

TUM

INSTITUT FÜR INFORMATIK

Avilus Tracking Contest 2010

16.-27. August 2010

Abschlussbericht

Dr. Bjoern Schwerdtfeger

Prof. Gudrun Klinker (Ph.D.)



TUM-I1102

Januar 11

TECHNISCHE UNIVERSITÄT MÜNCHEN

TUM-INFO-01-I1102-0/1.-FI

Alle Rechte vorbehalten

Nachdruck auch auszugsweise verboten

©2011

Druck: Institut für Informatik der
 Technischen Universität München

AVILUS Tracking Contest 2010

16. - 27. August 2010

Abschlussbericht

Björn Schwerdtfeger, Gudrun Klinker
Fachgebiet Augmented Reality, Technische Universität München

21. Dezember 2010

Technologiepartner:

- Fraunhofer Institut für Graphische Datenverarbeitung
(Harald Wüst, Folker Wientapper, Mario Becker)
- Metaio GmbH
(Selim Benhimane, Quintus Stiersdorfer, Georg Kuschik)
- Fachgebiet Augmented Reality, Technische Universität München
(Peter Keitler, Christian Waechter, Mahmood Bahaa, Gudrun Klinker)

Anwendungspartner und Szenariensteller:

- Volkswagen AG
(Fabian Doil, Ulrich Rautenberg)
- EADS Innovation Works
(Benjamin Becker, Holger Schmidt)
- Howaldtswerke-Deutsche Werft GmbH (HDW)
(Michael Riedel)
- Kuka Roboter GmbH
(Marinus Danzer)

Koordinator des Tracking Contests:

- Fachgebiet Augmented Reality, Technische Universität München
(Björn Schwerdtfeger, Gudrun Klinker)

Inhaltsverzeichnis

| | | |
|----------|---|-----------|
| 1 | AVILUS Tracking Contest | 4 |
| 2 | Teilnehmer und Technologien | 6 |
| 2.1 | Fraunhofer IGD | 6 |
| 2.1.1 | Fokus und Hintergrund der Technologie | 6 |
| 2.1.2 | Workflow während der Vorbereitung | 7 |
| 2.2 | metaio GmbH | 9 |
| 2.3 | Fachgebiet Augmented Reality (FAR) | 11 |
| 2.3.1 | Fokus und Hintergrund der Technologie | 11 |
| 2.3.2 | Workflow während der Vorbereitung | 13 |
| 3 | Testszenario: Volkswagen | 15 |
| 3.1 | VW Passat Exterieur | 15 |
| 3.1.1 | Versuchsszenario | 15 |
| 3.1.2 | Messergebnisse | 16 |
| 3.1.3 | Diskussion der Technologiepartner | 17 |
| 3.2 | VW Passat Motorraum | 18 |
| 3.2.1 | Versuchsszenario | 18 |
| 3.2.2 | Messergebnisse | 18 |
| 3.2.3 | Diskussion der Technologiepartner | 19 |
| 3.3 | VW Passat Interieur | 22 |
| 3.3.1 | Versuchsszenario | 22 |
| 3.3.2 | Messergebnisse | 23 |
| 3.3.3 | Diskussion der Technologiepartner | 25 |
| 3.4 | VW Genauigkeitsmessung (Schiene) | 27 |
| 3.4.1 | Versuchsszenario | 27 |
| 3.4.2 | Messergebnisse | 28 |
| 3.4.3 | Diskussion der Technologiepartner | 28 |
| 3.5 | Bewegtes VW Modellauto auf einem Kuka Roboter | 30 |
| 3.5.1 | Versuchsszenario | 31 |
| 3.5.2 | Messergebnisse | 31 |
| 3.5.3 | Diskussion der Technologiepartner | 32 |
| 3.6 | Diskussion des Anwenders | 33 |
| 4 | Testszenario: Tracking am Kuka Roboter | 36 |
| 4.1 | Versuchsszenario | 36 |
| 4.2 | Messergebnisse | 36 |
| 4.3 | Diskussion der Technologiepartner | 36 |
| 4.4 | Diskussion des Anwenders | 39 |
| 5 | Testszenario: Soll-Ist-Vergleich (HDW) | 41 |
| 5.1 | Versuchsszenario | 41 |
| 5.2 | Messergebnisse | 42 |
| 5.3 | Diskussion der Technologiepartner | 44 |

| | | |
|----------|---|-----------|
| 6 | Testszenario: EADS Flugzeugkabine | 46 |
| 6.1 | Versuchsszenarien | 46 |
| 6.1.1 | Versuchsszenario 1: AR Überlagerung eines 3D-Modells . . . | 46 |
| 6.1.2 | Versuchsszenario 2: Finden von 3D-Punkten | 47 |
| 6.1.3 | Versuchsszenario 3: Relative Genauigkeitsmessung | 47 |
| 6.2 | Messergebnisse | 47 |
| 6.2.1 | Ergebnisse FAR | 47 |
| 6.2.2 | Ergebnisse IGD | 49 |
| 6.2.3 | Ergebnisse metaio | 51 |
| 6.3 | Vergleich Sonderaufgaben: Wide-area Genauigkeit | 52 |
| 6.4 | Vergleich relative Genauigkeitsmessung | 53 |
| 6.5 | Diskussion der Technologiepartner | 53 |
| 6.6 | Diskussion des Anwenders | 56 |
| 7 | Testszenario: EADS Wellness-Bereich | 58 |
| 7.1 | Diskussion der Technologiepartner | 58 |
| 7.2 | Diskussion des Anwenders | 60 |
| 8 | Ergebnisbewertung, Potentiale und Forschungsbedarfe | 62 |
| 8.1 | Ergebnisbewertung | 62 |
| 8.2 | Einordnung der Ergebnisse auf der Technologie Roadmap | 63 |
| 8.3 | Forschungsbedarf | 64 |
| 9 | Resümee | 66 |

1 AVILUS Tracking Contest

Vom 16.-27. August 2010 wurde der AVILUS Tracking Contest bei EADS in Ottonbrunn durchgeführt. Die AVILUS-Partner VW, Kuka, HDW und EADS stellten dabei Szenarien zur Verfügung, an denen die Technologien von Fraunhofer IGD, metaio GmbH und Fachgebiet Augmented Reality (FAR) evaluiert wurden. Die Technologiepartner werden im Folgenden nur noch IGD, metaio und FAR genannt.

Dieser bisher umfangreichste Tracking Contest wurde aufgrund mehrerer Motivationen durchgeführt. Erstes Ziel war die Evaluation der Tracking Technologien, welche im Rahmen von AVILUS und AVILUS Plus entwickelt wurden. Weiteres Ziel war es, die Entwickler der Trackingtechnologien innerhalb eines konzentrierten Zeitraumes mit verschiedenen realistischen Szenarien der Anwender zu konfrontieren. Aus dieser Perspektive handelte es sich bei dem Contest also nicht nur um einen Wettbewerb im eigentlichen Sinne, sondern auch um eine Art "Entwickler Camp". Dieses führte letztendlich zu diversen Einsichten, Erfahrungen und technologischen Weiterentwicklungen bei den Technologiepartnern. Zudem konnten die beteiligten Anwender direkt sehen, wie gut die Technologien an ihren Szenarien funktionieren.

Der Fokus des Tracking Contests lag auf markerlosen Verfahren, also auf solchen Verfahren, die weitestgehend ohne eine Instrumentierung der Umgebung auskommen. Die weit verbreiteten Papiermarker, welche auf Basis eines schwarzen Rechtecks und eines Identifizierungscodes arbeiten, sollten in diesem Contest nur in Ausnahmefällen Anwendung finden. Auf vielen Abbildungen in diesem Dokument sind schwarz/weiße Rundmarken zu sehen. Diese dienten der Registrierung der Trackingsysteme zum Objekt. Während des eigentlichen Trackings waren diese Marken verdeckt.

Übersicht

Im Kapitel 2 werden zunächst die eingesetzten Trackingtechnologien der drei teilnehmenden Einrichtungen präsentiert. Danach werden die einzelnen Testszenarien samt Durchführung und Ergebnissen vorgestellt. Die Ergebnisse jedes Szenarios werden sowohl von den Technologieanwendern als auch von den Technologiepartnern diskutiert. Folgende Testszenarien waren Bestandteil des Tracking Contests:

- Kapitel 3: Diverse Szenarien von Volkswagen
 - am Exterieur
 - im Motorraum
 - im Interieur
 - Genauigkeitsmessungen mit einer Linearschiene
 - Tracking eines kontrolliert bewegten Miniaturautos
- Kapitel 4: Tracking am Kuka Roboter
- Kapitel 5: Soll-Ist-Vergleich bei HDW
- Kapitel 6: Diverse Szenarien im Mockup einer Flugzeugkabine bei EADS

- Überlagerung eines 3D Modells des Flugzeugs
 - Finden von 3D Punkten im Flugzeug
 - Relative Genauigkeitsmessung mit einer Linearschiene
 - Sonderaufgabe wide-area Genauigkeit
- Kapitel 7: EADS Wellnessbereich

In Kapitel 8 bewerten Technologiepartner die Ergebnisse des gesamten Tracking Contests und diskutieren mögliche Potentiale und Forschungsbedarfe.

2 Teilnehmer und Technologien

Zur Einführung wurden die drei Technologiepartner IGD, metaio und FAR gebeten, ihre Technologien zu beschreiben.

2.1 Fraunhofer IGD

2.1.1 Fokus und Hintergrund der Technologie

Für die Entwicklung des Trackingansatzes vom IGD war die Annahme entscheidend, dass sich der Einsatz von AR im industriellen Kontext insbesondere dann rentiert, wenn sich Skaleneffekte einstellen. Im Bereich von Montage-, Reparatur- und Fehlerdiagnosarbeiten ist dies der Fall, wenn einerseits ein und dieselbe AR-Applikation in hohem Maße mehrfach benutzt werden kann, und andererseits die Erstellung des Trackingmodells hierfür nur einmal durchgeführt werden muss. Mit der Häufigkeit der Anwendung amortisiert sich dann der Aufwand für das einmalige Einrichten der AR-Applikation.

Weiterhin ist es für die Akzeptanz von AR-Applikationen vorteilhaft, wenn beim Endanwender keine besonderen Kenntnisse hinsichtlich des Trackings bzw. der Benutzung der AR-Applikation vorausgesetzt werden müssen. Die AR-Applikation sollte den Endanwender bei der Erreichung seiner persönlichen Arbeitsziele unterstützen, ihm jedoch nicht ein zu hohes Maß an zusätzlichem Vorwissen in der Handhabung abverlangen. Seine Aufmerksamkeit sollte nicht darauf ausgerichtet sein, wie er sich verhalten muss, damit das Tracking funktioniert. Vielmehr sollte er durch AR schneller, einfacher und intuitiver an die benötigten Informationen gelangen.

Vor diesem Hintergrund kann beispielweise ein rein auf "Extensible Tracking" [3, 14, 6, 5] basierender Ansatz in zweifacher Hinsicht nachteilig sein. Neben den vom Endanwender durchzuführenden initialen Vorbereitungsschritten unmittelbar vor Beginn des Trackings (z.B. Platzierung eines Markers an einer definierten Stelle) - welche im Übrigen selbst mögliche Fehlerquellen darstellen können - spielt die Einschränkung der zulässigen Kamerabewegungen eine Rolle. Es muss nämlich stets genügend Translation in der Bewegung enthalten sein, damit die Umgebung rekonstruiert werden kann. Überwiegend rotatorische und ruckartige Bewegungen des Endanwenders führen zu Fehlern in der Rekonstruktion und somit fehlplatzierten Augmentierungen oder gar zum Aussetzen des Trackings.

Für die Entwicklung der Trackingtechnologie des IGD war daher der Gedanke maßgeblich, dass eine Vielzahl an Verarbeitungsschritten in eine Vorbereitungsphase ausgelagert werden kann, um damit wiederum während der Anwendungsphase ein schnelleres, robusteres und / oder genaueres Tracking zu ermöglichen. Ziel der Vorbereitungsphase ist dabei ein Trackingmodell, welches auf einer möglichst genauen Rekonstruktion der Szene basiert und dabei bereits den Großteil aller für das Tracking benötigten Informationen enthält bzw. alle während der Laufzeit zu erwartenden Blickwinkel abdeckt.

Aus technologischer Sicht eröffnet eine dedizierte (nur einmalig durchzuführende) Vorbereitungsphase den Spielraum für den Einsatz spezieller, offline-basierter Algorithmen. Im Speziellen werden beispielsweise maschinelle Lernverfahren eingesetzt, welche (im Gegensatz zu Deskriptor-basierten Matchingansätzen wie SIFT [21], SURF

[2], oder HOG [4] oder andere [19]) eine Featureerkennung mit breiter Baseline durch simultane Berücksichtigung mehrerer Ansichten aus der Vorbereitungsphase ermöglichen (Randomized Trees). Weiterhin wird das Bild-zu-Bild Tracking (narrow Baseline) über einen invers-kompositionalen, affinen KLT-Algorithmus realisiert (Optischer Fluss / Gauss-Newton basierte Referenzpatchverfolgung) [22, 1, 9, 27], welcher im Gegensatz zu Feature-basierten wide-baseline Matchingverfahren die Lokalisierung der 2D-Bildpunkte im Subpixelbereich ermöglicht. Um die Performanz des KLT-Verfahrens zu erhöhen, werden während des Trackings die zu verfolgenden Referenzpatches blickwinkelabhängig selektiert (Feature Management) [32]. Das Erlernen der blickwinkelabhängigen Sichtbarkeit der KLT-Patches ist ebenfalls Bestandteil der Vorbereitungsphase. Schließlich wird im Rahmen der Vorbereitung die Rekonstruktion der Feature Map durch (offline-basiertes) Bundle Adjustment [20, 7, 28] verfeinert. Online Rekonstruktionsverfahren (extensible Tracking) werden zwar zum Aufbau der Feature Map und zur Stabilisierung des Online Trackings eingesetzt, jedoch erlangen sie im Allgemeinen ihre Echtzeitfähigkeit nur durch Abstriche in der Genauigkeit.

Die Erstellung des Trackingmodells im Rahmen der Vorbereitung wird nicht als Aufgabe eines unbedarften Endanwenders angesehen, sondern eher als die eines geübten, mit der Trackinganwendung vertrauten Benutzers.

2.1.2 Workflow während der Vorbereitung

Rekonstruktion: Der Nutzer rekonstruiert die Feature Map, indem er eine Sequenz durch Kameraführung um das Objekt aufnimmt. Dabei werden Featurepunkte (KLT) automatisch initialisiert und grafisch in Echtzeit im Videobild angezeigt. Dadurch erhält der Nutzer während der Aufnahme ein Feedback darüber, in welchen Bereichen die Featurepunkte noch gering ist. Die initiale Rekonstruktion erfolgt automatisch über die verfolgten Featurepunkte mittels Epipolar-Geometrie [10]. Hierzu muss kein Marker zuvor in der Szene platziert werden. Weiterhin ist (abgesehen von der Kameraführung) keine Nutzerinteraktion erforderlich, d.h. die initiale Rekonstruktion erfolgt automatisch, sobald erkannt wurde, dass hinreichend Translation zwischen dem ersten und dem zuletzt aufgenommenen Bild vorhanden ist. Der Nutzer muss demnach nur darauf achten, dass er zur Rekonstruktionsinitialisierung eine translatorische Kamerabewegung und nicht nur Drehbewegungen ausführt. Die Erweiterung der Feature Map erfolgt über Online-Rekonstruktionsverfahren (SLAM / Extensible Tracking) [3], wobei eine "Ping-Pong" Strategie angewandt wird: Mit bereits rekonstruierten Features wird die Pose berechnet / mit bekannter Pose wiederum werden neue Features initial rekonstruiert und mittels Extended-Kalman-Filter-Methoden verfeinert. Währenddessen wird die Bildsequenz auf dem Rechner gespeichert.

Transformation: Da die vorherige Rekonstruktion ohne Vorinformationen (beispielsweise ohne Marker) durchgeführt wurde, befindet sich die Feature Map noch in einem willkürlich definierten Koordinatensystem. Der Abgleich (Transformation) zwischen diesem und dem Koordinatensystem, in welchem die Augmentierungen definiert sind, erfolgt interaktiv über manuelle Festlegung von 3D-3D Punktkorrespondenzen [29, 11]. Hierzu kann der Nutzer über einen Schieberegler durch die aufgenommene Sequenz navigieren und bekommt zu jedem Bild die 2D-Featurepunkte ange-

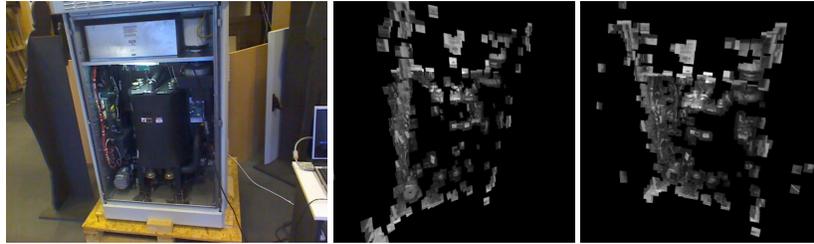


Abbildung 1: Rekonstruktion einer Feature Map aus einer gegebenen Bildsequenz

zeigt. Indem der Nutzer auf einen der angezeigten 2D-Punkte im Bild klickt, selektiert er implizit den jeweiligen 3D-Punkt im Koordinatensystem der rekonstruierten Map. Zusätzlich kann der Nutzer in diesem Koordinatensystem auch spezielle Punkte durch Anklicken manuell rekonstruieren, z.B. spezielle Referenz- oder Vermessungspunkte, sofern diese nicht bereits in der Map enthalten sind. Die Zuweisung der rekonstruierten 3D-Punkte zu den 3D-Punkten im (globalen) Koordinatensystem der Augmentierungen kann durch Anklicken eines 3D-Punktes auf der Oberfläche eines CAD-Modells oder durch manuelle Koordinateneingabe in einem Textfeld erfolgen. Mit Hilfe von mindestens drei nicht kollinear im Raum angeordneten 3D-3D Korrespondenzen kann so eine initiale Transformation (Translation, Rotation und Skalierung) berechnet werden. Abschließend wird die Rekonstruktion mit Hilfe eines Bundle Adjustments optimiert [28, 7, 20]. Hierbei fließen die definierten 3D-3D Korrespondenzen als Nebenbedingung in die Minimierungsprozedur ein, wodurch insbesondere nicht-rigide Deformationen der Feature Map aufgrund von Ausreißern (Outliers) oder Ungenauigkeiten in der Kamerakalibrierung kompensiert werden können. Mehr als drei Korrespondenzen erhöhen im Allgemeinen die Genauigkeit der Transformation.

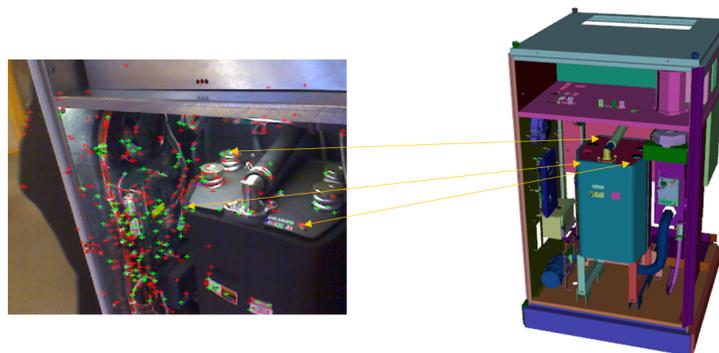


Abbildung 2: Manuelle Erstellung von 3D-3D Korrespondenzen mit Hilfe eines CAD-Modells

Sichtbarkeitsanalyse und Featureverwaltung: Um die Performanz des KLT-basierten Trackings zu erhöhen, wird eine statistische Auswertung über die Sichtbarkeit der Features in Abhängigkeit vom Blickwinkel durchgeführt. Während des Trackings können somit adaptiv die ressourcen-aufwändigen, bild-basierten Operationen auf diejenigen Features mit der höchsten Wahrscheinlichkeit hinsichtlich Wiedererkennbarkeit konzentriert werden (Featureverwaltung) [32]. Während der Vorbereitungsphase kann der Nutzer zur Auswertung die gleiche Sequenz verwenden, die bereits zur Rekonstruktion benutzt wurde, oder alternativ eine neue aufnehmen. Letzteres ist insofern von Vorteil, als der Aufwand für die Sichtbarkeitsanalyse nur linear mit der Länge der Sequenz ansteigt, wohingegen die Rekonstruktion (Bundle Adjustment) kubische Komplexität besitzt. D.h.: für die Sichtbarkeitsanalyse können im Allgemeinen deutlich mehr Blickwinkel herangezogen werden.

Feature Learning für Wide-Baseline Matching: Für die Wide-Baseline Initialisierung des Trackings wird das Verfahren von Lepetit und Fua (Randomized Trees / Ferns) herangezogen [19]. Zum Erlernen des Klassifikators werden jedoch nicht - wie im Originalansatz - texturierte CAD-Modelle oder projektiv verzerrte Planaransichten der Szene verwendet, sondern die zuvor aufgenommenen Realbildsequenzen. Auch hier bietet sich aufgrund der linearen Komplexität des Lernverfahrens die Möglichkeit, längere Sequenzen zu verwenden.

| Vorbereitungsschritt | Zeit |
|---|---|
| Initiale Rekonstruktion + Sequenzaufnahme | 2-4 min |
| Transformation | 1-2 min |
| Bundle Adjustment | 1-15 min (je nach Sequenzlänge: 200-500 Frames) |
| Sequenzaufnahme für Sichtbarkeitsanalyse + Featurelearning (optional) | 4-7 min |
| Sichtbarkeitsanalyse | < 1 min |
| Feature-Learning | 3-5 min |
| Gesamt | 12-35 min |

Tabelle 1: Aufschlüsselung der Vorbereitungsdauer (je Feature Map)

2.2 metaio GmbH

Beim von metaio eingesetzten Trackingsystem handelt es sich um eine Entwicklung, welche mit der Absicht der Kommerzialisierung vorangetrieben wird. Aus diesem Grund können leider nicht alle aus wissenschaftlicher Sicht interessanten Details beschrieben werden, jedoch kann die grundlegende Funktionsweise einzelner Schritte erklärt werden, um dennoch einige Charakteristika des Systems zu verstehen – auch im Hinblick auf andere im Tracking Contest eingesetzte Verfahren. Zu erwähnen ist, dass sämtliche während des Tracking Contests verwendeten Funktionalitäten in die Softwareplattform Unifeye integriert und für den Anwender verfügbar sind. Während des Tracking Con-

tests wurde keine Zeile Code geschrieben, sondern ausschließlich die in Unifeye vorhandene Funktionalität genutzt.

Desweiteren wurde in allen Szenarien derselbe Standard-Algorithmus mit identischen extern regulierbaren Parametern "out of the box" verwendet. Es erfolgte kein manuelles Umschalten der Algorithmen oder online Anpassungen der Parameter durch den Benutzer.

Das Trackingsystem besteht konzeptuell aus zwei Teilen, dem Erstellen einer Umgebungskarte und dem Tracken der Kamerapose bezüglich dieser Umgebungskarte. Die Umgebungskarte kann sowohl offline als auch online erstellt/erweitert werden. Zum Tracking muss lediglich eine Basiskarte verfügbar sein, welche aus rekonstruierten Featurepunkten und/oder Markern bestehen kann.

Offline Vorbereitung: Zur Erstellung der später zu verwendenden Featuremap (Punktwolke von markanten 3D Punkten) platziert der Anwender zunächst künstliche (ausgedruckte) Marker im Arbeitsbereich und nimmt mit einer Kamera Bilder aus den Perspektiven und Abständen auf, aus denen nachfolgend der Hauptteil des Trackings erfolgen soll.

Diese Kamerabilder werden zusammen mit der einmalig ermittelten und vorher bekannten Kamerakalibrierung als Eingabe dem Rekonstruktions-Tool übergeben (siehe Bild ??). Das Rekonstruktionstool extrahiert aus den einzelnen Bildern Merkmale welche gegenüber Beleuchtungsänderungen robust sind und die anhand ihrer Beschreibung über mehrere Skalenänderungen hinweg zugeordnet werden können. Die extrahierten Merkmale werden paarweise zwischen allen Bildern konsistent zugeordnet, anschließend werden diese paarweisen Zuordnungen auf Zusammenhangskomponenten hin untersucht und mittels u.a. geometrischer Kriterien in einen global konsistenten Zustand gebracht. Da die hierdurch automatisch generierte 3D Punktwolke noch in einem "willkürlichen" Koordinatensystem liegt, wird sie anschließend zu einem gegebenen Referenz-Koordinatensystem registriert. Dies erfolgt durch bekannte 3D-3D Korrespondenzen: Der Benutzer hat in Unifeye die Möglichkeit, 2D Bildmerkmale in den Kamerabildern zu markieren, z.B. die 2D Positionen der Referenzmarken. Zu diesen sind durch automatische Triangulierung die 3D Positionen im "willkürlichen" Koordinatensystem der Punktwolke bekannt, wodurch anhand der bereitgestellten, ausgemessenen 3D Koordinaten der Referenzmarken im globalen Koordinatensystem 3D-3D Korrespondenzen erstellt werden. Durch mindestens drei solcher Korrespondenzen kann eine erfolgreiche Registrierung in das Zielkoordinatensystem durchgeführt werden. Hiermit ist die Erstellung der Karte im Wesentlichen abgeschlossen, bei Bedarf kann der Nutzer die Karte manuell anpassen, etwa "versehentlich" rekonstruierte Punkte entfernen, welche nicht zum zu verfolgenden Objekt gehören.

Online Tracking: Nach dieser Registrierung kann die resultierende Punktwolke zum Tracking verwendet werden. Falls die Pose der Kamera nicht bekannt ist, werden aus dem aktuellen Kamerabild Merkmale extrahiert und denen Merkmalen der Punktwolke zugeordnet. Aus den so entstandenen 2D-3D Korrespondenzen wird unter Berücksichtigung eventueller Ausreißer/"Mismatches" eine Pose berechnet. Ist die Pose der Kamera ermittelt, werden die bestehenden 2D-3D Korrespondenzen mittels "sparse optical

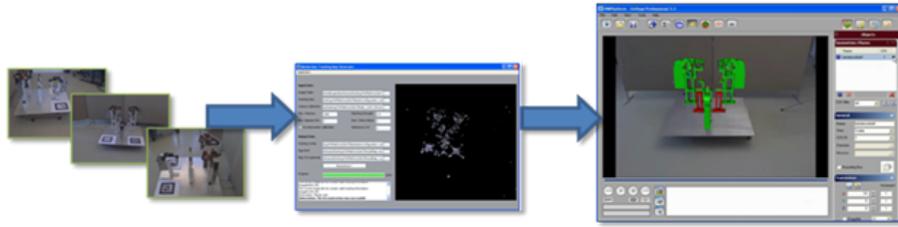


Abbildung 3: Rekonstruktionsprozess bei metaio

flow“ über mehrere Kamerabilder hinweg aufrechterhalten und um aus photometrischer und geometrischer Sicht geeignete Korrespondenzen aus der Punktwolke ergänzt. Das Ziel ist hierbei, immer eine möglichst hohe Anzahl an Korrespondenzen zur Berechnung der Kamerapose zur Verfügung zu haben, um eine möglichst hohe Genauigkeit zu erreichen.

Generell besteht einerseits die Möglichkeit, ausschließlich mit der zuvor berechneten, statischen Punktwolke zu arbeiten, andererseits kann diese Karte auch zur Laufzeit erweitert werden (extensible tracking). Durch letzteres wird das Tracking robuster gegenüber Veränderungen der Umgebungen, bzw. kann es in unbekannte Umgebungen fortgesetzt werden. In diesem Fall wird im Wesentlichen auf den bereits beschriebenen Ansatz zur 3D Rekonstruktion zurückgegriffen.

2.3 Fachgebiet Augmented Reality (FAR)

2.3.1 Fokus und Hintergrund der Technologie

Der Fokus des Fachgebiets Augmented Reality (FAR) der TU München liegt auf der Untersuchung von Ansätzen, um verschiedene optische oder andersgeartete Trackingkomponenten in fusionierter Weise online und dynamisch zur Verfügung zu stellen.

Grundsätzliche technische Ausrichtung (Ubitrack): Wenn möglich, nutzen wir bestehende Trackingkomponenten oder Algorithmen, in der Annahme, dass in konkreten Anwendungen aktuell dem State of the Art entsprechende Komponenten von akademischen oder kommerziellen Partnern erworben werden können. Nur wenn dies nicht möglich ist, oder wenn neuartige Ansätze integriert werden müssen, entwickeln wir eigene Trackinglösungen. Wir betrachten diverse egozentrische (inside-out-basierte) Technologien wie Mobilkameras, Inertialsensoren, odometrische Messgeräte, GPS- oder RFID-Leser an mobilen Displays (Handys, Laptops, HMDs) [26, 30, 12]. Wir ziehen aber auch exozentrische (outside-in-basierte) Technologien mit in Betracht, wie Weitwinkelkameras an der Decke [31], Infrarotkameras z.B. von A.R.T. oder einen Faro-Arm als lokal verfügbare spezielle Sensordienste in ubiquitär verteilten Sensorlandschaften in ”intelligenten, AR-fähigen Umgebungen“. Beispielsweise nutzen wir solche Sensorkombinationen einerseits für weiträumiges Tracking [30] und andererseits, um die Genauigkeit verschiedener Sensoren miteinander vergleichen zu können [17, 8, 15]. Diesbezüglich haben wir eine Softwarebibliothek, Ubitrack, entwickelt, die

diverse Methoden zur Fusion verschiedener Tracker, sowie zu deren Registrierung und Kalibrierung in einem einheitlichen, Komponenten-übergreifenden Konzept zur Verfügung stellt und automatisch zu echtzeit-fähigen Fusionsalgorithmen integriert [13]. Die teilweise recht komplexen räumlichen Beziehungen können durch das Konzept des "Spatial Relationship Graph" (SRG) abgebildet werden – eine graphische Repräsentation von Sensorkonfigurationen [23, 25, 24]. Einfache Veränderungen der Sensordaten lassen sich automatisch oder durch minimale manuelle Interaktion erreichen. Auch ein Umschalten zwischen komplexeren Szenarien ist ohne einen Neustart des Systems zur Laufzeit möglich [13, 30].

Ansatz beim AVILUS Tracking Contest (markerloses Tracking): Aufgrund der von den Anwendungspartnern gewählten Szenarien lag der spezielle Fokus beim AVILUS Tracking Contest auf optischen Verfahren mit markerlosem Tracking. Diesbezüglich konnten wir im Wesentlichen nur das in die Ubitrack Bibliothek integrierte PTAM (Parallel Tracking and Mapping) Verfahren von Georg Klein nutzen. Andere Trackingkomponenten der Ubitrack Bibliothek waren im gegebenen räumlichen und zeitlichen Kontext nicht praktikabel einsetzbar. Der Programmcode von PTAM basiert auf einem Algorithmus, den Georg Klein 2007 auf der ISMAR vorgestellt hat [18]. Der Source Code kann an der Oxford University für akademische Zwecke kostenfrei heruntergeladen werden. Das Verfahren arbeitet auf Punktwolken (auch Features genannt), die aus Kamerabildern der Arbeitsumgebung extrahiert werden, während die Kamera hin und her bewegt wird. Aus den Punktwolken werden jeweils die Kameraposition und -orientierung bestimmt. Gleichzeitig werden weitere neue Features in die Punktwolke aufgenommen und im Folgenden getrackt, sofern sie sich als zuverlässig erweisen. Die Trackingfunktionalität kann gleich nach dem Start des Systems genutzt werden. Von zwei anfänglichen sogenannten Keyframes (Kameraansichten) wird eine erste Punktwolke errechnet, die als Grundlage für das Tracking dient. Sobald eine genügend große Veränderung der Kameraorientierung registriert worden ist, wird ein neuer Keyframe zu der Menge hinzugenommen, und die Punktwolke so stets inkrementell erweitert. Eine Bündelblockausgleichung (Bundle Adjustment) wird von Zeit zu Zeit über die Menge aller Keyframes und der dazugehörigen Punktmerkmale ausgeführt, um die Gesamtpunktwolke und damit das Tracking zu optimieren. Dies geschieht parallel zum Tracking, damit die eigentliche Hauptanwendung, das Tracking, nicht unterbrochen werden muss. Dadurch entstehen zunehmend konsistente Feature Maps.

PTAM ist dafür bekannt, dass es dank seiner Aufteilung in zwei parallele Threads, einen für Echtzeittracking und einen zweiten separat davon für das langsamere, genaue Aufbauen einer Feature Map anhand gezielt ausgesuchter Keyframes, mehr Features (Tausende) und Keyframes (Hunderte) behandeln kann, als andere vergleichbare Algorithmen. Schwachpunkte des Systems liegen im aktuell gewählten Feature Detector und in Annahmen über die initiale Ausrichtung linearisierter Bereiche um ein gefundenes Feature, sowie der Behandlung von Outliers.

Generelles Ziel bei der Einbindung markerlosen Trackings in Ubitrack: In unserer Nutzung markerloser Verfahren gehen wir davon aus, dass sie in Kombination mit anderen, globaler oder robuster arbeitenden Verfahren eingesetzt werden und

einen nahtlosen Übergang zu diesen anderen Verfahren bieten [16]. Mögliche Anwendungsszenarien liegen in der industriellen Wartung, Kommissionierung und Produktion, in welchen Nutzer AR in weiträumig ausgelegten Arbeitsprozessen nutzen, und die Szenen sich dynamisch ändern. Dabei können Nutzer nicht während der Arbeit in große, wiederkehrende Vorbereitungsprozeduren involviert sein. Es wäre sehr aufwändig, wenn bei jeder Szenenänderung erneut ein komplexes, teil-manuelles Registrier- oder 3D-Rekonstruktionsverfahren durchgeführt werden müsste.

Dieser Vision entspricht das PTAM-Verfahren in idealer Weise, da es ermöglicht, dass parallel zum online Tracking kontinuierlich neue Features hinzu gelernt werden und sich somit die Kenntnis des Algorithmus über die Umgebung (3D-Rekonstruktion) kontinuierlich erweitert. Das Konzept findet zunehmend unter dem Schlagwort "Extensible Tracking" Beachtung. In diesem Sinne haben wir bei unseren Ubitrack-basierten Entwicklungen darauf geachtet, dass Registrierungsverfahren online durchführbar sind und somit keine Pause des Systems erzwingen. Als Konsequenz bieten wir nur minimale Werkzeuge an, um Punktwolken vorab zu generieren, zu registrieren und zu bearbeiten. Unser System erlaubt es, Punktwolken zu speichern und in einem späteren Prozess wieder hinzu zu laden. Aber wir haben keine Werkzeuge, um Punktwolken zu editieren. Dies erscheint in einigen der nachfolgenden Anwendungstests zum Nachteil, da wir nicht die Möglichkeit offen gelassen haben, Punktwolken zu bereinigen, um "saubere" Punktwolken zu erstellen oder eine genauere Registrierung zur Umwelt zu erzielen. Die Größe der Punktwolke spielt hier ebenfalls eine Rolle. Große Punktwolken werden ohne die manuelle Nachbearbeitung eher verzerrt (Drift) als kleinere. In kleinen Szenarien können wir allerdings ohne großen Zeitaufwand einen Bezug der Punktwolke zu der Referenz und somit zur Umwelt aufbauen.

2.3.2 Workflow während der Vorbereitung

PTAM kann online verwendet werden. Man kann aber auch vorab Punktwolken aufnehmen und speichern und diese dann später zum online Tracking hinzu laden. Dann sollte die Perspektive beim Aufnehmen der Daten ähnlich der Perspektive sein, aus der online gearbeitet wird. Kleinere Abweichungen sind akzeptabel. Falls eine Vorbereitung offline durchgeführt wird, ist dies der erste von zwei Vorbereitungsschritten, die Nutzer durchführen müssen.

Im zweiten Schritt muss die Punktwolke zu einem Weltkoordinatensystem in Bezug gesetzt werden, damit virtuelle Augmentierungen, die in diesem Koordinatensystem beschrieben sind, korrekt in die reale Umwelt eingebunden werden. Für diese einmalige Registrierung bieten wir zwei Varianten an.

- **Flachmarker-basierte Registrierung:** In der ersten Variante hilft ein von uns gleichzeitig geschalteter Markertracking-Algorithmus, die Skalierung zur Laufzeit fest zu legen. Zur Bestimmung müssen mehrere unterschiedliche Ansichten des Markers aufgenommen werden. Dabei sollten die Distanzveränderungen möglichst groß sein. Für eine exakte Bestimmung sollte der Marker gleichzeitig auch möglichst großflächig im Kamerafeld sichtbar sein, was die Bewegungsmöglichkeiten der Kamera wiederum einschränkt. Mehrere zueinander registrierte Marker können hier Abhilfe leisten. Über den Marker als Referenz können

wir nun den Bezug zur Realität herstellen, insofern der Marker zu seiner Umgebung registriert worden ist. Falls der Marker zu einem späteren Zeitpunkt abgenommen wird, kann die so erlernte und skalierte Punktwolke auch gespeichert werden und für ein späteres Tracking wieder geladen werden.

- **Registrierung durch manuelle Selektion von Referenzpunkten:** In der zweiten Variante wird die Beziehung der Punktwolke zu Objekten in der Realität hergestellt, ohne den Zwischenweg über einen Marker gehen zu müssen. Dieses Verfahren arbeitet ebenfalls zur Laufzeit und kann somit ohne Unterbrechung des Trackings eingesetzt werden. Es verlangt aber eine einmalige spezielle Interaktion des Benutzers mit dem System. Der Benutzer muss zu Beginn ihm bekannte Objektpunkte (z.B. aus einem CAD-Modell) dem Featuretracking zuweisen. Das Zuweisen erfolgt durch das Anklicken von beobachteten Punkten im Kamerabild, die bekannten Objektpunkten entsprechen. Die Reihenfolge der anzuklickenden Features ist dabei im Vorhinein festgelegt. Nach einer minimalen Zahl von drei Korrespondenzen kann eine erste Beziehung zwischen Feature Tracking und Objekt hergestellt werden. Mit jeder weiteren Korrespondenz steigert sich die Qualität der Abbildung der beiden Koordinatensystemen zueinander. Als problematisch hat sich erwiesen, dass schon vor dem Tracking eingeschätzt werden muss, ob ein Objektpunkt im Tracking als Feature erkannt werden kann. Oft werden in der Nähe eines markanten Objektpunktes auch mehrere Features zum Tracking erkannt, aus denen ein Punkt ausgewählt werden muss. Eine Änderung des Prozesses könnte hier Abhilfe leisten.

3 Testszenario: Volkswagen

Die Anforderungen von Volkswagen wurden zum einen an einem Passat und zum anderen an einem bewegten Modell eines Golfs evaluiert. Am Passat wurden Evaluationen am Exterieur (3.1), Motorraum (3.2), Interieur (3.3) und eine relative Genauigkeitsmessung (3.4) durchgeführt. Für das bewegte Szenario wurde ein Golf Modellauto auf einem Kuka-Roboter montiert (3.7), wodurch der Golf gleichmäßig bewegt werden konnte. Im Folgenden werden die Volkswagen Szenarien und Ergebnisse präsentiert.

3.1 VW Passat Exterieur

Die Exterieur-Szenarien wurden an einem beweglichen Passat durchgeführt (siehe Abbildung 4). Der Passat wurde zwischen Anlernphase und Tracking-Evaluationsphase bewegt (nicht während des Experimentes). So konnte verhindert werden, dass Features aus der Umgebung zum Tracken mitgelernt und benutzt wurden.



Abbildung 4: Beweglicher VW Passat

3.1.1 Versuchsszenario

Das Versuchsszenario bestand aus 2 Phasen. In der ersten Phase musste ein CAD-Modell der Fahrertür korrekt überlagert werden. In der zweiten Phase mussten, auf Basis der 3D-Koordinaten, insgesamt 4 Punkte an Nummernschild und Heckschriftzug gefunden und identifiziert werden (siehe Abbildung 5).



Abbildung 5: Passat Exterieur - Buchstaben, die aufgefunden werden mussten.

3.1.2 Messergebnisse

Überlagerung eines CAD-Modells Abbildungen 6 und 7 zeigen einige Beispiele der Überlagerung der Autotür.

FAR hatte Probleme mit der globalen Registrierung. Das Tracking funktionierte mittelmäßig bis schlecht (1 von 5 Punkten). Eine Annäherung an die Tür provozierte einen Abriss des Tracking.

Die Überlagerung von **metaio** (3 von 5 Punkten) war von schräg vorn oder schräg hinten befriedigend, von der Seite funktionierte das System nicht. Der Abstand zum Objekt musste relativ groß gewählt werden, damit das Tracking funktionierte. Bei Annäherung an das Objekt riss das Tracking schnell ab. Die Überlagerung schwankte zwischen sehr gut und 5cm Drift. Versuche mit einer "extensible Tracking" Variante lieferten keine positiven Ergebnisse.

Die Überlagerung vom **IGD** war gut und stabil (4 von 5 Punkten). Das Tracking funktionierte auch von der Fahrzeugseite aus, und man konnte sich an die Tür weit annähern. Wenn das Tracking allerdings abbricht, war es nötig, zur Initialisierungsposition zurückzukehren. Es wurden offline zwei Maps gelernt, von denen aus das Tracking online initialisiert werden konnte.



Abbildung 6: FAR - Probleme mit der globalen Registrierung



Abbildung 7: metaio (links) und IGD (rechts) - Bei beiden ist die Tür etwas zu weit nach hinten versetzt.

| Versatz | | | |
|--------------------|--------|-----|-------|
| | metaio | IGD | FAR |
| Nummernschild | ok | ok | ok |
| Schriftzug am Heck | 3cm | 5cm | 7cm |
| Schriftzug am Heck | 3cm | 3cm | 4,5cm |
| Nummernschild | ok | ok | ok |

Tabelle 2: Versatz bei Buchstaben am Passat Exterieur

Punktsuche Die Messergebnisse zur Punktsuche (siehe Abbildung 5) sind in Tabelle 2 dargestellt. Keiner der Technologiepartner konnte einen Buchstaben im Schriftzug am Heck eindeutig identifizieren. Die Buchstaben auf dem Nummernschild bereiteten keine Probleme.

3.1.3 Diskussion der Technologiepartner

Im Folgenden nehmen die Technologiepartner Stellung zu den Ergebnissen.

FAR: In diesem Testszenario nutzte FAR die Anlernphase des Vortages nicht. Die Punktwolke wurde mit PTAM erst kurz vor Testbeginn erstellt. Zur Registrierung mit dem virtuellen Koordinatensystem (des CAD-Modells der Tür und der Punkte auf dem Nummernschild und der Heckschrift) mit der PTAM-generierten Punktwolke wurde ein Flachmarker am Kühlergrill angebracht. Ausgehend von dieser frontalen Sicht auf das Auto, bei welchem die Überlagerungsaufgabe auf dem Nummernschild durchgeführt wurde, wurde das Auto dann Richtung Fahrertür umrundet, wobei die Kamera kontinuierlich auf das Auto sah. Die aufgetretenen Diskrepanzen zwischen der Überlagerung der Fahrertür und der realen Tür führen wir auf Drift während der Bewegung vom Kühlergrill zur Tür, sowie insbesondere auf die Existenz diverser Spiegelungen und Glanzlichter auf der sehr homogenen Lackierung der Karosserie zurück. Spiegelreflektionen sind ja bekanntermaßen nicht objektfest sondern hängen von der Blickposition ab und wandern also mit dem Betrachter mit. Zusätzlich zum systemati-

schen Drift war die Überlagerung in diesem Szenario relativ wackelig.

IGD: Auf der Außenseite des Fahrzeuges war die Anzahl von textur-basierten Punktmerkmalen begrenzt. Ein Großteil der Bildmerkmale befand sich auf den Rädern des Fahrzeuges. Reflektionen auf dem Lack erschwerten das Verfolgen und die Rekonstruktion von Punkten. Dennoch war es möglich, das Trackingsystem von vielen Blickbereichen erfolgreich einzusetzen. Die Transformation der Rekonstruktion in das Weltkoordinatensystem funktionierte an der Seite und im vorderen Bereich gut, da dort genügend Referenzpunkte vorhanden waren. Die Überlagerung der 3D-Geometrie war trotz der mangelnden Anzahl von verfolgbaren Bildmerkmalen zufriedenstellend. An der hinteren Seite wurde nur mit einem zur Laufzeit arbeitenden Rekonstruktionsverfahren von der Seite ausgehend der Kamerapfad verfolgt, da im hinteren Bereich keine Kalibrierpunkte vorhanden waren. Aufgrund der mangelnden Texturierung verlor das Tracking etwas an Robustheit, wenn die Anzahl an verfolgten Bildpunkten zu gering wurde. Dennoch war es mit Hilfe des Extensible Trackings möglich, sich der Fahrertür sehr nahe anzunähern, wobei fast keine Bildmerkmale der Offline-Map mehr im Bild zu finden waren und die Kameraposeberechnung nur auf online rekonstruierten, zum Teil sehr kurzlebigen Bildmerkmalen beruhte.

metaio: metaio war in der Lage aus nahezu allen Perspektiven zu initialisieren. Jedoch führte eine zu starke Annäherung an die Tür zu einem Abreißen des Trackings bzw. machte ein dortiges Reinitialisieren unmöglich. Dies ist darin begründet, dass das verwendete punkt-basierte Tracking auf der glänzenden und texturlosen Oberfläche keine markanten Merkmale finden konnte.

3.2 VW Passat Motorraum

In diesem Szenario von VW wurde markerloses Tracking im Motorraum eines Passats evaluiert (siehe Abbildung 8). Dafür wurde die Motorhaube geöffnet, und es mussten verschiedene Objekte gefunden werden. Die im Bild zusehenden Marker dienten nur zur Registrierung der Systeme in der Anlernphase. Beim Test selbst waren sie verdeckt.

3.2.1 Versuchsszenario

Unter verschiedenen Bedingungen (normal, starke Beleuchtung (2*1000W), und Entfernung einer Dekorabdeckung) mussten jeweils zwei Objekte gefunden werden. Beim ersten Objekt handelte es sich jeweils um ein großes Objekt (z.B. eine Öl-Abdeckung). Beim zweiten Objekt handelte es sich um eine kleine Sicherung < 1cm (siehe Abbildung 9).

3.2.2 Messergebnisse

Tabelle 3 zeigt einen Überblick über die gefundenen Punkte im Motorraum.

Das Verfahren von **FAR** scheiterte aufgrund geänderter Lichtbedingungen (sonnig hell gegenüber extrem bewölkt, siehe Abbildungen 8 und 10) zum Vortag, an welchem die Feature Map erstellt wurde. Es bestand am Testtag selbst nicht genügend Zeit, um eine neue Feature Map aufzunehmen.



Abbildung 8: Motorraum VW Passat

| metaio | IGD | FAR |
|--------|-----|-----|
| 6 | 5 | 0 |

Tabelle 3: Gefundene Punkte bei der Objektsuche im Motorraum (von 6 tatsächlichen Punkten)

Das Verfahren des **IGD** funktionierte unter hellen Lichtbedingungen nur, wenn die Initialisierung über eine lokale (KLT) Optimierung durchgeführt wurde und nicht über das wide-baseline Matching. Dieses bedingt, dass das Tracking immer ähnlich der Anlernperspektive initialisiert werden muss. Die globale Initialisierung aus beliebiger Perspektive funktioniert also nicht, die lokale dafür sehr gut.

Das Tracking von **metaio** funktionierte gut. Ein alternativer Versuch, den Sicherungskasten mit einem "extensible Tracking"-Ansatz zu tracken, produzierte einen großen Drift.

3.2.3 Diskussion der Technologiepartner

Im Folgenden nehmen die Technologiepartner Stellung zu den Ergebnissen.

FAR: Am Anlerntag herrschte eine diffuse Beleuchtung (trübes Wetter) vor; am Testtag schien die Sonne durch große Fenster in den Vorführraum. Darüber hinaus waren auch spezielle Scheinwerfer aufgebaut. Dadurch entstanden große Schlagschatten und auch Glanzlichter im Motorraum und auf der umliegenden Karosserie – und dementsprechend auch gänzlich neue Bildbereiche mit hohem Kontrast und entspre-



Abbildung 9: Kleine Sicherungen, die mit Hilfe eines AR Systems identifiziert werden mussten.

chend vielen neuen Features. Es ist nicht auszuschließen, dass diverse Features des Vortags wegen der hohen allgemeinen Beleuchtung und einer kamerainternen Anpassung der Blende (zur Adaption des dynamischen Messbereichs der Kamera) nicht mehr so prägnant zu erkennen waren. Als Konsequenz konnte die Feature Map des Vortages nicht mit den Features, die PTAM online finden konnte, in Einklang gebracht werden. PTAM war dann zwar in der Lage, eine neue Feature Map aufzubauen und damit zu tracken. Es fehlte aber die Zeit, um diese neu entstehende Punktwolke zum virtuellen Koordinatensystem zu registrieren, um so die virtuellen Objekte einblenden zu können. Hierfür hätte erneut der Flachmarker am Kühlergrill angebracht und eingemessen werden müssen. Der zweite in Abschnitt 2.3.2 vorgestellte Algorithmus, mit dem Nutzer eine Registrierung zwischen der Punktwolke und den virtuellen Objekten durch manuelle Selektion von Referenzpunkten in einem CAD-Modell durchführen können, war auch nicht anwendbar, da kein CAD-Modell des Motorraums vorlag.

IGD: Das Bestimmen der Kameraposition und -orientierung im Motorraum funktionierte sowohl beim Initialisieren als auch beim Tracken äußerst robust, was auch zu erwarten war, da in diesem Szenario genügend aussagekräftige Bildmerkmale erkannt und rekonstruiert werden konnten. Das Kalibrieren bereitete auch keine Probleme, da Kalibrierungspunkte zu genüge vorhanden waren. Bei stark veränderten Lichtverhältnissen bereitete das Initialisieren Schwierigkeiten. Die Parameter für den verwendeten Klassifizierungsansatz mit Randomized Trees waren für den Fall starker Be-



Abbildung 10: Starke Veränderung der Beleuchtungsbedingungen während des Versuchs

leuchtungsveränderungen zu restriktiv eingestellt. Eine alternative Initialisierungskomponente die mit Referenzbildern und einer Validierung mit KLT-Bildmerkmalen arbeitet, konnte das Tracking trotz der starken Beleuchtungsveränderungen besser initialisieren. Leider wurde diese Initialisierungsmethode während der Vorbereitung nicht tiefgreifend verfolgt, was zur Folge hatte, dass zur Laufzeit für die Initialisierungskomponente nur ein einziges Referenzbild im System zur Verfügung stand, und der Benutzer zum Initialisieren eine bestimmte Position aufsuchen musste. Das Bild-zu-Bild Tracking selber kam mit den Beleuchtungsveränderungen gut zurecht. Die Anzahl erfolgreich verfolgter Bildpunkte war zwar etwas geringer, aber dennoch stets groß genug um das Trackingergebnis nicht zu beeinträchtigen. Zusätzliche zur Laufzeit rekonstruierte Bildmerkmale stabilisierten den gesamten Trackingprozess. Bei einem der zu identifizierenden sehr kleinen Punkte deutete die Visualisierung eines sehr groben Fadenkreuzes genau zwischen zwei Sicherungen, und leider wurde der gesuchte Punkt knapp verfehlt. Eine klarere Visualisierung hätte hier von Vorteil sein können. Die Messdurchläufe wurden sowohl mit einer Streaminglösung auf einem Nokia N900 Smartphone durchgeführt.

metaio: Aufgrund der sehr gut texturierten Umgebung und der damit verbundenen vielen markanten Bildmerkmale handelt es sich um ein vergleichsweise einfaches Szenario, sodass sowohl die Initialisierung als auch das Tracking sehr genau und sehr robust funktionierte. Zu erwähnen ist, dass metaio in allen Szenarien und allen Beleuchtungssituationen denselben Tracking-Algorithmus mit identischen Parametern verwendet und keine manuellen Anpassung durch den Nutzer online vorgenommen wurden. Die Änderung der Beleuchtung hatte in diesem Szenario keinen erkennbaren Einfluß auf die Trackingqualität.

3.3 VW Passat Interieur

In diesem Szenario sollten Trackingverfahren im Inneren eines Autos evaluiert werden. Das Einsatzszenario umfasste die volle Breite des Innenfahrzeugs zwischen Fahrer- und Beifahrertür.



Abbildung 11: VW Interieur Szenario

3.3.1 Versuchsszenario

Das Interieur Szenario bestand aus mehreren Teilaufgaben. Zunächst sollten 6 Punkte (vorwiegend Knöpfe und Drehschalter) im Auto gefunden werden; davon 3 Punkte unter normalen Lichtbedingungen, und 3 Punkte mit eingeschalteten Standscheinwerfern (2x1000W).

Im zweiten Schritt sollte ein CAD-Modell dem Navigationssystem und dem Griff und Türöffner der Beifahrerseite überlagert werden. Diese Überlagerung wurde unter 4 verschiedenen Bedingungen getestet: 1) markerlos - ohne zusätzliches Licht, 2) markerlos - mit Zusatzscheinwerfer (2x1000W). Für die letzten beiden Überlagerungsvarianten wurde ein Marker für 20 Sekunden im Navigationsdisplay eingeblendet (siehe Abbildung 12). Während dieser Zeit mussten die Teilnehmer ad-hoc eine Feature Map aufbauen und danach ohne den Marker weiter tracken. Dies wurde unter normalen Bedingungen 3) Marker für 20 Sekunden ohne Zusatzlicht und unter erschwerten Bedingungen 4) Marker für 20 Sekunden mit Scheinwerfer (2x100W) getestet. Alle Tests wurden nicht von den Technologiepartnern selbst sondern von einer Person von VW durchgeführt, um somit festzustellen, ob die Technologie auch von Nichtexperten im markerlosen Tracking bedient werden kann.



Abbildung 12: Ein Marker erscheint für 20 Sekunden im Navigationsdisplay. In dieser Zeit musste ad-hoc eine Feature Map aufgebaut und registriert werden.

3.3.2 Messergebnisse

Alle Teilnehmer erkannten die sechs richtigen Punkte. Bei *metaio* war die gesamte Visualisierung um circa 1cm zu hoch registriert. Nach dem Hinweis durch den Versuchsleiter auf diesen Registrierungsfehler konnten alle Punkte richtig erkannt werden. Die Visualisierung zur versteckten Parkbremse (links hinter dem Lenkrad) sprang um circa 5 cm. IGD und FAR blendeten die CAD-Visualisierung des Navigationssystems nicht aus. Dadurch konnten Punkte / Knöpfe am Navigationsdisplay leichter erkannt werden. Im Detail wurden folgende Ergebnisse beobachtet:

Die Ergebnisse von **metaio** wurden unter den vier Versuchsbedingungen wie folgt beobachtet: 1) Ohne zusätzlichen Scheinwerfer funktionierte das Tracking gut. Es gab leichte Abweichungen in der Registrierung, welches insbesondere an der Tür zu sehen war. Bei schnellen Bewegungen riss das Tracking ab. Bei Stillstand war die Überlagerung gut bis sehr gut mit circa 1 cm Abweichung (4 von 5 Punkten). 2) Mit Scheinwerfer funktionierte das Tracking kaum. Es funktionierte selten, dann aber ganz gut (von einer bestimmten Pose in Fahrzeugmitte) (1 von 5 Punkten). 3) Wenn von einem Marker (ohne Zusatzlicht) gestartet wurde, war eine Lageabweichung der Türgeometrie schon während der Initialisierungsphase (Marker eingeblendet) zu erkennen. Während des reinen Trackings wurde die Geometrie des Navigationsdisplays gut überlagert; die Türgeometrie war verschoben (4 von 5 Punkten). 4) Wurde bei zusätzlichem Licht von einem Marker gestartet, war ein ähnliches Verhalten wie unter Bedingung 3 (mit Marker, ohne Zusatzlicht) zu beobachten. Die Überlagerung war leicht besser (4 von 5 Punkten).

FAR hatte zunächst einen schlechten Durchlauf, der nicht gewertet wurde. Die Er-



Abbildung 13: Überlagerungsbeispiele metaio.

gebnisse für den zweiten Durchlauf unter den vier Versuchsbedingungen waren wie folgt: 1) Ohne Zusatzlicht funktionierte das Tracking gut. Die Überlagerung war um circa 1 cm zu hoch registriert (4 von 5 Punkten) 2) Beim Einschalten der Beleuchtung benötigte das Verfahren eine Weile, um neue Features unter den neuen Bedingungen anzulernen, funktionierte dann aber wieder gut (4 von 5 Punkten). 3) Die Variante "Marker 20 Sekunden - ohne Zusatzlicht" funktionierte bei FAR erst im 4. Versuch, nachdem die Testperson von VW gewechselt wurde – in dem Fall aber dann wiederholt sehr gut. Für die Registrierung reichte es aus, wenn der Marker lediglich 10 (statt 20) Sekunden sichtbar war. Die Überlagerung des Navigationsdisplays war besser als bei metaio und IGD. Die Überlagerungen der Tür waren nicht so gut wie die des Navigationsdisplays (der Winkelfehler war im vorgegebenen Szenario an der weiter entfernten Tür sehr gut sichtbar). (4 von 5 Punkten). 4) Beim Einschalten des Lichts verhielt sich das Verfahren ähnlich wie schon unter Bedingung 2): es benötigte einige Zeit zum Anlernen neuer Features, lief dann wieder stabil (4 von 5 Punkten).

Die Visualisierung des **IGD** war um circa 1 cm zu weit nach links registriert. Folgende Beobachtungen wurden unter den verschiedenen Versuchsbedingungen gemacht. 1) Ohne Zusatzlicht funktionierte das Verfahren von IGD etwas besser als das von metaio (4 von 5 Punkten). Auch das Tracking im Nahbereich funktionierte gut, ohne dass das Tracking abbrach. 2) Mit zusätzlichem Licht konnte sich das Verfahren zunächst nicht initialisieren. Durch Umschalten von globaler auf lokale Initialisierung verhielt sich das Verfahren dann ähnlich gut wie ohne Scheinwerfer - mit dem Nachteil, dass sich das System nur aus bestimmten Posen initialisieren ließ (4 von 5 Punkten). 3) Die Variante "mit Marker - ohne Zusatzlicht" funktionierte beim IGD gut, schnell und robust – allerdings war die Überlagerung etwas schlechter (insbesondere an der Tür-Geometrie) als unter den ersten beiden Bedingungen (4 von 5 Punkten). 4) Die letzte Variante mit Marker und mit Licht verhielt sich ähnlich wie Variante 3 (4 von 5 Punkten).



Abbildung 14: Überlagerungsbeispiele FAR.

3.3.3 Diskussion der Technologiepartner

Im Folgenden nehmen die Technologiepartner Stellung zu den Ergebnissen.

FAR: Die online Registrierung funktionierte gut, weil der Marker gut einsehbar war und genügend Punktkorrespondenzen in den 10 Sekunden gesammelt werden konnten, um eine robuste Beziehung zwischen marker-basiertem Tracking und PTAM herzustellen. In diesem Test wurde keine vorher (offline) konstruierte Feature Map genutzt sondern alles online erstellt. Dies ist die ideale Situation des "Extensible Trackings", für welche PTAM konzipiert und entwickelt wurde ("PTAM at its best"). Seine Besonderheit, dass es mit sehr großen Mengen an Punkten über Keyframes sehr schnell umgehen kann, kam zur Geltung: Die Feature Maps waren schon nach der Hälfte der erlaubten Zeit so robust, dass markerlos weiter getrackt werden konnte. In diesem Fall konnte auch der negative Einfluss einer plötzlichen Beleuchtungsänderung abgefangen werden, indem adaptiv neue Features hinzu genommen wurden. Das heißt im Vergleich zum Motorraum-Szenario, dass PTAM mit Beleuchtungsänderungen umgehen kann, wenn diese zu einem Zeitpunkt geschehen, an dem die aktuelle Betrachterpose bekannt ist. Nur die Initialisierung der Pose (wenn die Feature Map vom Vortag stammt) in Kombination mit einer extremen Beleuchtungsänderung überfordert die automatisch-



Abbildung 15: Überlagerungsbeispiele IGD.

extensiblen Mechanismen des Systems.

Für eine gute Registrierung der Feature Map zum Marker – und die damit verbundene Überlagerung der Anzeige – ist es entscheidend, was für eine Bewegung der Nutzer ausführt. Die Geste sollte ein möglichst großes Volumen im Arbeitsraum umfassen, dabei aber nicht den Marker aus dem Blickfeld der Kamera verlieren (siehe Abschnitt 2.3.2). Dies mag der Grund dafür gewesen sein, dass ein Teil der Tests bei einer Testperson scheiterten, bei einer anderen Testperson dann aber sehr gut verliefen.

IGD: Die Qualität der Trackingergebnisse im Innenraum des Fahrzeugs war ähnlich wie beim Motorraum. Rekonstruieren, Kalibrieren, Initialisieren und Tracking funktionierten gut. Sowohl bei einer im Vorfeld rekonstruierten Map, als auch einer zur Laufzeit rekonstruierten Map, war die Überlagerung der CAD-Modelle zufriedenstellend. Bei sich ändernden Lichtverhältnissen wurde das Tracking zunächst etwas weniger robust, stabilisierte sich dann jedoch wieder. Das liegt daran, dass durch die online Rekonstruktion Bildmerkmale gesammelt werden, die den veränderten Lichtverhältnissen entsprechen.

Beim Initialisieren gab es ähnliche Probleme wie beim Motorraum. Das Randomized-Trees-basierte Initialisierungsverfahren hatte mit den Beleuchtungsveränderungen Schwierigkeiten, das Referenzbild-basierte Verfahren funktionierte nach Einschalten der Schein-

| | metaio | IGD | FAR |
|---|--------|-----|-----|
| 1) ohne Licht | 4 | 4.5 | 4 |
| 2) mit Licht | 1 | 4.5 | 4 |
| 3) Marker 20 Sekunden sichtbar, ohne Licht | 4 | 4 | 4 |
| 4) Marker 20 Sekunden sichtbar, mit Licht | 4 | 4 | 4 |

Tabelle 4: Bewertung der Überlagerungsqualität am Navigationsdisplay und am Türgriff (1-5 Punkte)

werfer deutlich besser. Zu bedauern war wiederum, dass während der Vorbereitung für eine allgegenwärtige Benutzung nicht genügend Referenzbilder abgespeichert wurden. Dennoch sollte man das Potential von Klassifizierungsverfahren wie Randomized-Trees nicht zu gering schätzen, da zu erwarten ist, dass mit großzügigeren für starke Beleuchtungsänderungen angepassten Parametereinstellungen sowohl für die Bildmerkmalsdetektion als auch die Klassifizierung auch gute Ergebnisse möglich sind.

3.4 VW Genauigkeitsmessung (Schiene)

Mit Hilfe einer 1.443 Meter langen Linearschiene wurden relative Genauigkeitsmessungen durchgeführt.

3.4.1 Versuchsszenario

Es wurde eine Kamera auf der Schiene angebracht und mit Blick auf einen Passat vor und zurück bewegt. Dabei musste durch Tracking des Autos und der Umgebung die vorab unbekannte Länge der Schiene bestimmt werden. Für das Tracking durfte vor dem Versuch mit frei beweglicher Kamera eine Feature Map erstellt werden, welche dann während des Versuchs zur Verfügung stand.

Der Versuch wurde unter drei verschiedenen Kameraorientierungen auf der Schiene (um 0, 45 und 90 Grad um die "roll-Achse" gedreht) durchgeführt (siehe Abbildung 17). In der ersten Versuchsreihe wurde kontinuierlich über alle drei Kameraorientierungen hinweg getrackt; in einer zweiten Versuchsreihe wurde vor jedem neuen Start einer Vor- und Zurückbewegung entlang der Schiene (d.h. bei jeweils geänderter Kameraorientierung) das Tracking unterbrochen. So sollte die Fähigkeit des Systems getestet werden, die Feature Map bei verschiedenen Anfangsorientierungen erfolgreich zu lokalisieren.

Zusammen mit den Technologiepartnern wurde beschlossen, die Motorhaube des Fahrzeugs zu öffnen, da ansonsten der weiße Lack wenige kaum trackbare Features geboten hätte.



Abbildung 16: Die Länge einer Linearschiene musste dadurch bestimmt werden, dass der Passat von einer Kamera, die auf der Schien hin und her fuhr, getrackt wurde.

3.4.2 Messergebnisse

Die Messergebnisse sind in Tabelle 5 zu sehen. Die Länge der Schiene betrug 1443 Millimeter. Die Werte *avg* sind gemittelt über drei "Hin-und-Zurück" Fahrten. *Relnit* bedeutet, dass das Tracking zwischen einer "Hin-und-Zurück" Fahrt neu initialisiert wurde. *P(+/-)* bezeichnet jeweils die Präzision (Wiederholungsgenauigkeit) – also die maximale Abweichung zum Mittelwert der Messung *avg*. *G(avg)* bezeichnet die mittlere Genauigkeit – also die mittlere Abweichung zwischen Soll und Ist. Konnte das Tracking nicht initialisiert werden, war ein Neustart erlaubt. Konnte eine Messung nicht abschließend durchgeführt werden, ist diese mit "-" markiert.

3.4.3 Diskussion der Technologiepartner

Im Folgenden nehmen die Technologiepartner Stellung zu den Ergebnissen.

FAR: Die wiederholbaren Ergebnisse der Genauigkeitsmessung zeigen ein stabiles Tracking. PTAM findet angelernte Punktmerkmale präzise wieder und ordnet diese zuverlässig zu. Die Genauigkeit der Ergebnisse der Distanzmessung zeigen, dass unser Registrierverfahren in diesem Szenario sehr gut zur Anwendung gebracht werden konnte. Die korrekte Bestimmung der Skalierung der Translation des Featuretrackings hängt von dieser Registrierung ab.

Die besten Ergebnisse wurden allerdings mit einer auf die geöffnete Motorhaube gerichteten Kamera erzielt. Bei anderen Neigungswinkeln (0/90 Grad) der Kamera

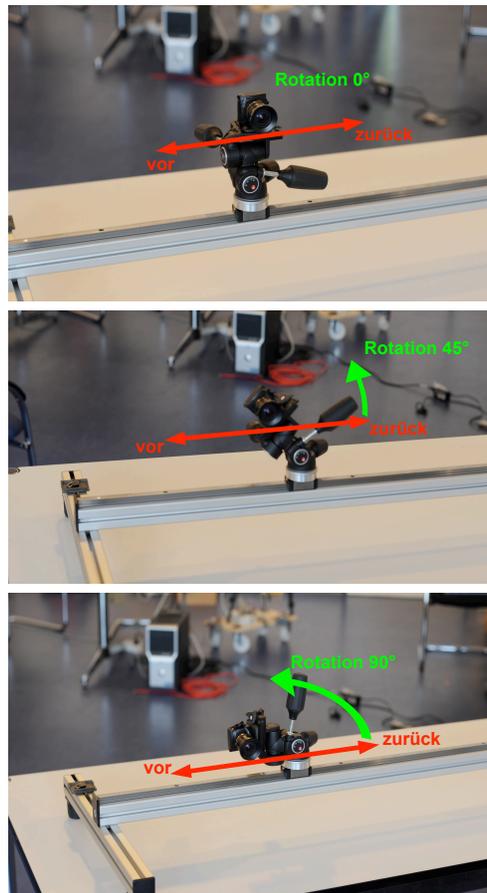


Abbildung 17: Der Versuch wurde unter drei Drehwinkeln der Kamera, relativ zur Schiene, durchgeführt: 0, 45 und 90 Grad.

wurden leicht andere Translationswerte erzielt. Zunächst einmal bot die initiale Blickrichtung (45 Grad) die meisten Features. Darüber hinaus wurde diese Blickrichtung bereits zum Initialisieren der Feature Map sowie zur Bestimmung deren korrekter Skalierung verwendet, während für die anderen Blickrichtungen wesentlich stärker auf die durch das Extensible Tracking hinzugelernen Features zurückgegriffen werden musste. Dies legt den Schluss nahe, dass es beim Extensible Tracking trotz globaler Bündelblockausgleichung zu einer Abdrift von Features und Keyframes in der Feature Map kommt, welche sich in Richtung der randständigen Keyframes verstärkt.

Als Grund für die fehlgeschlagene Reinitialisierung des Feature Trackings bei einem Drehwinkel um 90 Grad ist vermutlich die geringe Anzahl von Features verantwortlich, die in diesem Bereich sichtbar waren. Im Wesentlichen befand sich dort eine weiße Wand und nur wenige markante Strukturen. Außerdem wurde die Fähigkeit zur Reinitialisierung in dieser Blickrichtung beim Anlernen der Map nicht explizit getes-

| Win- kel | Re- Init | IGD | | | FAR | | | metaio | | |
|-------------|-------------|------|------------|------------|------|------------|------------|--------|------------|------------|
| | | avg | P (+/-) | G (avg) | avg | P (+/-) | G (avg) | avg | P (+/-) | G (avg) |
| 0 | | 1436 | 2 | -7 | 1441 | 6 | -2 | 1447 | 8 | 4 |
| 45 | | 1434 | 26 | -9 | 1440 | 2 | -3 | 1441 | 5 | -2 |
| 90 | | 1399 | 22 | -44 | 1457 | 3 | 14 | 1440 | 5 | -3 |
| 0 | x | | | | 1438 | 4 | -5 | 1447 | 10 | 4 |
| 45 | x | | | | 1443 | 2 | 0 | 1437 | 16 | -6 |
| 90 | x | | | | - | - | - | 1437 | 6 | -6 |

Tabelle 5: Ergebnisse Schiene (VW Szenario).

tet, ansonsten hätte die Map vorab durch Einbeziehung zusätzlicher Kamerapositionen vermutlich besser trainiert werden können.

IGD: Bei der Genauigkeitsmessung war ein systematischer Fehler erkennbar, da bei allen Messungen ein relativ beständiger, aber zu kleiner Versatz gemessen wurde. Erklärbar wäre diese Tatsache dadurch, dass sich bei der Kalibrierung ein Skalierungsfehler eingeschlichen hat. Wahrscheinlicher ist jedoch, dass die intrinsischen Kameraparameter, vor allem die radiale Verzerrung, vor der Rekonstruktion nicht präzise genug bestimmt wurden.

Getracked wurde mit einer Menge von Punkten, die zuvor ohne zusätzliche Hilfsmittel (Marker) und lediglich mit einer einfachen Webcam rekonstruiert wurde. Das Registrieren dieser Punktwolke im Bild lieferte letztendlich das Trackingergebnis. Eine Unterscheidung mit und ohne Re-Initialisierung bewirkte daher keinen Unterschied in der Kameraposebestimmung. Auf einen zweiten Testdurchlauf mit Re-Initialisierung wurde daher verzichtet. Dass eine Rotation um die Kamerablickrichtung eine Auswirkung auf das Messergebnis hat ist verwunderlich, da es sich praktisch nur um eine Rotation im Bildbereich handelt. Es könnte sein, dass während der Messungen ungünstiger Weise schlecht rekonstruierte Bildmerkmale in die Punktwolke mit aufgenommen wurden, die das Messergebnis negativ beeinträchtigten. Mit einem Verzicht, zusätzliche Bildmerkmale zur Laufzeit zu rekonstruieren, könnte man diese Fehlerquelle vermeiden.

metaio: Für die Durchläufe mit aktiviertem Tracking ergab sich eine bessere Genauigkeit als für Reinitialisierungs-Durchläufe. D.h. an den Initialisierungsalgorithmen besteht noch Potential für Verbesserungen. Oder während des Anlernens der Feature Map, hätten noch mehr Bilder aus den entsprechenden Perspektiven aufgenommen werden müssen.

Die Veränderung des Blickwinkels der Kamera hatte im Vergleich zu dem Winkel, in welchem diese kalibriert wurde, keinen Einfluss auf die Genauigkeit.

3.5 Bewegtes VW Modellauto auf einem Kuka Roboter

Ziel dieses Szenarios war die Evaluierung der Trackingverfahren bei bewegten Objekten.

3.5.1 Versuchsszenario

Es wurde ein Miniaturmodell eines Golf VI auf einem Roboter von Kuka befestigt. Es wurden zwei Laufbahnen programmiert, auf denen der Roboter das Automodell bewegte. In der ersten Schwierigkeitsstufe wurde das Auto nur in einer Ebene rotiert. In der zweiten Stufe wurde das Auto frei in allen Freiheitsgeraden bewegt. Zur Bewertung der Trackingqualität sollte auf der grauen Platte an definierter Stelle ein Avilus-Logo eingeblendet werden.

In der Vorbereitungsphase stellte sich heraus, dass das Fahrzeug aufgrund der Lackierung nur sehr schwer getrackt werden konnte. Darum wurden sämtliche Fahrzeurtüren geöffnet, um so mehr Features zur Verfügung zu stellen.

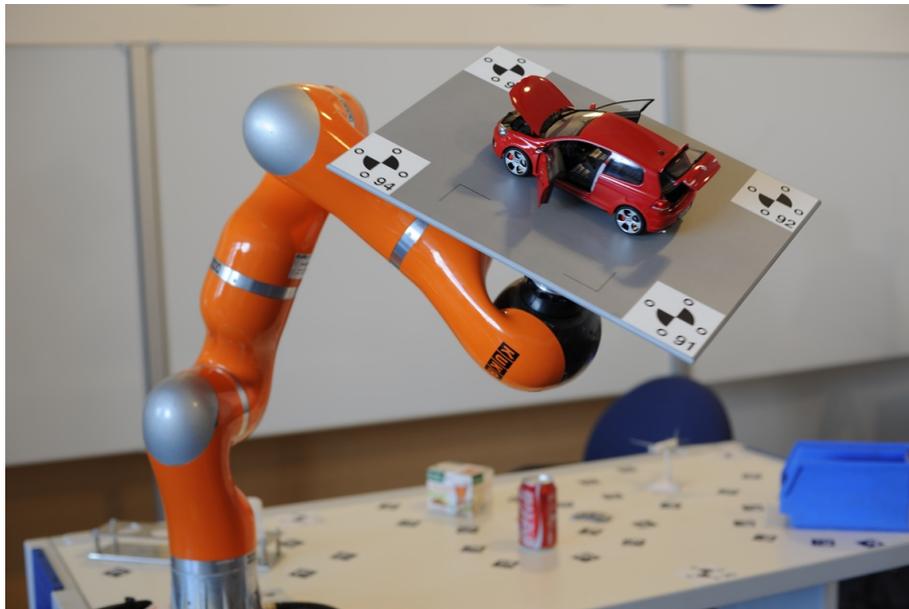


Abbildung 18: Miniaturmodell eines VW Golf, das von einem Kuka Roboter kontrolliert bewegt wird.

3.5.2 Messergebnisse

In der ersten Stufe, der Verfolgung reiner Rotationsbewegungen, konnte nur von **metaio** eine gute Überlagerung erzielt werden. Es kam nur manchmal zu Aussetzern oder zu einer falsch platzierten Augmentierung.

Das **IGD** konnte das Modell von einer Seite gut tracken. Die anderen Seiten konnten aus zeitlichen Gründen nicht mehr erneut eingelernt werden.

Das Tracking von **FAR** funktionierte gar nicht, da kein Werkzeug zum Entfernen fälschlich gelernter Hintergrundfeatures existierte. Darum stand die Augmentierung nahezu fest im Raum und bewegte sich nicht mit dem Auto mit.



Abbildung 19: Tracking des bewegten VW Modells. Der Monitor auf der rechten Seite zeigt das augmentierte AVILUS Logo.

In der zweiten Stufe konnte das Modell lediglich von **metaio** getrackt werden – und dieses auch nur in den seltenen Fällen, in denen die Perspektive günstig war, und das Auto sich gerade mal nicht bewegte.

3.5.3 Diskussion der Technologiepartner

Im Folgenden nehmen die Technologiepartner Stellung zu den Ergebnissen.

FAR: Das Objekt konnte nicht getrackt werden, weil nur ein kleiner Teil der gefundenen Punktmerkmale auf dem zu trackenden Auto lagen. Viele weitere Punktmerkmale wurden auf stationären Hintergrundobjekten im Sichtbereich der Kamera gefunden. Da wir im Rahmen des Tracking Contest nur ein reines Online Verfahren zur Verfügung gestellt haben, und keine Werkzeuge zur manuellen Entfernung von Punkten aus der Punktwolke, welche stationären Szenenbereichen zuzuordnen waren, hatten wir keine Einflussmöglichkeit auf die Selektion der Features, die zum Tracken benutzt worden sind. Somit wurde die Pose der Kamera in Bezug zum stationären Gesamtszenario und nicht in Bezug zum sich bewegenden Modell bestimmt. Prinzipiell wäre es durchaus möglich, automatische Verfahren zur Segmentierung der Punktwolken hinsichtlich ihrer zeitlichen Veränderung einzusetzen. Dieses wird in der Automobilindustrie bei der Nutzung von Sensoren für Fahrerassistenzsysteme gemacht, wenn zwischen verschiedenen Verkehrsteilnehmern mit unterschiedlicher Geschwindigkeit und Fahrtrichtung unterschieden werden soll. Auch die Daten der neuen Microsoft Kinect werden dementsprechend analysiert. Des Weiteren kommen solche Ansätze beim People Tracking mit stationären Einzelkameras zum Einsatz [31]. Zurzeit sind sie aber nicht mit den von PTAM erzeugten Tracks der Punktwolken integriert.

IGD: Zur Rekonstruktion der Feature Map wurde eine Sequenz mit rotierendem VW aufgenommen. Während der Vorbereitung war jedoch eine automatisierte Rekonstruktion ohne Marker (wie oben beschrieben durch automatische Translationsdetektion) nicht möglich, da zu viele Features im Hintergrund dies verhinderten. Durch manu-

ellen Eingriff (interaktive Initialisierung von Features ausschließlich auf dem Modell) konnte das Problem behoben werden. Featurepunkte auf den Kalibrierpunkten wurden manuell aus der Rekonstruktion entfernt, da diese während des Testlaufes verdeckt waren.

Während der Evaluierung / beim Tracking zeigten sich Probleme bei der Initialisierung des Trackings. Der wahrscheinliche Grund war wohl, dass sich die Features auf dem Modell untereinander zu ähnlich waren (insbesondere in Anbetracht der Symmetrie des Objekts) bzw. nicht stabil genug (aufgrund von Reflektionen / Schattenwurf). Das Anlernen des Objektes von nur einer Seite, ließ im zweiten Durchlauf eine Initialisierung zu. Eine Initialisierung basierend auf Referenzbildern (statt auf Feature-basis) könnte besser funktionieren, vor allem wenn man eine große Menge an Referenzbildern in Betracht zieht. Aus zeitlichen Gründen wurden aber leider keine alternativen Initialisierungsansätze ausprobiert. Einmal initialisiert funktionierte das Bild-zu-Bild Tracking mittels KLT einigermaßen, setzte aber bei höheren Drehgeschwindigkeiten aus.

Während der Evaluierung / beim Tracking zeigten sich Probleme bei der Initialisierung des Trackings. Der wahrscheinliche Grund war wohl, dass sich die Features auf dem Modell untereinander zu ähnlich waren (insbesondere in Anbetracht der Symmetrie des Objekts) bzw. nicht stabil genug (aufgrund von Reflektionen / Schattenwurf). Das Anlernen des Objektes von nur einer Seite, ließ im zweiten Durchlauf eine Initialisierung zu. In der Nachbetrachtung könnte eine Initialisierung basierend auf Referenzbildern (statt auf Feature-basis) besser funktionieren. Einmal initialisiert funktionierte das Frame-to-Frame Tracking mittels KLT einigermaßen, setzte aber bei höheren Drehgeschwindigkeiten aus.

Allgemein lässt sich sagen, dass ein solches Szenario wie das Miniaturmodell eine VW Golfes mit Punktbasierten Trackingmethoden nur sehr schwierig zu verfolgen ist. Andere Kontur-basierten Verfahren, wie [32] sind deutlich vielversprechender, benötigen jedoch als Input ein CAD-Modell.

metaio: Zunächst war eine Initialisierung und Tracking nur in seltenen Fällen erfolgreich. Erst nach der Änderung des Szenarios und der Öffnung sämtlicher Fahrzeugtüren war dies einigermaßen zuverlässig möglich.

Je schneller die Bewegungen des Objekts ausgeführt wurden, desto schlechter wurde jedoch auch die Trackingqualität. Dies kann zum Einen durch den entstehenden motion blur in den Kamerabildern erklärt werden, und zum Anderen dadurch, dass sich bei schnellerer Bewegung auch die zeitlich benachbarten Kamerabilder im Videostrom stärker unterscheiden und ein sogenanntes Interframe-Tracking dadurch erschweren.

Was sich für metaio ausgezahlt hat, war die in Unifeye integrierte Möglichkeit der Nachbearbeitung einer angelernten Feature Map durch den Benutzer. Dadurch war es möglich, aus der während des Anlernens generierten Punktwolke grob all jene Punkte zu entfernen, die nicht zum Trackingobjekt gehören.

3.6 Diskussion des Anwenders

Im Folgenden bewertet Volkswagen die Ergebnisse auf Basis eines Fragenkatalogs.

Ergebnisbewertung, Potentiale und Forschungsbedarfe

F: *Wie bewerten Sie die Ergebnisse des Tracking Contests für Ihr Szenario?*

A: Die Ergebnisse aus dem Tracking Contest sind für die Volkswagen AG von großem Interesse. Sie lassen einen deutlichen Fortschritt im Technologiefeld Markerloses Tracking erkennen. Die Ergebnisse zeigen aber auch deutlich, dass die Technologie noch nicht produktiv eingesetzt werden kann.

F: *Was funktioniert im Vergleich zu vorherigen Erfahrungen und Ergebnissen? Wie gut funktioniert es?*

A: Im Vergleich zu vergangenen Projekten konnten alle Teilnehmer durch robuste Systeme überzeugen (vor allem IGD und metaio). Änderungen in der Umgebung wie beispielsweise das Entfernen von Bauteilen oder die Veränderung der Beleuchtungssituation können sehr gut kompensiert werden. Auch die Nutzung von selbstlernenden Algorithmen ermöglicht Flexibilität, unbekannte Objekte ohne großen Vorbereitungs-aufwand zu erfassen.

Bezogen auf das Fahrzeugszenario (Szenarien "Motorraum" und "Interieur") ist die Genauigkeit der Systeme ausreichend (sehr gut, metaio - gut, IGD). Die kontinuierliche Überlagerung beim Bewegen der Kamera konnte nur vom IGD zufriedenstellend dargestellt werden.

Sehr interessant ist das Erfassen von bewegten Objekten bzw. das Trennen von Tracking-Objekt vom Hintergrund. Im Szenario "bewegtes Fahrzeug Modell" konnte nur metaio überzeugen. Es konnte eine einfache Drehbewegung erfasst werden. Komplexe Bewegungen konnten von keinem Partner erfasst werden.

F: *Was funktioniert noch nicht?*

A: Die kontinuierliche Überlagerung konnte nur vom IGD zufriedenstellend dargestellt werden. Bei den anderen Partnern führen Unterbrechungen im Tracking, bzw. Fehlerkennungen zu Sprüngen in der Überlagerung.

Das Szenario "Fahrzeug Exterieur" konnte von allen Partnern nicht in der geforderten Genauigkeit bearbeitet werden. Bewegte Objekte können bisher nur sehr eingeschränkt erfasst werden (siehe oben).

Die Verwendung von selbstlernenden Algorithmen ist leider noch nicht ohne Einschränkung möglich.

F: *Welche Entwicklungen sind nötig, damit ein Endanwender diese Technologie wirklich einsetzen kann?*

A: Für die Akzeptanz der Anwender ist eine kontinuierliche Überlagerung sehr wichtig – vor allem in Training- bzw. Marketinganwendungen. Das Erfassen von bewegten Objekten und die Abdeckung des Szenarios "Exterieur" ist eine Grundvoraussetzung zur Nutzung der Technologie in der Fertigung.

Für einen Einsatz im Bereich Werkerassistenz sind zudem selbstlernende Algorithmen zwingend erforderlich. Eine Flexibilität in der Bewegung des Mitarbeiters und der gering zu haltende Vorbereitungs-aufwand sind dabei entscheidende Erfolgsfaktoren.

Wirtschaftliche Bedeutung und Potentialanalyse

F: *Wie hoch schätzen Sie den Umsatz pro Jahr in Mio Euro (gemittelt über einen Zeitraum von 5 Jahren ab Einsatzjahr) für dieses oder vergleichbare Einsatzszenarien in Deutschland? Welches Potential steckt in der Technologie?*

A: Zur Zeit keine Aussage seitens der Volkswagen AG möglich.

Einordnung der Ergebnisse auf der Technologie Roadmap

F: *Wo befindet sich die Technologie auf der Technologie Roadmap? Wann wird diese Technologie den Technologie Readiness Level 6 (TLR6) erreicht haben (d.h.: wann wird es einen Prototypen in Einsatzumgebung geben)? Wie (un-)sicher ist diese Abschätzung?*

A: Zurzeit handelt es sich um einen erweiterten Forschungsdemonstrator. TLR6 könnte Ende 2012 erreicht werden. Diese Abschätzung ist noch recht unsicher.

Forschungsbedarf

F: *Was muss Ihrer Meinung nach passieren, damit die Technologie wirklich eingesetzt werden kann? Welcher Forschungsbedarf lässt sich hieraus ableiten? Welche Probleme müssen gelöst werden? Wie viel weiteres Potential steckt aus technischer Perspektive noch in dieser Technologie? Welches sind die notwendigen Schritte, um die Technologie weiter voranzubringen?*

A: Aus der Sicht von Volkswagen steckt ein sehr großes Potenzial in der Technologie in Bezug auf die Unterstützung von Werkerassistenz und Fahrzeug Service Anwendungen. Der Schlüssel zum Erfolg ist ein robustes System mit minimalem Referenzierungsaufwand, welches sich auf unterschiedliche Fahrzeugvarianten und Umgebungsbedingungen einstellt.

4 Testszenario: Tracking am Kuka Roboter

Kuka ist an der Computer Vision gestützten Steuerung ihrer Roboter interessiert. Zum Beispiel soll ein Roboter "grob" an ein zu greifendes Objekt herangeführt werden. Eine Kamera am Kopf des Roboters soll dann die Pose des Teils exakt bestimmen, so dass der Roboter das Teil greifen kann. Kuka möchte insbesondere Erfahrungen sammeln, wie gut die AVILUS-Trackingtechnologien sich hierfür eignen, da bisher weitestgehend herkömmliche (modell-basierte) Computer Vision Verfahren zu solchen Zwecken benutzt werden.

4.1 Versuchsszenario

Für den Versuch wurde eine Kamera am Kopf des Roboters befestigt (siehe Abbildung 20). Während des Versuchs wurde der Roboter an verschiedene Positionen relativ zu Objekten gefahren, die auf einer Tischplatte unter dem Roboter lagen. Aus den verschiedenen Perspektiven (Abbildung 21) musste vom Trackingsystem die Kamerapose relativ zu den Objekten ermittelt werden.

Die Versuchsreihe wurde in zwei Schritten durchgeführt. Im ersten Schritt musste die Pose mit Hilfe der jeweiligen Trackingtechnologie geschätzt werden, während die schwarzen Referenzmarken abgedeckt waren. Im zweiten Schritt wurde dann ein Bild mit freiliegenden Referenzmarken, aus gleicher Perspektive, aufgenommen. Für diese Stellung des Roboterarms wurde zusätzlich die "Soll-Pose" der Kamera mit Hilfe einer Offline-Messsoftware ermittelt. Für die Evaluierung wurden in der nachfolgenden Auswertung die kamera-getrackten Ist-Werte mit den gemessenen Soll-Werten in allen 6 Freiheitsgeraden verglichen.

4.2 Messergebnisse

Die Auswertung der Messergebnisse führte leider bei keinem der Teilnehmer zu einem verwendbaren Resultat. Bei den Demonstrationen von **metaio** und **FAR** lagen die ermittelten Pose-Werte weit außerhalb des Toleranzbereiches – d.h.: es war kein Tracking der Objekte möglich.

Bei der Demonstration des **IGD** wurde die Auflösung der Kamera oder die Kamera selbst zwischen Kalibrierenaufnahmen und Versuchsdurchführung versehentlich geändert. Somit konnten die Ergebnisse nicht ausgewertet werden.

Die Technologiepartner hatten Schwierigkeiten, die Objekte überhaupt tracken zu können. An dieser Stelle sei angemerkt, dass während der Versuchsplanung noch realistischere ("schwierigere") Objekte für den Test geplant waren. Das Tracking solcher Objekte wurde jedoch von den Technologiepartnern als nahezu unmöglich eingestuft. Die feature-basierten Verfahren, welche von den Technologiepartnern bevorzugt benutzt wurden, eignen sich somit nur sehr bedingt für das Tracking kleinerer, texturloser Objekte.

4.3 Diskussion der Technologiepartner

Im Folgenden nehmen die Technologiepartner Stellung zu den Ergebnissen.

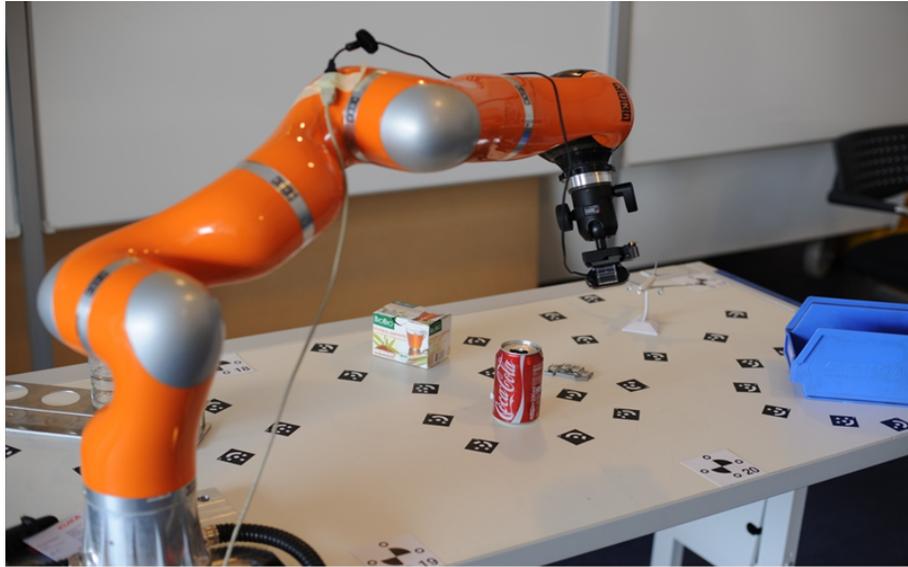


Abbildung 20: Versuchsaufbau zum Tracking am Kuka Roboter. Am Kopf befindet sich eine Kamera.

FAR: Die erzeugte Punktwolke konnte nicht ausreichend zur Umgebung registriert werden. Trotz des hohen Schwierigkeitsgrades (viel Symmetrie und wenig Textur der Objekte) konnte zwar eine Punktwolke aufgebaut werden, allerdings nur sehr spärlich. Die Kamerabewegung konnte zwar getrackt werden, aber die Registrierung war aufgrund der geringen Anzahl an Punktmerkmalen nicht gut. Normalerweise stabilisieren Punkte im Hintergrund das Tracking. Diese konnten in diesem Szenario aufgrund des homogenen Hintergrundes nicht gefunden werden. Möglicherweise wurde die Punktwolke durch die geringe Anzahl an Punkten nur verzerrt erkannt.

IGD: Während der Vorbereitung wurde für jedes der sechs Objekte eine eigene Featuremap erstellt, wobei zur Kalibrierung jeweils der gesamte Tisch rekonstruiert wurde, so dass auch alle vier bekannten Referenzpunkte mindestens einmal sichtbar waren (die übrigen Referenzpunkte hatten – obwohl ihre Koordinaten nicht bekannt waren – einen stabilisierenden Effekt auf die Rekonstruktion). Während des Durchlaufs wurden dann alle sechs Feature Maps durchprobiert, bis die entsprechende Map initialisierte.

Die Objekte waren unterschiedlich schwer zu tracken. Am einfachsten erkennbar waren die Tee-Box sowie die Klemme (gut texturiert, keine Reflexionen). Bei der Cola-Dose musste das Logo sichtbar sein, da ansonsten zu wenig Textur vorhanden war. Schwieriger war der blaue kastenförmige Behälter (wenige Features an den Ecken, dafür waren diese stabil), sowie das Flugzeug (besonders von der Vorderansicht: stabile Features nur am Cockpit, an den Düsen sowie am Sockel). Die Plastikflasche mit Metallfassung bot kaum stabile Features (Reflexionen und viele Verdeckungen).

Während des Durchlaufs konnten alle Objekte erkannt werden (wobei dies bei der



Abbildung 21: Perspektive der Kamera, aus der die Objekte getrackt werden sollten (Abb. 4.2)

Plastikflasche fast an ein Wunder grenzte). Auch konnte die Richtigkeit der Pose qualitativ festgestellt werden, da an den Referenzpunkten (sofern sichtbar) entsprechende Augmentierungen korrekt platziert wurden. Dass die Auflösung der Kamera zwischen Kalibrierung und Versuchsdurchführung verändert worden sein soll, ist für das IGD-Team nicht nachvollziehbar (Es wurden beide Mal 640x480 Bilder mitgegeben).

metaio: Für einige Objekte mit homogen-texturierter, "feature-ärmer" Oberfläche war die Detektionsrate schlecht und somit der Schwachpunkt in diesem Szenario (Modellflugzeug und blaue Kiste).

Potential besteht hierbei im Bereich der Kantenerkennung, da der verwendete Algorithmus im Grundprinzip nach markanten Punkten sucht und diese trackt. Eine Kombination mit Kantenerkennungsalgorithmen dürfte qualitative Verbesserungen hinsichtlich schlecht texturierter Objekte erzielen.

Die anderen Objekte konnten in den Durchläufen durchweg erkannt werden, und

anhand der richtigen Augmentierung bekannter Referenzpunkte dürfte die Pose einigermaßen korrekt gewesen sein.

4.4 Diskussion des Anwenders

Im Folgenden bewertet Kuka die Ergebnisse auf Basis eines Fragenkatalogs.

Ergebnisbewertung, Potentiale und Forschungsbedarfe

F: *Wie bewerten Sie die Ergebnisse des Tracking Contests für Ihr Szenario? Was funktioniert im Vergleich zu vorherigen Erfahrungen und Ergebnissen? Wie gut funktioniert es? Was funktioniert noch nicht? Welche Entwicklungen sind nötig, damit ein Endanwender diese Technologie wirklich einsetzen kann?*

A: In Anbetracht der Messergebnisse, laut denen die Messwerte weit außerhalb des Toleranzbereiches lagen, war somit ein Tracking zum Zweck der Objektwiedererkennung und Lagebestimmung unmöglich.

Wie beim Contest zu sehen war, konnten nur Objekte mit Textur getrackt werden. Meist besitzen nur für Verkauf bestimmte Konsumartikel eine solche Textur. Eine große Klasse von Objekten, wie Halbzeuge in der Fertigung, besitzen hingegen keine.

Um das Verfahren allgemein für die Objekterkennung in der Robotik einsetzen zu können, sollten die Genauigkeit, die Robustheit und die Klasse der "trackbaren" Objekte verbessert bzw. vergrößert werden.

Wirtschaftliche Bedeutung und Potentialanalyse

F: *Wie hoch schätzen Sie den Umsatz pro Jahr in Mio Euro (gemittelt über einen Zeitraum von 5 Jahren ab Einsatzjahr) für dieses oder vergleichbare Einsatzszenarien in Deutschland? Welches Potential steckt in der Technologie?*

A: Der Umsatz ist schwierig anzugeben, da es für die genannten Aufgaben bereits Verfahren gibt (industrielle 2D-Bildverarbeitung), die jedoch nicht die Flexibilität des Trackings in Kombination mit dem automatischen Einlernen bieten. Zudem ist die Technologie im Umfeld der Robotik eine Hilfstechnologie, deren Erfolg von weiteren Fortschritten auf dem Gebiet der Robotik und deren Zusammenwirken, wie der Mensch-Maschine-Kooperation und der Flexibilität der Greifertechnologie, abhängig ist.

Einordnung der Ergebnisse auf der Technologie Roadmap

F: *Wo befindet sich die Technologie auf der Technologie Roadmap? Wann wird diese Technologie den Technologie Readiness Level 6 (TLR6) erreicht haben (d.h.: wann wird es einen Prototypen in Einsatzumgebung geben)? Wie (un-)sicher ist diese Abschätzung?*

A: Das grundlegende Verfahren zur Nutzung markerlosen Trackings für die Objekterkennung in der Robotik ist beschrieben, und derzeit wird am Nachweis der Funktionsfähigkeit gearbeitet. Daher kann man es in TLR3 einordnen. Das Erreichen von TLR6 hängt, wie schon erwähnt, von mehreren Faktoren ab, so dass eine Einschätzung derzeit nicht möglich ist.

Forschungsbedarf

F: *Was muss Ihrer Meinung nach passieren, damit die Technologie wirklich eingesetzt werden kann? Welcher Forschungsbedarf lässt sich hieraus ableiten? Welche Probleme müssen gelöst werden? Wie viel weiteres Potential steckt aus technischer Perspektive noch in dieser Technologie? Welches sind die notwendigen Schritte, um die Technologie weiter voranzubringen?*

A: Die Technologie muss robuster und genauer werden.

Das Einlernen der Trackingobjekte muss automatisch ablaufen.

Getrackte Features sollten sich nicht auf die gesamte sichtbare Szene beziehen, sondern auf einzelne Objekte.

Sowohl Tracking (Objekt- und Lageerkennung) als auch Einlernen müssen schnell ablaufen. Im ein-, bzw. zweistelligen Sekundenbereich.

Angesichts der Fortschritte des Trackings in den letzten Jahren gehen wir davon aus, dass die Möglichkeiten dieser Technologie bei weitem noch nicht ausgeschöpft sind.

Erste Pilot-Anwendungen der Trackingtechnologie im Robotik-Bereich der Programmierung flexibler Aufgaben würden mit der Erweiterung der Möglichkeiten in jenem Bereich auch die Nachfrage nach Trackingtechnologien aus dem Bereich steigern.

5 Testszenario: Soll-Ist-Vergleich (HDW)

HDW stellte ein Szenario zum Soll-Ist-Vergleich als Unterstützung der Qualitätssicherung im U-Bootbau zur Verfügung. Dabei sollte das in der Fertigung auftretende Problem gelöst werden, Unterschiede zwischen einer geplanten Installation entsprechend einem CAD-Modell (Soll) und der tatsächlichen Realisierung (Ist) zu finden. Dafür wurde ein circa 0.5m*1.0m*0.5m großes (fehlerhaftes) Mockup eines U-Bootbauteils erstellt (siehe Abbildung 22). Der Versuchsort ist den Umgebungsbedingungen in einer Schiffsbauhalle nachempfunden (wenig Kontrast, Verdeckung durch andere Baumaterialien).

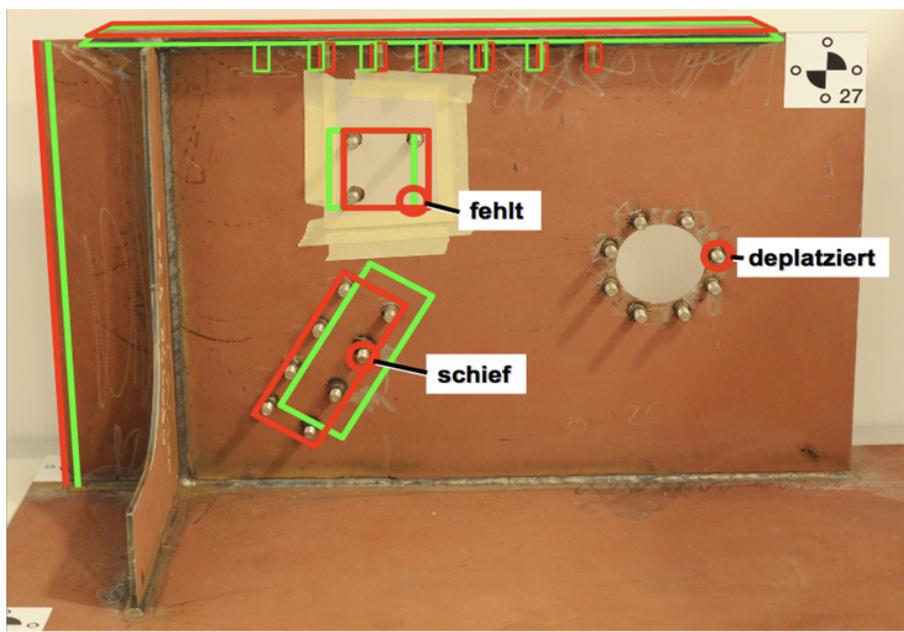


Abbildung 22: Abweichungen zwischen Mockup und CAD

5.1 Versuchsszenario

Das Versuchsszenario bestand aus zwei Aufgaben. Für beide Aufgaben sollte dem fehlerhaften Mockup ein CAD-Modell überlagert werden. In der ersten Aufgabe sollten neun Abweichungen (siehe Abbildung 22) zwischen CAD-Modell und Mockup entdeckt werden. In der zweiten Aufgabe sollte das Mockup unter erschwerten Bedingungen (Schattenwurf, Verdeckungen) getrackt werden (Abbildung 23).



Abbildung 23: Tracking des Objektes unter Normalbedingungen, unter Scheinwerferbestrahlung (starker Schattenwurf), und bei Verdeckungen durch Objekte und Personen

5.2 Messergebnisse

Folgende Anzahl an Abweichungen zwischen CAD-Modell und Mockup wurden korrekt entdeckt:

Die Qualität des Trackings wurde in Tabelle 7 zusammen gefaßt. Dabei ist zu beachten, dass die Verfahren von **FAR** und **IGD** vorübergehend nicht so gut funktionierten, wenn sich die Bedingungen während des Tests änderten (z.B. beim Einschalten eines Lichts). Die Systeme passten sich aber an die neue Situation an, indem sie neu aufgetretene Features automatisch hinzu lernten. Dadurch erhöhte sich die Tracking-Qualität dann längerfristig wieder.

Das Tracking von **metaio** war unter normalen Bedingungen stabil (4 von 5 Punkten) und sprang nur teilweise. Das CAD-Modell war circa 0.5-1cm zu weit rechts re-

| metaio | IGD | FAR |
|--------|-----|-----|
| 6 | 5 | 6 |

Tabelle 6: Gefundene Abweichungen zwischen CAD-Modell und Mockup im U-Boot Szenario (von 9 tatsächlich vorhandenen Abweichungen).

| Bedingung | metaio | IGD | FAR |
|------------------------|--------|-----|-----|
| normal | 4 | 5 | 4.5 |
| Übergangsphase | - | 2.5 | 3 |
| Nach erneutem Anlernen | 3.5 | 3 | 4.5 |

Tabelle 7: Trackingqualität beim Uboot Szenario (1-5 Punkte).

gistriert. Mit Licht war die Überlagerung leicht verzerrt. Die Pose sprang teilweise hin und her, aber die Überlagerung befand sich immer in einem robusten Toleranzbereich (nicht genau überlagert, aber keine wirklichen Ausreißer) (3,5 von 5 Punkten). metaio verwendete als Einziger die gleiche Feature Map unter den verschiedenen Bedingungen. Dadurch war der benötigte Aufwand zum Anlernen des Szenarios kleiner und ein manuelles Umschalten der Tracking Algorithmen und Feature Maps bei sich ändernder Beleuchtung im Vergleich zu anderen Trackern nicht nötig. Das Tracking stieg erst bei einer sehr starken Verdeckung durch Objekte und Personen aus.

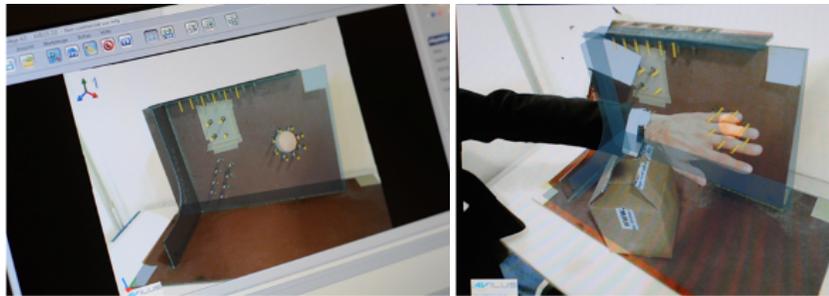


Abbildung 24: Tracking Beispiele metaio

Unter normalen Bedingungen war das **FAR** Tracking sehr stabil (4,5 von 5 Punkten). Es war aber schlecht registriert. Darum war die Überlagerung ungenau. Wurden die Bedingungen verändert (Licht oder verdeckende Objekte), benötigte das Verfahren einige Zeit, um die neuen Features anzulernen. Danach lief das System dann aber wieder genau so stabil wie unter normalen Bedingungen (4,5 von 5 Punkten).

Beim **IGD** war unter normalen Bedingungen die Überlagerung bei der Verwendung vieler Features nahezu perfekt (5 von 5 Punkten). Durch die Verwendung vieler Features war das Verfahren jedoch langsamer als die von FAR und metaio. Wurden weniger Features verwendet, war das Verfahren des IGD ähnlich performant und robust wie dass von metaio. Wurden die Bedingungen verändert (Licht/verdeckte Objekte),

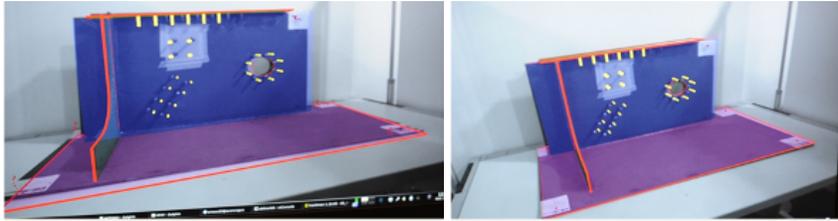


Abbildung 25: Tracking Beispiele FAR

verhielt sich das Verfahren, wie das von FAR: es mussten neue Features angeleitet werden, bis das Tracking wieder stabil lief. Insgesamt funktionierte das Verfahren des IGD unter den schwierigeren Bedingungen am schlechtesten (3 von 5 Punkten).

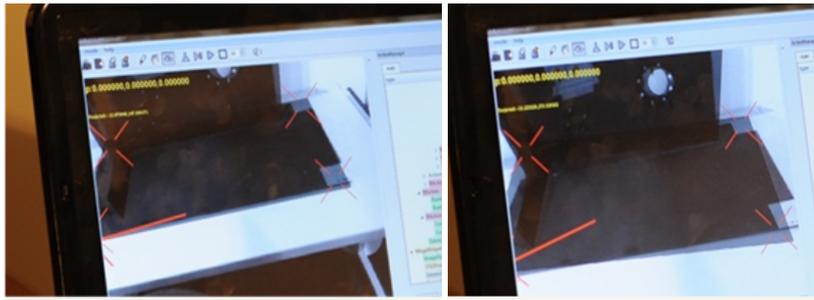


Abbildung 26: Tracking Beispiele IGD

5.3 Diskussion der Technologiepartner

Im Folgenden nehmen die Technologiepartner Stellung zu den Ergebnissen.

FAR: Wie in vorangehenden Szenarien (z.B. Kapitel 3) wurde die Registrierung zwischen Feature Map und Objekt über einen Flachmarker hergestellt, der während der Vorbereitungsphase auf das Objekt gelegt wurde. Während des Testlaufs erwies sich das Featuretracking als extrem robust unter verschiedenen Bedingungen (Beleuchtung / Bewegung / Verdeckung). Das System war in der Lage, einige der Fehler zuverlässig aufzudecken. Es zeigte sich jedoch ein perspektivenabhängiger systematischer Verzerrungseffekt. Während aus frontalen Betrachtungspositionen das Modell gut registriert erschien, verschlechterte sich die Registrierung mit seitlicheren Betrachtungspositionen, was die Beurteilung der Korrektheit einiger Bauteilpositionen erschwerte. Dieses Phänomen ist uns bereits in der Vorbereitungsphase aufgefallen. Als Kompromiss wurde die Registrierung deshalb auf eine frontale Betrachtungsposition hin optimiert. Grund für diese Ungenauigkeiten sind vermutlich abermals Verzerrungseffekte in der Feature Map, welche sich in Richtung randständiger Keyframes verstärken, ein Effekt der bereits in den Längenmessungen in Kapitel 3.4 offenbar wurde.

IGD: Das Mockup von HDW bestand zwar nicht aus vielen texturierten Flächen, aber die erkennbaren Punkte reichten aus, um eine Punktwolke zu rekonstruieren und das Szenario damit erfolgreich zu tracken. Ein paar Abweichungen waren offensichtlich zu erkennen. Andere Abweichungen, wie z.B. die obere horizontale nach hinten versetzte Platte, waren in der Überlagerung nicht erkennbar, was auch an der Perspektive des Betrachters gelegen hat. Dass nicht alle Unterschiede erkannt werden konnten, lag an einer unvoreilhaftigen Visualisierung und einem nicht so geschultem Blick der Test-Durchführer. Das Tracking an sich bereitete keine Schwierigkeiten.

Es wurde eine einzige Map in einem Vorverarbeitungsschritt rekonstruiert. Sowohl die Initialisierung mit Randomized Trees als auch das KLT-Feature-Tracking funktionierte damit einwandfrei bei konstanten Lichtverhältnissen. Auch Verdeckungen beeinträchtigten die Trackingqualität nicht. Bei starken Beleuchtungsänderungen gab es wiederum Schwierigkeiten mit der Randomized-Tree-basierten Initialisierungskomponente. Jedoch war es möglich mit dem Extensible Tracking Modul zusätzliche Bildmerkmale zu rekonstruieren und damit das Tracking zu stabilisieren. Die initiale im Vorfeld rekonstruierte Map war jedoch immer dieselbe.

Eine Überlegung war auch, mit dem vorhandenen CAD-Modell ein kontur-basiertes Trackingmodul zu verwenden [citeWuestEtAl07a](#), welches sich für derartige Szenarien mit vielen Kanten und starken Beleuchtungsänderungen sehr gut eignen würde. Da für solch ein Verfahren jedoch angenommen werden muss, dass das CAD-Modell genau mit der Realität übereinstimmt, was in diesem Test nicht der Fall war, könnte man diese Trackingkomponente zwar verwenden, jedoch würden die Fehler vom Modell direkt zu Fehlern in der Kameraposeberechnung führen.

metaio: Wie in allen anderen Szenarien auch, verwendet metaio eine einmalig angelegte Map (angelernte 3D-Punktwolke von Features) und verwendet diese bei allen unterschiedlichen Beleuchtungsvarianten. Ein manuelles Umschalten der Algorithmen/Karten findet nicht statt.

Potential sieht metaio im Bereich der Kantendetektion, da der verwendete Algorithmus im Grundprinzip nach markanten Punkten sucht und diese trackt. Eine Kombination mit Kantenerkennungsalgorithmen dürfte qualitative Verbesserungen erzielen.

Für realitätsnähere Tests sollte zukünftig beachtet werden, dass das zu verwendende Szenario den gesamten Kamerabereich der Tracker ausfüllt, so dass ausgeschlossen wird, dass bei derartigen "feature-armen" Szenarien Punkte entlang der Kanten des Mockups verwendet werden.

Einige der zu erkennenden Fehler lagen in sehr geringen Millimeter-Abweichungen, wodurch die Qualität der globalen Registrierung bereits zu schlecht war, um diese Fehler zu erkennen.

Was sich als besonders einflußreicher Faktor in diesem Szenario herausgestellt hat, war die Art der Visualisierung des zu überlagernden CAD-Modells. D.h. unabhängig von den Trackingtechnologien sollte bei derartigen Szenarien (Soll-Ist-Abgleich) auch Wert auf die Aufbereitung und Darstellung der CAD-Modelle gelegt werden.

Gegenüber partiellen Verdeckungen und Veränderungen des Objekts erwies sich der verwendete Algorithmus als sehr robust (siehe Abbildung 24). Allerdings wurde die Qualität des Trackings dennoch mit zunehmendem Verdeckungsgrad instabiler.

6 Testszenario: EADS Flugzeugkabine

Ziel des EADS Szenarios war die Evaluierung von Marker-unterstützten markerlosen Tracking Verfahren im Flugzeugbau. Dafür wurde ein Mockup einer A340 Kabine herangezogen (siehe Abbildung 27). Den Technologiepartnern wurde eine Referenzplatte zur Verfügung gestellt, welche mit einem Marker bestückt war (siehe Abbildung 28). Diese Platte konnte an definierten Messpunkten im Flugzeug befestigt werden. Sie ermöglichte ein sehr schnelles Einmessen in das Flugzeugkoordinatensystem durch eine Platzierung auf einer von 14 bekannten Messpunkten.



Abbildung 27: Mockup einer Flugzeugkabine bei EADS

6.1 Versuchsszenarien

EADS stellte 3 sehr eng miteinander verbundene Szenarien zur Verfügung, die in den Versuchen der Technologiepartner in einer Einheit bearbeitet wurden, und für die jeweils die gleiche Feature Map verwendet wurde.

6.1.1 Versuchsszenario 1: AR Überlagerung eines 3D-Modells

Es sollte ein VRML Modell als Augmented Reality Visualisierung aus verschiedenen, fest vorgegebenen Perspektiven im Mockup angezeigt werden (siehe Abbildung 29). Die Kamera konnte dabei langsam bewegt werden. In einer schwierigeren Variante sollte das 3D-Modell kontinuierlich überlagert werden, während sich die Kamera frei bewegte. Dafür wurden vier Laufwege definiert (siehe Abbildung 30).

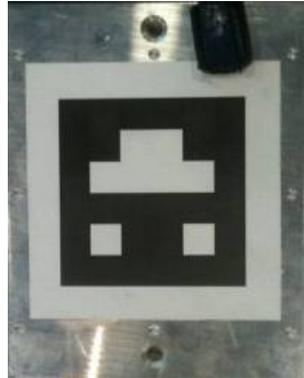


Abbildung 28: Marker in der Flugzeugkabine

6.1.2 Versuchsszenario 2: Finden von 3D-Punkten

Um die Genauigkeit der Verfahren im Mockup zu evaluieren, wurden Punkte, die im Mockup mittels eines Theodoliten vermessen worden waren, mit einem UV-Stift markiert. Für den Versuch erhielten die Technologiepartner 3D-Koordinaten, welche mit einem Klebepfeil markiert werden mussten. Die Genauigkeit der Punkte wurde unter Zuhilfenahme einer UV-Lampe und eines Maßbandes überprüft (siehe Abbildung 31).

6.1.3 Versuchsszenario 3: Relative Genauigkeitsmessung

Mit Hilfe einer 1.443 Meter langen Linearschiene wurden relative Genauigkeitsmessungen durchgeführt. Dieses Szenario entspricht vom Prinzip her dem Genauigkeitsszenario von Volkswagen (siehe Kapitel 3.4). Allerdings ist die visuelle Umgebung anders und stellt dadurch auch andere Anforderungen.

Es wurde eine Kamera auf der Schiene in der Mitte des Flugzeugmockups angebracht und vor und zurück bewegt. Dabei musste durch Tracking des Flugzeuginterieurs die vorab unbekannte Länge der Schiene bestimmt werden. Der Versuch wurde unter drei verschiedenen Blickwinkeln durchgeführt (siehe Abbildung 32).

6.2 Messergebnisse

Im Folgenden werden die Messergebnisse für die EADS Szenarien präsentiert. Die Aufschlüsselung erfolgt nach FAR, IGD und metaio. Tabelle 8 gibt einen vergleichenden Überblick über die Genauigkeit der Punktmessungen der einzelnen Technologiepartner. Punkte in einer Reihe entsprechen dem gleichen Schwierigkeitsgrad.

6.2.1 Ergebnisse FAR

Bei FAR wurde zunächst in einem ausgiebigen Prozess eine Feature Map angelernt. Dafür wurde die Kamera circa 15 Minuten lang durch das MockUp bewegt. Durch dieses Vorgehen sollte die Map sehr stabil werden. Erst nachdem die Map vollständig

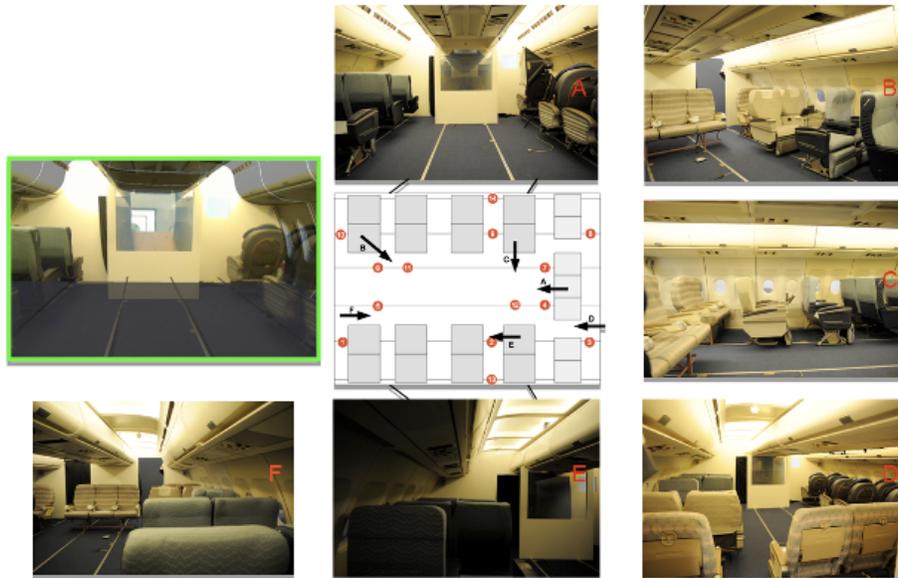


Abbildung 29: Perspektiven A-F in der EADS-Kabine. Aus den 6 Perspektiven soll ein VRML Modell überlagert werden (siehe grüner Kasten). Die roten Punkte markieren die möglichen Positionen der Markerplatte.

erlernt war, wurde sie zum Flugzeugkoordinatensystem registriert. Diese Registrierung fand über eine Befestigung der Markerplatte an zwei Positionen (vorne und hinten im Flugzeugmockup) statt und benötigte circa 2-5 Minuten. FAR verwendete als einziger Teilnehmer eine Industrie-Kamera (im Gegensatz zu Webcams bei IGD und metaio). Die Verwendung einer Industriekamera ermöglichte den Einsatz von höherwertigen und verzerrungsfreien Optiken. Der Nachteil war das sehr scharfe Bild, welches für die feature-basierten Algorithmen, nicht so gut geeignet war. Diese kamen besser mit einem unscharfen Bild zurecht. Darum verschlechterte FAR das Bild künstlich, um dadurch die Bildanalyse des Algorithmus optimal zu nutzen.

Das Verfahren von FAR, die Map zunächst aufzubauen und später zu registrieren,

| Abweichungen der Messpunkte | | | | | |
|-----------------------------|--------|-----|--------|--------|--------|
| FAR | | IGD | | metaio | |
| S2 | 44 cm | S3 | 50 cm | S3 | 4 cm |
| S4 | 55 cm | S5 | 1 cm | S6 | 2 cm |
| S7 | 19 cm | S8 | 9.5 cm | S9 | 0,5 cm |
| S10 | 46 cm | S12 | 10 cm | S11 | 2,5 cm |
| S14 | 8,5 cm | S13 | 6 cm | S15 | 2,5 cm |

Tabelle 8: Abweichungen der Messpunkte im EADS Mockup

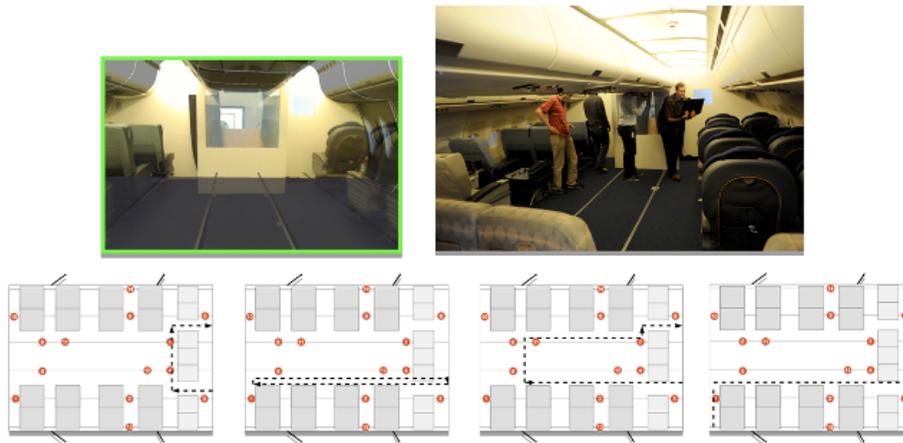


Abbildung 30: Vier verschiedene Laufwege, auf denen das VRML Modell kontinuierlich überlagert werden soll.

erwies sich als ungünstig. In dem großen Messvolumen akkumulierte sich der Drift-Fehler innerhalb der Map. Dieser Fehler konnte nicht mehr eliminiert werden. Dadurch war auch die Gesamtregistrierung der Map zu virtuellen Informationen sehr ungenau. Dieses schlug sich erheblich im Ergebnis nieder.

Die Ergebnisse zum ersten Experiment, der statischen Überlagerung, zeigten die Registrierungsprobleme (siehe Abbildung 33). Die Ergebnisse zur dynamischen Überlagerung waren ähnlich und können in den beigefügten Videos begutachtet werden.

Aufgrund der schlechten Registrierung waren auch die Ergebnisse der Genauigkeitsuntersuchung schlecht, wie die ersten beiden Spalten in Tabelle 8 und Abbildung 34 zeigen.

6.2.2 Ergebnisse IGD

Das IGD baute in der Lernphase vier lokale Feature Maps für das Flieger MockUp auf (eine in jede Himmelsrichtung). Diesen dienten zur Laufzeit als Startpunkte für markerloses Tracking. Das heißt, dass sich der Benutzer zum Starten, bzw wenn das Tracking einmal abbricht, immer wieder in eine der vier Richtungen bewegen musste. Die vier lokalen Feature Maps wurden jeweils über mindestens drei der 14 Kalibrierpunkte zum Flugzeug registriert. Die Registrierung war in das Bundle Adjustment der eigentlichen Feature Map integriert. Befand sich ein Marker im Sichtbereich, wurde die Feature Map zur Laufzeit korrigiert. Zur Laufzeit war zu beobachten, dass die Benutzer die Kamera mit einer "pumpenden" Bewegung durch den Raum bewegten. Dieses diente der Stabilisierung der Feature Map. In Arealen, in denen so gut wie gar keine Feature vorlagen, wurden ad-hoc bunte Papierblätter ausgestreut um Laufzeit Features zu erzeugen.

Die von IGD benutzte Visualisierung war nur zweidimensional und sehr schlecht



Abbildung 31: Messpunkt, der mit einer UV-Lampe sichtbar gemacht wurde. Die Klebpeile markieren die von den Trackern geschätzte Position.

gewählt. Darum wurde zur Verifizierung der anzusteuern Messpunkte (also zum Kleben der Pfeile) viel Zeit verbraucht, da diverse Perspektiven angesteuert werden mussten.

Die Ergebnisse zum ersten Experiment, der statischen Überlagerung, zeigten die relativ gute Registrierung (siehe Abbildung 35). Die Ergebnisse zur dynamischen Überlagerung waren ähnlich und können in den beigefügten Videos begutachtet werden.

Die Ergebnisse der Genauigkeitsuntersuchung sind in den mittleren zwei Spalten von Tabelle 8 und Abbildung 36 dargestellt.

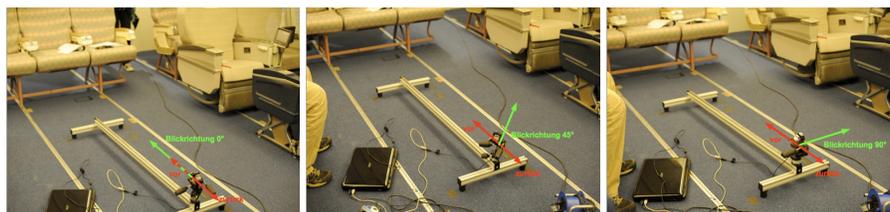


Abbildung 32: Die Länge einer Linearschiene musste dadurch bestimmt werden, dass das Flugzeuginterieur von einer Kamera auf der Schiene getrackt wurde. Dieser Versuch wurde unter drei Blickwinkeln durchgeführt.

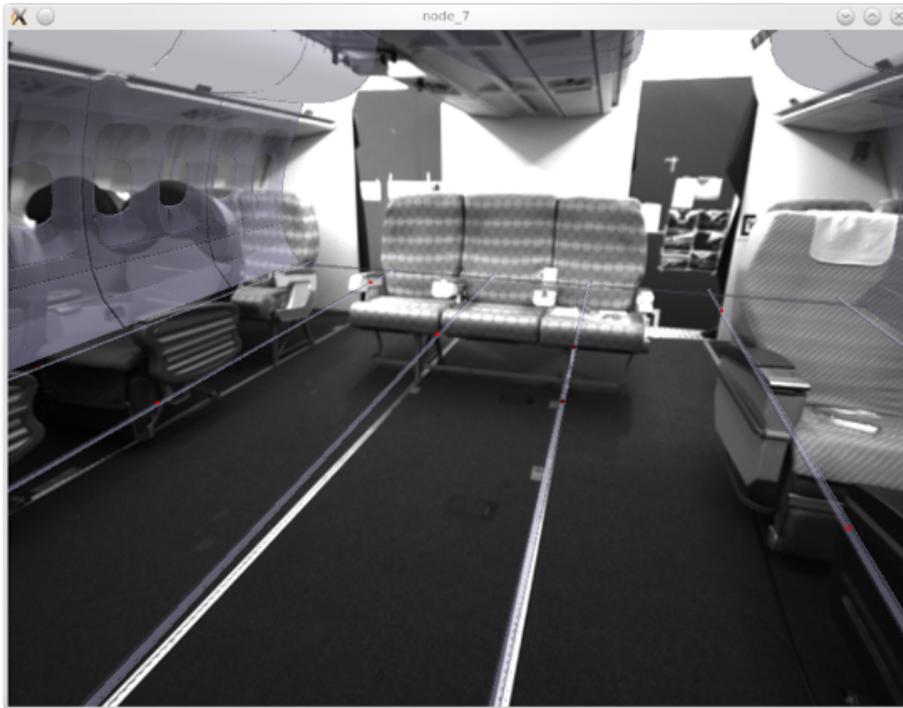


Abbildung 33: Überlagerung des VRML Modells, Beispiel FAR.

6.2.3 Ergebnisse metaio

metaio unterstützte den Aufbau der Feature Map mit Hilfe ihres Marker-Tracking. Zunächst wurden im ganzen Arbeitsbereich viele Marker platziert. Im ersten Schritt wurden Photos mit einer hochgenauen Spiegelreflexkamera von den Markern gemacht. Hierdurch konnte eine Inter-Marker-Kalibrierung (Bündelblockausgleich) auf hochauflösenden und nahezu verzerrungsfreien Bildern durchgeführt werden. Nach diesem Prozess wurde mittels der Tracking Kamera (herkömmliche Webcam), die Feature Map aufgenommen. Dabei befanden sich die Marker immer noch in der Szene. Dieses Vorgehen ermöglichte das Erstellen einer sehr guten (verzerrungs- und driftfreien) Feature Map. Das metaio System bot die Möglichkeit, nachträglich (zum Beispiel am nächsten Tag unter anderen Lichtbedingungen) weitere Features aufzunehmen und zu registrieren, ohne dass sich hierzu Marker im Bild befinden mussten. Es konnte jedoch nicht eindeutig beobachtet werden, dass letzteres Vorgehen die Qualität erhöht hat.

Die Ergebnisse zum ersten Experiment, der statischen Überlagerung, zeigen die relativ gute Registrierung (siehe Abbildung 37). Die Ergebnisse zur dynamischen Überlagerung sind ähnlich und können in den beigefügten Videos begutachtet werden.

Die Ergebnisse der Genauigkeitsuntersuchung sind in den rechten beiden Spalten von Tabelle 8 und Abbildung 38 dargestellt. Deutlich ist zu sehen, dass metaio die Punkte am genauesten trifft.



Abbildung 34: FAR Ergebnisse der Punktesuche.

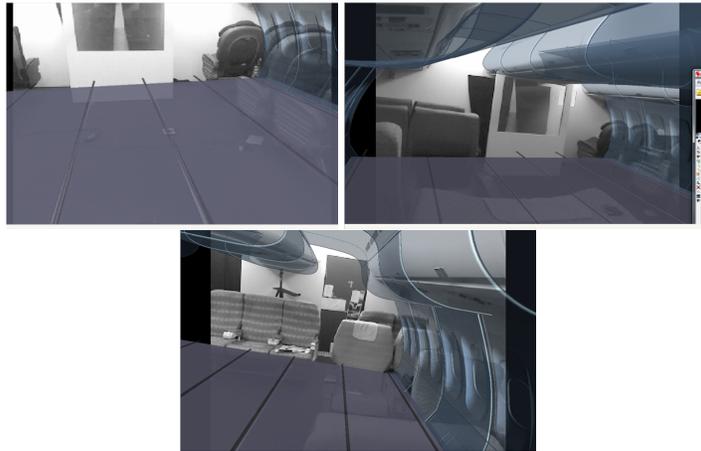


Abbildung 35: Überlagerung des VRML Modells, Beispiele IGD.

6.3 Vergleich Sonderaufgaben: Wide-area Genauigkeit

Zusätzlich wurde eine Sonderaufgabe gestellt, bei der aus dem "gelernten" Flugzeugkoordinatensystem ein 10-15 Meter entfernter Punkt in der "ungelernten" Werkshalle angesteuert werden musste. Diese Aufgabe wurde von IGD und metaio durchgeführt.

IGD baute dafür, wie gefordert, eine ad-hoc/extensible Feature Map ohne zusätzliche Hilfsmittel (Referenzmarken) auf. Der Punkte wurde um nur 1 Meter verfehlt. Dieses ist für den aktuellen Stand der Technologie eine sehr gute Leistung.

metaio löste die Sonderaufgabe marker-basiert (eigentliches Ziel wäre es gewesen, Tracking in der "unbekannten" Werkshalle ad-hoc/extensible zu testen). Zur Ansteuerung des Punktes wurden 6 Marker auf dem Weg verteilt. metaio gelang es damit, zwei Punkte auf wenige Zentimeter genau anzusteuern. Dieses trifft zwar keine Aussage über das Extensible Tracking, zeigt aber, dass die Verfahren zur Inter-Marker-

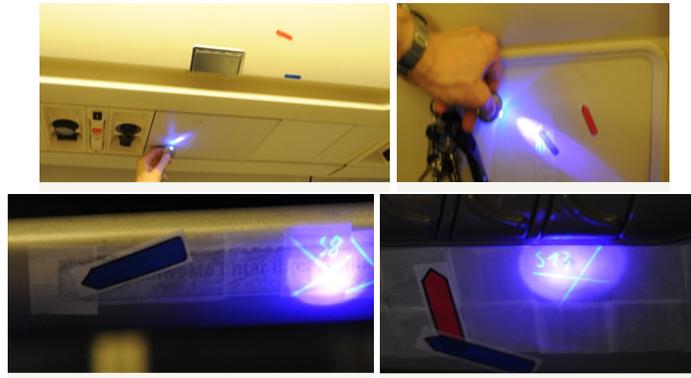


Abbildung 36: IGD Ergebnisse der Punktesuche.



Abbildung 37: Überlagerung des VRML Modells, Beispiele metaio.

Kalibrierung sehr gut sind.

6.4 Vergleich relative Genauigkeitsmessung

Die Ergebnisse zur Messung der relativen Genauigkeit innerhalb des EADS Mockups sind in Tabelle 9 zu sehen. Die Schiene hat eine Länge von 1443 Millimetern. Die Werte *avg* sind gemittelt über drei "Hin-und-Zurück" Fahrten. *ReInit* bedeutet, dass das Tracking zwischen einer "Hin-und-Zurück" Fahrt neu initialisiert wurde. *P(+/-)* bezeichnet jeweils die Präzision (Wiederholungsgenauigkeit) – also die maximale Abweichung zum Mittelwert der Messung *avg*. *G(avg)* bezeichnet die mittlere Genauigkeit – also die mittlere Abweichung zwischen Soll und Ist. Alle Angaben sind in Millimetern.

6.5 Diskussion der Technologiepartner

Im Folgenden nehmen die Technologiepartner Stellung zu den Ergebnissen.

FAR: Für die Überlagerungen in der EADS Kabine wurde eine einzige, sehr große Punktwolke erzeugt. Hier hat sich der schon früher erwähnte Verzerrungseffekt, der bei

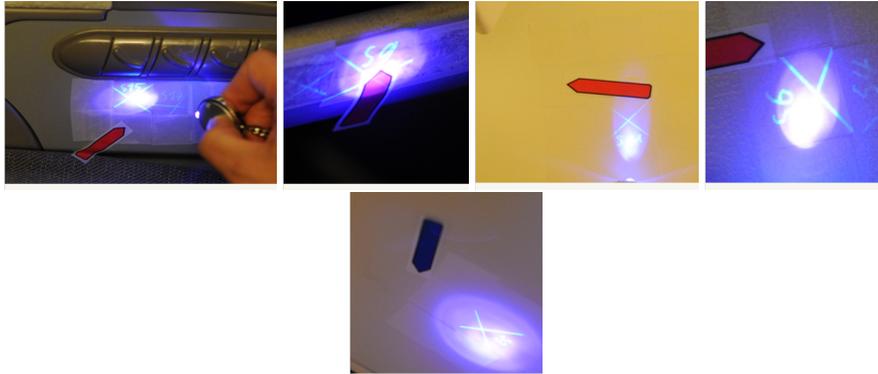


Abbildung 38: IGD Ergebnisse der Punktesuche.

| Win- kel | Re- Init | IGD | | | FAR | | | metaio | | |
|-------------|-------------|------|------------|------------|------|------------|------------|--------|------------|------------|
| | | avg | P (+/-) | G (avg) | avg | P (+/-) | G (avg) | avg | P (+/-) | G (avg) |
| 0 | | 1432 | 6 | -11 | 1480 | 1 | 37 | 1431 | 8 | -12 |
| 45 | | 1440 | 2 | -3 | 1444 | 2 | 1 | 1442 | 20 | -1 |
| 90 | | 1451 | 9 | 8 | 1461 | 1 | 18 | 1442 | 20 | -1 |
| 0 | x | | | | | | | 1435 | 11 | -8 |
| 45 | x | | | | | | | 1458 | 7 | 15 |
| 90 | x | | | | | | | 1458 | 7 | 15 |

Tabelle 9: Ergebnisse EADS Linearschiene

dem Featuretracking in großen Bereichen auftritt, extrem bemerkbar gemacht. Eventuell könnten mehrere kleinere Feature Maps Abhilfe schaffen, wenn diese geschickt im Raum verteilt sind. Diese müssten dann allerdings jeweils registriert werden.

Die sehr guten Wiederholgenauigkeiten bei den Genauigkeitsmessungen (Schiene) lassen auf ein sehr präzises Tracking schließen. Das sehr gute Ergebnis bei 45° zeugt erneut von einer sehr guten Kalibrierung des Featuretrackings zur Realität für diese Kameraansicht. Offensichtlich ergeben sich bei unterschiedlicher Neigung der Kamera aber keine akuraten Ergebnisse mehr. Dies könnte eventuell einer je nach Richtung unterschiedlich skalierten Punktwolke geschuldet sein, wobei der Fehler beim Aufbauen der Punktwolke erzeugt worden ist.

Der schon oben erwähnte Verzerrungseffekt kam bei der Genauigkeitsmessung (Punkte) ebenfalls zu tragen, so dass die anzufahrenden Punkte nicht exakt lokalisiert werden konnten.

IGD: In der Kabine wurde sowohl ein marker-basiertes Tracking mit anschließender Online-Rekonstruktion, als auch ein Tracking mit zuvor rekonstruierten Referenzdaten getestet. Die vom Marker ausgehende Online-Rekonstruktion funktionierte ausrei-

chend gut, solange der Benutzer die Kamera wohlwollend bewegte. Mit in einem Vorverarbeitungsschritt präparierten Trackingdaten ist der Benutzer deutlich flexibler. Insgesamt wurden vier Referenzdatensätze erstellt, für jede Richtung einer, wobei Überlappungen vorhanden waren. Die statischen Überlagerungen funktionierten gut. Bei dem Gang durch die Kabine konnte das Modell größtenteils auch lagerichtig dargestellt werden.

Bei der Genauigkeitsmessung auf der Schiene gab es ähnliche Probleme wie beim VW. Ein systematischer Fehler in der Kalibrierung verursachte einen ständig zu klein gemessenen Abstand.

Beim Finden der gesuchten Punkte waren Abweichungen zu erkennen. Einerseits lag es daran, dass die zusätzliche Online-Rekonstruktion Drift in die Feature Map mit aufnahm, andererseits wurden die Referenzdaten nicht sehr genau vom IGD kalibriert, da in diesem Szenario keine subpixel-genaue Vermessung von Referenzpunkten im Kalibrierungsschritt durchgeführt wurde.

metaio: Seitens metaio waren die Ergebnisse sehr zufriedenstellend, was an den äußerst geringen Abweichungen der aufzufindenden Messpunkte ersichtlich ist.

Der Ablauf zum Anlernen der zu verwendenden Feature Map war derselbe wie in den anderen Szenarien. In der Airbus Kabine wurden einige Marker (hier 6) verstreut platziert und Kamerabilder von dieser Szenerie aufgenommen. Diese Kamerabilder wurden zusammen mit der einmalig ermittelten und lange vorher bekannten Kamera-kalibrierung als Eingabe dem Rekonstruktions-Tool übergeben. Da die hierdurch automatisch generierte 3D-Punktwolke noch in einem "willkürlichen" Koordinatensystem lag, wurde sie anschließend zum Kabinen-Koordinatensystem registriert. Dies erfolgte durch bekannte 3D-3D Korrespondenzen. Der Benutzer hat in Unifeye die Möglichkeit, 2D-Bildmerkmale in den Kamerabildern zu markieren (z.B. die 2D-Positionen der Referenzmarken). Zu diesen sind durch automatische Triangulierung die 3D-Positionen im "willkürlichen" Koordinatensystem bekannt, wodurch anhand der bereitgestellten, ausgemessenen 3D-Koordinaten der Referenzmarken im globalen Koordinatensystem 3D-3D Korrespondenzen existieren. Durch mindestens drei solcher Korrespondenzen kann eine erfolgreiche Registrierung in das Zielkoordinatensystem durchgeführt werden.

Nach dieser Registrierung kann die resultierende Feature Map zum Tracking verwendet werden.

Die Zusatzaufgabe, aus der Kabine heraus einen auf dem Laufweg ca. 20m entfernten Punkt anzusteuern (eine blaue 5cm x 10cm grosse Schublade in einem Werkzeugkasten von ca. 50 identisch blauen Schubladen), konnte mit dem Extensible Tracking nicht in der geforderten Genauigkeit realisiert werden. Daher beschloss metaio, diese Zusatzaufgabe mittels Markertracking zu lösen, da Unifeye hierfür eine sehr genaue Möglichkeit der Intermarker-Kalibrierung anbietet. In diesem Fall wurden auf dem Weg aus der Kabine heraus bis zum Ziel 6 Marker "willkürlich" verteilt, sodass aus bestimmten Perspektiven immer zwei benachbarte Marker zusammen in einem Bild zu sehen waren. Durch die automatische Intermarker-Kalibrierung konnte sich das Markertracking ausgehend von der bekannten Startordinate bis hin zum aufzufindenden Ziel "durchhangeln". Die Genauigkeit lag dann immer noch in einem Bereich, der es

ermöglichte, die korrekte Schublade anzugeben.

6.6 Diskussion des Anwenders

Im Folgenden bewertet EADS die Ergebnisse auf Basis eines Fragenkatalogs.

Ergebnisbewertung, Potentiale und Forschungsbedarfe

F: *Wie bewerten Sie die Ergebnisse des Tracking Contests für Ihr Szenario? Was funktioniert im Vergleich zu vorherigen Erfahrungen und Ergebnissen? Wie gut funktioniert es? Was funktioniert noch nicht? Welche Entwicklungen sind nötig, damit ein Endanwender diese Technologie wirklich einsetzen kann?*

A: Die Lösungsvorschläge der drei Tracking Technologie Anbieter ermöglichen eine gute Abschätzung, unter welchen Bedingungen eine Industrialisierung zweckmäßig ist. Das zeitintensive (>20min) Aufbauen einer "Map" stellt aktuell insbesondere durch die Sensitivität hinsichtlich Änderungen an der Kabine (Sitzposition, Gurte, Lichtbedingungen) keine zweckmäßige Umsetzung in Aussicht. Bei der Inbetriebnahme müssen die Features der Szene aktualisiert werden. Zudem ist eine Skalierbarkeit auf eine gesamte Flugzeugkabine nicht gewährleistet und die Problematik der repetitiven Strukturen im Flugzeug für eine zuverlässige Initialisierung hinderlich.

Die ad-hoc Verfahren, die ohne vorheriges Einmessen mit Hilfe des Referenzmarkers starten, sind diesbezüglich praktikabler. Hinsichtlich eines Einsatzes durch einen ungeschulten Anwender war es jedoch offensichtlich, dass fehlendes Verständnis für die Eigenschaften eines monokularen Trackings zum Versagen des Systems führten (schnelle Perspektivenwechsel, zwischen Sitze, Rotation ...). Die Robustheit eines solchen Systems ist natürlich etwas eingeschränkt, da initial keine große Feature Map existiert, um eine Initialisierungsprozedur zuverlässig laufen zu lassen. Schwerpunkt der Entwicklung eines solchen Systems muss es sein, Features beim Verlust der Referenz weiterhin aufzubauen, und die dadurch entstehenden getrennten Maps bei Überlappung zu vereinigen. Zudem sollten Visualisierungen geprüft werden, die das Anwenderverhalten innerhalb der ersten 30 Sekunden dahin steuern, eine möglichst gute initiale Feature Map aufzubauen.

Wirtschaftliche Bedeutung und Potentialanalyse

F: *Wie hoch schätzen Sie den Umsatz pro Jahr in Mio Euro (gemittelt über einen Zeitraum von 5 Jahren ab Einsatzjahr) für dieses oder vergleichbare Einsatzszenarien in Deutschland? Welches Potential steckt in der Technologie?*

A: Bei der Unterstützung von Wartungs- und Fertigungsprozessen ist die Technologie vermutlich in 2-3 Jahren einsetzbar. Zahlreiche Einsatzszenarien sind hinsichtlich der voranschreitenden Digitalisierung der Fertigungsprozesse gegeben. Der Vorteil einer positions- und ortsabhängigen Visualisierung (Tracking) gegenüber klassischen Digitale-Fabrik Anwendungen ist vermutlich zuerst nur bei komplexen Aufgaben gegeben. Durch den noch verhältnismäßig hohen Fehler (2-3mm) lässt sich das Einsatzszenario auch nur für einfache qualitative Fehlererkennung nutzen.

Sollten Fehler frühzeitig erkannt und einfacher behoben werden, ließen sich bei der Einrüstung von Flugzeugen circa 50 Prozent des Zeitaufwandes für die Kontrol-

le sparen. Dies ist jedoch für den Fertigungsprozess unkritisch (Parallelarbeit). Jedoch reduziert sich die Wahrscheinlichkeit einer Verzögerung entsprechend. Bei einer Verzögerung der Produktion um 24h entstehen Kosten von circa 100.000 bis 500.000 Euro (Zinsfuß 12 Prozent p.A auf die Kosten der unterschiedlichen Flugzeugmodelle). Das Potential der Technologie liegt vermutlich bei 10 Prozent dessen – angenommen die Technologie ist hinsichtlich der Anforderungen der Anwendung ausgereift. (Inbetriebnahme, Mobilität, Genauigkeit und Qualität ...)

Einordnung der Ergebnisse auf der Technologie Roadmap

F: *Wo befindet sich die Technologie auf der Technologie Roadmap? Wann wird diese Technologie den Technologie Readiness Level 6 (TLR6) erreicht haben (d.h.: wann wird es einen Prototypen in Einsatzumgebung geben)? Wie (un-)sicher ist diese Abschätzung?*

A: Der Versuch, einen Prototyp zu installieren, beginnt nach aktueller Planung 2011. Eine Erreichbarkeit von TRL6 für die Trackingtechnologie für das Szenario sollte bis Ende 2012 machbar sein. Die TRL Bewertung der Anwendung liegt bei TRL1-2. Ein TRL 5 würde bereits klare Konzepte zur Integration in die Fertigungsprozesse erfordern – und dies ist vorerst nicht notwendig. TRL6 für die Anwendung ist 2014+.

Forschungsbedarf

F: *Was muss Ihrer Meinung nach passieren, damit die Technologie wirklich eingesetzt werden kann? Welcher Forschungsbedarf lässt sich hieraus ableiten? Welche Probleme müssen gelöst werden? Wie viel weiteres Potential steckt aus technischer Perspektive noch in dieser Technologie? Welches sind die notwendigen Schritte, um die Technologie weiter voranzubringen?*

A: Untersuchung zum Nutzen einer Positionserfassung (Tracking) bei der Visualisierung von Bauunterlagen.

Konzepte zur Visualisierung bei der Initialisierung/Start des Systems, um dem Anwender die notwendigen Bewegungsmuster zu vermitteln.

Erweiterung des Systems, um auch ohne Referenz bzw. bei Verlust der Referenz die Karte aufzubauen. Sofern die Referenz gegeben ist, sollten sich die Feature Maps registrieren/mergen. Eine automatische Identifikation von Überlappungen der Bereiche sollte zudem diese entsprechend zusammenführen können.

7 Testszenario: EADS Wellness-Bereich

Das Szenario des EADS Wellness-Bereichs zielt darauf ab, den vollständigen Workflow zu beobachten der durchgeführt werden muss, um ein unbekannte Szene anzulernen, so dass diese markerlos getrackt werden kann. Im einem zweiten Schritt soll dann ein CAD-Modell zu der Szene registriert und überlagert werden. Abbildung 39 zeigt ein Mockup des Wellnessbereichs, welches in diesem Szenario mit einem CAD-Modell überlagert werden sollte.



Abbildung 39: EADS Wellness Bereich

Die Aufgabe konnte von allen Teilnehmern erfolgreich gelöst werden. In den folgenden zwei Abschnitten werden die unterschiedlichen Vorgehensweisen der Teilnehmer diskutiert.

7.1 Diskussion der Technologiepartner

Im Folgenden nehmen die Technologiepartner Stellung zu den Ergebnissen.

FAR: Unser zweites Registrierungsverfahren konnte in diesem Szenario sehr gut zur Anwendung gebracht werden. Nach Anklicken der vorher bekannten Referenzpunkte hat sich die Geometrie gut mit der Realität überlagert. Auch ein Hineinzoomen in den Nahbereich des Objektes hat gut funktioniert.

IGD: Dieses Szenario war recht dünn mit gut texturierten Punktmerkmalen besetzt, konnte aber dennoch erfolgreich rekonstruiert und getrackt werden. Zur Kalibrie-

ung wurden manuell Korrespondenzen zwischen der rekonstruierten Punktwolke und Punkten auf dem gegebenen 3D-Modell verwendet.

Das Erstellen eines lauffähigen Trackingsetups lässt sich wie folgt beschreiben: Der Benutzer schwenkt seine Kamera über das Szenario, wobei ähnliche Kamerapositionen benutzt werden sollen, welche dann auch in der AR-Applikation von Bedeutung sind. Anhand der Bilddaten wird eine genaue 3D Punktwolke von Bildmerkmalen von dem betrachteten Szenario erstellt. Nach erfolgreicher Rekonstruktion besteht für den Benutzer die Möglichkeit durch die aufgenommene Bildsequenz zu browsen und die rekonstruierte Punktwolke manuell nachzubearbeiten. So kann den Benutzer durch klicken in nur einem einzigen Bild Punkte löschen und hinzufügen. Anhand der gespeicherten Bilder wird die 3D Position eines hinzugefügten Punktes automatisch rekonstruiert. In dem EADS Szenario war eine manuelle Nachbearbeitung der Punktwolke jedoch nicht nötig, da bereits genügend Punkte an aussagekräftigen Stellen rekonstruiert wurden. Zur Transformation der Punktwolke in das Modellkoordinatensystem müssen nun manuell mindestens 3 Korrespondenzen zwischen Rekonstruktion und 3D Modell bestimmt werden. Dies geschieht durch einfaches selektieren der Punkte mit der Maus. Es ist auf jeden Fall zu empfehlen mehrere Punktkorrespondenzen zu selektieren, da der durchschnittliche Fehler der beim Klicken zustande kommen, dann geringer ist. Nach erfolgreicher Transformation muss die Feature-Map lediglich abgespeichert werden und kann direkt für eine AR-Applikation genutzt werden.

Für die Erstellung der Rekonstruktion war keine Manipulation der Szene, wie das Anbringen von Markern, nötig. Die Randomized-Trees-basierte Initialisierung funktionierte in diesem Szenario gut. Mit dem Extensible Tracking Ansatz war es möglich die Kamera an Positionen zu bewegen, die nicht zur Maperstellung beigetragen haben. So konnte man z.B. sich sehr nahe in das Mockup-Modell hineinbewegen, ohne dass das Tracking abbriss.

metaio: Die Besonderheit dieses Szenarios war, dass keine Referenzmarken zur Registrierung der Koordinatensysteme vorhanden waren. Die Arbeitsschritte des metaio Workflows waren wie folgt:

Zwei Marker wurden neben dem Objekt platziert und 5 Fotos aufgenommen. In diesen Fotos wurden 4 ausgewählte Punkte des Objekts in den 2D-Bildern vom Benutzer markiert, sodass aufgrund der bekannten Posen (mittels Markertracking) eine automatische Rekonstruktion dieser 4 Punkte erfolgen konnte.

Die Registrierung der Koordinatensysteme erfolgte im Weiteren dadurch, dass im bereitgestellten CAD-Modell dieselben 4 Punkte vom Benutzer markiert wurden. Dadurch wurde automatisch eine Transformation generiert, welche als Eingabe für das folgende Tracking verwendet wurde.

Dieses Tracking erfolgte aufgrund der Natur des Szenarios ohne eine vorher angelegte Feature Map, sondern ähnlich dem VW Interieur Szenario des eingblendeten Markers im Navigationssystem.

D.h. anhand der platzierten Marker wurde das Extensible Tracking verwendet um eine Online Map des Objekts zu generieren, und die Marker nach ca. 10 Sekunden wieder entfernt, da die bis dahin generierte Map ausreichend war, um das Extensible Tracking ab diesem Punkt selbsttragend fortzuführen.

Der Gesamtaufwand dieser Arbeitsschritte lag unter 10 Minuten.

7.2 Diskussion des Anwenders

Im Folgenden bewertet EADS die Ergebnisse auf Basis eines Fragenkatalogs.

Ergebnisbewertung, Potentiale und Forschungsbedarfe

F: *Wie bewerten Sie die Ergebnisse des Tracking Contests für Ihr Szenario? Was funktioniert im Vergleich zu vorherigen Erfahrungen und Ergebnissen? Wie gut funktioniert es? Was funktioniert noch nicht? Welche Entwicklungen sind nötig, damit ein Endanwender diese Technologie wirklich einsetzen kann?*

A: Die grundsätzliche Zielsetzung, möglichst schnell und einfach zu einer AR Präsentation für das ausgesuchte 1:5 Design-Modell einer Toilette zu kommen, wurde erfüllt. Die unterschiedlichen Konzepte waren jedoch von der Anwendbarkeit recht unterschiedlich. Das intuitivste System war von metaio präsentiert und erlaubte das Selektieren einzelner Positionen im 3D-Modell sowie auf 2 Bildern des Objektes. In den anderen Systemen musste eine 3D-Koordinate des Modells mit einem explizit selektierten Feature aus der Map verbunden werden - was natürlich nicht immer möglich ist. Das Tracking lief bei allen Systemen zuverlässig und korrekt, wobei das IGD System am besten mit kleinen Ausschnitten zurecht kam - wenn die Kamera also weit in das Modell bewegt wurde. Das System von metaio hingegen war hinsichtlich der Reinitialisierung am besten, verwarf aber etwas häufiger eine gültige Schätzung als zu unsicher (Frame-to-Frame Tracking). Das System von FAR war am robustesten hinsichtlich schneller Bewegungen (Rotationsschätzung aus dem Motion-Blur) und konnte auch verhältnismäßig zuverlässig und einfach initialisieren. Der Registrierungsprozess von FAR war nur durch einen Experten möglich und nur rudimentär tool-gestützt.

Grundsätzlich lässt sich sagen, dass es auch einer in AR ungeschulten Person nach einer kurzen Einführung möglich ist, eine solche AR-Präsentation zu halten - zumindest mit dem Verfahren von metaio.

Wirtschaftliche Bedeutung und Potentialanalyse

F: *Wie hoch schätzen Sie den Umsatz pro Jahr in Mio Euro (gemittelt über einen Zeitraum von 5 Jahren ab Einsatzjahr) für dieses oder vergleichbare Einsatzszenarien in Deutschland? Welches Potential steckt in der Technologie?*

A: Zur Präsentation neuer Design-Konzepte werden häufig Modelle angefertigt. Diese haben jedoch den Nachteil, dass aus Kostengründen häufig Teile der Funktion nicht abgebildet werden können. Mittels AR lassen sich im vollen Umfang die Funktionen auch im Kontext des finalen Einsatzortes darstellen (z.B. innerhalb der Flugzeugkabine). Das gebaute Modell lässt sich natürlich nicht hinsichtlich möglicher Neuerungen am Konzept nachpflegen. Mittels AR lassen sich Varianten des Produktes sowie mögliche Entwicklungsschritte einfach visualisieren und vergleichen. Des Weiteren ist es mit AR möglich, einem speziellen Kunden ein an seine Farbwahl (Corporate Identity, CI) angepasstes Modell zu präsentieren.

Eine seriöse Abschätzung der Kostenersparnis oder des Mehrwertes ist für dieses Szenario nicht möglich.

Einordnung der Ergebnisse auf der Technologie Roadmap

F: *Wo befindet sich die Technologie auf der Technologie Roadmap? Wann wird diese Technologie den Technologie Readiness Level 6 (TRL6) erreicht haben (d.h.: wann wird es einen Prototypen in Einsatzumgebung geben)? Wie (un-)sicher ist diese Abschätzung?*

A: Die Technologie ist für diesen Zweck anwendbar. Bis auf kleinere Bereiche bei der Anwendbarkeit (IGD und vor allem FAR!) sowie bei der Robustheit (wie Frame-to-Frame Tracking bei metaio oder Reinitialisierung beim IGD) ist das Szenario abzudecken. Bei entsprechendem Interesse sollte sich TRL6 innerhalb von einem Jahr umsetzen lassen.

Forschungsbedarf

F: *Was muss Ihrer Meinung nach passieren, damit die Technologie wirklich eingesetzt werden kann? Welcher Forschungsbedarf lässt sich hieraus ableiten? Welche Probleme müssen gelöst werden? Wie viel weiteres Potential steckt aus technischer Perspektive noch in dieser Technologie? Welches sind die notwendigen Schritte, um die Technologie weiter voranzubringen?*

A: Der Forschungsbedarf bei diesem Thema ist sehr differenziert für die unterschiedlichen Technologielieferanten.

metaio: Frame-to-Frame Tracking, Motion-Blur für schnelle Rotationen

IGD: Reinitialisierung, Motion-Blur für schnelle Rotationen

FAR: Registrierungsprozess/Setup

8 Ergebnisbewertung, Potentiale und Forschungsbedarfe

Die Technologiepartner wurden gebeten, Fragen zur Ergebnisbewertung, zur Einordnung der Technologie auf der Technologie Roadmap zu beantworten und daraus Forschungsbedarfe abzuleiten.

8.1 Ergebnisbewertung

Wie bewerten Sie die Ergebnisse des Tracking Contest insgesamt? Was funktioniert im Vergleich zu vorherigen Erfahrungen und Ergebnissen? Wie gut funktioniert es? Was funktioniert noch nicht? Welche Entwicklungen sind nötig, damit ein Endanwender diese Technologie wirklich einsetzen kann? Welches waren die wichtigsten Erfahrungen, die Sie gemacht haben?

FAR: Die Ergebnisse zeigen insgesamt ein hohes Potential der Feature Tracking Algorithmen für das Tracking, auch in schwierigen Arbeitssituationen. Bevor ein Feature Tracking robust zur Anwendung gebracht werden kann, bedarf es allerdings noch eines hohen Zeitaufwands und Expertenwissens in der Vorbereitung. Meist muss der Bezug zur Realität, speziell die Skalierung, noch in Offline Verfahren hergestellt werden, und die Realität darf sich in dem Zeitraum zwischen Aufnahme der Daten und dem Starten des Trackingvorgangs selber nicht allzu sehr verändert haben. Es ist leider auch nicht immer vorherzusagen, was sich in der Realität verändern darf und was nicht, so dass momentan die Einschränkung gilt, dass sich möglichst gar nichts innerhalb des Zeitraums ändert.

IGD: Die Ergebnisse der Tracking Contests zeigten, dass ein sehr großer Teil der Anforderungen der Industriepartner im Bereich des kamera-basierten Trackings erfüllt werden konnten. Im Vergleich zu vorhergehenden Tracking Contests (ISMAR) konnten deutlich Fortschritte erzielt werden. Die Rekonstruktion von Referenzdaten, die Transformation in ein gewünschtes Koordinatensystem, und das Verfolgen des Kamerapfades funktionierte in vielen Fällen recht zufriedenstellend. Der Prozess zur Erstellung von Referenzdaten hat sich vereinfacht, was zu schnelleren Setupzeiten geführt hat. Selbst bei starken Lichtveränderungen war es möglich, die rekonstruierten Referenzdaten zum Tracken zu benutzen. Allerdings war auch deutlich zu erkennen, dass bei veränderten Beleuchtungsverhältnissen, das Trackingsystem nicht so robust arbeitet. Vor allem bei der Initialisierung gab es Schwierigkeiten. Geringfügige Manipulationen der Szene (weniger als 50% Veränderung zum angelernten Zustand) bereiteten jedoch keine Probleme.

Schwierigkeiten gab es bei Objekten, welche aus wenig texturierten Flächen bestanden. Das bewegte VW-Modell auf dem Kuka Roboter, sowie die Modelle des anderen Kuka Szenarios boten für Punktmerkmals-basierte Trackingansätze zu wenig Struktur.

Damit ein Endanwender die Technologie benutzen kann, müsste der Erstellungsprozess noch etwas einfacher gestaltet werden. Zudem müssen einem Endanwender besser aufbereitete Informationen gegeben werden, um zu erkennen, ob das Tracking mit einer erstellten Feature Map gut funktioniert, oder ob an bestimmten Bereichen der

Referenzdaten noch Nachholbedarf für eine zusätzliche Rekonstruktion besteht.

metaio: Im Vergleich zum vorhergehenden Tracking Contest (ISMAR 2009) konnten viele Verbesserungen erzielt werden. Hierzu zählen insbesondere das Extensible Tracking, die verbesserte 3D-Rekonstruktion der Feature Map, sowie die Registrierung in die Zielkoordinatensysteme über 3D-3D Korrespondenzen, wodurch im Minimalfall lediglich drei solche Korrespondenzen bekannt sein müssen.

Zu erwähnen ist, dass sämtliche verwendeten Funktionalitäten in die Softwareplattform Unifeye integriert und für den Anwender verfügbar sind. Bemerkenswerterweise wurde während des Tracking Contests keine Zeile Code geschrieben, sondern ausschließlich die in Unifeye vorhandene Funktionalität genutzt.

Probleme wurden beim Tracking sich schnell bewogender Objekte erkannt (Szenario: Kuka Roboter mit VW Modell), bei schlecht texturierten Objekten (VW Passat Exterieur - weiße Fahrzeugtür mit Spiegelungen), sowie teilweise bei sich ändernder Beleuchtung. In letzterem Fall verschlechterte sich die Qualität der Initialisierung beim Szenario VW Passat Interieur.

Generell zeigte sich, dass ein Ansatz, der rein auf punkt-basiertem Featuretracking arbeitet, für schlecht texturierte Objekte nicht ausreichend ist. Hierfür wird eine Kombination mit Kantentracking und/oder Dense-Tracking wahrscheinlich eine signifikante Verbesserung erzielen.

8.2 Einordnung der Ergebnisse auf der Technologie Roadmap

Wo befindet sich die Technologie auf der Technologie Roadmap? Wann wird diese Technologie den Technologie Readiness Level 6 (TLR6) erreicht haben (d.h.: wann wird es einen Prototypen in Einsatzumgebung geben)? Wie (un-)sicher ist diese Abschätzung?

FAR: Keine Einschätzung. FAR verweist auf die Diskussion in Abschnitt 8.3.

IGD: Um die Technologie des bild-basierten Trackings vollständig im industriellen Umfeld einsetzbar zu machen, müssen die Probleme in weniger Tracking-freundlichen Umgebungen noch gelöst werden. Da es auch vorkommen kann, dass nur geringfügig texturierte Bereiche auch augmentiert werden sollen, macht es für manche industriellen Einsätze erst dann Sinn, ein AR-System zu benutzen, wenn es universell eingesetzt werden kann. Für gutmütigere Umgebungen mit wohldefinierten Lichtverhältnissen könnte man die Technologie schon einsetzen, allerdings gäbe es im Bereich der Benutzbarkeit und der Handhabung des System noch Verbesserungsmöglichkeiten.

metaio: Für Anwendungen im industriellen Umfeld sind nach wie vor Vorkenntnisse des Benutzers erforderlich. Diese nötigen Vorkenntnisse wurden in der Softwareplattform Unifeye durch einfach zu bedienende GUI-Elemente versucht so gering wie möglich zu halten, jedoch kommt ein Anwender nicht umhin, bestimmte Vorkenntnisse vor allem im Bereich der Koordinatentransformationen zu besitzen und einen geringen Aufwand zur Erlernung der Bedienung zu investieren.

Eine allgemeine Verwendung der optischen Trackingalgorithmen für jedwede Art von Szenarien ist derzeit noch nicht möglich. Ein zuverlässiges Tracking (im Rahmen industrieller Anwendungen) ist derzeit nur bei ausreichend texturierter und beleuchteter Umgebung möglich.

Für schlecht texturierte Umgebungen, ungenügend beleuchtete Szenarien, sowie stark reflektierender Lack-/Glasoberflächen besteht derzeit noch ein zu leistender Ent-

wicklungsaufwand. Vielversprechende Möglichkeiten liegen hierbei in der Sensorfusion des rein feature-basierten Ansatzes mit zusätzlichen Systemen (Infrarotkamera, Laserscanner, 3D-Kamera, Gyroskop, Accelerometer, Dense-Tracking, Kanten-Tracking, ...).

8.3 Forschungsbedarf

Was muss Ihrer Meinung nach passieren, damit die Technologie wirklich eingesetzt werden kann? Welcher Forschungsbedarf lässt sich hieraus ableiten? Welche Probleme müssen gelöst werden? Wie viel weiteres Potential steckt aus technischer Perspektive noch in dieser Technologie? Welches sind die notwendigen Schritte, um die Technologie weiter voranzubringen?

FAR: Der Tracking Contest hat gezeigt, dass das Tracking an und für sich meistens funktioniert. Die Probleme liegen heutzutage im Tracken von schwierigen Objekten, die nicht genügend oder ungenaue Features erzeugen. Eventuell könnten kombinierte Verfahren, die nicht auf Punktdaten alleine basieren, robustere Ergebnisse liefern.

Desweiteren liegt noch enormes Potential im automatischen Aufbereiten und Bewerten von Trackingdaten (beispielsweise Punktwolken) zur Laufzeit, so dass ein System auch ohne Expertenwissen und lange Vorlaufzeit benutzt werden kann. Allerdings kann nicht jedes Wissen automatisch aus den gewonnenen Daten abgeleitet werden, so dass es immer noch manueller Eingriffe bedarf. Diese sollten allerdings auch ohne Expertenwissen erfolgen können. Eine einfache, möglichst automatische Aufbereitung der Feature Maps ist jedoch die Grundvoraussetzung für die effiziente Integration von Featuretracking in industrielle Prozesse.

Der auf dem Contest betriebene Aufwand für die Aufbereitung der Feature Maps ist im Hinblick auf deren Volatilität (beschränkte Wiederverwendbarkeit unter leicht veränderten Bedingungen wie Umgebung, Beleuchtung, ...) für den produktiven Einsatz meist nicht hinnehmbar. Der von FAR verfolgte Ansatz zur automatischen ad-hoc Registrierung ohne manuelles Editieren der Feature Maps versuchte, diesen Aufwand zu vermeiden. Der Ansatz funktionierte passabel – jedoch nur in den kleineren Szenarien (HDW, VW Interieur, EADS Wellness-Bereich). Wie oben bereits erwähnt, besteht jedoch in diesem Ansatz noch Optimierungspotential. Die Referenzmessungen aus dem Markertracking dürfen nicht nur zur rigiden Registrierung (7 Freiheitsgrade) der Welt mit einer verzerrten Feature Map verwandt werden. Vielmehr müssen die Referenzmessungen in die Bündelblockausgleichung mit einbezogen werden. Mittels dieser Referenzmessungen könnte die Abdrift randständiger Keyframes in der Feature Map verhindert werden. Die Map würde stabilisiert. Wie im bereits bestehenden System auch, könnte dies in vielen Fällen durchaus automatisch im Rahmen eines ad-hoc Verfahrens zum Anlernen und Registrieren der Map geschehen.

Das Ergänzen der Systeme durch weitere Hardware (Time-of-Flight, Gyroskope, etc.) könnte zu weiteren Verbesserungen im Tracking führen, da Nachteile der optischen Verfahren ausgeglichen werden könnten. Es könnte auch von Vorteil für die optischen Verfahren sein, wenn man die Einstellungen der Linse (beispielsweise die Einstellung der Blende) vom Programm aus steuern könnte, um so mehr Kontrolle bei wechselnden Lichtbedingungen zu haben. Dies ist bei den meisten heutigen Kameras nicht der Fall.

Eine weitere Einschränkung ist die bisherige Verwendung von Referenzmarken, die für eine hohe Genauigkeit erst eingemessen werden müssen. Dieser Prozess nimmt ebenfalls einiges an Vorbereitungszeit ein und erfordert spezielles Wissen und spezielle Geräte. Am Beispiel der EADS Kabine kann man erkennen, wie solche Referenzmarken geschickt gestaltet werden können, damit der unbedarfte Nutzer auf schon gewonnene Referenzdaten zurückgreifen kann. Damit ein kamera-basiertes Tracking schnell eingesetzt werden kann, müssen auch für andere Szenarien ähnliche Lösungen gefunden werden, um zuverlässige Referenzen zur Umwelt zu liefern.

IGD: Obwohl im Bereich der bild-basierten Trackingverfahren enorme Fortschritte gemacht worden sind, und die Technologie in gutmütigen Szenarien schon einsatzbereit ist, gibt es noch einigen Forschungsbedarf, um ein AR-System universell einsetzen zu können.

Derzeit gibt es noch Probleme bei glatten, reflektierenden Flächen, wie sie in industriellen Szenarien oft vorkommen. Eine Verbesserungsmöglichkeit wäre die zusätzliche Nutzung von anderen Arten von Bildmerkmalen wie Linien, homogene Flächen oder Farben.

Starke Beleuchtungsveränderungen bereiten ebenfalls noch Schwierigkeiten. Bildmerkmale mit einer mehrdeutigen Beschreibung von Eigenschaften könnten helfen das Trackingsystem in derartigen Szenarien deutlich robuster zu machen.

Oft kommt es insbesondere bei Montageszenarien vor, dass die Umgebung stark verändert wird. Auch hierfür wäre es notwendig, dass das Trackingsystem nicht nur von einem einzigen visuellen Aussehen der Szene ausgeht, sondern dass ein und dieselbe Stelle mehrere Erscheinungsformen haben kann.

Bei einer Anwendung im industriellen Einsatz kommt es auch vor, dass Ummengen von Referenzdaten nötig sind, damit das Trackingsystem einwandfrei arbeiten kann. Hierfür ist es nötig eine geeignete Infrastruktur zu erfinden, welche dieser Skalierungsproblematik gerecht wird.

metaio: metaio verweist an dieser Stelle auf ihre Antwort in Abschnitt 8.2: Einordnung der Ergebnisse auf der Technologie RoadMap.

9 Resümee

Dieser Tracking Contest hat gezeigt, dass die Technologien der AVILUS-Entwickler, sich in den letzten Jahren erheblich weiterentwickelt haben. Die Technologien haben oft auch noch unter erschwerten Bedingungen wie starken Veränderungen der Lichtverhältnisse, großflächigen Verdeckungen, Bewegungen in großen Räumen, und Initialisierungen in größeren Volumen funktioniert. Es waren aber auch oft die Grenzen der Systeme zu sehen, insbesondere wenn wenig texturierte oder glänzend lackierte Objekte getrackt werden sollten. Deutlich wurden zudem die Schwächen der Systeme, sobald sie von Anwendern und nicht von den Technologiepartnern selbst benutzt wurden – Eine Benutzung durch den Anwender führte oft ziemlich schnell zu einem Abriss des Tracking. Zusammenfassend kann festgehalten werden, dass die Technologien immer besser werden. Für einen breiten Einsatz in der Praxis besteht aber noch erheblicher Forschungsbedarf.

Die Szenarien von VW am Passat und von EADS in der Flugzeugkabine haben sich als sehr geeignete Testszenarien herausgestellt, in denen die Trackingsysteme in all ihren Facetten sehr gut ausgetestet werden konnten. Szenarien dieser Art sollten ein fester Bestandteil in weiteren Tracking Contests werden. Das Szenario des bewegten VW-Modell auf dem Kuka Roboter konnte nur teilweise gelöst werden. Zum einen war dieses Szenario schwierig, zum anderen hätten einige Problem mit etwas mehr verfügbarer Zeit möglicherweise gelöst werden können. Ein Szenario mit bewegten Objekten sollte in jedem Fall Bestandteil zukünftiger Evaluierungen bleiben. Das Tracking am Kuka Roboter hat nicht funktioniert, und die Tracker haben probiert, die ganze Szene zu tracken und nicht die einzelnen Objekte. Die weitere Evaluation von Systemen zum Tracken einzelner, kleiner Objekte macht nur Sinn, wenn andere oder angepasste Verfahren benutzt werden. Das HDW Soll-ist-Vergleich Szenario war für einen ersten Test sehr gut geeignet. In weiteren Evaluationen sollte jedoch ein "richtiges" Einsatzumfeld gewählt werden, welches das ganze Kameraumfeld ausfüllt. In solchen kleinen Setups ist nicht auszuschließen, dass künstliche, gute Features (z.B. U-Bootmodell-Kante zum Tisch) genutzt werden. Zudem kann bei kleinen Modellen eine kleine detailreiche Feature Map gewählt werden, welches in größeren Setups unter Umständen nicht mehr performant nutzbar ist. Das Szenario am EADS-Wellness-Bereich, war nicht nur für die Beobachter von Interesse, sondern auch für die anderen Technologiepartner (Tracking), die so sehen konnten, wie die Tool-Chain der Konkurrenz funktioniert. Dieses Szenario war nicht sehr umfangreich (Aufbau/Durchführung), jedoch lieferte es für alle Beteiligten einen sehr guten Einblick in die Technologie der anderen.

Ein zukünftiger Tracking Contest sollte einen noch stärkeren Fokus auf die Benutzung durch Teilnehmer legen. Wie auch in diesem Tracking Contest zu sehen war, besteht ein sehr großer Unterschied zwischen der Benutzung eines Trackingsystems durch den Technologiepartner und der Benutzung durch den (geschulten) Anwender. Ein Tracking Contest ist eine gute Möglichkeit, die Technologiepartner direkt mit den Problemen von Anwendern im Umgang mit Ihrer Technologie zu konfrontieren. Aus diesem Grund sollte in zukünftigen Evaluationen ein Großteil der Versuchsdurchläufe von Anwendern und nicht von den Technologie-Entwicklern durchgeführt werden. Für einen zukünftigen Tracking Contest wurde ebenfalls angeregt, den Aufbau (insbeson-

dere den Umfang des Anlernens) als Faktor mit in die Bewertung aufzunehmen. Oft machten sich die "schlechten" Visualisierungen der Teilnehmer bemerkbar. Nicht selten benötigten die Teilnehmer mehrere Anläufe um ihre eigenen Visualisierungen zu interpretieren, also um herauszufinden, auf welches Objekt die Visualisierung zeigt. Die Visualisierungen haben nicht genug 3D-Tiefeninformationen vermittelt.

Aufgrund der positiven Ergebnisse dieses Tracking Contests, insbesondere der technologischen Entwicklungen, die direkt während solch eines Contests stattfinden und dem praxisnahen Austausch zwischen Technologieentwicklern und Anwendern, empfehlen wir die Durchführung eines jährlichen Tracking Contests als festen Bestandteil in die Planung von AVILUS Nachfolgeprojekten mit aufzunehmen.

Literatur

- [1] S. Baker and L. Matthies. Lucas-Kanade 20 Years On: A Unifying Framework. *International Journal on Computer Vision (IJCV)*, 56(3):221–255, 2004.
- [2] H. Bay, T. Tuytelaars, and L. VanGool. SURF: Speeded up robust features. In *Proc. European Conference on Computer Vision (ECCV)*, pages 404–417, 2006.
- [3] G. Bleser, H. Wuest, and D. Stricker. Online camera pose estimation in partially known and dynamic scenes. In *Proc. IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 56–65, 2006.
- [4] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 886–893, 2005.
- [5] A. Davison, I. Reid, N. Molton, and O. Stasse. MonoSLAM: Real-Time Single Camera SLAM. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 29(6):1052–1067, 2007.
- [6] A. Davison, M. W.W., and D. Murray. Real-time localization and mapping with wearable active vision. In *Proc. IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 18–27, 2003.
- [7] C. Engels, H. Stewenius, and D. Nister. Bundle Adjustment Rules. *Photogrammetric Computer Vision (PCV)*.
- [8] L. Gruber, S. Gauglitz, J. Ventura, S. Zollmann, M. Huber, M. Schlegel, G. Klinker, D. Schmalstieg, and T. Höllerer. The City of Sights: Design, Construction and Measurement of an Augmented Reality Stage Set. In *Proceedings of the 9th International Symposium on Mixed and Augmented Reality (ISMAR)*, Oct. 2010.
- [9] C. Harris and M. Stephens. A combined corner and edge detector. In *Proc. Alvey Vision Conference*, pages 147–151, 1988.
- [10] R. Hartley and A. Zisserman. *Multiple View Geometry, 2nd edition*. Cambridge University Press, 2003.
- [11] B. Horn. Closed-form solution of absolute orientation using unit quaternions. *Journal of the Optical Society of America (JOSA)*, 4(4):629–642, 1987.
- [12] M. Huber, B. Becker, and G. Klinker. Location aware computing using RFID infrastructure. *Int. J. Autonomous and Adaptive Communications Systems*, 3(1):23–38, 2010.
- [13] M. Huber, D. Pustka, P. Keitler, E. Florian, and G. Klinker. A System Architecture for Ubiquitous Tracking Environments. In *Proceedings of the 6th International Symposium on Mixed and Augmented Reality (ISMAR)*, Nov. 2007.

- [14] B. Jiang and U. Neumann. Extendible tracking by line auto-calibration. In *Proc. IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 97–103, 2001.
- [15] P. Keitler, B. Becker, and G. Klinker. Management of tracking for industrial ar setups. In *Proceedings of the 9th International Symposium on Mixed and Augmented Reality (ISMAR)*, October 2010.
- [16] P. Keitler, F. Pankratz, B. Schwerdtfeger, D. Pustka, W. Rödiger, G. Klinker, C. Rauch, A. Chathoth, J. Collomosse, and Y.-Z. Song. Mobile augmented reality based 3d snapshots. In *Proc. Sechster Workshop Virtuelle und Erweiterte Realität der GI-Fachgruppe VR/AR*, Braunschweig, Germany, November 2009.
- [17] P. Keitler, D. Pustka, M. Huber, F. Ehtler, and G. Klinker. *Engineering of Mixed Reality Systems*, chapter Management of Tracking for Mixed and Augmented Reality Systems. 2009.
- [18] G. Klein and D. Murray. Parallel tracking and mapping for small AR workspaces. In *Proc. Sixth IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR'07)*, Nara, Japan, November 2007.
- [19] V. Lepetit and P. Fua. Keypoint recognition using randomized trees. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 28(9):1465–1479, 2006.
- [20] M. Lourakis and A. Argyros. SBA: A software package for generic sparse bundle adjustment. *ACM Transactions on Mathematical Software (TOMS)*, 36(1).
- [21] D. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal on Computer Vision (IJCV)*, 2(60):91–110, 2004.
- [22] B. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *Proc. International Conference on Artificial Intelligence (IJCAI)*, pages 674–679, 1981.
- [23] J. Newman, M. Wagner, M. Bauer, A. MacWilliams, T. Pintaric, D. Beyer, D. Pustka, F. Strasser, D. Schmalstieg, and G. Klinker. Ubiquitous Tracking for Augmented Reality. In *Proceedings of the 3rd IEEE/ACM International Symposium on Mixed and Augmented Reality*, pages 192–201, 2004.
- [24] D. Pustka. Construction of Data Flow Networks for Tracking in Augmented Reality Applications. In *Proc. Dritter Workshop Virtuelle und Erweiterte Realität der GI-Fachgruppe VR/AR*, Koblenz, Germany, September 2006.
- [25] D. Pustka, M. Huber, M. Bauer, and G. Klinker. Spatial Relationship Patterns: Elements of Reusable Tracking and Calibration Systems. In *Proc. IEEE International Symposium on Mixed and Augmented Reality (ISMAR'06)*, October 2006.
- [26] D. Pustka and G. Klinker. Dynamic Gyroscope Fusion in Ubiquitous Tracking Environments. In *Proceedings of the 7th International Symposium on Mixed and Augmented Reality (ISMAR)*, September 2008.

- [27] J. Shi and C. Tomasi. Good features to track. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 693–600, 1994.
- [28] B. Triggs, R. McLauchlan, R. Hartley, and A. Fitzgibbon. Bundle Adjustment - A Modern Synthesis. In *Proc. International Workshop on Vision Algorithms: Theory and Practice*, pages 298–372, 2000.
- [29] S. Umeyama. Least-Squares Estimation of Transformation Parameters Between Two Point Patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 13:376–380, 1991.
- [30] C. Waechter, M. Huber, P. Keitler, M. Schlegel, D. Pustka, and G. Klinker. A Multisensor Platform for Wide-Area Tracking. In *9th IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR 2010)*, Seoul, Korea, October 2010.
- [31] C. Waechter, D. Pustka, and G. Klinker. Vision based people tracking for ubiquitous augmented reality applications. In *8th IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR 2009)*, pages 221–222, Orlando, USA, October 2009.
- [32] H. Wuest, A. Pagani, and D. Stricker. Feature Management for Efficient Camera Tracking. In *Proc. Asian Conference on Computer Vision (ACCV)*, pages 769–778, 2007.