten in Ausstellungsbüros mit vielen Mitarbeiter\*innen beschäftigt sind, wird hier zumeist interdisziplinär gearbeitet. Gleichzeitig stehen große Budgets zur Verfügung. Aus diesem Grund konnten die Befragten nicht beurteilen, ob das Tool für deren Arbeit Vorteile bietet, da für die konkrete Umsetzung interaktiver Medien entweder Programmierer\*innen aus dem Team oder externe Spezialist\*innen eingesetzt werden. Hegemann sah jedoch für Argus Vision einen „großen Markt [...] für einfache [...] Installationen, die dann gerade für eine Wechselausstellung oder für ein Event gedacht sind“ (Hegemann, S. 66).

Zusätzlich konnten sich mehrere der Befragten vorstellen, dass ein Ausstellungsgestaltungsteam mit Argus Vision arbeiten könnte (Hegemann, S. 66; Hundt, S. 102) – „Da muss man kein Informatiker sein, um das bedienen zu können, das ist ja das Wichtigste“ (Schlag, S. 85). Dennoch sollten in einer weiteren Studie kleinere Ausstellungsgestaltungsteams – welche aufgrund ihres Budgets und ihrer Kapazität auf Standardlösungen beim Einsatz interaktiver Medien angewiesen sind – dahingehend befragt werden.

Als interessant wurde das Bündeln der Trackingfunktionalität in einer Bedienoberfläche (Hegemann, S. 66; Diederichs, S. 92) sowie das Triggerzone-Prinzip bezeichnet (Hegemann, S. 66; Hundt, S. 101). Drei der fünf Befragten konnten sich abstrakte Anwendungsfälle für die Verwendung von Triggerzones vorstellen (Hegemann, S. 68; Diederichs, S. 91 ; Hundt, S. 101) – MESO entwickle sogar gerade eine Installation, welche auf diesem Prinzip beruhe, so Oschatz (Oschatz, S. 74).

Hundt sah zusätzlich die Möglichkeit Triggerzones in der Prototyping-Phase einzusetzen: Insbesondere, da die „[...] thematische Ausrichtung zuerst kommt [...] wäre es natürlich interessant, wenn das Ausprobieren, Programmieren und Simulieren von solchen Umgebungen einfacher und schneller wird“ (Hundt, S. 101).

Neben der Evaluation von Argus Vision war ein Ziel der Interviews, Weiterentwicklungsmöglichkeiten für das Tracking Tool zu generieren. Diese lassen sich in die bereits beschriebenen Verbesserungen der Bedienoberfläche, in die Verbesserung des Trackings sowie in Vorschläge zur Anbindung von Argus Vision an andere Tools gliedern.

Als Weiterentwicklungsmöglichkeiten zum bereits bestehenden Tracking wurde das Verbinden mehrerer Kinects zu einer virtuellen Kamera genannt, um „ein fortlaufendes Tracking“ zu ermöglichen sowie, um zwei Kinects „von zwei Seiten einzusetzen, damit man die Tiefe differenzierter einsetzen kann“ (Hegemann, S. 66). Diese Funktion war bereits Teil des ursprünglichen Konzepts von Argus Vision und findet sich in der Bedienoberfläche wieder. Das Drücken des Camera Calibration-Buttons (vgl. Abbildung 10) soll in einer späteren Version einen Kalibrierungsmodus öffnen, in welchem die einzelnen Kinects in eine räumliche Beziehung zueinander gebracht werden können<sup>43</sup>. Zusätzlich dazu müsse das Tracking verbessert werden, damit eine konstante, fehlerfreie ID-Erhaltung möglich sei (Hegemann, S. 66; Schlag, S. 85). Um Argus Vision dann beispielsweise in Spielen einsetzen zu können, muss das Tracking Tool eine Gesten- oder Skeleterkennung bieten (Schlag, S. 86). Ein erster Schritt ist, das proprietäre Skelett-Tracking der Kinect in Argus Vision einzubinden und die Skelettdaten vereinfacht verfügbar zu machen. Über einen Button in der Bedienoberfläche sollte das Skelett-Tracking aktivierbar sein, sodass dieses nur verwendet wird, wenn die Ausrichtung der Kinect parallel zum Boden ist.

Mehrere der Befragten merkten zudem an, dass bei Installationen ein „schnelles und präzises Feedback“ (Hundt, S. 101) für die Besuchenden essentiell sei (Schlag, S. 81). Hier betonten sowohl Hundt als auch Schlag, dass Argus Vision für einen Einsatz in Ausstellungen die Trackingdaten ohne Verzögerung übermitteln muss, damit die Besuchenden die Lust an der Interaktion nicht verlieren (Schlag, S. 81; Hundt, 101). In Abschnitt 5.2.4 wurde bereits erläutert, dass Benutzer\*innen die Möglichkeit gegeben werden soll, die zu übertragenen Daten in Argus Vision manuell einzustellen, um eine verlust- und verzögerungsfreie Übertragung zu gewährleisten.

Um Argus Vision einfach an andere Tools anzubinden, können, wie in Abschnitt 5.1.4 beschrieben, Plug-Ins für häufig verwendete Tools geschrieben werden. Eine alternative Möglichkeit ist die Erstellung eines Frontend-Editors, in dem Ausstellungsgestaltende die Trackingdaten direkt verarbeiten können (Diederichs, S. 91). Dieser müsse aber, so Diederichs, um in ausreichendem Maße verwendbar zu sein, „sehr aufwendig und vielseitig“ sein (ebd.).

Oschatz hingegen sieht Argus Vision als „Baustein in einem Ökosystem für Ausstellungstechnologie“, aus dem sich, „wenn Szenografen erstmal verstehen, was der in einer Ausstellung leisten könnte, [...] sensationelle Möglichkeiten“ ergeben. Er sieht die zukünftige Arbeit darin, „strategisch und längerfristig die Bausteine zu isolieren, die man in einem Museum der Zukunft braucht“ (Oschatz, S. 77). Ziel fortführender Arbeiten muss also die Entwicklung weiterer DIY Tools sein, die einen modularen Aufbau einer möglichst großen Bandbreite interaktiver Medien ermöglichen.

---

<sup>43</sup> Eine Möglichkeit eine solche Kalibrierung vorzunehmen findet sich beispielsweise in Dias & Jorge 2015.

## 5.4 Zusammenfassung der Ergebnisse

Für die Evaluation von Argus Vision wurden in einem ersten Schritt zwei Fallstudien durchgeführt, in welchen das Tracking Tool in unterschiedlicher Art und Weise eingesetzt wurde.

In der ersten Fallstudie wurde das Tracking Tool verwendet, um anhand der getrackten Centroids der Besuchenden eine Animation in einer Installation zu steuern. In der zweiten Fallstudie wurden hingegen Triggerzones verwendet, um Bilder und Videos an einer Wandprojektion genau dann anzeigen zu lassen, wenn Personen in bestimmten Bereichen im Raum stehen. In beiden Fällen war die Anbindung von Argus Vision erfolgreich.

Jedoch zeigte sich auch, dass die Kommunikation zwischen Argus Control und Argus Kinect sowie zwischen Argus Control und einem Client verbessert werden muss. Ein Fehler in der Übertragung der Point Cloud-Daten von Argus Kinect zu Argus Control, der durch eine zu große übertragene Datenmenge entstanden war, konnte zwischen den beiden Fallstudien behoben werden. In der zweiten Fallstudie wurde überdies ersichtlich, dass die Menge der übertragenen Daten allgemein zu groß ist und, dass dies bei einer Verwendung mehrerer Argus Kinects zu Datenverlust führen kann. Eine Lösungsmöglichkeit für dieses Problem ist, den\*die Benutzer\*in von Argus Vision das zu verwendende Netzwerkprotokoll und die zu übertragene Personen- und Triggerzone-Daten selbst auswählen zu lassen. So lässt sich die Datenmenge erheblich reduzieren. Gleichzeitig erleichtert dies die Analyse der übertragenen Daten auf der Seite des Clients.

In einem zweiten Schritt wurden Interviews mit Ausstellungsgestaltenden – der Zielgruppe von Argus Vision – geführt. Hier war das Ziel zum einen die Bewertung des Tools, zum anderen die Erzeugung von Weiterentwicklungsmöglichkeiten.

Die Ausstellungsgestaltenden hatten nicht die Möglichkeit, das Tool selbst zu benutzen, sondern bewerteten Argus Vision anhand einer Präsentation, welche während der Interviews gehalten wurde. Die Bedienoberfläche und deren Verständlichkeit wurden von allen Befragten positiv bewertet. Auch die dem Tracking zugrundeliegende Logik sowie das Triggerzone-Prinzip wurde als verständlich bewertet. Für letzteres konnten sich mehrere Ausstellungsgestaltende abstrakte Anwendungsfälle vorstellen – ein Ausstellungsgestalter erwähnte, dass sie eine solche Funktion gerade in einer Ausstellung einsetzen.

Mehrere der Befragten konnten sich zudem vorstellen, dass kleinere Ausstellungsgestaltungsteams das Tool nutzen können. In einer weiteren Studie sollte dies überprüft werden, indem man solchen Teams Argus Vision für die tatsächliche Verwendung zur Verfügung stellt.

Die Ausstellungsgestaltenden nannten zusätzlich Möglichkeiten zur Weiterentwicklung der Bedienoberfläche und des Trackings sowie Vorschläge, wie Argus Vision an andere Tools angebunden werden könnte.



## 6 Fazit und Ausblick

Argus Vision ist ein für zeitgenössische Ausstellungen entwickeltes Tracking Tool, welches Ausstellungsgestaltende bei der Umsetzung interaktiver Medien mit Kamera-Tracking unterstützt. Gegenstand dieser Arbeit war die Beschreibung des Designs und die Evaluation von Argus Vision.

Zunächst wurde aufgezeigt, dass Ausstellungen heutzutage zunehmend komplex – unter anderem durch Mittel der Architektur, der Gestaltung sowie durch interaktive Medien – inszeniert werden. Um derartige Inszenierungen zu konzipieren und umzusetzen, müssen sich Ausstellungsgestaltungsbüros demnach aus Vertreter\*innen einer Vielzahl unterschiedlicher gestalterischer sowie technischer Disziplinen zusammensetzen. Einige Forschungsarbeiten der MCI beschäftigten sich bereits mit der Unterstützung der Arbeit von Ausstellungsgestaltenden im Design-Prozess. Hier lässt sich jeweils ein inter- und transdisziplinärer Ansatz erkennen: Das Ziel des interdisziplinären Ansatzes ist die Verbesserung der Kommunikation durch die Definition von Mustersprachen, also durch das Schaffen eines gemeinsamen Vokabulars aller Disziplinen. Im zweiten Ansatz hingegen sollen Ausstellungsgestaltungsteams über den gesamten Design-Prozess hinweg disziplinübergreifend arbeiten. Dazu werden unter anderem DIY Tools benötigt, welche in der Anwendung so einfach sein müssen, dass sie von allen Beteiligten verwendet werden können.

Mit Argus Vision wurde ein solches Tool für die Anwendung von Kamera-Tracking in Ausstellungen geschaffen. Um die technisch anspruchsvolle Programmierung von Kamera-Tracking in eine leicht verständliche und bedienbare Applikation zu übersetzen, wurden zunächst mögliche Tracking-Einsatzbereiche in Museen vorgestellt. Die daraus generierten Anforderungen an ein Tracking Tool für Ausstellungsgestaltende wurden darauffolgend präsentiert. Der anschließende Vergleich dieser mit bereits bestehenden Tracking Tools ergab, dass keines der Tools diesen Anforderungen in ausreichendem Maße genügte – Argus Vision musste daher vollständig neu entwickelt werden.

Die Bedienoberfläche und Funktionen des Tracking Tools wurden im Anschluss an die Anforderungsanalyse ausführlich beschrieben und bilden zusammen mit der Evaluation das Kernstück dieser Arbeit. Argus Vision selbst teilt sich in zwei Einheiten auf – beliebig viele Argus Kinect-Instanzen und eine Argus Control-Instanz. Dabei wird mit jeder Argus Kinect-Instanz eine Kinect an einem eigenen PC betrieben. Diese werden über LAN mit Argus Control verbunden, welches sämtliche Daten der Argus Kinect-Instanzen bündelt und an einen beliebigen Client weiterleitet. Benutzende können die einzelnen Kinects über die Bedienoberfläche von Argus Control überwachen sowie sämtliche wichtige Einstellungen in Argus Control vornehmen. Um die Verwendung des Kamera-Trackings zusätzlich einfacher zu gestalten, wurde das Triggerzone-Prinzip eingeführt. Hierbei handelt es sich um frei erstellbare, dreidimensionale Zonen im Raum, welche bei Berührung von Personen aktiviert werden und in ihrer Funktion damit komplexeren Lichtschranken gleichen.

In der Evaluation wurden zwei Ziele verfolgt: Die Bewertung der Bedienoberfläche und Funktionalität des Tools in der zum Abschluss des Projektberichts vorliegenden Form sowie der Generierung von Weiterentwicklungsmöglichkeiten für dieses. Dafür wurden in einem ersten Schritt zwei Fallstudien mit Argus Vision durchgeführt.

Zunächst wurde Argus Vision für eine Installation im Rahmen einer Ausstellung eingesetzt. In dieser wurden die Personendaten von Argus Vision verwendet, um virtuelles Glas bei Bewegung von Besuchenden der Ausstellung an- und zerbrechen zu lassen. Dabei wurde Argus Vision über einen Zeitraum von 44 Tagen fehlerfrei betrieben. Bei der Integration von Argus Vision in die Installation konnte ausgearbeitet werden, dass die Kommunikation sowohl von Argus Kinect und Argus Control als auch von Argus Control zu einem Client verbessert werden muss. Ersteres ließ sich dabei auf einen Fehler in der Programmierung zurückführen, welcher zwischen den beiden Fallstudien behoben werden konnte.

In einer zweiten Fallstudie wurde Argus Vision an smartPerform angebunden. Dies ist ein Tool zur einfachen Erstellung von Präsentationen auf interaktiven Tischen und Displays, das häufig in Ausstellungen eingesetzt wird. Hier wurden im Gegensatz zur ersten Fallstudie Triggerzonen eingesetzt, um Personen beim Betreten bestimmter Zonen Bilder und Videos zu präsentieren. Zusätzlich dazu wurden die Daten über zwei Argus Kinect-Instanzen in Argus Control gebündelt und an einen Client, smartPerform, weitergeleitet. Es ließen sich hierbei weitere Probleme in der Kommunikation feststellen. Diese konnten auf das verwendete Netzwerkprotokoll sowie die Syntax und Menge der übertragenen Daten zurückgeführt werden. Als Lösungsmöglichkeit dafür wurde vorgeschlagen, den Benutzenden die Möglichkeit zu geben, selbst das Protokoll sowie die zu übertragenen Daten auszuwählen.

In einem zweiten Schritt wurden Interviews mit Ausstellungsgestaltenden führender Ausstellungsbüros geführt. Diese beurteilten Argus Vision anhand einer – während des Interviews durchgeführten – Präsentation des Tools. Die Bedienoberfläche, deren Verständlichkeit, die dem Tracking zugrundeliegende Logik sowie das Triggerzone-Prinzip wurden dabei durchgehend positiv bewertet. Zusätzlich brachten die Ausstellungsgestaltenden zahlreiche Vorschläge zur Weiterentwicklung der Bedienoberfläche, der Bedienlogik sowie zur Verbesserung des Trackings von Argus Vision ein.

Obschon mehrere der Befragten sich vorstellen konnten, dass das Tool für kleinere Ausstellungsbüros nützlich sein könnte, bleibt dies in einer weiteren Studie zu prüfen. Argus Vision sollte demnach Ausstellungsgestalter\*innen solcher Büros zur Verfügung gestellt werden, um zu prüfen, ob diese das Tracking Tool tatsächlich in Projekten einsetzen können.

Argus Vision lässt sich als ein Modul für die Umsetzung von Kamera-Tracking in einer Reihe von Technologien ansehen, aus welchen interaktive Medien in Ausstellungen geschaffen werden. Betrachtet man Argus Vision als ein solches, so muss das Ziel fortführender Arbeiten folglich die Entwicklung von weiteren DIY Tools sein – von Modulen also, welche leicht zu verwenden sind, nahtlos ineinandergreifen und damit sämtliche Bausteine abbilden, die für die Entwicklung interaktiver Medien in zukünftigen Ausstellungen benötigt werden.

## Literaturverzeichnis

- avedition (Hrsg.). (2015). *Tamschick Media+Space: Immersive narrative installations*. Ludwigsburg, Germany: avedition.
- Ballendat, T., Marquardt, N. & Greenberg, S. (2010). Proxemic interaction: designing for a proximity and orientation-aware environment. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*. Saarbrücken, Germany: ACM. S. 121-130.
- Barthelmes, C. & den Oudsten, F. (2011). Prolog. In Atelier Brückner GmbH (Hrsg.): *Scenography : making spaces talk; projects 2002 – 2010*. Ludwigsburg, Germany: avedition. S. 11-56.
- Baur, R. (2001). Zwischen vielen. In Schwarz, U. & Teufel, P. (Hrsg.): *Museografie und Ausstellungsgestaltung*. Ludwigsburg: avedition. S. 130-143.
- Bertron, A., Schwarz, U. & Frey, C. (2012). *designing exhibitions – ausstellungen entwerfen 2nd edition*. Basel, Switzerland: Birkhäuser
- Borchers, J. O. (2000). A Pattern Approach to Interaction Design. In *Proceedings of the 3rd conference on Designing interactive systems: processes, practices, methods, and techniques*. New York City, New York, USA: ACM. S. 369-378.
- Buchanan, A. (2008). Predicting User Behavior – The Creation of the Immersion Installation. *Computers in Entertainment – Media Arts 6(3)*, New York City, New York, USA: ACM. S. 34:1-34:9.
- Bullivant, L. (2005). ICE, Klein Dytham Architects, Toshio Iwai, Bloomberg Headquarters, Marunouchi, Tokyo, 2003. *Architectural Design 71(1)*, S. 12-13.
- Burger, D. (2011). *Computergestützter organisationaler Wissenstransfer und Wissensgenerierung – Ein Experteninterview basierter Forschungsansatz*. Wiesbaden, Germany: Springer VS.
- Cafaro, F., Panella, A., Lyons, L., Roberts, J. & Radinsky, J. (2013). I see you there!: developing identity-preserving embodied interaction for museum exhibits. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Paris, France: ACM. S. 1911-1920.
- Derks, H. (2000). Ausstellungsgestalter gesucht. In Compania Media (Hrsg.): *Handbuch Museumsberatung: Akteure – Kompetenzen – Leistungen*. Bielefeld, Germany: transcript. S. 199-214.
- Dernie, D. (2006). *Ausstellungsgestaltung: Konzepte und Techniken*. Ludwigsburg, Germany: avedition.
- Dias, J. & Jorge, P. M. (2015). People tracking with multi-camera system. In *Proceedings of the 9th International Conference on Distributed Smart Cameras*. Seville, Spain: ACM. S. 181-186.
- Gestalten & Hanschke, V. (Hrsg.) (2011). *A Touch of Code – Interactive Installations and Experiences*. Berlin, Germany: Gestalten.
- Greenberg, S., Marquardt, N., Ballendat, T., Diaz-Marino, R. & Wang, M. (2011). Proxemic interactions: the new ubicomp?. *ACM Interactions 18(1)*, S. 42-50.
- Hakvoort, G. (2013). The immersive museum. In *Proceedings of the 2013 ACM international conference on Interactive tabletops and surfaces*. St. Andrews, Scotland: ACM. S. 463-468.
- Hall, E.T. (1966). *The Hidden Dimension*. New York City, New York, USA: Doubleday.

- Hartson, R. & Pyla, P. (2012). *The UX Book: Process and Guidelines for Ensuring a Quality User Experience*. San Francisco, California, USA: Morgan Kaufmann.
- Hassenzahl, M., Burmester, M., & Koller, F. (2003). AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In Ziegler, J. & Szwillus, G. (Hrsg.): *Mensch & Computer 2003. Interaktion in Bewegung*, Stuttgart, Leipzig: B.G. Teubner. S. 187-196.
- Hespanhol, L., Brown, O., Cao, J. & Tomitsch, M. (2013). *Evaluating the effectiveness of audio-visual cues in immersive user interfaces*. In *Proceedings of the 25th Australian Computer-Human Interaction Conference: Augmentation, Application, Innovation, Collaboration*. Adelaide, Australia: ACM. S. 569-572.
- Hughes, P. (2010). *Exhibition Design*. London, UK: Laurence King.
- John, H. (2001). Spielen wir noch in der Champions-League? Oder: Plädoyer für professionelle Ausstellungsplanung im Museum. In Schwarz, U. & Teufel, P. (Hrsg.): *Museografie und Ausstellungs-gestaltung*. Ludwigsburg, Germany: avedition. S. 38-59.
- Kaiser, R. (2014). *Qualitative Experteninterviews – Konzeptionelle Grundlagen und praktische Durchführung*. Wiesbaden, Germany: Springer VS.
- Kilger, G. (2011). Einleitung. In Kilger, G. (Hrsg.): *Szenografie in Ausstellungen und Museen*. Essen, Germany: Klartext. S. 18-23.
- Klinkhammer, D. (2009). Blended Museum – Steigerung von Besuchererfahrungen durch Interaktions- und Informationsdesign. Verfügbar auf: [http://hci.uni-konstanz.de/downloads/BM/Masterarbeit\\_BlendedMuseum\\_Klinkhammer.pdf](http://hci.uni-konstanz.de/downloads/BM/Masterarbeit_BlendedMuseum_Klinkhammer.pdf), Abgerufen am 30.04.16
- Kortbek, K. J. & Grønbaek, K. (2008). Communicating art through interactive technology: new approaches for interaction design in art museums. In *Proceedings of the 5th Nordic conference on Human-computer interaction: building bridges*. Lund, Sweden: ACM. S. 229-238.
- Kossmann, H. & de Jong, M. (2010). *Engaging Spaces: Exhibition Design Explored*. Amsterdam, Netherlands: Frame.
- Kuflik, T., Lanir, J., Dim, E., Wecker, A., Corra', M., Zancanaro, M. & Stock, O. (2011). Indoor positioning: challenges and solutions for indoor cultural heritage sites. In *Proceedings of the 16th international conference on Intelligent user interfaces*. Palo Alto, California, USA: ACM. S. 375-378.
- Lazar, J., Feng, J. H. & Hochheiser, H. (2010). *Research Methods in Human-Computer Interaction*. Chichester, UK: John Wiley & Sons.
- Locker, P. (2011). *Ausstellungsdesign*. Munich, Germany: Stiebner.
- Lorenc, J., Skolnick, L. & Berger, C. (2007). *What is Exhibition Design?*. Hove, UK: RotoVision.
- Marquardt, N., Diaz-Marino, R., Boring, S. & Greenberg, S. (2011). The proximity toolkit: prototyping proxemic interactions in ubiquitous computing ecologies. In *Proceedings of the 24th annual ACM symposium on User interface software and technology*. Santa Barbara, California, USA: ACM. S. 315-326.
- Maye, L. A., McDermott, F. E., Ciolfi, L. & Avram, G. (2014). Interactive exhibitions design: what can we learn from cultural heritage professionals?. In *Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational*. Helsinki, Finland: ACM. S. 598-607.

- McDermott, F., Avram, G. & Maye, L. (2014). Co-Designing Encounters with Digital Cultural Heritage. Verfügbar auf: <http://mesch-project.eu/co-design/>, Abgerufen am: 30.04.16
- McDermott, F., Clarke, L., Avram, G. & Hornecker, E. (2013). The Challenges and Opportunities Faced by Cultural Heritage Professionals in Designing Interactive Exhibits. In *Proceedings of NO-DEM 2013 Beyond Control*. Stockholm, Sweden: ICT. S. 19-26
- Monaci, G., Gritti, T., Vignoli, F., Walmink, W. & Hendriks, M. (2011). Flower Power. In *Proceedings of the 19th ACM International Conference on Multimedia*. Scottsdale, Arizona, USA: ACM. S. 909-912.
- Nestler, S., Artinger, E., Coskun, T., Yildirim-Krannig, Y., Schumann, S., Maehler, M., Wucholt, F., Strohschneider, S. & Klinker, G. (2011). Assessing qualitative usability in life-threatening, time-critical and unstable situations. *GMS Med. Inform. Biom. Epidemiol.* 7(1), o.S.
- Nebeling, M., Teunissen, E., Husmann, M. & C. Norrie, M. (2014). XDKinect: development framework for cross-device interaction using kinect. In *Proceedings of the 2014 ACM SIGCHI symposium on Engineering interactive computing systems*. Rome, Italy: ACM. S. 65-74.
- Oexle, J. (1999). polyphone vielsprachigkeit: magd und knecht – ein rückblick. In Knyrim, P., Nowak, S., Teufel, P. (Hrsg.): *einszueins – positionen zum ausstellen*. Freiburg, Germany: modo. o.S.
- Oppenheimer, F. (1976). Everyone is you ... or me. *Technology Review* 78(7), S. 7-14
- Pece, F., Kautz, J. & Weyrich, T. (2011). Adapting standard video codecs for depth streaming. In *Proceedings of the 17th Eurographics conference on Virtual Environments & Third Joint Virtual Reality*. Nottingham, UK: Eurographics Association, S. 59-66.
- Petrelli, D., Ciolfi, L., Dijk, D. v., Hornecker, E., Not, E. & Schmidt, A. (2013). Integrating material and digital: a new way for cultural heritage. *interactions* 20(4), S. 58-63.
- Schaeffer, J. & R. Lindell (2016). It Could Just as Well Have Been in Greek: Experiences from Introducing Code as a Design Material to Exhibition Design Students. In *Proceedings of the TEI '16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction*. Eindhoven, Netherlands: ACM. S. 126-132.
- Schwarz, U. (2001). Entstehungsphasen einer Ausstellung. In Schwarz, U. & Teufel, P. (Hrsg.): *Museografie und Ausstellungsgestaltung*. Ludwigsburg, Germany: avedition. S. 16-37.
- Seyed, f., Azazi, A., Chan, E., Wang, Y. & Maurer, F. (2015). SoD-Toolkit: A Toolkit for Interactively Prototyping and Developing Multi-Sensor, Multi-Device Environments. In *Proceedings of the 2015 International Conference on Interactive Tabletops & Surfaces*. Madeira, Portugal: ACM. S. 171-180.
- Snibbe, S. S. & Raffle, H. S. (2009). Social immersive media: pursuing best practices for multi-user interactive camera/projector exhibits. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Boston, USA: ACM. S. 1447-1456.
- Simon, N. (2010). *The Participatory Museum*. Santa Cruz, California, USA: Museum 2.0.
- Teufel, P. (2001). Museografie, Ausstellungsgestaltung und Szenografie. In Schwarz, U. & Teufel, P. (Hrsg.): *Museografie und Ausstellungsgestaltung*. Ludwigsburg, Germany: avedition. S. 10-15.
- Wang, N. (2015). Enhancing User Experience to Design Enjoyable Exhibition Events. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*. Seoul, Republic of Korea: ACM. S. 239-242.

- 
- Weiser, M. (1991). The Computer for the 21st Century. *Scientific American* 265(3), S. 94-104.
- Wolf, K., Abdelhady, E., Abdelrahman, Y., Kubitz, T. & Schmidt, A. (2015). meSch: tools for interactive exhibitions. In *Proceedings of the Conference on Electronic Visualisation and the Arts*. London, United Kingdom: BCS. S. 261-269.
- Yilmaz, A., Javed, O. & Shah, M. (2006). Object tracking: A survey. *ACM Computing. Surveys* 38(4), S. 13:1-13:45, 2006.

Anhang

## Interview-Leitfaden

### Phase 1 (Aufwärmphase):

- Was sind Ihre typischen Aufgaben im Rahmen der Ausstellungskonzeption?
- In welche Projekte waren Sie zuletzt eingebunden?
- Wie war der Ablauf dieser Projekte?
- Welche Aufgaben werden an andere Unternehmen delegiert?
- Was sind aktuelle Trends in der Ausstellungsgestaltung?
- Wie viele interaktive Exponate wurden von Ihnen umgesetzt?
- Wie wurden Besucher in diese Installationen einbezogen?
- Wie viele interaktive Installationen wurden von Ihnen umgesetzt?
- Wie wurden Besucher in diese Installationen einbezogen?
- Wie wird versucht, einem Raum verschiedene Wirkungen zu geben?
- Wie schätzen Sie das Verständnis von Besuchern für szenografische Änderungen ein?

### Phase 2 (UI-Interface):

- Wie würden Sie die Software einsetzen?
- In welchen Phasen würden Sie die Software einsetzen?
- Welche gezeigten Funktionen sind interessant?
- Wie gefällt Ihnen das Design der Software?
- Wie praktisch ist die Software?
- Wie verständlich ist die Bedienung der Software?
- Welche Bereiche der Software sind missverständlich?
- Welche Funktionen erwarten Sie von der Software?
- Warum ist die Software interessant?
- Was ist innovativ an der Software?

### Phase 3 (Tracking Konzept):

- Wie wird Tracking in bisherigen Ihnen bekannten Installationen eingesetzt?
- Inwiefern hätte das Triggerzone-Konzept in bisherigen Projekten eingesetzt werden können?
- Wie nützlich finden Sie das Triggerzone-Konzept?

### Phase 4 (Design Space):

- Wie könnte das Triggerzone-Konzept in Exponaten eingesetzt werden?
- Wie könnte das Triggerzone-Konzept in Installationen eingesetzt werden?
- Wie könnte das Triggerzone-Konzept für Szenografie eingesetzt werden?
- Wozu könnte das Tool noch eingesetzt werden?

### Phase 5 (Future Work):

- Warum kann das Tool nicht für bestimmte Trackingzwecke eingesetzt werden?
- Wo fehlen Funktionen?
- Wie beurteilen Sie die Wichtigkeit von ID Erhaltung?
- Wie beurteilen Sie die Wichtigkeit von Skeletterkennung?



## Interview mit Dominik Hegemann

Teile des Interviews wurden auf Bitte des Befragten nicht veröffentlicht, da dort über aktuelle, unter Verschluss gehaltene Projekte gesprochen wurde. Diese sind mit [...] gekennzeichnet. Passagen, die vornehmlich von allgemeinem und nicht fachspezifischem Charakter waren, wurden nicht transkribiert und sind mit (...), kurze Redepausen mit (..) gekennzeichnet.

Das Interview mit Dominik Hegemann, Assoziierter bei Atelier Brückner wurde am 15.02.16 im Atelier Brückner in Stuttgart geführt:

*Was machen Sie eigentlich im Rahmen der Ausstellungskonzeption? Was nehmen Sie für eine Rolle ein?*

Das ist schwer zu sagen, grundsätzlich bin ich seit gut 16 Jahren beim Atelier Brückner und zu Zeiten des Beginns haben wir als Projektleiter alles gemacht, also vom Wettbewerb oder der Konzeptionserstellung bis zur Übergabe an den Kunden. Dazu zählte auch die Detailplanung und Bauüberwachung und manchmal auch Mängelbeseitigung über mehrere Jahre hinweg, wenn es sein muss. Am Anfang habe ich die Konzeption von großen Messeständen geleitet oder als Thema überhaupt erst eingebracht. Die übersteigen oft das Budget von Museen, werden aber nach zwei bis zehn Tagen komplett demontiert. Dann ging es so langsam über in die – weil es ja auch großer Part der Firmenkommunikation ist – Firmenmuseen. Dort bin ich dann auch hängen geblieben, abgesehen von ein paar Zwischenspielen. Bei den Museen - wobei auch Expo Pavillons dazugehören – bin ich jetzt eigentlich seit dem BMW Museum nur noch im fernerem Ausland, also beispielsweise Asien oder Arabien. Da ist mein Schwerpunkt hauptsächlich mediale Ausstellung. Ich weiß aber nicht, ob das auf Dauer so sein wird, weil die haptische Komponente natürlich immer wichtiger wird, je mehr die Kids mit Computern aufwachsen. Letztendlich geht es aber immer um ein Zusammenspiel – auch bei der Interaktion – zwischen der haptischen und rein digitalen Komponente.

*Was waren denn Ihre letzten Projekte?*

[...]

*Nochmal zum Tracking. Was haben Sie denn sonst noch für Projekte umgesetzt, in denen Tracking verwendet wurde?*

Also das geht eigentlich schon lange zurück. Eigentlich hat das Tracking angefangen, da gab es noch keine Kinect. Das erste großflächig angelegte Trackingsystem haben wir versucht mit ART+COM im BMW Museum umzusetzen – da hat es nicht wirklich funktioniert. Da war allerdings die komplette Fläche unter einer LED Fassade durchgehend getracked mit Infrarotkameras. Ein großes Problem beim Tracking an sich ist: Wenn es keiner vormacht, merken die Besucher gar nicht, dass da eine Interaktion möglich ist. Und das ist grundsätzlich bei allen interaktiven Elementen der Hauptpunkt. Wie kriege ich den Besucher dazu zu interagieren, ohne dass da jemand ist, der vortanzt? Das kann von den kleinsten Touchscreens bis zu den größeren Trackinganwendungen passieren. Mit dem Aufkommen der Kinect war das erste Projekt, das ich gemacht habe, 2010 in China. Damals war es schon mit Infrarotkameras ausgeschrieben und wurde aus Kostengründen dann auf ein chinesisches

Modell umgearbeitet. Das war eine Fläche von 16m<sup>2</sup> in der Ebene und in der vertikalen 20m<sup>2</sup>, die abgedeckt wurde. Da waren glaube ich 16 Kinects im Einsatz, sodass man wirklich die gesamte Fläche tracken konnte, also nicht nur den Boden, sondern auch die Wand. Davor gab es noch ein Zwischending, da waren wir aber nicht wirklich involviert. Das erste Kinect-tracking, das ich gesehen habe – allerdings mit irgendwelchen Prototypen – war auf der EXPO in Zaragoza. Ich weiß nicht mehr welcher Pavillon, aber er war von Marc Tamschick. [...]

*Verwenden Sie standardisierte Software für Ihre Installationen?*

Meistens wird es für den konkreten Raum entwickelt. Also für die Parameter, die uns da wichtig sind, gibt es nicht wirklich standardisierte Software.

*Gibt es sowas auch aus Konkurrenzgründen in dem Bereich nicht?*

Ich glaube, dass die Kollegen, die das programmieren, schon schauen, dass sie die weiteren Aufträge, die so eine Anlage generieren kann auch bei sich behalten und entsprechend auch die Programmierung darauf auslegen. Da steht dann kein Open Source Gedanke dahinter, schon eher Business.

*Ich habe mit [Ausstellungsgestalter] gesprochen und ihm von meinem Programm erzählt. Er meinte auch, dass er wisse, dass mehrere Unternehmen in dem Bereich ihre Eigenentwicklung vorantreiben.*

Ja, es ist sicher ein Bereich, der auch für die günstigeren Varianten von Trackingsystemen so eine Programmierung gutheißen würde. Gerade in Projekten, die wir umsetzen, kostet so eine Installation zwischen 500.000€ und 1.000.000€. Aber das sind natürlich alles große Installationen. Bei einer Ausstellung, bei der ich zum Beispiel nur eine oder zwei Kinects, also z.B. eine von vorne, eine von hinten einsetze – um die Raumentiefe besser abzubilden – wäre es sicherlich interessant, wenn man dann eine Standardprogrammierung hätte, auf die man mit möglichst geringem Aufwand aufsatteln könnte.

*Haben Sie jemals ein dynamisches Raumkonzept umgesetzt, bei denen sich Dinge wie Licht und Ton verändert haben?*

Das hatten wir auf jeden Fall. Also gerade der Stage Grid Pavillon auf der EXPO in Shanghai. Da gab es eigentlich keine Interaktionsmöglichkeit, aber das war ein immersiver Raum, der ganz bespielt war. Das war ein Würfel mit LED Bespielung und man ist auf einer Glasbrücke in diesen Würfel getreten. Da war allein der Sound so komprimiert, dass man den kompletten Raum nur über den Sound drehen konnte. Dafür haben wir 21 Kanäle plus zwei große Bassarrays verwendet. Verstärkt wurde das natürlich durch den Film. Ein paar Szenen mussten sogar entschärft werden, weil es den Leuten schlecht geworden ist. Wir haben den Raum also auch filmisch gedreht. Man fliegt dann zum Beispiel einen Staudamm runter und das wurde dann auch durch den Ton unterstützt. Wir haben das erst im Kleinen, als 3m<sup>3</sup> Würfel gebaut, um es zu überprüfen. Letztendlich hat man die Wirkung erst vor Ort gemerkt. Im Großen 15 m<sup>3</sup> Würfel fand ich es dann ziemlich mau. Manchen wurde es aber immer noch schlecht. Da war es dann aber auch wirklich erstaunlich, dass die Komponente Ton das

Ganze da stark zurückholte, wo die visuellen Effekte nachgelassen haben. Das war für mich der intensivste Eindruck – was man dann mit Ton machen kann.

Grundsätzlich hätte man da auch Interaktion mit reinbringen können, aber das Hauptproblem ist: Je mehr Leute, desto weniger Interaktion ist möglich. Oder eben nur grobe Interaktion. Und dann ist es umso schwieriger den Personen zu verstehen zu geben, dass und wie sie interagieren können. Das funktioniert dann wirklich nur noch als Gruppe und nicht mehr als Einzelperson. Die erste Interaktionsgeschichte, die wir gemacht haben, war im Haus der Geschichte in Stuttgart: eine interaktive Karte im Eingang. Das verstehen die Besucher bis heute nicht. Die funktioniert auch wirklich nur höchstens mit zwei Personen. Also rein technisch würde sie mit mehr Personen funktionieren, aber das klappt vom Besucher aus nicht. Wenn er es überhaupt versteht.

Und seitdem verfolgt einen das. Also die Interaktion mit Touch- und medialen Tischen ist inzwischen gelernt. Das kann mittlerweile jedes Kind. Beim Interface ist es dahingehend auch wirklich wichtig auf grundlegende Apple-Prinzipien zurückzugehen. Da geht es tatsächlich nur um die Programmierung des Interfaces. Und so ist es auch das größte Problem bei vielen Interfaces, dass es dann von reinen IT-Nerds programmiert wird, die das Interfacedesign nicht wirklich im Vordergrund haben und zu kompliziert denken. Das ist alles logisch, aber sie beziehen den normalen User nicht mit ein. Das ist ein größeres Problem: dass oft das, was hinten aufgesetzt ist, viel zu komplex ist für das, was man eigentlich bräuchte. Und da sind auch einfache Lösungen, die zusätzlich einfach programmiert sind, sehr gewünscht.

Die ganzen touch-sensitiven Tische haben natürlich den Vorteil, dass das über längere Zeit läuft ohne, dass man da hinschreiben muss: Hier drücken, um zu starten. Hier drücken, um zu wischen. Wenn man nicht auf etwas Gelerntes aufsetzen kann wie wischen und blättern, dann ist es nicht so einfach. Ganz schlimm war es dieses Jahr auf der EXPO im deutschen Pavillon. Da gab es viel zu viele verschiedene Interaktionsmöglichkeiten. Da musste wirklich ein Tourguide die ganze Zeit mit durch die Ausstellung laufen und einem sagen, wie man interagieren musste. Da gab es Exponate, die zwei bis drei unterschiedliche Interaktionsmöglichkeiten hatten. Einmal musste man dann einen Regenschirm aufmachen, zumachen, drehen und dann hatte man diese (unverständlich), die man in unterschiedlichen Höhen auf unterschiedliche Ebenen schalten konnte. Aber dann haben nur drei der Ebenen an dem Exponat wegen des Trackingsystems funktioniert und alle anderen mussten dann anders interagieren. Das versteht halt kein Mensch.

Wenn man eine Interaktionsmöglichkeit hat, dann ist es natürlich gut, wenn man vom Anfang bis zum Ende nur eine durchs Museum oder die ganze Ausstellung durchführen kann. Denn für uns ist das wichtigste das Raumerlebnis. Und da ist die Technik völlig nebensächlich oder rückt in den Hintergrund. Man sollte den Raum erleben und intuitiv die unterschiedlichen Wissens Ebenen durch ein System erschließen, das eigentlich nebensächlich sein soll. Deswegen wollen wir eigentlich auch nicht irgendwelche Screens zeigen, sondern immersive Raumerlebnisse schaffen.

Es ist dann natürlich ein kleiner Schritt von interaktiven Tischen zu einem gesamten Raumtrackingsystem. Aber das ist noch nicht so ausgereift, dass der Besucher sich die Informatio-

nen selber rausziehen kann. Und überall Schilder aufhängen, wo man interagieren kann, ist auch nicht der richtige Weg. [...]

(DEMO)

*Inwiefern ist das Tool nützlich?*

Es ist sicher nützlich, aber es gibt sicherlich auch weitere Entwicklungsmöglichkeiten. Also gerade ein fortlaufendes Tracking, zum anderen die Tiefe darzustellen, also die Kinect von zwei Seiten einzusetzen, damit man die Tiefe differenzierter einsetzen kann. Und dass die Konturen natürlich auch weiterverfolgt werden, wenn sie sich überschneiden – also, dass die IDs beibehalten werden. Das ist dann natürlich wesentlich komplexer. Aber grundsätzlich, wie schon gesagt, gibt es gerade für einfache oder einfach zu handelnde, schnelle Installationen, die dann gerade für eine Wechsausstellung oder für ein Event gedacht sind, einen großen Markt – meiner Meinung nach. Wenn die Standardlösung denn dann auch einfach und für mehrere Anwendungen umsetzbar ist.

*Welche der Funktionen, die denn jetzt schon drin sind finden Sie denn interessant?*

Interessant (..), also Tracking ist ja nichts neues, daher am ehesten die Triggerzones. Interessant natürlich auch, dass man das in einem Interface vor sich hat und eben sieht, was passiert.

*Wie läuft die Prototypingphase bei Ihnen ab?*

Ja, das wird dann meistens bei den ausführenden Firmen gemacht, denn das sind oft größere Hallen, die benötigt werden. Also damals bei Stage Grid wurde die LED Wand beispielsweise durch Projektoren ersetzt. Das heißt, ich brauche einen Projektionsabstand und dann ist schnell eine große Fläche belegt. Gut, so eine Halle muss man sich dann sowieso anmieten. Bei den Summen muss man natürlich auch testen, ob das wirklich und wie es funktioniert. Das sind am besten eins-zu-eins-Tests oder eben skalierte Tests.

Für ein Trickfilmfestival haben wir damals das Ganze auch mal im Mercedes Museum im Foyer aufgebaut. Da musste es dann auch so funktionieren, dass die Besucher es nachvollziehen konnten. Aber es ist sehr wichtig, dass man das dann alles vorher testet und, dass man auch Interaktionsmöglichkeiten testet. [...] aber insgesamt ist es natürlich so, dass das Verständnis vom Kunden für die digitalen Hintergründe da nicht so da ist. Das heißt, wenn was gebaut ist, dann muss es eigentlich schon funktionieren. Dann gibt es eine kleine Testphase, aber da wird natürlich nicht akzeptiert, dass irgendwas nicht tut. Deshalb muss alles, was vorher getestet und eingestellt werden kann, auch schon eingestellt sein. Das ist unabdingbar. Eigentlich ist es notwendig, dass man das eins-zu-eins nachbaut.

*Wie übersichtlich finden Sie das Interface?*

Ja, ich finde es schon sehr übersichtlich. Design hin oder her - das ist dann Geschmackssache. Aber letztendlich muss es halt übersichtlich sein, in der Anwendung verständlich und intuitiv sein und dann ist das auch grundsätzlich kein Problem. Man kann sich in das Design dann sicherlich schnell reindenken. Das ist bei uns auch ein wichtiger Punkt, dass wir dann irgendwann eine Ausstellung an irgendjemand übergeben müssen, den wir nicht kennen und von dem wir nicht wissen, was der für ein Verständnis hat. [...] Dann gilt natürlich: Umso

einfacher und simpler es ist, desto mehr verstehen es die Leute, die es nachher betreiben und warten müssen. In Ihrem Fall ist es natürlich umso wichtiger, wenn Sie das als Tool für andere zur Verfügung stellen, die dann daran weiter programmieren. Bei uns ist das selten der Fall, da die Firmen die Weiterprogrammierung dann auch beauftragt haben wollen und das Ganze so komplex bearbeitet wird, dass sich da nicht jeder einloggen kann.

*Finden Sie denn irgendwas missverständlich?*

(..) Nein, das ist eigentlich schon alles klar.

*Inwiefern haben Sie ID erhaltendes Tracking, auch über mehrere Kinects, eingesetzt?*

[...]

*Sie haben vorhin erwähnt, dass das Kameratracking im Rückfluss sei (..)*

Das nicht wirklich, aber was wir nicht absehen können, ist, wie das Kameratracking – welches ja aus der Bastlerseite gekommen und standardmäßig ein billiges Tool ist, das noch nie eine Profi 24/7 Anwendung geschafft hat – sich über die Dauer hinweg verhält. Wir haben bei allen anderen Sachen, ob das Monitore oder Touchfolien sind oder das ganze System, genügend Profianwendungen, die dann meist um das zehnfache teurer sind als die Consumer-Variante. Man hat halt sonst keine Gewährleistung, wenn die dann auslesen, dass das Ding schon viele Stunden auf dem Buckel hat. Und so ist es bei der Kinect auch. Da ist es dann so, dass letztendlich keine Daten über einen Einsatz über sechs oder zehn Jahre vorliegen. Und für uns ist es halt wichtig, dass, wenn wir sechs Jahre gewährleisten, es auch hält. Und da gibt es halt noch keine Profihardware.

(...)

*Wo sehen Sie gerade die Trends in der Ausstellungsgestaltung?*

Personalisierung ist sicherlich wichtig. Vielleicht auch das Voranzutreiben, denn der Besucher beschäftigt sich sicher viel intensiver mit etwas, wenn er eine direkte Rückkopplung hat und, wenn er auch einen Effekt sieht. Wenn man etwas wirklich verankern oder einen Aha-Effekt haben will, dann geht das eigentlich nur, wenn man die Besucher direkt anspricht. Der Idealfall wäre, dass man die einzelnen Menschen durch einen Gesichtsscann oder irgendetwas, das automatisch funktioniert, begleiten kann und andere Ebenen für jede Person anbietet. Also das ist für mich sicherlich ein Trend, gerade, wenn man nicht diese Massen haben will oder wird, wie in Asien. In Deutschland sind selbst in gut besuchten Ausstellungen vielleicht 50, allerhöchstens 100 Leute gleichzeitig in einer Ausstellung. Selbst in Körperwelten sind nicht mehr. In China hat man dann 800 bis 1.000 die auf einer Fläche zur gleichen Zeit unterhalten werden müssen. Für mich, obwohl ich jetzt auch 2000 an der EXPO das noch wunderbar gefunden habe, hat das Zukunft, was das Analoge mit dem Digitalen verbindet. Ohne haptische Rückkopplung glaube ich, geht heutzutage nicht mehr viel. Apple macht das mit seinem ForceTouch zum Beispiel. Das wichtigste für die Besucher ist nunmal die Rückkopplung, ob das jetzt mechanisch ist oder nicht, die muss es geben. Und bloß mit einem Feedbacksound, der einem sagt: „Ok, ich hab dich verstanden“ ist es nicht getan. Die meisten Interaktionen werden mit diesem Feedbacksound gemacht. Das ist ein großes Problem. Das zweite große Problem ist, dass man bei den interaktiven Geschichten die Leute dazu bewe-

gen muss. Wenn man sich die Besucherschicht anschaut, sind das je nach Anwendung Kids, die gezwungen werden in eine Ausstellung zu gehen oder Rentner zwischen 60 bis 65, die die Zeit haben in die Ausstellung zu gehen. Aber dann muss man natürlich auch diese Spanne abdecken. Das ist zur Zeit und in den nächsten 30 bis 40 Jahren noch so, dass die Kids zwar als Digital Natives aufwachsen, die Alten aber trotzdem mitgenommen werden müssen und Spaß haben müssen an der Ausstellung. Da ist vielleicht auch das Interesse an den unterschiedlichen Vertiefungsebenen erst da. Die muss man dann aber auch bedienen können. Deswegen ist es auch wichtig, dass es bei nicht irgendeiner Reaktion bleibt, die halt technisch möglich und ganz nett ist und mit der man kurz spielen kann. Es muss dann auch einen Mehrwert für die Kommunikation bilden und so zugänglich sein, dass es auch benutzt wird. Das ist, glaube ich, das Wichtigste.

[...]

(...)

*Gibt es diese Angst vor dem Interagieren bei Besuchern noch?*

Bei den Alten schon. Aber wenn es rein digital ist, dann kann man ja auch nichts kaputt machen. Ich glaube, wenn, dann ist das Problem: Ich mache mich vor meiner Kaffeefahrtreisegruppe zum Affen. Aber sonst glaube ich nicht. Und gerade die Jungen suchen ja jede Möglichkeit zu interagieren. Aber wie gesagt, da ist es sinnvoll, dass, wenn man eine Interaktionsmöglichkeit gelernt hat, diese nicht dauernd wechselt. Diese sollte einen auch durch das Museum begleiten.

*Abschließend möchte ich noch zur Future Work für mich einige Fragen stellen. Inwiefern ist das Triggerzone Konzept nützlich?*

Ja, das ist sicherlich nützlich. Es gibt auch genug Anwendungen, gerade im kleineren Bereich. Da glaube ich schon, dass da eine Zukunft liegt. Dadurch, dass es die Dinge auch vereinfacht und zugänglicher macht, kann man auch einen größeren Markt bespielen.

*Was fehlt denn im Tool?*

Eigentlich sehe ich nicht wirklich was.

ENDE

## Interview mit Sebastian Oschatz

Passagen, die vornehmlich von allgemeinem und nicht fachspezifischem Charakter waren, wurden nicht transkribiert und sind mit (...), kurze Redepausen mit (..) gekennzeichnet.

Das Interview mit Sebastian Oschatz, Geschäftsführer von MESO Digital Interiors, wurde am 29.02.16 bei MESO Digital Interiors in Frankfurt am Main geführt:

*Bevor ich etwas zu meinem Tool sage: Ich hatte letzte Woche einen Interviewtermin bei einem eher klassischen Ausstellungsbüro, einem also, welches vom Anfang bis zum Ende der Konzeption durchgehend involviert ist. Nun sind Sie ja eher digital unterwegs. Wie sieht das bei Ihnen aus? Sind Sie auch direkt im Wettbewerb mit dabei oder werden Sie erst angesprochen, wenn es um die konkrete Umsetzung geht?*

Um etwas Innovatives zu erschaffen, muss man sowohl etwas von Gestaltung als auch von Technologie verstehen. Nur so kommt man zu etwas, das funktioniert, aber auch für den Menschen wahrnehmbar, verständlich und schön ist. In der tatsächlichen Ausschreibungspraxis ist das gelegentlich schwer, weil man sich schnell entscheiden muss, ob man sich in einem Wettbewerb zum Thema Konzeption oder Medientechnik bewirbt. In der Tat landen wir oft vertraglich auf der umsetzenden Seite. Wir haben dann gelegentlich das Gefühl, dass wir nicht das einbringen können, was wir eigentlich am besten können. Gleichzeitig beobachten wir immer wieder, dass eine strategische Interaktionsplanung das ist, womit man anfangen müsste.

*Was genau meinen Sie damit?*

Traditionell werden Ausstellungen von Architekten gemacht und zunehmend natürlich von Szenografen. Die kommen aber häufig aus dem ganz klassischen musealen Kontext. Ich glaube aber, beginnend mit einem klugen, digital gedachten Konzept, kommt man auf andere Ansätze, was eine Ausstellung für Besucher überhaupt leisten kann bzw. generell was die Rolle einer Ausstellung ist im Zeitalter, in dem jeder sein Smartphone sowieso dabei hat. Ich glaube, das muss jetzt etwas Anderes sein als bisher. Das geht ja an allen Stellen ans Eingemachte: Die Rolle von Sammlungen, von Vermittlung, von Inszenierung, von Forschung, von Nutzungsrechten, von Authentizität usw. Alle reden immer davon, dass man die authentischen Objekte im Museum erfahren soll und dann geht man in so ein Museum und drückt sich die Nase an Glas platt. Dann ist das Alles schön daher geredet aber irgendwie erfahre ich das Objekt ja auch nicht, wenn es in einer Vitrine drin ist. Da kann eine gute digitale Umsetzung durchaus eine größere Aura bekommen.

*Ist es denn so, dass das Interaktionskonzept oft im Nachhinein in das eigentliche Konzept gepresst wird?*

Das gibt es oft. Es gibt zunehmend mehr Leute, die das auch als Problem wahrnehmen. Aber ich glaube das Thema Technologie in Ausstellungen ist heute hauptsächlich ein Organisationsproblem bei den Ausrichtenden. Und kein technisches.

*Was haben Sie denn für Projekte in letzter Zeit durchgeführt?*

Es sind sehr viele sehr unterschiedliche Projekte. Wir arbeiten gerade an einer Ausstellung in der Unternehmenszentrale von KUKA (Ich denke, das ist in dem Kontext hier interessant, weil da auch ein Location Tracking System zum Einsatz kommt.) In der Ausstellung gibt es mehrere Exponate mit Industrierobotern und an der Stelle ergibt es sich natürlich schnell, dass man auch bei strategischen Fragen nicht nur gestalterisch, sondern auch sehr technisch diskutieren muss. Tisch13 aus München haben das Projekt akquiriert und konzeptioniert aber die eigentliche Exponatentwicklung ist zu großen Teilen bei uns gelandet.

*Sie haben das Location Tracking erwähnt. Was genau tut es und wie funktioniert es?*

In der Ausstellung ist das Tracking primär ein magischer Effekt. Wir nutzen eine Bluetooth Technik, die über Angle-Of-Arrival funktioniert. Wir haben ca. 10 Antennen unter der Decke und damit können wir bis auf 10cm Genauigkeit die Position der Besucher bestimmen. Das dient in der Ausstellung primär dazu, dass die Exponate zum Leben erweckt werden. Man geht also in einen relativ dunklen Raum hinein und beim Betreten folgen Lichtspots dem Gehpfad. Zusätzlich erwachen die Exponate, wenn man sich ihnen nähert. Das ist eigentlich ein ganz schöner Effekt. Wir haben viel überlegt, wie wir das technisch umsetzen. Wir dachten erst an Kameras, das Problem dabei war aber, dass wir für den verwinkelten Raum relativ viele Kameras gebraucht hätten. Deswegen haben wir das Bluetooth System eingesetzt, was im nächsten Schritt auch die Möglichkeit bieten könnte, jeden Besucher zu identifizieren. Das heißt, wir könnten zu jedem Zeitpunkt genau wissen, wer wo ist. Da wird es spannend – wenn man im ganzen Interaktionsdesign von Exponaten keinerlei Log-In-Prozesse mehr bräuchte. Aber das sind alles Sachen, die im ersten Schritt erst noch das Konzept und den Kunden überfordern.

*Inwiefern decken Sie Multiuser (Rucksack) Szenarien in dem Konzept ab?*

Das ist meines Erachtens gar nicht so relevant. Ich verstehe folgendes als das Thema: „Ich will etwas mitnehmen, was ich in der Ausstellung erlebt habe“. Denn letztlich ist es ja so, dass die Leute eh schon alles zuhause haben und im Internet nachschauen hätten können. Warum will ich denn als jemand, der eine Ausstellung macht, nicht sowieso alles, was ich produziere auch online stellen? Warum will ich das filtern auf die Zielgruppe, die in der Ausstellung zufällig an dem Ding hängen geblieben ist, das sie sowieso interessiert hat?

Daher war das Thema Rucksackprinzip in der KUKA Ausstellung bisher noch gar kein großes Thema. Im Moment erwachen die Exponate zum Leben und animieren sich in die Richtung, aus der die Leute kommen. Das angesprochene Rucksackprinzip ist aktuell eine zweite Stufe, die wir noch einmal sauber hinterfragen möchten.

Eine weitere in dem Projekt diskutierte Nutzungsmöglichkeit des Tracking-Systems war zum Beispiel, dass eine Kamera, die ein Foto von mir macht, dorthin schaut, wo ich gerade stehe – mich also in der Ausstellung verfolgt, um Fotos machen zu können. Das sind mögliche Szenarien aber auch hier muss man sich fragen, inwieweit die Besucher das auch als positiv wahrnehmen. Das muss sicherlich konzeptionell/strategisch gut begleitet sein.

*Das Ganze, was Sie jetzt beschrieben haben, geht natürlich jetzt sehr stark in eine szenografische Richtung. Sie haben eine interaktive Lichtinstallation namens Gravity in Aarau umgesetzt. Inwiefern verstehen die Besucher solche Installationen?*



Das ist eine gute Frage. Es ist ja sehr stark eine Frage, was man damit erreichen will. „Interaktiv“ ist ja oft nur ein Buzzword in einem Sales Pitch: Man glaubt es wäre besser, weil es interaktiv ist. In der Praxis ist die interessantere Frage: „Wieviel Mehrwert bringt das Verstehen der Installation eigentlich?“

Wir haben 1998 ein entfernt ähnliches Projekt mit dem Lichtkünstler Achim Wollscheid in Bochum umgesetzt. Das war eine Fußgängerbrücke die Lichtpanele und Radarsensoren eingebaut hatte. Die Idee dabei war, dass diese Sensoren das Licht auf der Brücke in Abhängigkeit der Geschwindigkeit der Bewegung aktivieren. Das heißt als schnell fahrender Fahrradfahrer kann man die Brücke fast vollständig illuminieren und sieht dann auch entsprechend vor sich schon Licht. Spaziergänger erhalten nur ein kleines Licht um sich herum. Bleibt man auf der Brücke stehen, dimmt sich das Licht komplett ab. Wartet man einige Minuten in der Stille, beginnt ein Lichtspiel, das wieder erlischt in dem Moment, in dem man sich bewegt. Das fand ich insofern smart, da die Besucher mit dieser Interaktion anfangen konnten, zu spielen: Mit hin- und herlaufen auf der Brücke kann man die Elemente aktivieren. Verschiedene Verkehrsteilnehmer erzeugen unterschiedliche Effekte. Ein Pärchen, das auf der Brücke anfängt zu knutschen steht irgendwann im Dunkeln und kann auf das Lichtspiel warten. Die Interaktion spart Energie, erhöht die Sicherheit und hat noch eine Spielenebene. Es gibt also ein paar Ebenen, wo diese Interaktion tatsächlich einen konkreten Vorteil bietet. Bei der Installation in Aarau, von der Sie gesprochen haben, ist das anders – da bewegt sich im Wesentlichen die Grafik und spiegelt den Besucherfluss in sehr unterschiedlichen Weisen. Das sieht schön aus und funktioniert auch sehr schön aber irgendwie ist die genaue Interaktion auch ein wenig egal. Alles was man macht ist toll. Ein bisschen wie ein Selfie zu machen. Es ist mehr ein Effekt, der den Tunnel ein wenig belebt.

*Wie relevant ist es, dass Besucher solche Installationen verstehen? Insbesondere im Bereich Szenografie?*

Aus unserer persönlichen Haltung heraus ist das sehr relevant. Das ist etwas, wozu wir den Kunden oft raten. Diese treten häufig an uns heran und wollen etwas Interaktives haben. Dann ist oft die Frage: warum überhaupt? Denn natürlich ist eine ordentliche choreografierte, eine realistisch zufällig programmierte oder eine sauber im Loop laufende Geschichte deutlich besser gestaltbar, als eine, die interaktiv ist. Es ist manchmal die einfachere Variante, etwas interaktiv zu machen, weil man dann nicht so viel Aufwand mit der Gestaltung hat. Man lässt es dann einfach passieren und übergibt die Gestaltung dem Besucher. Oder man kann in der Prototypingphase besser arbeiten. Aber Interaktion wird natürlich erst dann interessant, wenn die Installation dadurch einen zum Beispiel didaktischen oder spielerischen Mehrwert hat. Und dazu muss sie natürlich verstanden werden. Das muss sauber geplant werden. Bei vielen Dingen, die man in Ausstellungen zu erleben versucht, hat man meines Erachtens nicht genug nachgedacht, ob die Interaktion auch wirklich verständlich ist. Dann ist es letztlich nur ein Novelty-Effekt.

*Was haben Sie sonst noch für Installationen oder Exponate umgesetzt, die mit einem Tracking-Konzept funktionieren?*

Für MINI haben wir z.B. ein Exponat mit Besucher-Tracking entwickelt. Die Agentur Meire+Meire hatte eine grobe Idee von einem Exponat, bei dem man durch gestische Bewegun-

gen ein Auto anmalen sollte. Sie zeigten uns ein Video von dem Medienkünstler Memo Akten, bei dem Leute vor einer Wand stehen, mit den Händen wedeln und damit wilde Farbspritzer produzieren. Das ist konzeptionell genau wie eben gesagt: Alles, was man macht, ist toll. Die erste Arbeitsidee der Agentur war, dass man genau so vor der Wand wedelt und damit eben wild ein Auto anmalt. In den ersten Gesprächen kamen wir dann natürlich auf den Punkt, dass das zu wenig ist. Dazu propagieren wir ein sehr mächtiges vieldimensionales und hochwertiges Eingabegerät, obwohl die Aufgabenstellung eigentlich gar nicht sehr komplex ist. Von der Aufgabenstellung hätte ja ein Touchscreen gereicht, um ein Auto anzumalen. Das hätte jeder sofort begriffen, man hätte ein Auto bemalen können und es wäre eine einfache Lösung gewesen. Es hätte nur eben keinen Spaß gemacht, weil man das gleiche im Prinzip mit seinem Handy ja auch machen könnte, wenn man sich daran erinnern wollte, mit welcher App man das am besten machen würde. So kamen wir dann darauf, ein professionelles optisches Motion Capture System als Eingabegerät einzusetzen, und dazu Sprühdosen zu entwickeln, die man in die Hand nehmen kann und deren Position und Orientierung wir hochpräzise bestimmen konnten. Mit diesem System wurde aus dem einfachen Konzept des Auto-anmalen eine extrem interessante Geschichte, denn wir hatten eine superhohe Bandbreite im Interaktionskanal: es war schnell, es war präzise, es hat nicht gerauscht, man hatte keine Verzögerungen etc. Damit konnten wir Dinge entwickeln, die über die reine Notwendigkeit hinaus Spaß machen. Was ja auch bei den Markenwerten von MINI genau vorkommt. Dass es also zum Beispiel einen Sound gibt, der sich verändert, wenn man die Sprühpistole hin- und herschwenkt. Dass die Entfernung zum Bildschirm wichtig wird. Dass beide Dosen einen unterschiedlichen Effekt haben. Dass man auf die Farbeimer deutet und damit die Farbe wählt, dass man mit dem Hundnapf die Farbe abschüttelt, usw. Das ist dann eine sehr gute Lösung geworden, die sich aber aus der Entscheidung für ein hochwertiges Interface abgeleitet hat. Wir haben sehr dafür gekämpft, dass wir nicht eine Kinect einsetzen mussten, bei der man quasi immer nah an der Grenze ist von dem, was gerade noch ausreichend funktioniert.

*Nochmal zurück zum Szenografischen. Inwiefern haben Sie Räume so inszeniert, dass sie sich verändern, wenn sich irgendwelche Daten verändert haben? Also beispielsweise Trackingdaten oder Tageszeit?*

Das ist natürlich ein ganz großes Feld. Auch da ist das Gleiche zu sagen wie zur Interaktion. Denn auch das versteht man zu großen Teilen nicht. Das funktioniert in dem Moment, in dem man ein wiederkehrendes Publikum hat oder die Daten sehr trivial sind. Für die Fraport Konzernzentrale haben wir eine große Medienwand entwickelt. Die zeigt z.B. internationales Wetter in einer verständlichen und lesbaren Form. Das ist sehr gestaltet und abstrahiert dargestellt, damit es nicht zu platt wirkt, aber jemand, der da jeden Tag daran vorbeiläuft, versteht es ab irgendeinem Punkt.

Die generelle Frage von Trackingdaten oder anderen Datenquellen ist immer: Mache ich da mit viel Aufwand etwas, was ich auch durch einen Zufallsgenerator simulieren könnte? Verstehen die Besucher das? Das ist eine wichtige Frage. Für uns ist es sehr wichtig, dass man die Daten verstehen kann oder zumindest eine Geschichte darüber erfinden kann. Ich persönlich finde es relativ witzlos, Daten einfach in eine Maschine reinzugeben, irgendwie zu processieren und irgendwie wieder anzuzeigen. Das muss schon konkret Sinn ergeben. Oder es

ist eben nur ein Effekt, so wie ein bestimmter Vorhangsstoff, ein Material. Vielleicht ein strategisches Ornament.

*Was für Trends sehen Sie im Ausstellungsbereich, sei es im interaktiven oder klassischen Kontext?*

Ich bin noch ein wenig gespannt auf die Konsolidierung. Es gibt immer noch kaum Ausstellungen, die das Thema Smartphone wirklich auf den Punkt gebracht haben und die Möglichkeiten voll ausgeschöpft haben. Mein Gefühl sagt mir, dass da wahnsinnig viel möglich ist. Das ist nicht unbedingt ein Trend, aber eine Hoffnung von mir, dass das zu einem wird.

Der offensichtliche Megatrend ist natürlich, dass die Technik immer kleiner und günstiger wird. Man braucht keine riesigen Rechnerräume und Kabelwege mehr. Alles wird günstiger und dadurch auch leichter zu installieren. LEDs, Bildschirme und Drucker kosten kaum noch etwas. Aber das heißt natürlich nicht, dass es konzeptionell einfacher wird. Oder gestalterisch.

Und man kann mit einem innovativen inhaltlichen Konzept einen ganzen Entwurfsprozess durcheinanderbringen, weil das Konzept plötzlich Arbeitspakete generiert, die in dem Prozess gar nicht vorgesehen sind. Für das Goethe Nationalmuseum in Weimar haben wir eine Rauminszenierung über Goethes Faust entwickelt: Man blättert mithilfe eines kleinen Interface durch eine Liste aller Substantive und sieht in einer Projektionsskulptur alle Textstellen, in denen dieses Wort im Faust vorkommt. Zunächst einmal eine einfach klingende Idee, die wir eingebracht haben, als wir auf der Suche nach einer Funktion waren, in welcher der zentral geplante „Faust“-Raum in dem Museum seine Nutzung findet. Das Konzept hat sofort großen Zuspruch gefunden, aber interessanterweise hat es doch eine Komplexität im Prozess losgetreten. Man könnte denken, dass die Datenlage zum Faust ziemlich eindeutig ist, und der „Sourcecode“ von Goethe frei verfügbar ist. Aber es hat doch zwei Mitarbeiter über mehrere Monate beschäftigt, eine wissenschaftlich geeignete Edition herauszusuchen, die wir in einen maschinenlesbaren Text umsetzen konnten, Substantive aufzulisten, diese in die zur Goethe-Zeit relevante nominativ/singulare Schreibweise zu übersetzen, die Liste zu konsolidieren, die Quellen zu referenzieren usw. Das ist natürlich etwas, worauf in einem wissenschaftlichen Museum geachtet werden muss. Das ist hier wunderbar gelungen. Aber es ist etwas, was jeden Projektleiter, der nach seinen Projektphasen durch sein Projekt laufen will, in den Wahnsinn treibt, wenn man aufgrund „nur“ aufgrund eines Interaktionskonzepts mehrere Monate Arbeit mit Wissenschaftlern generiert.

Eigentlich müsste man aber so anfangen: Was sind wirklich interessante, wissenschaftliche Fragestellungen und mit welchen Interaktionskonzepten können wir diese darstellen? Du hattest vorhin gefragt, wie ein Ausstellungskonzept beginnen kann. Ich hatte daraufhin geantwortet, dass oft Architekten oder Szenographen damit anfangen. Und für den Architekten hat sich das Problem relativ einfach dargestellt: Wir haben in dem Raum eine Skulptur und auf die muss eine Projektion drauf. Da kann man als Architekt ein paar Schnitte zeichnen und ist fertig. Wir haben hinterfragt, was die Projektion überhaupt bedeuten soll. An dieser Stelle haben sich dann Forschungsfragen ergeben, die wiederum für die Wissenschaftler interessant waren. Insofern bin ich immer ein Freund von dem Ansatz, mit interessanten

Forschungsfragen anzufangen und dann zu überlegen, wie man diese digital/interaktiv gut darstellen kann. Das ist besser für den Projektplan.

*Wie relevant ist das Thema Personalisierung, beispielsweise Backpacking?*

Ich finde es kein sonderlich interessantes Thema. Denn letztlich will man als Ausstellungsgestalter seine Inhalte unter die Leute bringen. Das Ganze jetzt zu personalisieren ist selten relevant, denn letztlich ist es nicht gesetzt, dass man genau die Sachen, die man schon im Museum gesehen hat, zu Hause nochmal sehen will oder ob man nicht eher die Sachen sehen will, die man im Museum nicht gesehen hat. Oder man erinnert sich nur vage, dass man etwas gesehen hat, oder hört nur von jemandem davon und wünscht sich dann, dass man das noch schnell findet.

Vielleicht gibt es für das Thema Backpacking eine Option, wie man das leidige Thema Nutzungsrechte handhaben kann. Dass man mit dem Eintritt ins Museum ein besonderes Nutzungsrecht am Content erlangt. Also zum Beispiel: Ja, ich habe dieses Exponat in Echt gesehen, ich hätte es fotografieren können, also darf ich es mir auch runterladen. Aber ich bin kein Rechtsanwalt, ich improvisiere hier.

Ansonsten ist Backpacking in allen Bereichen relevant, wo Kreativitäts- oder „Selfie“-Spiele gemacht werden, oder man ein Spiel oder einen Wettkampf absolvieren kann. In welchen „klassischen“ Museen das Sinn macht, ist eine andere Frage. Für Kindermuseen oder Erlebnisausstellungen mag das Sinn machen, für eine klassische, beispielsweise naturkundliche Ausstellung glaube ich das eher nicht.

(DEMO)

*Was setzen Sie gerade mit der Kinect um?*

Im Prinzip genau solche Triggerzones, bzw. den Use Case, den Sie auch gezeigt haben. Für einen Showroom der Deutschen Telekom in Bonn haben wir zum Thema Future Retail eine kleine Inszenierung entwickelt, die aus einer Theke besteht, auf der verschiedene Produkte stehen. Mit zwei Kinects können wir tracken, ob der Besucher etwas angefasst, in der Hand gehabt oder sich angeschaut hat oder ob er nur davorsteht. Das ist dort als ein Arbeitsprototyp für zukünftige innovative Retail-Szenarien gedacht. Dort haben wir eine ähnliche Logik mit verschiedenen Trigger-Cubes gemacht, die dann Tracking-Nachrichten an mögliche Visualisierungen schicken. Das Ganze haben wir aktuell um eine Point Cloud Visualisierung ergänzt, um den Besuchern assoziativ-technisch zu illustrieren, wie die Daten getrackt werden und wie man sie mit Daten aus Warenwirtschaftssystemen kombinieren kann. Aber da ist auch viel Show dabei.

Ihr Programm hier [Anm.: Argus Vision] ist durchaus interessant, weil das Hauptthema natürlich das Aufsetzen und die Usability ist. Es ist wünschenswert so etwas zum Beispiel in einem Standalone-System am Laufen zu haben, sodass man nur noch einen Browser braucht, um das Ganze Remote zu administrieren. Das ist etwas, was fehlt – letztlich eine überschaubare Aufgabe, aber etwas, das im Projektgeschäft schwer umzusetzen ist. Denn klar – auch wenn Argus Vision jetzt hier als Abschlussarbeit sehr respektabel ist und länger gedauert hat,

ist es konzeptionell natürlich kein wirklich großes Projekt. Es ist ein kleiner Baustein in einem Ökosystem für Ausstellungstechnologie.

Das ist immer etwas, was mich so ein wenig ärgert. Man bekommt Ausstellungsausschreibungen, für welche zum Beispiel ein innovatives Besuchertracking gefragt ist, aber es ist wie eine Stand-Alone-Lösung ausgeschrieben, die man einfach aus der Schublade ziehen könnte. Man muss es fast zur Produktreife entwickeln, aber die Ausschreibung erfordert keine Nachnutzung in anderen Kontexten. Ich wünsche mir, dass es mal ein richtig großes Museum gibt, das sagt: Jetzt bauen wir nicht nur für zig Millionen so ein Gebäude, sondern wir geben ein paar Prozent davon aus, um unsere digitale Infrastruktur für die Ausstellungen der nächsten zehn Jahren mit der gleichen Detailliebe mit zu planen.

Da fehlt mir in der Tat immer noch eine Institution, die die Weitsicht hat zu sagen: Wir müssen auch mal einen signifikanten Teil unseres Budgets in solche Systeme investieren. Also nicht nur Vitrinensysteme und Lichttechnik anzuschaffen und sauber klimatisierte Ausstellungenräume für physische Sammlungen aufzubauen, sondern auch zu überlegen, wie digitale Sammlungen erlebbar gemacht werden können. Das ist sicher an vielen Stellen auch noch strategische Konzeptionsarbeit, aber an vielen Stellen auch nur ein Zusammenkehren und Vereinheitlichen von bekannten Ansätzen. Und es muss natürlich so angelegt sein, dass Sie das in fünf Jahren noch einem Szenografen zeigen können. Das geht auch – aber dann braucht man Viewer, Weboberflächen, standardisierte Datenprotokolle, Schnittstellen usw. Ich muss sagen können, wie genau Trigger funktionieren können, mit was ich sie koppeln kann, mit Licht, mit Bildschirmen, Audioguides, mit Smartphones, mit Datenbanken etc.

Da fehlt mir die Vision bei Museumsgestaltungen.

*Wenn wir bei Visionen sind: Sie sagten bereits, dass Remotesteuerung wichtig ist. Wo sehen Sie sonst noch Verbesserungsmöglichkeiten?*

Genau, mir ginge es primär um die Frage, wie ich das System konfiguriere, warte und monitore im täglichen Prozess. Und wie bekomme ich es hin, dass ein Szenografenteam das in einer zukünftigen Ausstellung einfach und selbstverständlich benutzen kann? Letztlich klassische Usability Fragestellungen. Ich denke, hier sind viele Sachen schon genau richtiggemacht. Es hat ein übersichtliches Menü und eine relativ klare Funktionalität. Es ist sauber gestaltet und ich glaube, man findet sich schnell ein. Aber klar, Teamviewer [Anm.: für Remotesteuerung von Control] ist eigentlich ein Showstopper, Windows wird generell zum Showstopper.

*Es läuft auch auf Mac und Linux-Systemen.*

Ja, Mac ist auch ein Showstopper. Ich denke die Zukunft in Ausstellungen liegt in kleinen standardisierten Mini-Plattformen, zum Beispiel den Raspberry PIs, die komplett standardisiert sind, nicht konfiguriert werden müssen, von der Speicherkarte booten und ihre eigene kleine Aufgabenstellung richtig gut machen. Wenn es nicht funktioniert, wirft man entweder die Speicherkarte oder den PI weg, aber man ist reproduzierbar nachher wieder dort, wo und wie es vorher funktioniert hat. Und insgesamt ist die Aufgabenstellung von Argus Vision ja einfach: Ein System, das Daten sensiert und weiterleitet.

*Im Grunde genommen ist dieser Control Teil Middleware. Dort laufen also alle Daten zusammen und diese Daten werden dann kumuliert über OSC an eine bestimmte IP und Port gesendet.*

Was ich wesentlich interessanter finde als OSC sind Pub-Sub-Frameworks wie MQTT. Vorteil ist, dass Sender und Empfänger nicht gleichzeitig aktiv sein müssen. Wenn ein Client mal weg ist und wiederkommt und der Master zwischendurch wiederkommt, dann bekommt er trotzdem den letzten Stand und muss nicht auf den Client warten, da es einen zentralen Broker gibt, an dem man sich subscriben kann. Das ist deutlich hilfreicher als diese Punkt-zu-Punkt-OSC Verbindungen, wo du hier und da eine IP Adresse eingeben musst und der eine immer nur mit dem anderen sprechen kann. Mit MQTT hat man einfach einen Raum von Kommunikationskanälen und jeder kann jede Art von Daten publizieren oder sich abonnieren. Aber das ist ein technisches Thema.

(...)

Die Triggerzones sind Bounding Boxen. Und die zeigen im Prinzip nur, ob ich drin oder nicht drin bin?

*Es gibt aus wer drin ist und wieviel Punkte der Person in der Bounding Box liegen.*

Ah, ok. Das heißt, Sie können im Prinzip auch Theremin-artige Anwendungen machen.

*Was auch gut wäre, ist, wenn man 3-Dimensional clippen könnte. Also man streckt die Hand rein und bekommt dann nur die Kontur der Hand.*

mhm (bejahend)

*Haben Sie jemals eine Anwendung implementiert, in der eine ID-Erhaltung über mehrere Kameras hinweg wichtig war?*

Nein, denn ich glaube Auftraggeber verstehen noch nicht so richtig, was man mit der ID-Erhaltung wirklich machen kann. Und dann ist es darüber hinaus noch schwierig mit mehreren Kinects, wenn sie eine große gemeinsame Fläche analysieren sollen. Ich habe ehrlich gesagt länger nicht verfolgt, ob das irgendwelche Leute schon gebaut haben – mein Gefühl ist, dass man ziemlich viele Daten zwischen ziemlich vielen Kinects austauschen muss, um das auf den Punkt zu bringen. Das ist, glaube ich, nicht einfach.

*Es ist nicht einfach. Im Grunde genommen kriegt man die kompletten Daten an einer Stelle rein. Das heißt, man hat keine Kommunikation mehr unter den Kinects. Es geht also nur noch darum, die Daten hier intelligent zusammenzuführen. Dann muss man am Anfang verstehen, wie die Kinects in Relation zueinanderstehen und dann kriegt man auch ein Tracking über mehrere Kinects hin. Das ist natürlich das Schwierige. Was es zum Beispiel gibt ist MultiKinect, ein C++ Tool, das eher für Unity Anwendung geschrieben wurde.*

Genau. Mit der Kinect2 ist man sowieso relativ schnell an der Grenze von dem, was man in den Rechner reinkriegt. Da ist die Frage, wieviel Kinects kriegt man sinnvollerweise an einen Rechner?

*Bei mir eine Kinect. Es gibt Workarounds, dass man mehrere Kinects anbinden kann, aber das funktioniert schon schlecht wegen den USB Ports und dann noch schlechter wegen der Daten. Das Programm hier läuft auch schon am Limit.*

Letztlich werden aber diese Tiefenbildkameras immer billiger. Für die Chipsätze in den Kameras zahlt man inzwischen nur noch 5\$, die heißen dann nur nicht Kinect. Es gibt von Texas Instruments zum Beispiel riesige Toolkits, mit denen man sich seine eigenen Time-Of-Flight Sensoren konstruieren kann. Für BMW haben wir gerade ein Projekt umgesetzt, wo wir diesen Gestensensor, der im neuen 7er verbaut ist, integriert haben. Da haben BMW im Dachhimmel (also da wo die Schiebedachsteuerung und die Leselampe ist) eine Tiefenbildkamera integriert, mit der der Fahrer das Entertainment-System steuern kann. Die funktioniert auch im Fahrzeug bei Sonnenlicht. Noch ist das nur in der Oberklasse, aber letztlich sind das inzwischen Standardkomponenten in der Industrie. Insofern ist es bestimmt interessant, weiter darüber nachzudenken, was man mit einer dedizierten Hardware hinbekommen kann, auf die man ein gutes, anwenderorientiertes Interface aufsetzt. Also soweit, dass es natürlich auch für Diebstahlschutz im Museum total hilfreich ist. Da gibt es im Prinzip ja ähnliche Aufgabenstellungen.

*Was gefällt Ihnen am Interface?*

Ich finde, das ist schon alles sehr schlüssig. Es erschließt sich glaube ich sofort. Natürlich hängt das ein wenig davon ab, was ich vorhabe: Wenn ich mir vorstelle, dass ich wirklich eine Ausstellung mit 30 Kinects habe, dann würde ich sagen, dass man das Thema zentraler administrieren müsste. Und nicht über einen Processing Sketch auf jedem Rechner, sondern über ein gemeinsames Interface und am besten über den Browser bedienen. Plug & Play ist natürlich auch ein gutes Feature – also, dass man sich das komplette Kalibrieren und Adressieren spart. Die andere Frage ist natürlich: Wie teuer darf es eigentlich sein? Wer ist die Zielgruppe für das Produkt? Und was passiert mit den ganzen Daten im nächsten Schritt? Denn beim Tracking das Geld zu sparen, um dann nachher doch die gleiche aufwendige Terminalprogrammierung zu machen, ist nur halb interessant. Da ist es in der Tat wirklich wichtig strategisch und längerfristig die Bausteine zu isolieren, die man in einem Museum der Zukunft braucht. Denn Argus ist ganz klar nur ein Baustein.

Wenn Szenografen erstmal verstehen, was der in einer Ausstellung leisten könnte – also über einem Tisch mit Exponaten oder anfassbaren Dingen oder technologischen Artefakten, Tafeln oder Steine, etc. ergeben sich sensationelle Möglichkeiten. Man muss sich dann die Contentproduktion natürlich auch antun wollen. Irgendwas muss dann passieren, wenn irgendwas angefasst wird. Höre ich dann mit meinem Audioguide einen Track zu jedem Stein, den ich anfasse oder sehe ich auf einem Bildschirm irgendwelche Fotos oder läuft ein Film oder steuere ich die große Inszenierung hier, was auch immer? Ist es nur ein Teil einer komplexeren Bedienlogik, die ich dann damit aufbaue? Und das müsste man noch ein bisschen rauskitzeln.

*Dann sind wir auch wieder bei dem Punkt: Ok, ich kann irgendwo hinfassen und irgendwas passiert. Aber es ist natürlich schwer, das bei Tracking begreifbar zu machen, außer man klebt irgendwo ein Hinweisschild hin.*

Ja, eigentlich muss man argumentieren, dass es in der Ausstellung immer überall so ist. Nur so macht man das verständlich. Dass man im Museum alles anfassen kann und überall passiert etwas. Das meine ich auch mit der strategischen Konzeption einer Ausstellung. Im Moment ist es halt eher so, dass gesagt wird: „Ja, das ist so ein Gimmick, das probieren wir mal und das ist für die jungen Leute, damit die auch was haben, da machen wir mal was Digitales.“ Aber ich glaube es wird erst wirklich interessant, wenn in einer Ausstellung jedes Exponat so ist. Also so wie jedes Exponat ein gutes Licht hat, hat auch jedes Exponat ein Tracking, damit ich es anfassen und erleben kann, dann funktioniert das auch. Aber dazu muss es noch günstiger werden.

*Vier Kinects in unserer Ausstellung waren 600€ pro PC und 200€ pro Kinect.*

Naja, klar. Aber der typische technische Leiter wird zurecht hellhörig, wenn man sagt, dass man einen PC installieren will. Wer administriert uns den? Der ist dann vielleicht nach zwei Jahren kaputt, braucht Updates, fängt an zu pfeifen, Microsoft kündigt das Betriebssystem etc. Das ist dann gleich wieder Servicevertrag; Servicetechniker, Anfahrtspauschale usw.

(...)

*In welchen Projekten haben Sie Skeletterkennung benutzt?*

Wir haben ein Projekt für Nokia gemacht. Da haben wir in einem Showroom einen digitalen Spiegel konstruiert, in dem man sich selbst in einer Wolke sieht. Da haben wir Skeletterkennung tatsächlich benutzt.

Ansonsten hat Marcel Wiessler, ein Freelancer von uns, gerade an der Uni eine total schöne Geschichte gemacht, da ging es um den menschlichen Körper. Die haben ein 3D-Modell von den Knochen, Muskeln und Organen des Körpers gemacht und das dann an ein Kinect-Skelett drangehängt. Du stehst davor und siehst genau, wo welche Organe sind usw. Wenn man näher rangeht, blättern sich verschiedene Ebenen des Körpers ab. Das, was die gemacht haben, ist schon relativ spooky, wenn man von sich zum Beispiel nur noch seinen Muskelapparat oder nur die inneren Organe sieht. Das ist natürlich ein Thema, wo es sensationell gut passt. Sonst ist es immer ein wenig mühsam mit der Skeletterkennung. Thema 1 ist, dass die Kinect immer an der falschen Stelle ist. Die Kunden denken, dass die Kinect ja irgendwo hängen kann, aber in Wirklichkeit muss sie ja ideal genau da hängen, wo auch der Bildschirm sein muss. Thema 2 ist, dass die Besucher wissen müssen, wo man sich hinstellen muss und wohin man schauen soll. Also nicht zu nah, nicht zu weit und möglichst nicht in großen Gruppen und das geht dann doch meistens schief. Und Bodenmarkierungen mag selten jemand haben.

*Inwiefern kann Argus Vision in einer Prototyping Phase eingesetzt werden? Wie läuft diese Phase überhaupt bei Ihnen ab?*

Prototyping ist wahnsinnig wichtig, denn sonst bekommt man die Sachen nicht verkauft und auch nicht entwickelt. Denn auch Entwickler verstehen vieles nicht, wenn sie es nicht in Funktion sehen. Da sind wir gut aufgestellt, weil wir das alles Inhouse machen und mit der Software vvvv einen relativ soften Weg von dem ersten Konzeptprototypen bis zur finalen Applikation etabliert haben. Insofern können wir uns von der ersten Skizze, über einen funk-



tionierenden Prototypen, bis zur finalen Applikation und zur Software, die es dann im Betrieb hat, weiterhangeln. Das ist ein Prozess, bei dem man durchaus viel wieder wegwirft, aber im Wesentlichen ist es derselbe Code, der von Anfang an bis Betrieb dann läuft. Die Tools sind also in allen Projektphasen die gleichen.

Ob uns sowas [Anm.: Argus Vision] dann dabei helfen würde? Uns glaube ich nicht wirklich, denn wir würden uns das eher selber zusammenbauen. Ich glaube ein Szenario wäre, wenn zum Beispiel ein „klassischer“ Szenograf den Lead hat, man diesem solche einfachen, verständlichen Tools geben kann und sie damit spielen und Ideen entwickeln lassen kann. Man stellt denen fünf solcher Trackingsysteme hin und kann sagen: Macht die Ausstellung wie ihr wollt, aber benutzt dieses System.

Früher haben alle Museen diese riesigen schwarzen Röhrenmonitore von Sony gehabt und es war allen Szenographen klar, dass wenn sie eine Videoarbeit zeigen wollen, man halt so einen Sonymonitor auf ein Podest stellen muss. So ähnlich könnte man das damit im Prinzip auch heute noch machen. Ob Szenographen das heute noch so akzeptieren würden, weiß ich natürlich nicht (Lachen).

(...)

Was ich in der Tat interessant finde ist das Thema Kinect auf Raspberry Pi.

(...)

ENDE

## Interview mit Prof. Eberhard Schlag

Passagen, die vornehmlich von allgemeinem und nicht fachspezifischem Charakter waren, wurden nicht transkribiert und sind mit (...), kurze Redepausen mit (..) gekennzeichnet.

Das Interview mit Prof. Eberhard Schlag, Partner beim Atelier Brückner wurde am 02.03.16 im Atelier Brückner in Stuttgart geführt:

*Was machen Sie denn gerade für ein Projekt in China?*

In China machen wir gerade ein Projekt für einen großen Immobilieninvestor: Eine 360° Bespielung in einer Shopping Mall. Es geht darum, dass es sich die Stadt auf die Fahnen geschrieben hat, sich zu einer Smart City zu entwickeln. Das soll der Bevölkerung in diesem Raum mit der Bespielung näher gebracht werden.

*Ich war am Montag bei MESO in Frankfurt, die eher digital sind. Dagegen würde ich Sie als klassischer bezeichnen. Sie sind in der Konzeption von Anfang bis zum Ende drin, das heißt sind Sie diejenigen, die eine Ausstellung komplett durchkonzipieren?*

Das ist richtig.

*Jetzt sprachen Sie 360° Bespielungen an. Was ist Ihre Rolle zum Beispiel in diesem Kontext?*

Wir haben einen sehr umfassenden Gestaltungsanspruch. Das heißt, wir machen in diesem Projekt erst einmal alles: Wir entwickeln die Idee, das Konzept, überlegen uns dann, wie das alles technisch umgesetzt werden kann und treten in der Regel gegenüber unseren Kunden auch als Generalplaner oder als klassische Agentur auf, die alle Dienstleistungen übernimmt. Dann schauen wir was wir selber machen können. Das geht in der Regel bei Medienproduktionen bis zum Storyboard und der Entwicklung einer groben Struktur. Im Anschluss beauftragen wir normalerweise eine Medienproduktionsfirma, die dann die Produktionsleistungen übernimmt. Wir selbst sind keine Filmproduktion. Und was die technische Fachplanung angeht, da arbeiten wir auch mit Partnern zusammen, die sagen: Dafür brauchen wir die-und-die Projektoren und genau die Sensorik und so weiter.

*Wie akquirieren Sie Ihr technisches Know-How? Wie bleibt man am Ball, wenn man das Ganze schlussendlich im technischen nicht umsetzt?*

Das funktioniert über verschiedene Kanäle. Zum einen hat man natürlich ein Interesse an der ganzen Thematik. Wir waren ja ganz, ganz früh im medialen Bereich unterwegs. Wir haben schon 2000 auf der Expo gemeinsam mit der Filmakademie Ludwigsburg ganz große Medieninstallationen gemacht. Wir haben eine sehr große Medienaffinität und sind da sehr breit aufgestellt. Wir gehen mit offenen Augen durch die Welt und schauen, was es so an aktuellen Entwicklungen gibt. Dann sind wir auch viel auf Konferenzen, besuchen regelmäßig Ausstellungen. Ganz oft bekommt man seine Inspirationen auch eher aus künstlerischen oder rein technologiegetriebenen Projekten. Es gibt ja diese beiden Richtungen: Also zum Beispiel das, was in Linz am Ars Electronica Center entsteht. Das sind oft Dinge, bei denen man denkt: „Das ist eine tolle Technik, aber was sie damit gemacht haben ist eigentlich noch nicht gut“. Das sind Sachen, die für uns interessant sind. Da sehen wir, es gibt neue Ideen,

die sind aber bisher nur technologisch gedacht. Wir versuchen diese dann inhaltlich zu denken und räumlich zu integrieren. Dann kommen spannende, neue Lösungen dabei heraus.

*Was für Trends sehen Sie in aktuellen Ausstellungen oder Installationen?*

Also ich glaube ein großer Trend ist nach wie vor das Thema Partizipation. Und Partizipation ist mittlerweile auch ein Trend, der das klassische Museum erreicht hat. Es geht immer stärker um die Frage: „Wie involviere ich den Besucher tatsächlich? Wie wird er Teil einer Inszenierung? Wie kriege ich ihn aus der klassischen Betrachterrolle heraus und mache ihn zu einem Teil der Ausstellung?“ Ich glaube das ist ganz elementar. Man kannte das bisher eher von Science Centern oder Mitmach-Ausstellungen und dort ist das eigentlich schon ein alter Hut. Aber die klassischen Museen haben jetzt auch erkannt, dass sie sich damit ganz andere Zielgruppen erschließen können und dass die Besucher auch mit einem ganz anderen Gefühl aus so einer Ausstellung herausgehen, wenn sie selber etwas gemacht haben und nicht nur Betrachter waren. Das geschieht aktuell auf zwei Ebenen: Ebene eins ist eine Interaktion durch interaktive Medien – das ist sicherlich schon ein paar Jahre alt – und aktuell auch immer mehr über Spiele.

*Inwiefern?*

Über Spiele, die dann auch teilweise moderiert sind und die auch analog sind. Rollenspiele zum Beispiel. Diese haben wir jetzt schon in einigen Projekten umgesetzt, weil sie in der Gruppe noch viel intensiver ein Erleben möglich machen. Im Parlamentarium in Brüssel haben wir das zum Beispiel eingesetzt oder in einer Ausstellung über das Altern der Gesellschaft. Das ist, glaube ich, der aktuellste Trend. Da mischen sich analoge und digitale Formen relativ selbstverständlich. Überhaupt würde ich sagen, dass sich gerade der digitale Bereich aus einer Phase der kritischen Distanz, den ja viele Museen eingenommen hatten, über einen Prozess der Aneignung mittlerweile, in einen ganz selbstverständlichen Umgang gewandelt hat. Es ist immer ein ganz gutes Zeichen, wenn es nicht mehr so stark hinterfragt wird. Früher gab es die Frage, wieviel Medieneinsatz wir in Museen und Ausstellungen wollen. Wie digital darf das Ganze sein? Das ging hin bis hin zu Prozentangaben. Völlig absurd. Heute stellt man sich eher die Frage, was eigentlich die richtige Lösung für das Problem ist. Dann kann eine analoge Lösung genauso richtig sein, wie eine digitale. Da ist man jetzt deutlich weniger dogmatisch unterwegs.

*Sie haben gerade schon gesagt, dass Technologie bei Besuchern mittlerweile mehr Akzeptanz findet. Mich haben ja sehr stark leichte szenografische Änderungen im Raum interessiert. Herr Hegemann meinte bereits, dass die LED Wand im BMW Museum nicht so ganz funktioniert hätte, da die Leute das nicht verstanden hätten.*

Die Besucher haben zunächst nicht gemerkt, dass sie mit der Wand interagieren. Ich glaube, das ist ein ganz wichtiger Aspekt: Dass es immer einen direkten Response gibt, den der Besucher versteht. Ich glaube, das ist ganz elementar für jede Form der Interaktion. Gerade bei interaktiven Räumen muss der Besucher verstehen, dass er überhaupt die Möglichkeit hat zu interagieren und wie das Ganze funktioniert. Und wenn etwas passiert, dass er das auch nachvollziehen kann, dass er das war und, dass er das so-und-so ausgelöst hat. Das hatten wir damals nicht immer erreicht. Gut, das ist auch schon ein paar Jahre her. Da hatten wir eigent-

lich die Idee, dass man daran vorbeiläuft und beim Vorbeilaufen verändert sich die ganze Wand. Das funktioniert natürlich aus der Distanz, aber wenn ich gerade an der Wand vorbeilaufe und runterschaue, was da unten für ein tolles Auto steht, dann merke ich gar nicht, dass ich etwas verändert habe. Dann war es vielleicht auch nicht deutlich genug, dass die Besucher von gegenüber das gesehen haben. Wenn sich immer alles bewegt, dann erkenne ich natürlich auch nicht unbedingt, dass sich dann plötzlich etwas anders bewegt. Ich glaube, das ist oft ein Problem, dass wir Dinge viel zu kompliziert denken. Weil wir so tief in den Projekten drin stecken. Mit mir ist meine ich jetzt nicht unbedingt uns, sondern uns als Gestalter allgemein. Für uns ist dann Schritt Eins, Zwei, Drei, Vier schon klar, aber für den Besucher überhaupt nicht. Der kommt dahin, hat keine Ahnung und muss das was er sieht intuitiv verstehen und nachvollziehen können. Deswegen funktionieren die einfachsten Dinge eigentlich immer am besten. Im ZKM [Anm.: ZKM Karlsruhe] gab es eine Installation, die mir bis heute in Erinnerung geblieben ist: Da sind einfach nur Seifenblasen runtergefallen, eine einfache Projektion mit Seifenblasen. Auf dem Boden sind sie zerplatzt und als Besucher konnte ich sie in der Luft halten. Total simpel. Eigentlich total banal, aber jeder hat das sofort verstanden.

*In welchem Kontext war das?*

Das war nur Spielerei (Lachen).

*Haben Sie sonst Exponate oder Installationen umgesetzt, die eher „schön“ waren, als dass tatsächlich Wissensvermittlung im Vordergrund war?*

Da muss ich mal kurz überlegen. Wir haben relativ viele Konzepte für interaktive Räume gemacht. Aber tatsächlich umgesetzt haben wir relativ wenig. Also viele klassische Installationen natürlich, wo es um das Erklären von Zusammenhängen geht. Aber diese intelligenten, interaktiven Räume, das sind Projekte, da gibt es viele Ideen und Konzepte dazu, aber relativ wenig Möglichkeiten, das so wie gedacht umzusetzen.

*Können Sie mir dazu ein wenig mehr erzählen?*

Ja, wir haben einen großen Wettbewerb gewonnen für die Expo in Kasachstan 2017, haben das Projekt allerdings nicht bekommen. Lange Geschichte. Da haben wir tatsächlich für einen Pavillon, der unter dem Oberbegriff „Game“ lief, ein großes interaktives Spiel für die Besucher entwickelt, die dann in Gruppen interagieren konnten. Die Idee war, dass immer 17 Besucher in ein Spielfeld kommen und dann mit Hilfe eines Moderators ein Spiel spielen. Mit Gesten und Bewegungen hätten Sie dieses Spiel steuern und beeinflussen können.

*Sie haben das Konzept dann nicht weiterverfolgt, weil Sie den Auftrag nicht bekommen haben?*

Ja.

*Hatten Sie dazu schon Ideen, wie Sie das technisch umsetzen wollten?*

Noch nicht konkret. Wir machen keine Vorschläge, von denen wir nicht wissen, ob wir sie technisch hinbekommen. Aber genau mit welcher Technik man das umsetzt, so weit geht man im Wettbewerb nicht.

*Was für Projekte haben Sie sonst noch mit Tracking umgesetzt? Sie können auch gerne von Konzepten berichten.*

Bei einem ganz spannenden Konzept ging es um die Idee einen komplett interaktiven Raum zu gestalten, der sich durch die Interaktion mit dem Besucher auch physisch verändern kann. Das ist ein Konzept für einen kinetischen Raum, der mit variablen Wänden funktioniert. Es gab verschiedene Einflussfaktoren, die den Raum verändert haben, unter anderem die Besucher, aber auch Ton, Geräusche und so weiter. Die Idee war, dass es ein vollständig getrackter Raum ist, bei dem die Wände aus einzelnen Modulen bestehen, die dann verschiebbar sind, sodass der Raum nicht nur seine Bespielung, sondern auch seine Physis verändern kann. Das ist ein Projekt, was schon länger in der Schublade liegt, aber noch keinen konkreten Anwendungsfall hat. Aber genau das interessiert uns: der Weg vom kinetischen Objekt zum kinetischen Raum. Das sind Dinge, an denen wir gerade arbeiten. Da gibt es zum Beispiel ein interaktives Aquarium, in dem der Besucher mit Fischeschwärmen interagieren kann. Da ist ein klassisches Tracking im Einsatz. Die digitalen Fische reagieren auf den Besucher und interagieren mit ihm. Das ist in einem Naturkundemuseum in Saudi-Arabien. Aber das ist auch schon konkret in Planung und Umsetzung.

*Haben Sie jemals ein Projekt umgesetzt, bei denen sich der Raum in Abhängigkeit von Daten, wie zum Beispiel Wetter oder Twitterfeeds, verändert hat?*

Die Deutsche Börse in Frankfurt kennen Sie ja, die ist ja jeden Tag in den Nachrichten. Da haben wir das versucht. Das ist ein Raum, der sich aus Livedaten versorgt. Also erstmal gibt es dort eine abstrakte Weltkarte, die den Tag-Nacht-Verlauf auf der Welt zeigt. Das ist natürlich vorprogrammiert, denn man weiß ja (Lachen), wie sich das verändert. Und in dieser Karte sind dann immer die aktuellen Börsendaten der Handelsplätze weltweit eingebettet. Je nachdem, ob der Kurs fällt oder sinkt, verändert sich die Farbe. Der Wert ist natürlich auch ablesbar. Das Gleiche gibt es nochmal in der Besuchergalerie auf dem Boden. Da wird das über farbige Kacheln gezeigt, die in Abhängigkeit von den Kursen entweder weiß, rot oder grün leuchten. Das ist natürlich jeweils ein Livefeed. Der ganze Raum ist auch über das Tageslicht gesteuert. Das heißt morgens hat man eher tageslicht-ähnliches Licht und abends wird es dann immer wärmer. Es ist ein Raum, der sich komplett verändert. Teilweise vorprogrammiert und teilweise mit Echtzeitdaten gefüttert.

*In Ausstellungen haben Sie sowas aber nie umgesetzt? Bei der deutschen Börse hat man wahrscheinlich eher nicht so das Besucherprinzip. Da hat man den Vorteil, dass die Leute jeden Tag da sind und das wahrscheinlich deshalb besser verstehen.*

Wobei die Börse ja zwei Facetten hat. Da gibt es natürlich den Handelssaal für die Insider. Aber es gibt auch den Besucherbereich, wo jeden Tag im Stundentakt Besuchergruppen durchgeführt werden. Klar, die laufen nicht ungeführt durch. Die bekommen Informationen dazu und können das dadurch besser nachvollziehen. Aber es richtet sich schon an beide Zielgruppen.

(DEMO)

Und Sie können auch Personen unterscheiden?

*Ja, das funktioniert allerdings noch nicht perfekt.*

Und wie kommt diese Unterscheidung zustande?

*Wenn eine Person das Bild betritt, bekommt sie eine ID zugewiesen. In jedem darauffolgenden Kameraframe wird dann geprüft, ob sich der Mittelpunkt der Person stark bewegt hat. Wenn der Punkt jetzt wo ganz anderes ist, dann kann ich davon ausgehen, dass das höchstwahrscheinlich nicht dieselbe Person ist. Wenn der Punkt sich nur wenig bewegt hat, kann ich davon ausgehen, dass das dieselbe Person ist. Das hat den Nachteil, wenn sich zwei Personen kreuzen. Dann verliere ich oft die ID. Ansonsten funktioniert es stabil.*

mhm (bejahend) Muss die Person denn immer in diesem Feld bleiben oder kann die auch mal raus- und wieder reinkommen und er erkennt, dass es die gleiche Person ist.

*Das geht nicht. Das erkennt er nicht*

Er erkennt es also nur solange die Person im Trackingbereich ist.

*Was ein Plan war, den ich am Anfang verfolgt habe, war dieser Button hier, der im Moment noch gar nichts tut. Da ging es darum, dass man mehrere Kameras zu einer verbindet. So, dass man vier Kameras in alle Ecken eines Raums hängt und damit ein komplettes Raumtracking umsetzen könnte.*

Das heißt, Sie benutzen nur eine Kamera?

*Man kann beliebig viele Kameras anbinden. Die senden ihre Daten jeweils an die Middleware, die Kameras wissen aber in dem Fall nichts voneinander. Im Moment kann ich also die Kameras alle zentral ansteuern, sie werden aber nicht als eine große Einheit gesehen.*

Das wäre dann der nächste Schritt.

*Ja, dann könnte man auch so weit gehen und diese Point Cloud Ansicht dreidimensional zusammenfügen und dann hätte man eine komplette Ansicht des Raums.*

Ja, das wäre eigentlich interessant, oder? Das ist dann die Masterarbeit (Lachen). Und was ist jetzt das Neue daran? Oder was unterscheidet das von bisher existierenden Systemen?

*Neu daran ist prinzipiell im Trackingbereich nichts. Neu daran ist, dass es meines Wissens nach keine einheitliche Lösung dafür gibt. Es gibt nichts, das so funktioniert, dass man sich einfach ein Programm runterlädt, die Kinect anschließt und sie funktioniert, vor allem nicht aus allen Winkeln. Das Ganze über eine Oberfläche anzusteuern, das so runterzuberechnen und das Triggerzoneprinzip gibt es auch nicht.*

Können Sie mir das Triggerzoneprinzip nochmal erklären? Ich bin ja Laie (Lachen).

*Hier kann ich in einen dreidimensionalen Raum Zonen einzeichnen und dann nicht mehr über eine Kontur Trackinginformationen herleiten, sondern einfach nur sagen: „Wenn jemand durch eine bestimmte Stelle im Raum durchläuft, dann löse ich etwas aus“.*

Ok, und wie fein kann man das machen?

*Die Kinect ist – das schwankt, je nachdem wie groß der Störfaktor in dem Bereich ist – 5 bis 50 mm genau. Diese Zonen kann ich im Grunde genommen so groß oder klein machen, wie ich will.*

Ok.

(...)

Also was ich interessant finde, ist das Thema Personen erkennen und zuordnen. Das halte ich für extrem interessant, deshalb habe ich auch nachgefragt. Kann ja mal sein, zum Beispiel bei einem Spiel, dass der eine kurz raus und wieder reingeht. Aber das ist vielleicht eine Schwierigkeitsstufe, die man in einem zweiten Schritt lösen könnte. Aber diese Zuordnung ist natürlich gerade für solche Spielanwendungen – ich rede jetzt nicht von einem Computer-game, sondern einem interaktiven Spiel – an denen mehrere Teilnehmer mitmachen, interessant. Das ist schon etwas, was wir in den Konzepten immer häufiger überlegen. Da ist es natürlich wichtig, dass jeder seine Rollen übernimmt und dass die Rollen klar erkannt werden.

*Haben Sie denn schon ein Spielprinzip mit einem Trackingsystem umgesetzt?*

Nein, das sind alles Sachen, die wir in letzter Zeit in Wettbewerben vorschlagen und an denen wir gerade dran sind.

*Wenn Sie solche Konzepte entwerfen, bauen Sie dann auch selbst Prototypen innerhalb des Ateliers?*

Ja, das machen wir. Also nicht im Wettbewerb, dafür ist die Zeit nicht da. Eigentlich ist das immer zweistufig: Phase Eins ist, dass wir Sachen selbst ausprobieren. Da rufen wir ICT an oder irgendeinen anderen Partner und basteln uns dann selber was zusammen. Wenn man der Meinung ist, dass das funktioniert, dann gibt es meistens einen Probeaufbau für den Kunden, der natürlich auch ein wenig Geld kostet. Bei den großen internationalen Projekten gibt es teilweise sechstellige Prototyping-Budgets, weil die natürlich nichts kaufen, was sie nicht vorher probiert haben und von dem sie wissen, dass es 100% funktioniert.

*Wie gefällt Ihnen das Design der Software?*

Ich wollte gerade fragen, ob Sie das selber gemacht haben.

*Das habe ich selber gemacht, ja.*

Ja? Ok. Finde ich gut.

*Schön (Lachen).*

Sie haben ja nur die Schrift ausgeliehen, oder?

*Ja.*

Das Design ist gut. Das sieht erstmal so aus, als hätte da auch ein Gestalter mitgewirkt. Wenn Sie das alleine gemacht haben, dann bin ich beeindruckt. Das sieht auch so aus, als würde man das verstehen. Da muss man kein Informatiker sein, um das bedienen zu können,

das ist ja das Allerwichtigste. Sonst wendet das ein Gestalter nicht an, wenn er da irgendwelche Codezeilen eingeben muss oder es so aussieht, als bräuchte man ein abgeschlossenes Studium, um es verwenden zu können.

*Sie finden es also verständlich?*

Es sieht zumindest so aus, ja.

(...)

*Sie meinen ja bereits, dass ID-Erhaltung wichtig ist. Sehen Sie sonst noch Verbesserungsmöglichkeiten? Wie wichtig ist beispielsweise die Skeletterkennung?*

Da haben wir auch schon ein paar Sachen gemacht im musealen Kontext. Aber ich finde am interessantesten ist im Interaktionsbereich die Interaktion im Raum mit mehreren Personen. Dass Personen zuverlässig erkannt werden und dass sie bestimmte Dinge tun, die dann auch zuverlässig zugeordnet werden können. Ob man dafür ein Skelett braucht weiß ich nicht. Sie machen das ja aber über diese Triggerzones, oder? Die kann man ja beliebig im Raum anlegen. Wie ist das mit Gestenerkennung?

*Das funktioniert gar nicht.*

Das wäre noch interessant natürlich. Aber wichtig ist bei solchen Systemen die Zuverlässigkeit. Wenn so etwas nicht richtig funktioniert ist das im Ausstellungskontext eine Katastrophe. Das muss wirklich total zuverlässig und robust sein.

*Zuverlässig im Kontext von (..)?*

Man überlegt sich ja Interaktionsmöglichkeiten, die dann auch im Rahmen der Software möglich sind. Aber die müssen auch absolut zuverlässig erkannt werden. Und wenn es dann nur jedes zweite Mal funktioniert, dann ist das so ziemlich das Schlechteste, was passieren kann. Dann hat keiner Lust, damit zu interagieren. Dann kippt so etwas.

*Was für Anwendungsfälle sehen Sie für das Triggerzonekonzept?*

Ich meine, für uns ist die Technologie eigentlich nicht so wichtig, sondern, dass es zuverlässig ist und – das ist natürlich ein Thema – kostengünstig. Ist es denn eine kostengünstige Technologie? Man könnte ja Triggerzones einfach dadurch realisieren, dass man mit mehreren Infrarotsensoren arbeitet.

*Klar, man kann natürlich auch unendlich viele Infrarotsensoren aufhängen. Das ist wahrscheinlich sogar immer noch kostengünstiger, da diese nichts kosten. Was meiner Einschätzung nach teurer wird, ist die Programmierung. Also bei mir ist es im Grunde genommen Plug&Play. So eine Sensorik muss jedes Mal programmiert werden, da muss jedes Mal eine Logik dahinterstehen und die sieht jedes Mal anders aus.*

Aber das heißt auch, dass ich jedes Mal im Raum die Kameras kalibrieren muss? Das heißt, ich muss den Raum aufnehmen, zeichne dann im Raum diese Felder auf und weise diesen dann bestimmte Dinge zu.

*Ja, das macht man natürlich einmal.*



Aber das muss man immer im fertigen Raum machen?

*Diese Zonen behalten natürlich ihre x,y,z Koordinaten. Das heißt, bewege ich die Kamera, dann stehen diese auch am falschen Fleck. Wenn ich die in einem Raum aufbaue, der identisch ist, dann stehen die natürlich auch wieder an der richtigen Stelle. Was man machen kann, ist, in einem kleinen Raum Zonen einzuzeichnen und zumindest schon was passiert an die Zonen anzupassen. Im richtigen Raum muss man die Zonen nur noch kurz verschieben.*

mhm (bejahend). Wie groß sind so Räume, die man damit erfassen kann?

*Das hängt natürlich von dem Abstand von der Kinect zum Raum ab. Meine Einschätzung wäre, für eine noch gute Erkennung, ca. sechs Meter Entfernung in der Länge und sieben Meter in der Breite.*

(...)

*Wie ist das Verständnis von Besuchern für solche Trackingkonzepte?*

Ich glaube, wenn Sachen nicht vollständig selbsterklärend sind, dann braucht es einfach eine Handlungsanweisung. Dass das [Anm.: Das dritte Obergeschoss in der Ausstellung „Tell Genderes“] ein interaktiver Raum ist, das hat man dem Raum ja nicht angesehen. Das hätte auch eine Filmproduktion auf drei Leinwänden sein können. Man hätte sich auch hinstellen können und warten bis der Film endlich losgeht. Als Gestalter will man natürlich nicht gerne solche Erklärungen oder Handlungsanweisungen hinhängen. Aber oft ist es notwendig, weil es sonst nicht verstanden wird. Und ich finde das auch nicht schlimm. Schöner ist es natürlich, wenn man ohne auskommt, aber dann muss es wirklich völlig klar sein. Dann muss, wenn ich denn Raum betrete, schon irgendwas passieren. Dann merke ich: „Jetzt bin ich da reingekommen, jetzt hat sich irgendwas verändert. Und wenn ich dann nochmal was tue, dann verändert sich nochmal was.“

ENDE

## Interview mit Christoph Diederichs

Passagen, die vornehmlich von allgemeinem und nicht fachspezifischem Charakter waren, wurden nicht transkribiert und sind mit (...), kurze Redepausen mit (..) gekennzeichnet.

Das Interview mit Christoph Diederichs, Interaktionsgestalter bei Atelier Markgraph wurde am 03.03.16 im Atelier Markgraph in Frankfurt am Main geführt:

*Ich habe auf der Internetseite gesehen, dass Sie als Interaktionsgestalter aufgeführt sind. Was genau kann man sich darunter vorstellen?*

Das kommt von meinem Studium. Ich habe Digitale Medien studiert. Da lernt man relativ viel Interaktionsgestaltung, also wie bereite ich Software designmäßig so auf, dass es für den Benutzer einfach, intuitiv ist und Spaß macht. Ich habe mich dann hier im Laufe der Zeit zum Interaktions- und Produktgestalter für Exponate weiterentwickelt und mache in einem kleinen Team – wir sind drei bis vier Leute – hauptsächlich Exponate, also alle die Sachen, die man so anfassen kann. Vom Touchscreen bis zu irgendwelchen haptischen Interfaces.

*Das heißt Sie entwickeln das auch tatsächlich im Haus?*

Also wir entwickeln auch selbst. Im Prinzip ist es so, dass wir bis auf Ausnahmen gestalterisch alles selbst machen. Das hängt immer ein wenig von der Gesamtauslastung ab. Wir geben die Programmierung dann meistens an Partner ab, die spezialisiert sind auf gewisse Dinge. Aber hin und wieder machen wir auch Inhouse-Produktionen, wenn es die Zeit erlaubt.

*Was waren denn die letzten Produkte, die Sie umgesetzt haben?*

Wir haben vor kurzem, das ist noch nicht ganz öffentlich, für eine Dachgesellschaft einer Bank, eine Ausstellung gemacht, die die Grundlage der Geldanlage erklären soll. Da kann man an ein paar Exponaten lernen, dass man sein Risiko streuen muss, Geduld haben muss am Aktienmarkt, den Zinseszins ausnutzen muss. Und das wurde mit haptischen, relativ einfach zu verstehenden Anwendung umgesetzt. Ansonsten arbeiten wir viel für Mercedes Benz. Da gibt es auch immer relativ viele Exponate. Das ist so das Brot und Butter Geschäft hier.

*Für Mercedes Benz sind es dann eher immer Messen?*

Ja, hauptsächlich. Wir entwickeln immer eine Art Gesamtsprache bei den Exponaten, die man wiederholt anwenden kann, weil wir nicht die Gesamtästhetik jedes Jahr bei jedem Exponat neu denken müssen. In etwa wie ein Grundgerüst. Die Sachen, die dann oben drauf sind, das sind immer Neuentwicklungen. Zuletzt war das ein intelligentes Sicherheitssystem, da erfährt man, wie diese funktionieren. Da haben wir eine Installation gemacht, an der man mit seinen Händen um das Auto rumgehen kann und das Auto erkennt dann mit Infrarotsensoren wo die Hand ist. An der Stelle leuchten dann LED-Streifen unter dem Auto auf. Zusätzlich wird auf einem Monitor angezeigt, welches Radarsystem jetzt aktiv ist. Dann kann sich das der Besucher ein wenig spielerisch ertasten.

*Das heißt im Grunde genommen haben Sie dann, abgesehen von dem Licht, nur die Technologie des Autos verwendet, um zu zeigen was es kann?*

Ja, wir haben das mit Infrarotsensoren gemacht. Die sind so im Auto nicht drin, aber wir haben eben eine Adaption genutzt, die dem entspricht, wie es im Auto auch Verwendung findet.

*Warum haben Sie es mit Infrarotsensoren gelöst? Warum kein Kamerasystem oder ähnliches?*

Das war ein Platzproblem. Wir wollten das möglichst so machen, dass man die Sensoren gar nicht sieht, genau wie beim Auto. Diese Infrarotsensoren sind relativ klein und werden in eine kleine Haube gebaut, die infrarotdurchlässig ist, aber durch die man ansonsten nicht schauen kann. Das ist relativ elegant eingelassen und das war die Entscheidung, dass damit der Abstand besser gemessen werden kann. Wir waren letztendlich nicht ganz zufrieden mit der Lösung, aber haben auch keine bessere in der Zeit gefunden.

*Warum waren Sie nicht zufrieden?*

Weil die sehr fehleranfällig sind, diese Sensoren. Die haben eine sehr große Varianz in den Werten, die sie ausgeben. Das heißt, man muss sehr viel filtern und je mehr man filtert, desto höher ist die Trägheit von so einem System. Natürlich soll so schnell wie möglich alles reagieren. Aber wenn man erst mal 20 Werte speichern muss und dann den Mittelwert errechnen muss, um einen sauberen Wert rauszukriegen, dann ist das kontraproduktiv. Das war nicht ganz zu meiner Zufriedenheit. Aber wie gesagt, alle anderen Technologien... Also wir haben uns auch überlegt mit Kinect oder der Leap Motion was zu machen, aber bei der Leap Motion war das Problem, dass man nicht mehrere miteinander in einem System verbinden kann, dann hätten wir wieder mehr Rechner gebraucht.

*Wie haben Sie das System dem Besucher verständlich gemacht?*

Das war relativ schwierig. Wir haben kleine Bildschirme vor dem Fahrzeugmodell, auf denen eine Animation als Interaktionsaufforderung zu sehen ist. Also „Legen sie ihre Hände um das Fahrzeug“. Aber wir stellen schon fest, dass es einige Leute gibt, die das nicht direkt verstehen. Der Standard ist dann, dass sie hingehen, auf den Monitor drücken. Die lesen das gar nicht, sondern drücken erstmal. Und dann lesen sie und verstehen Hände und dann versuchen sie es mit dem Monitor (Lachen). Aber es gibt auch welche, die gehen direkt hin oder sehen, dass es jemand anderes macht und verstehen es dann direkt. Aber das ist sicher eine Herausforderung von solchen Sachen, die mit Gesten usw. arbeiten. Wie bringt man das dem Besucher bei, was er eigentlich machen muss ohne ein Schild aufzuhängen? So ca. 2010 war das noch so, dass man Hinweise auf Touchscreens aufbringen musste, damit die Leute verstehen: hier kann ich was drücken. Mittlerweile ist es eher so, dass man Hinweise auf normalen Videoscreens anbringen, dass die Leute hier nicht touchen können. Das hat sich in den paar Jahren dann doch stark geändert.

*Haben Sie sonst Exponate oder Installationen umgesetzt, bei denen Sie Tracking verwendet haben?*

Ja, eigentlich immer wieder. Wir haben relativ großflächig auf der IAA ein Tracking benutzt, das eine Eigenentwicklung von ein paar Leuten war, die für uns gearbeitet haben. Wir haben die ganze Festhalle getracked. Mit einer Spidercam, wie vom Fußball, wurde die Show gefilmt und dann wurden zusätzlich dazu Augmented Reality Inhalte angezeigt. Dazu haben wir vorher mit einem Laser die Festhalle vermessen und das in 3D übertragen. Damit konnten wir immer sagen wo die Spidercam ist und wenn man das auf ein 3D Modell überträgt, konnte man entsprechenden 3D Content anzeigen. Davor haben wir auch schon mit Kinects so zwei bis drei Sachen gemacht. Für die B-Klasse haben wir 2011 eine Selfie-Station gebaut, bei der man ein Selfie machen konnte und zusätzlich sind zum Beispiel Bs von oben runtergefallen. Dann konnte man damit spielen, wie die auf einen runterfallen oder konnte sie sammeln oder nur auf der einen Seite runterfallen lassen. Da hatten die Leute auch Spaß gehabt. Da war die Kinect auch noch relativ neu, da stand sie noch nicht in jedem Wohnzimmer. 2012 in Paris hatten wir auch eine Kinectanwendung. Da konnte man sich mit der Kinect kurz filmen lassen und dann wurde aus einem eine Art Punktwolke generiert und die ist dann in die Livebespielung der Bühnenshow mit rein geflossen. Das waren dann ca. drei Meter hohe Punktwolken, wo man sich selbst sehen konnte.

*Ich fand es immer spannend, wenn ein Besucher irgendwas nicht unbedingt merken muss, vielleicht auch nicht unbedingt versteht, was aber gleichzeitig eine dynamische Raumveränderung verursacht hat. Haben Sie so etwas einmal umgesetzt?*

Dass sich virtuell der Raum verändert hat?

*Nicht unbedingt virtuell. Das kann auch einfach eine Veränderung der Beleuchtung sein. Oder als banales Beispiel der Trigger eines Soundschnipsels.*

Weniger. Wir haben Projection Mapping Sachen gemacht. Natürlich arbeiten wir auf Messen schon viel mit Licht oder mit Veränderung von Raumstimmungen. Jetzt aber weniger auf Besucherinteraktion.

*Diese Veränderung der Raumstimmung ist dann vorgefertigt?*

Ja.

*Haben Sie so etwas umgesetzt, wie Veränderungen über – banales Beispiel – RSS Feed?*

So Echtzeitgeschichten quasi. Ich glaube im Messekontext nicht. Vielleicht haben wir sowas im Museumskontext eher schon mal gemacht, aber da fällt mir jetzt gerade kein konkretes Beispiel ein.

(DEMO)

*Wie gefällt Ihnen das Design?*

Ich finde es ansprechend. Es ist slick, ein bisschen technisch. Und ähnelt ja schon von der Einfachheit so ein bisschen vvvv. Das hat so ein wenig eine techy Optik. Für mich würde es jetzt passen.

*Inwiefern könnte ein Gestalter mit wenig Programmierkenntnissen das Programm benutzen?*

Gut, ich habe es natürlich nicht selbst ausprobiert. Aber erstmal erschien mir es von der Handhabung relativ nachvollziehbar oder einfach, ja. Es macht sicher Sinn so eine Art Tooltipmodus beim ersten Mal zu haben, damit man ein wenig navigiert wird. Man hat immer einen Rechner pro Kinect?

Ja.

(...)

Und hier unten sind, wenn man mehrere Räume anlegt, die Räume drin?

*Genau. Der Button hier unten hat im Moment noch keine Funktion. Da war natürlich die Idee, dass, wenn ich einen kompletten Raum abdecken will, die Kameras eng zusammenfügen und dann ein Point Cloud Stitching machen. Das ist das Ziel, mehrere Kameras zu einer großen Kamera kombinieren zu können.*

Und diese Triggerpunkte, die setzt man im Point Cloud Modus?

*Die kann man prinzipiell in jedem Modus setzen. Beim Einzeichnen in der Point Cloud hat man allerdings das Problem der Perspektive nicht.*

Ja, das ist immer das Problem bei 3D. Da braucht man schon das Drehen der Kamera um die Sachen gut positionieren zu können. Da wäre vielleicht eine Überlegung noch eher so eine Art Setup Modus zu haben, wo man die Sachen einstellen und kontrollieren kann und eine Art Edit Modus, wo man die Triggerpunkte setzen kann. Dass man das ein wenig unterteilt zwischen Einstellungsmodus und Triggermodus.

*Was genau meinen Sie mit Einstellungsmodus?*

Naja, das hier [Anm.: zeigt auf andere Streams] sind ja eher Analysemodi, oder? Es ging mir quasi darum, dass man für das Positionieren nur den Triggermodus braucht. Das macht in diesem 3D Modus am meisten Sinn. Vielleicht macht es auch von der Denkrichtung Sinn, dass man zuerst die Punkte setzt, bevor ich an andere Einstellungen gehe.

*Was für Anwendungsfälle können Sie sich denn für so ein Prinzip vorstellen?*

Ich kann mir absolut Anwendungsfälle vorstellen. Also gerade Projekte, wo es eher um einfachere Interaktion oder Fälle geht, wo man den Raum tracken will und Sachen auslösen will. Da kann ich mir das sehr gut vorstellen. Ideal wäre es natürlich, wenn dieses Frontend Services dazu anbieten würde. Also wenigstens noch eine Node in vvvv. Oder einen Frontend-Editor, wo ich zum Beispiel Sounds rein lade und triggern lasse oder Videosequenzen. Oder, was gibt es denn noch, mit Farbflächen oder sowas arbeitet. Oder ein kleines Layersystem, sodass man zwischen zwei Layern hin- und herwechseln kann. Bild eins ist das, und Bild zwei ist das und sobald jemand den Trigger aktiviert, togglet das sozusagen zwischen den beiden Layern hin und her. Klar, wahrscheinlich ist das am Ende schon so, dass man ein bisschen mehr Freiheit braucht und dann wird entweder das Frontendtool sehr aufwendig und vielseitig oder man sagt lieber: „OK, ich stell eine API für vvvv oder Processing zu Verfügung und die Programmierer entwickeln dann ihr eigenes Zeug“.

*Genau. Finden Sie denn sonst noch Funktionen vom Tool interessant?*

Naja, an sich halt, dass man das in so einem Frontend hat.

*Und das Tool im Allgemeinen? Finden Sie das interessant?*

Ja, schon. Also gerade in Museen ist das Budget immer eine Sache und da einfache, kostengünstige Lösungen zu haben, ist schon gut. Ansonsten sehe ich den Einsatz jetzt eher bei kleineren Sachen.

(...)

*Inwiefern haben Sie ID erhaltendes Tracking eingesetzt?*

Noch gar nicht. Aber wir denken immer wieder über solche Sachen nach. Bis jetzt gibt es aber noch keine gute Lösung dafür, außer dass man den Leuten eine Karte mit RFID in die Hand drückt und dann irgendwo einen Anmeldeprozess durchläuft.

*Sehen Sie bestimmte Trends im Ausstellungs- und Messebereich?*

Ja, schon so in die Individualisierungsrichtung. Das ist ein Thema, das wichtig wird in nächster Zeit. Dass die Leute, sei mit ihren Smartphones oder in einer anderen Art und Weise, ein personalisiertes Erlebnis bekommen. In welcher Form auch immer. Das ist momentan schon noch das Problem, dass man sich zum Beispiel die App runterladen muss, um an die Inhalte zu kommen. Aber wenn das ein wenig automatisierter geschehen könnte, ist das sicher ein Mehrwert für die Unternehmen und wenn es gut gemacht ist, auch für die Besucher. Also, dass ich zum Beispiel einen Plan vom Museum bekomme, wo ich im Voraus sage, was ich sehen will zum Beispiel. Das ist glaube ich etwas, womit man sich beschäftigen kann. Ansonsten erleben wir gerade eher so eine Art Plateau, was technologische Weiterentwicklung angeht. Vor sechs Jahren war es zuletzt das iPad, das so einen riesen Hype ausgelöst hat. Und dann war es ein wenig facebook und social media, aber das ist auch schon selbstverständlich geworden. Ansonsten waren technologisch gesehen die Kinect und die Oculus Rift die letzten technologisch interessanten Neuentwicklungen, die aber im Messekontext immer relativ schwer einzusetzen sind. Wir nennen es ja Begegnungskommunikation und wenn ich den Leuten eine Oculus Rift aufsetze, dann hat das nicht mehr sonderlich viel mit Kommunikation zu tun. Und der Kunde gibt ja auch relativ viel Geld dafür aus, um ein schönes Erlebnis in einer Messe oder Ausstellung zu schaffen. Und den Leuten dann in eine virtuelle Realität zu versetzen ist unserer Meinung nach eher kontraproduktiv, aber es gibt sicher auch Anwendungsfälle, wo es Sinn machen kann.

(...)

*Sie hatten bereits das Frontend als Verbesserungsmöglichkeit genannt. Fallen Ihnen sonst noch Verbesserungsmöglichkeiten ein?*

Da muss ich mich selbst erstmal in die Lage versetzen, wenn ich selbst etwas dafür entwickeln würde. Vielleicht könnte man noch einen kleinen Sound abspielen, wenn jemand eine Triggerzone betritt. Also gerade wenn man sich im Raum bewegt und nicht perfekten Blick auf den Monitor hat und ich versuche etwas zu aktivieren, dass ich dann wenigstens ein „Ding“ höre und dadurch Audiofeedback bekomme. Für das Frontend wäre noch eine Möglichkeit, dass ich Text eingeben kann. Also, dass ich Texte im Raum setzen kann und die

dann irgendwie projiziert werden können. Aber das ist natürlich weniger Backend. Eine Speicherfunktion wäre noch praktisch, dass ich also zwei Projekte gleichzeitig bearbeiten kann und die dann jeweils reinladen kann.

*Wenn Sie Exponate und Installationen umsetzen, sind bei Ihnen immer PCs in Verwendung? Also gerade bei Infrarotsensoren und Lichtsteuerungen könnte ich mir auch vorstellen, dass man das direkt auf einem Raspberry PI laufen lässt.*

Also normalerweise ist der Standard schon Windows PCs. Ansonsten hängt das davon ab, was wir für eine Technologie einsetzen. Diese Selfieanwendung beispielsweise, da haben wir es auch so gemacht, dass man die Bilder selbst auf so kleine Kreditkarten ausdrucken konnte. Das haben wir getrennt, dass die eigentliche Anwendung, auf dem man das Selfie macht, auf einem Rechner läuft und es einen zusätzlichen Printserver gibt, auf dem dann die ganze Druckgeschichte läuft. Dass, wenn das eine hängt, wenigstens der Rest noch funktioniert. Ansonsten haben wir die Infrarotgeschichte mit Arduinos umgesetzt. Ein anderes Projekt mit Controllino, das ist quasi die professionelle Form von Arduino.

*Wie verstecken Sie die Technik im Messebereich?*

So gut es geht. Das ist ein eigener Anspruch von uns. Gerade bei Premiumkunden ist es ein NoGo irgendwelche Kabel oder Kabelkanäle zu sehen. Da achten wir schon sehr drauf. Spaßeshalber bezeichnen wir unseren Beruf als Produktdesigner für Touchmonitor-Gehäuse. Das wird oft angefragt und dann müssen wir uns überlegen, wie wir ein möglichst schönes Gehäuse dafür bauen können. Da achten wir darauf, dass es möglichst schlank und slick ist. Was schon allein daran liegt, dass viele Leute mittlerweile zuhause sehr slicke Fernseher haben und wenn man dann auf die Messe geht und ein fettes Ding sieht mit riesen Rahmen, dann sieht das auch nicht zeitgemäß aus. Das versuchen wir schon so gut wie möglich zu lösen.

ENDE

## Interview mit Prof. Thomas Hundt

Passagen, die vornehmlich von allgemeinem und nicht fachspezifischem Charakter waren, wurden nicht transkribiert und sind mit (...), kurze Redepausen mit (..) gekennzeichnet.

Das Interview mit Prof. Thomas Hundt, Gründer und Geschäftsführer von jangled nerves wurde am 20.04.16 an der Hochschule für Technik Stuttgart geführt:

*Was ist Ihre Aufgabe bei jangled nerves?*

Ich bin einer von zwei Gründern und Geschäftsführern, wobei wir uns beide eigentlich eher als Kreativdirektoren verstehen. Die organisatorisch notwendigen Strukturen werden dann eher nach und nach durch Mitarbeiter gebildet. So schaffen wir uns die Freiheit, dass wir vor allem gestalterisch, kreativ tätig sein können. Es sind jetzt inzwischen achtzehn Jahre seit dem Diplom – da kommen viele Erfahrungen zusammen, sowohl im kulturellen Bereich, als auch in der Industrie, im Marketing und im Museum, die wir nun in den Projekten einbringen.

*Was waren denn die letzten Projekte, die Sie umgesetzt haben?*

Bei uns ist für die gesamte Entwicklung entscheidend, dass wir etwa zu gleichen Teilen in kulturellen Bereichen und für die Industrie arbeiten. Für die Industrie war die Großentwicklung der letzten Jahre, dass wir seit 2012 die Mercedes A-Messen realisieren – also die großen Messeauftritte unter anderem in Detroit, Las Vegas, Genf. Das größte Projekt, was wir in diesem Gesamtzusammenhang umgesetzt haben, war die IAA im Sommer 2015 für Mercedes auf 10000 m<sup>2</sup>. Parallel dazu laufen große Kulturprojekte, also beispielsweise das Stadtmuseum Stuttgart oder das Museum der Bayerischen Geschichte in Regensburg. Das würde ich auch als Hochpunkte der Kreativphase bezeichnen, bei dem zum Beispiel die gesamte Ausstellung im Detail aus der bayerischen Geschichte um Regensburg entwickelt wird. Dann gibt es natürlich viele Dinge dazwischen, wie Museen für einen Industriekunden, bei denen beide Welten zusammenfließen.

*Ich hatte bereits Interviews mit Atelier Brückner und MESO geführt. Das Atelier Brückner würde ich eher als klassisches Ausstellungsbüro bezeichnen. MESO hingegen ist ein Büro, was sehr technikorientiert ist. Wo passen Sie da rein?*

Wir versuchen da die Balance zu halten. Das war im Prinzip unsere Gründungsidee: Das Physische, Bauliche, das Digitale und Analoge als gleichwertig zu betrachten. Da sehen wir natürlich immer wieder andere kulturelle Prägungen in anderen Büros. Da wird dann doch oft das ein oder andere ausgeblendet. Tendenziell wird dann beispielsweise erstmal ein Raum gemacht und anschließend die Inhalte rein gefüllt. Oder man kommt aus einer digitalen Inszenierung und braucht noch den dazu passenden Raum. Wir versuchen hier eine Gleichzeitigkeit umzusetzen. Obwohl ich ausgebildeter Architekt bin, ist Architektur eben nicht die Königsdisziplin, die vor allem anderen kommt, sondern kann im nächsten Moment auch nur Medium sein und hat dann wieder eine andere Funktion.

*Würden Sie im Allgemeinen sagen, dass es nah wie vor oft so ist, dass die Architektur am Anfang steht und dann im Hintergrund die Medien aufgesattelt werden?*



Das hat zwei wesentliche Faktoren. Man kann natürlich seine eigene Prägung nur schwer ablegen – das ist ja auch gut. Wenn man eine Neigung zur Architektur hat, dann wird man natürlich auch eine gewisse Architekturprägung mit sich tragen. Da ist sicherlich der Vorteil bei uns, dass wir die ganze Schlussphase des Studiums bewusst abtrünnige Architekten waren. Das andere, was uns dann natürlich stark prägt, ist die Kultur und die Ausrichtung, die man in dem eigenen Büro etabliert. Die ersten Angestellten waren bei uns allesamt nicht Architekten, sondern kamen aus den Medien, aus der Technik, aus dem Film, aus dem Grafik- und Mediadesign. Wir haben also bewusst nicht erst das Architekturbüro um uns gebaut, sondern uns mit ganz anderen Themenfeldern provoziert. Ich würde behaupten, wenn man bei den Büros in die Verteilung der Professionen reinschaut, dann hat man auch ein gewisses Indiz dafür, wie gedacht wird. Wenn 70 Prozent Architekten sind, dann wird die Denkrichtung auch erstmal eine architektonische sein. Deshalb war es für uns von Anfang an sehr wichtig, bewusst heterogen aufgestellt zu sein.

*Ist es dann eher ein interdisziplinäres Arbeiten, bei dem die Spezialisten zu einem bestimmten Zeitpunkt ihre Arbeiten beginnen? Oder arbeitet man an einem Tisch und redet bei allen Entscheidungen mit?*

Das ist ganz wesentlich, dass es da keine Tabuzonen gibt. Das ist ja gerade das Spannende, dass jemand aus dem Design heraus eine architektonische oder ein Architekt eine filmische Idee hat.

*Die schlussendliche Realisierung – ist die dann oft in-house oder geschieht das über Externe?*

Wir merken eher, dass häufig die Entwicklung der spezifischen Software, des Storytellings, der Inhalte und das Aufbereiten der Inhalte viel wesentlicher ist, als die Interaktion oder auch die Möglichkeit alles permanent selber ändern zu können. Dann fällt natürlich die ganze hochkomplexe Funktionalität in so einer Software weit nach hinten. Deshalb setzen wir schon Tools von anderen Firmen ein, haben aber lange davor ein für den Kunden spezifisches Konzept entwickelt und greifen dann natürlich gerne auf verfügbare Technologien zurück. Wir machen also erstmal eine sehr spezifische, persönliche Entwicklung: „Was wollen die denn eigentlich erzählen? Was ist es wert zu erzählen? Wie könnte man das erzählen, was sie erzählen wollen?“ und dann folgt räumliches Setup – Hardware, Software – dieser Konzeptidee.

*Aber Prototypen setzen Sie auch in-house um?*

Ja. Also wir programmieren da schon selbst. Da gab es Phasen, in denen wir mehr gemacht haben und dann gab es auch Phasen, in denen wir mehr auf Tools zurückgegriffen haben, teilweise wirklich out-of-the-box, teilweise auch mit einfachen Anpassungen. In einigen Projekten entwickeln wir aufgesetzt auf bestehenden Tools Softwarelösungen, die dann spezifisch sind.

*Haben Sie Projekte umgesetzt, die Trackingtechnologien einsetzen?*

Ja, ein sehr großes Projekt läuft gerade in Wien für die Erste Bank. Das wird nun im Mai fertig und ist eine begehbbare, simulative Welt, in der die Bank und das Finanzwesen erlebbar

werden. Jeder Besucher bekommt ein digital Wallet, ein Pad, mit dem man individuell rumreisen kann und trotzdem immer wieder Dinge kollektiv getan werden. Zu viel Interaktion ist häufig ein Problem – es isoliert dann eher, anstatt Gruppen kollektiv zusammenzubringen. Dieses Isolierende „Ich hänge mich in mein Gerät und schaue immer auf mein kleines Device“ haben wir glaube ich alle genug zu Hause und an anderen Orten. Ich sehe das nicht pessimistisch, es ist einfach eine neue Kulturtechnik, die wir noch nicht ganz erlernt haben. Aber wir versuchen natürlich in Ausstellungen bewusst die Leute davon wegzubringen: Jeder hat zwar sein individuelles Device, indem ich mein Ergebnis habe und meine Dinge tun kann, aber idealerweise wird es so eingesetzt, dass man nicht flüchten kann. Das heißt, es kann nicht – wie ein Smartphone – immer alles, sondern nur etwas, was in der Situation im Raum gerade benötigt wird. Ich kann also vielleicht den Raum damit steuern oder der Raum hat eine Wirkung auf mein Device. Zusätzlich ist es auch gekoppelt mit den anderen Devices, die sich um dieses herum befinden. Es gibt also Stationen, bei denen vier Leute gegeneinander antreten müssen, wie bei Trimm-dich-Geräten – um zu zeigen, dass Geld arbeitet und um etwas zu vermehren, musst du auch mehr reinstecken. So wird das Device zu einem Display an einem Fahrradtrainer. Dann muss ein Besucher etwas machen, die anderen feuern ihn an, schließen Wetten ab, und plötzlich entsteht eine erlebbare Systematik, die die Presenter daraufhin wieder erklären können: Genauso funktioniert Börse, einer strampelt, der andere bemüht sich und kommt nicht vorwärts, einige wetten darauf, dass der eine gewinnt, andere, dass er verliert. So lässt sich das Ganze abstrahieren. In der Ausstellung gibt es jede Menge unterschiedlichste Trackingsysteme. Wir haben teilweise auch ganz einfache Dinge gemacht, wie große grafische Wände, bei denen nur die Kamera im iPad verwendet wird, um das Tracking zu machen und Orte zu erkennen, die der Besucher dann wieder auswählen kann. Wir haben immer wieder die Erkenntnis gewonnen, dass die einfachen Dinge extrem gut funktionieren. Kompliziertes Tracking und Richtungsbestimmung – das braucht es häufig gar nicht. Bei dem Projekt Zeche Zollverein hatten wir ein riesiges Gelände. Der Tracker weiß zwar per GPS ziemlich genau, wo die Person steht, nicht aber, in welche Richtung diese schaut. Nun gibt es auf dem Gelände aber Fördergerüste, Schloten, Gebäude, alles sieht anders aus und bietet wunderbare Landmarks. Wir haben einfach nur 3D-Panoramen installiert, der Besucher dreht sich dann von alleine, wenn er etwas auf dem Bildschirm entdeckt, bis er das Landmark gefunden hat. Oder er fährt durch das Panorama, bis er das gerade angepeilte findet. Da muss man gar nichts erklären. Am Anfang haben wir uns furchtbar damit abgemüht, wie wir die Blickrichtung hinbekommen. Dabei bekommen das die Besucher auch selber hin. Für uns ist das durchaus eine Grundhaltung: Wir spielen gerne mit technischen Sachen, versuchen es aber auch gerne bewusst vieles low-tech zu halten. Am Ende haben die Leute sonst auch keinen Spaß daran. Zusätzlich sind diese komplizierteren Dinge auch noch fehlerträchtig und wenn Fehler entstehen, macht es gar keinen Spaß mehr und die Leute lassen es liegen.

*Das ist ja auch oft das Problem bei Gestenerkennungen.*

Sicherlich. Gestikererkennung sieht man ja oft bei Ars Electronica. Da finde ich es noch ok. Aber die Sachen sind teilweise von der Interaktion hanebüchen. Sodass dann zwei Betreuer neben einem Ausstellungsbesucher stehen und ihm erklären müssen, wo er sich hinstellen muss. Das ist ja hochgradig anstrengend, wie etwa bei Photographie kurz nach dessen Erfindung, bei der man drei Minuten belichten muss und Leute Drahtklammern hinten an den

Kopf angebracht bekommen haben, damit diese nicht wackeln. Heute will ich einfach nur direkt ein Foto machen.

*Haben Sie sonst noch Tracking in Projekten umgesetzt?*

Das Stadtmuseum ist gerade im Werden. Das Museum ist am Ende eine Art artifizielle, verkleinerte Stadt mitten in der Stadt. Die Wechselwirkung zwischen realer Stadt und der Ausstellung soll auch über ein Guidesystem erfolgen. Das ist aber im Entstehen, da ist also vieles noch offen. Die Idee ist aber, dass man viele Dinge im Museum einsammeln und bewusst mit raus in die Stadt tragen kann. Zusätzlich besteht natürlich die Idee, möglichst viel über „Bring Your Own Device“ zu machen.

*Dieses Backpackingprinzip ist ja definitiv etwas, was sich häufiger sehen lässt in Ausstellungen. Sehen Sie sonst noch Trends in der aktuellen Ausstellungsgestaltung?*

Ich glaube schon, dass ein großes Diskussionsthema ist, was ich für Devices anbiete. In Wien haben wir uns beispielsweise aus gutem Grund dagegen entschieden, dass jeder sein Handy mitnehmen kann. Denn sonst flüchten die Besucher in ihre Devices. Die Zielgruppe ist dort eher zwischen Mitte Grundschule bis Ende weiterführende Schulen. Und das ist dann natürlich eine Zielgruppe, bei der Ablenkung gefährlich ist. Wir glauben, dass der Wunsch erst gar nicht aufkommt, da sie auf den Geräten, die sie bekommen, so viel machen können. Ansonsten ist die Tendenz eher, wie beispielsweise im Stadtmuseum, dass möglichst viel auf dem eigenen Device gemacht wird. Die Tendenz, das aus dem Museum zu verbannen und Störsender aufzubauen oder zu verhindern, dass Leute irgendwas fotografieren und weiter-schicken – das werden sich Museen die nach vorne wollen, nicht leisten können.

*Wie schwierig ist es, Leuten interaktive Elemente begreifbar zu machen? Und wie groß ist auch die Bereitschaft von Besuchern?*

Also Besuchern oder Museen? Denn ich habe den Eindruck, dass man mit den Besuchern gar kein Problem hat, sondern eher mit den Museen. Denn sehr häufig sind in Museen Entscheider, welche Berührungspunkte haben. Und diese – das gilt aber nicht nur für interaktive Stationen – unterschätzen den Besucher vollständig. Dann hört man sowas wie „Das versteht doch dann niemand“ oder „Das machen die Leute nicht“. Und gerade letzteres bedeutet für mich, dass ich den Leuten die Sicherheit geben muss, dass sie sich nicht zum Idioten machen und dann machen die Leute gerne alles. Wir sind ja alle verspielt und haben Lust, Dinge zu entdecken und auszuprobieren. Natürlich ist das ab einem gewissen Alter bis zu einem gewissen Grad „wegkultiviert“. Der eine macht es dann freiwilliger und der andere benötigt eine gewisse Sicherheit, um etwas zu tun. Aber da hatte ich noch nie den Eindruck, dass man von Leuten Dinge erwartet, von denen sie ganz genau wissen, dass sie es nicht schaffen. Wenn man diese Angst vor einem Fehlverhalten nimmt, dann sind alle bereit alles auszuprobieren und zu machen. Es ist auch häufig eine räumliche Frage. Für meinen Geschmack – da spricht natürlich auch der Architekt aus mir – hat ein Großteil der schlecht oder nicht interaktiven Stationen damit zu tun, dass sie räumlich falsch situiert sind. Ein Beispiel ist da für mich der Geldautomat. Denn ich stehe, wenn ich an mein eigenes Geld rangehe, mit dem Blick Richtung Wand und dem ungeschützten Rücken Richtung Öffentlichkeit. Da habe ich einen kompletten Kontrollverlust und fast alle Geldautomaten funktionieren so. Und so ist es

häufig das Problem, dass solche Medienstationen unabhängig von der räumlichen Situation entwickelt oder später dazu addiert werden von Leuten, die sich gar nicht mit dem Raum beschäftigen. Und dann stehen die nachher rum, wie Bankautomaten. Manchmal ist es wirklich nur das einfache Umdrehen, was einem die Souveränität gibt. So wird der Besucher auch zum Helden, zum Akteur in der Ausstellung. Für das Richard Wagner Museum haben wir ein riesiges Partitur-Blätterbuch gemacht. Und da ist es auch so: Ich komme in den Raum rein und schaue quasi in dieselbe Richtung wie das Orchester. Das Buch ist also erstmal weg von mir. Und das benötigte auch viel Überzeugungsarbeit. Für uns war klar: „Ich sehe das Dirigentenpult und das Buch und jetzt werde ich selber zum Dirigenten und habe vorher schon gesehen, dass dahinter der Raum zu Ende ist. Ich habe also den Rücken komplett frei. Dann liegt vor mir das Buch, ich blättere es um und dann wird irgendetwas ausgelöst. Da muss ich gar nicht viel machen und dann fangen auch schon einige fast an zu dirigieren. Wenn nun der nächste Besucher in den Raum kommt, dann erlebt dieser mich mit einem leichten Licht von oben, mit dem Glanz aus dem magischen Buch. Ich bin als derjenige, der da steht, der Held. Der andere sieht dann, wie ich damit blättere und kann vorsichtig schauen oder Signale geben, dass er das auch gerne mal probieren würde. Das hat auch den Effekt, dass ganz schnell Dialoge zwischen den Besuchenden entstehen. Und das hat ganz stark erstmal mit der räumlichen Inszenierung zu tun und nicht mit der Interaktion. Natürlich muss es dann mit der Interaktion weitergehen, denn, wenn ich keine spannenden Contents habe, wenn ich nicht an der richtigen Stelle interagieren kann, funktioniert es auch nicht.

*Das heißt derjenige, der mit dem Pult interagiert ist auch derjenige, der den Raum inszeniert?*

Genau, denn das andere, was in dem Raum drin ist, ist sehr guter Ton. Darum geht es ja auch an der Stelle.

*Haben Sie noch mehr Projekte umgesetzt, in denen der Besucher den Raum inszeniert hat, also beispielsweise eine Raumänderung ausgelöst hat?*

Ja, viel. Ein Beispiel dafür ist das Besucherzentrum im europäischen Parlament, das wir zusammen mit dem Atelier Brückner gemacht haben. Da war von Anfang an die Idee, bewusst dreiundzwanzigsprachig zu sein, also alle Sprachen Europas gleichwertig verfügbar zu machen. Und man kann dann natürlich auch zwischen den Sprachen wechseln. Denn selbst wenn ich vielleicht Deutsch hören möchte, kann es ja sein, dass ich eine Mittelrand-Rede im Originalton hören möchte. Dort gibt es immer wieder Räume, die überflutet sind, mit Texten und Zitaten aus allen Sprachen. Das heißt, ich sehe einfach nur babylonisches Chaos. Wenn ich den Raum betrete, dann weiß der Raum, dass ich bei mir deutsch eingestellt habe und dann rutschen alle anderen Sprache zur Seite und ich bekomme den Einführungstext auf Deutsch. Wenn ein Franzose dazukommt, dann setzt sich der französische Text daneben. Rein theoretisch, wenn Vertreter aller dreiundzwanzig Sprachen gleichzeitig in dem Raum wären, dann wäre es wieder genauso babylonisch wie am Anfang. Die Chance, dass alle dreiundzwanzig Nationen da drinstehen ist relativ gering. Und bei fünf oder sechs verschiedenen Sprachen – wo es auch schwer lesbar wäre – muss halt vielleicht einer raus oder die sollen sich so unterhalten. Das ist doch toll. Rein den Effekt wahrzunehmen löst ja auch etwas aus. Einerseits dient es vordergründig dem Raum. Tatsächlich macht es natürlich viel mehr mit den Menschen. Wenn ich drinstehe und meinen Text gerade auf Deutsch lese und

jemand anderes kommt rein und es kommt Französisch dazu, dann kann ich auch einfach mal „Bonjour“ sagen. Das heißt, ich kann das auch als Anknüpfungspunkt zur Kommunikation, zum Interagieren zwischen Menschen benutzen.

*Technisch funktioniert das über RFID?*

Ja. Das ist auch dadurch bedingt, dass man auf den einzelnen Geräten die Sprache wählt. Die Idee lieferte der Titel „Leave your ID and get a new ID.“ Man muss also seinen Ausweis als Pfand abgeben und bekommt dann seine europäische Identität. Da wurde bewusst gesagt: „Alle sind Europäer, es ist nur die Frage, in welcher Sprache du sprechen, bzw. hören möchtest.“ Dadurch funktioniert das Gerät in beide Richtungen. Mal hilft mir das Gerät nur zur Vertiefung, gleichzeitig ist das Gerät aber auch häufig ein Schlüssel, um eine interaktive Installation mit meiner Sprache zu versehen. Dazu muss ich mein Gerät in eine Schale in der Installation legen. Und dann gibt es den umgekehrten Fall, bei dem alle Sprachen einen Raum beeinflussen. So gibt es die unterschiedlichsten Möglichkeiten mit den beiden Richtungen zu spielen.

*Gab es auch Projekte, bei denen die Besucher nicht verstanden haben, dass sie etwas ausgelöst haben?*

Meistens ist die Angst davor – wenn man das so theoretisch durchspielt – unbegründet. Beim Richard Wagner Museum ist es auch so, dass während der Festspiele Horden durch die Ausstellung laufen. Da war uns schon klar, dass man bei diesen Besuchermassen nicht alles vollständig und mit allen Interaktionsmöglichkeiten wahrnehmen kann. Das wird so sein. Da muss man auch einfach manchmal Abstriche machen. Spannend ist es natürlich, wenn man in so einem Museum nicht nur eine, sondern unterschiedliche Sachen bietet. Dann erwischt man die Leute schon ganz gut. Wenn man mit einem Übererfüllungsanspruch – jeder muss alles erleben – an die Sache rangeht, dann wird man sowieso scheitern. Man erschrickt, wenn Museumswissenschaftler überprüfen, wie wenig Wissen tatsächlich aus Ausstellungen mitgenommen wird und die Leute trotzdem total glücklich und zufrieden sind. Also gerade bei den Berliner Museen ist das extrem. Da gehen natürlich viele Touristen rein und die gehen natürlich nicht, um über Nofretete viel zu erfahren, sondern deshalb, weil es zu einem touristischen Erlebnis gehört. Ob die danach viel gelernt oder gesehen haben – damit muss man manchmal auch lockerer umgehen. Das heißt natürlich nicht, dass man nicht Sorgfalt für die Inhalte walten lässt, aber andererseits darf man es auch nicht überbewerten und diesem Zwang verfallen, zu sagen: „Nein, die müssen das alles wissen und können“. Viele gehen auch einfach rein und wollen unterhalten werden und Spaß haben.

*Was haben Sie in Messen und Ausstellungen für technische Projekte umgesetzt?*

Das gilt jetzt sowohl für klassische Museen, als auch für die marketinggetriebenen Messen, die es für Industriekunden gibt. Unsere Erfahrung ist, dass es hilfreich ist, wenn man sich auf ein Prinzip konzentriert. Also da, wo das Museum bei einem zweistündigen Besuch wirklich auch davon profitiert, dass es unterschiedliche Dinge gibt und Tools einen auch überraschen und etwas anderes bietet, hat es sich für Ausstellungen häufig bewährt, mit einem durchgängigen Prinzip zu arbeiten. Zum Beispiel die erste Weltkriegsausstellung im Haus der Geschichte Baden-Württemberg – da haben wir über Pepper's Ghost Egodokumente einge-

spielt. Das kam auch direkt aus der Aufgabenstellung heraus. Das hat damit angefangen, dass wir – im Gegensatz zu vielen Ausstellungen – unglaublich viele Egodokumente zur Verfügung hatten. Briefe in die Heimat, Tagebuchaufzeichnungen, Verarbeitungen in der Literatur. Dann haben wir uns die Frage gestellt, was wir damit machen. Der eine hat zum Beispiel 1500 Briefe an seine Frau geschrieben. Wie will ich sowas ausstellen? Wie hole ich mir die wichtigen Sätze da raus? Wenn ich diese als Grafik präsentiere, dann verlieren sie komplett ihre Wirkung, das Mystische. Irgendwann kam dann die Idee auf, diese geisterhaft über den Exponaten schweben lassen. Damit kann auch die Original-Handschrift aus dem Dokument rausfliegen und in lesbare Schrift transkribiert werden. Wir alle hatten sofort den Eindruck, dass das ein Volltreffer ist. Das Tolle ist, dass die gesamte Ausstellungsgestaltung, die Architektur, die Form der Vitrinen, sich nur aus dieser Idee heraus entwickelt hat. Alles ist dieser Idee gefolgt. Bis hin zu der Frage der drei Orte: Ich brauche die Front, die Etappe und die Heimat, denn das waren die wesentlichen Orte, an denen diese Briefe geschrieben wurden. Da gibt es Briefe, die wirklich im Schützengraben geschrieben sind, bei denen dann fast schon ein Kanonenschlag visualisiert wird, weil es einen großen Tintenklecks gibt. Der Verfasser schreibt dann weiter, dass eben das Tintenfass umgefallen ist, weil eine Bombe explodiert ist. Und dann habe ich in der Etappe diese unendliche Langeweile. Tausend Briefe – jeden Tag werden ein bis zwei Briefe geschrieben, weil nichts passiert, über Monate hinweg. In der Heimat sind es Schreiben, in denen klar gemacht wird, dass der Mann oder Sohn gefallen ist. Das heißt auch diese drei Orte hatten mit dieser einfachen medialen Inszenierung zu tun, die aber viel wirkmächtiger war, als man denkt. Hier war gerade der Entzug der Interaktion das Tolle. Dass ich nicht schnell durchblättern kann, sondern zum Leserhythmus gezwungen werde. Die Direktorin sagte auch, dass noch nie so viel in einer Ausstellung gelesen wurde.

Bei Industrieausstellungen ist es häufig so, dass man stärker von Dingen motiviert ist, die mit dem Produkt zu tun haben. Mit Mercedes haben wir zum Beispiel eine kleine App gemacht, mit der man Inhalte vertiefen konnte. Da konnte man zum Beispiel mit dem SLR Pedal spielen.

*Wie lange waren denn diese Abfolge von Texten in der Weltkriegsausstellung?*

Ich denke länger als drei bis vier Minuten war nichts. Einzelne Dokumente, bei denen es viel zu erzählen gab, da haben wir uns auch bewusst dafür entschieden, längere Passagen zu zeigen.

(DEMO)

*Wie gefällt Ihnen das Interface?*

Also es wirkt jetzt erstmal aufgeräumt und erweckt den Eindruck, dass sowohl ich, als auch Mitarbeiter, die ein bisschen mehr Ahnung haben, damit loslegen können. Wie viele Menschen im Museum sich an sowas heranwagen ist fragwürdig. Aber auch das wird sich ändern mit der Zeit. Ich meine, wenn man noch vor ein paar Jahren Leuten erzählt hätte, dass diese ihre Apps über „If This Then That“ miteinander verbinden können, dann hätten diese auch gesagt, dass sie nicht programmieren können und wollen. Aber solche Tools haben ja nichts mit programmieren zu tun, sondern mit normalen Abhängigkeiten, wie sie jeder Mensch

kennt und verwendet. Daher kann ich mir auch vorstellen, dass es mit der Zeit in Museen junge Kuratoren und Mitarbeiter gibt, die sich an sowas herantrauen. Deshalb denke ich, dass es von der Oberfläche gut überschaubar ist. Da gilt natürlich: Draufschauen ist das eine, Bedienen ist nochmal eine andere Sache. Die Logik dahinter finde ich allerdings direkt nachvollziehbar. Es kann natürlich sein, dass der Grafiker das gut findet, der Programmierer hingegen das gar nicht braucht. Wobei auch da ja viele Programmierer mittlerweile sehr gerne mit grafischen Oberflächen und Tools jenseits des Quellcodes arbeiten.

*Können Sie sich Anwendungsfälle für das Triggerzoneprinzip vorstellen oder kennen Sie Projekte bei denen es hätte eingesetzt werden können?*

Also häufig haben wir Anwendungsfälle, bei denen wir einen Museumsraum haben, in den gelegentlich jemand reinkommt. Und dann starte ich eine Abfolge von Dingen. Das kann ich natürlich mit einfachsten Bewegungsmeldern schon hinbekommen. Da bräuchte ich die Logik nicht. Aber prinzipiell ist das auf jeden Fall interessant, dieses „mit dem Körper was auslösen“. Da bin ich aber immer ein wenig vorsichtig, denn sobald mehrere Menschen drin sind, geht es doch ganz schnell Richtung Chaos. Und häufig ist es natürlich so, dass Menschen gemeinsam in eine Ausstellung gehen. Dann ist es furchtbar anstrengend, wenn ich einer Schulklasse sagen muss, dass sie sich verteilen sollen oder dass nur eine bestimmte Zahl damit interagieren darf. Das war für uns häufig ein Grund, dass wir auf solche Dinge nicht eingegangen sind. Wo ich es mir hingegen sehr gut vorstellen kann, sind Industrieanwendungen. Also wenn ich mir etwas konfiguriere, etwas zusammenstelle. Ich könnte dann einen Katalog um mich herum haben, bei denen ich bestimmte Dinge zusammenziehe und mir anschauen kann oder wechsle die Farbe eines Autos, das ich mir gerade anschau. Das über einfache Gesten steuern zu können ist sehr interessant. Also interessant finde ich es, ich habe aber keinen direkten Anwendungsfall. Es ist ja auch mehr so rum, dass die thematische Ausrichtung zuerst kommt. Dementsprechend wäre es natürlich interessant, wenn dann das Ausprobieren, Programmieren und Simulieren von solchen Umgebungen einfacher und schneller wird.

*Wir sprachen ja bereits ein paar Mal über ID-erhaltende Technologien. Wie wichtig ist das im Bereich Kameratracking?*

Das ist bei mir wahrscheinlich eine Mischung aus eigenem Empfinden und Erfahrung aus Ausstellungen. Was ich unglaublich wichtig finde, ist ein schnelles und präzises Feedback. Denn sonst habe ich immer den Eindruck, dass die Leute den Spaß verlieren. Darum bin ich immer vorsichtig bei Installationen, die mehrere Personen gleichzeitig unterstützen. Ich kenne auch viele Installationen, die mit großem Aufwand versuchen festzustellen, wo, wann, wie Leute durchgegangen sind. Dann wird irgendwas Kaleidoskopisches in den Raum rein-geleuchtet und dir gesagt, dass das die Bewegungsmuster sind, die du ausgelöst hast. Aber wenn ich da mit irgendwelchen Zufallszahlen in After Effects was raus rendern würde und einfach behaupte, dass das aus den Bewegungsdaten generiert wurde, dann sind alle genauso glücklich. Umgekehrt haben wir auch für Mercedes schon eine Ausstellung gemacht, in der vier freifahrende Roboter rumgefahren sind und innovative Technikfilme gezeigt haben. Die hatten definitiv eine feste Choreographie. Aber 99 Prozent der Besucher haben Stein und Bein geschworen, dass die Roboter auf sie reagiert haben. Daher würde ich als Beispiel bei einer Fassadenbespielung sagen: lieber bekomme ich kurz ein Zeitfenster, um zu interagieren

und etwas zu zeichnen und den sehe ich direkt und weiß auch, dass ich das war. Oder ich wähle Orange als Farbe und dann wird alles Orange, sobald ich an der Reihe bin. Das macht dann auch Spaß. Wenn ich aber sage, dass durch eine Gaußsche Normalverteilung aller Farbwünsche aller Leute auf diesem Platz die Software die ganze Zeit Braun zeigt, dann ist das eher uninteressant. In dem Moment, in dem ich etwas Interaktives mache, möchte ich auch einen Moment haben, in dem ich sehe, dass ich etwas beeinflusst habe. Um es zusammenzufassen: Ich finde es interessant und eigentlich sollten dann bei mir im Büro noch andere Leute darüber schauen, um einzuschätzen, ob wir das einsetzen können, ob es Spaß macht oder uns was bringt. Ich kann mir aber durchaus vorstellen, dass das für andere Büros, die jetzt nicht diesen technischen Background haben, sehr interessant sein kann. Die Gefahr ist natürlich immer die, dass – in dem Moment, in dem etwas dann so vermeintlich einfach wird – die Chance, dass dadurch andere Fehler gemacht werden, riesengroß wird. Da wird zum Beispiel der falsche Beamer reingehängt, der nach zwei Wochen in die Knie geht, weil der nicht für den Ausstellungsbetrieb geeignet ist.

ENDE



## USB-Stick

Auf dem beiliegenden USB-Stick finden sich folgende Ordner und Dateien:

- `Abbildungen_Originale`: Die Originale der Point Cloud Abbildungen.
- `ArgusVision_Code`: Der Quellcode von Argus Control und Argus Kinect.
- `ArgusVision_Jar`: Ausführbare Versionen von Argus Control und Argus Kinect.
- `Präsentation_Interviews`: Die Präsentation, welche im Rahmen der Interviews gehalten wurde als Keynote- und PDF-Datei.
- `Schriftliche_Arbeiten`: Seminar- und Projektarbeit sowie Bachelorthesis als PDF-Dateien.
- `smartPerform_Fallstudie`: Der vvvv Patch und die smartPerform Präsentation, welche in Fallstudie 5.2 verwendet wurden.
- `Video`: Ein Präsentationsvideo, welches die Funktionen von Argus Vision zeigt.

