



Hochschule für Angewandte Wissenschaften Hamburg
Hamburg University of Applied Sciences

Thesis Outline

Hosnia Najem

“Modellbasiertes Suchen von Objekten
in einer Smart-Home-Umgebung“

Hosnia Najem
“Modellbasiertes Suchen von Objekten
in einer Smart-Home-Umgebung“

Thesis Outline eingereicht im Rahmen der Seminar Ringvorlesung
im Masterstudiengang Informatik
am Department Informatik
der Fakultät Technik und Informatik
der Hochschule für Angewandte Wissenschaften Hamburg

Betreuender Professoren :
Prof. Dr. rer. nat. Kai von Luck
Prof. Dr.Ing. Gunter Klemke

Abgegeben am 29. Februar 2012

Inhaltsverzeichnis

1	Einleitung	4
2	Vorarbeiten zur Masterthesis	7
2.1	AW1 und Projekt 1	7
2.2	AW2 und Projekt 2	10
3	Zukünftiges Vorhaben in der Masterthesis	14
3.1	Bestimmung der Raumposition	14
3.1.1	Stereoanalyse	14
3.2	Die Kamerasteuerung	17
4	Fazit	19
	Literaturverzeichnis	20

1 Einleitung

Das Living Place Hamburg ist seit 2009 eine im Aufbau befindliche, ca. 130 m^2 große Einpersonenzwohnung, und stellt die intelligente Wohnung (Smart-Home) an der HAW Hamburg dar, in der die verschiedenen Ansätze des modernen Lebens untersucht werden.

Das Living Place Hamburg ([LP-HH](#)) integriert bereits verschiedenste Anwendungssysteme, die auf die verschiedenen Sensor-Daten, über die sogenannte „Sensor-Wolke“ zugreifen, um „Intelligenz“ in der Wohnung zu schaffen.

Ein integriertes Anwendungssystem im Living Place Hamburg ist „das intelligente Bett“ (vgl. [Hardenack, 2011](#)). Die eingebauten kapazitiven Sensoren im Bett erfassen Bewegungen und stellen diese Sensor-Daten dann der Wohnung bzw. „der Sensor-Wolke“ zur Verfügung. Die gelieferten kapazitiven Sensor-Daten, kombiniert mit den weiteren Sensor-Daten aus weiteren Gegenständen in der Wohnung, wird vom Anwendungssystem „intelligente Bett“ analysiert und steuert bei Erkennen eines bestimmten Kontexts das Licht passend zu dem Gemütszustand des Bewohners an.

Mit dieser Masterthesis soll das Living Place Hamburg bzw. „die Sensor-Wolke“ mithilfe der Kamera mit den optischen Sensor-Daten ergänzt werden. Diese Ergänzung führt dazu, dass die Wohnung sich besser an die Bedürfnisse des Bewohners anpassen und ihm mehr Services zur Unterstützung seines Alltags anbieten kann.

Motivation Die zentrale Fragestellung für die Motivation zu der Masterthesis „Modellbasiertes Suchen von Objekten in einer Smart-Home-Umgebung“ ist:

Wie können die bereits gewonnenen Vorkenntnisse, Erfahrungen im Bereich der Objekterkennung und des Objekt-Trackings aus der Bachelorarbeit [Najem \(2008\)](#) mit dem Thema „Realisierung einer Bibliothek für die schnelle Objektverfolgung und Objekterkennung in Bildsequenzen“ im Living Place Hamburg eingesetzt werden, um die Wohnung zu einer Wohnung der Zukunft zu machen?

Hierdurch entstand die Vision, das Living Place Hamburg (vgl. [LP-HH](#)) mit einem Auge auszustatten. Die Grundlage der Vision bilden die im Living Place Hamburg integrierten Pen-Tell-Zoom Dome Kameras ([PTZ-Kameras](#)), welche die optischen Sensoren in der intelligenten Wohnung darstellen und das menschliche Auge unterstützen und ergänzen soll.

Zielsetzung in der Masterthesis Das Ziel der Masterthesis „Modellbasiertes Suchen von Objekten in einer Smart-Home-Umgebung“ ist, die Grundlagen für die Interaktionen der integrierten Pen-Tell-Zoom Dome Kameras mit dem Bewohner und anderen integrierten sensorenbasierten Technologien im Living Place Hamburg zu legen.

Dabei sollen ein kameragestütztes System entwickelt und implementiert werden, das dem Bewohner beim Suchen und Wiederfinden von bestimmten Objekten im Living Place Hamburg anhand zuvor erstellter Referenzmodelle unterstützt.

Die grundlegenden Gesichtspunkte der Realisierung eines kameragestützten Systems wurden bereits im Rahmen des Masterstudiums in Hausarbeiten und den Projekten erarbeitet. In dieser Thesis Outline soll auf der Basis bisher erlangter Erfahrungen aus den Masterprojekten die Vorgehensweise der Masterthesis (Masterarbeit) vorgestellt werden.

Anwendungsszenarien Um die Ziele in konkreten Prozessen abbilden zu können, wurde im Rahmen des Masterstudiums (vgl. Najem, 2009) (vgl. Najem, 2011, seite 6) ein Szenario erstellt, das der Umsetzung zugrunde liegt und welches im Living Place realisiert werden soll.

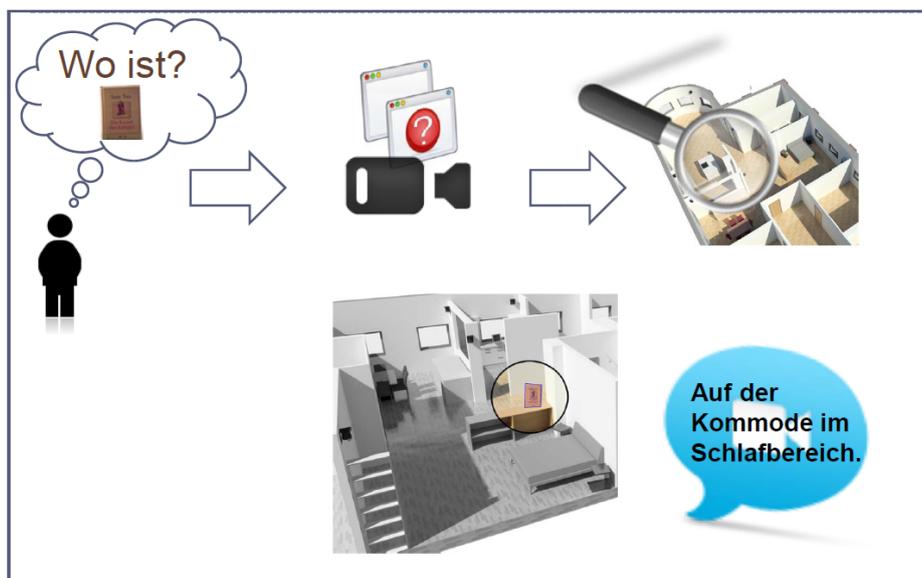


Abbildung 1.1: Kameragestützte Suche des Buches

Die Abbildung 1.1 zeigt ein Szenario, in dem der Bewohner des Living Place Hamburg, das kameragestützte System nutzt, um eine Serviceanfrage - hier im Beispiel „Wo ist mein Buch?“ - an die Wohnung zu stellen.

Über ein Interface selektiert der Bewohner das zu suchende Referenzmodell und aktiviert damit den Suchprozess des kameragestützten Systems.

In der Abbildung 1.1 dargestellten Szenario wurde das gesuchte Referenzmodell „Buch“ von der Kamera im Schlafbereich auf der Kommode wiedergefunden.

Sobald eine der Kamera im Living Place Hamburg das gesuchte Objekt in dem Video-Stream lokalisiert hat, wird die Lage des Objektes auf einem der Monitore in der Wohnung dargestellt.

Über einen weiteren internen Dienst wird ermöglicht, dem Bewohner die Position des gefundenen Objektes und seine relative Lage zu den anderen Gegenständen in der Wohnung verbal in einem dem menschlichen Sprachgebrauch entsprechenden Satz zu formulieren:

„Das Buch ist im Schlafbereich auf der Kommode.“

In den Vorarbeiten zur Masterarbeit in den Masterprojekten Najem (Feb 2011) Najem (2011) wurde für die Realisierung des Szenario die Teilkomponenten Object Teaching, Object Searching und Object Localisation entwickelt.

Jeder Teilbereich wurde unter bestimmten Fragestellungen im Rahmen der Masterprojekten untersucht, um Lösungen zu entwickeln, die die Anforderung eines Systems im Kontext einer intelligenten Wohnung erfüllen.

Forschungsfragestellungen Die Teilkomponente Object Teaching befasst sich mit den Fragen, welche Techniken und Methoden der Bildverarbeitung ermöglichen es, Objekte in den Bildern zu erfassen und Modelle im Gesamtsystem zu hinterlegen, um diese effizient und robust in den Kamerabildern der Wohnung wiederzufinden.

Die Teilkomponente Object Searching thematisiert die Frage, wo sich das gefundene Objekt im Bild in der realen Welt befindet. Mit welchen mathematischen Verfahren kann mithilfe der Bildkoordinate die Raumkoordinate bestimmt werden?

Die Ergebnisse dieser Teilkomponenten werden in Abschnitte 2.1 dargestellt.

Der letzte Teilkomponente Object Localisation integriert die sprachliche Lokalisierung des gefundenen Objektes im Raum. Diese thematisiert unter anderem die Fragestellungen:

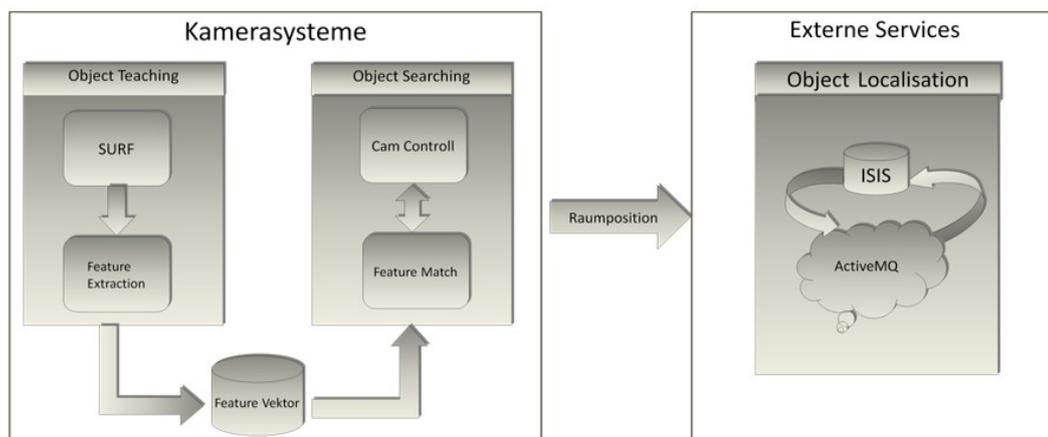
- Wie können räumliche Informationen gewonnen werden?
- Wie können die räumlichen Informationen semantisch interpretiert werden?
- Wie kann ein lokalisierender Antwortsatz definiert werden, sodass der Bewohner die Position des gefundenen Objektes im Raum erkennt.

Die Lösungsansätze dieser Fragestellungen wird im dem Abschnitt 2.2 erläutert.

2 Vorarbeiten zur Masterthesis

Die Grundlagen zu den theoretischen Vorüberlegungen wurden im Rahmen des Masterstudiums in Anwendung 1 (AW1) [Najem \(2009\)](#) und Anwendung 2 (AW2) [Najem \(2010\)](#) erarbeitet.

Im Rahmen der Projekte wurden hierfür die ersten Implementierungen und praktische Umsetzungen in das reale Umfeld durchgeführt. In diesem Abschnitt werden einige Projektergebnisse und die gewonnenen Erkenntnisse aus den praktischen Experimenten in der realen Umgebung erläutert. Die [Abbildung 2](#) soll ein Einblick über das Zusammenspiel der einzelnen Funktionalitäten für das kameragestützte System und die Integration des externen Services gegeben werden.



2.1 AW1 und Projekt 1

Im Rahmen der AW1-Hausarbeit ([Najem, 2009](#)) und des Projekts PO1 ([Najem, Feb 2011](#)) wurde für die automatische Erstellung und Speicherung von Referenzmodellen ein Vergleich von verschiedenen Verfahren und Algorithmen der OpenCV-Bibliotheken durchgeführt.

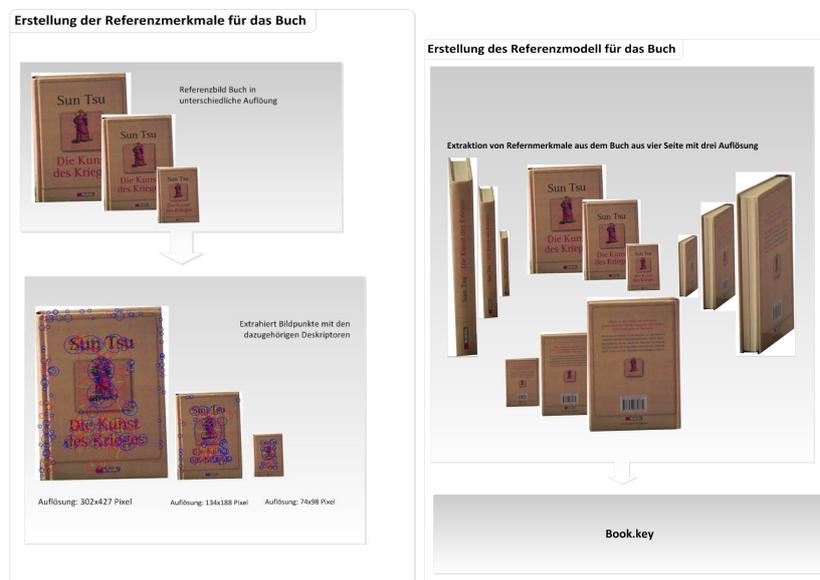
Object Teaching Die erste Umsetzung in Projekt 1 war die Erstellung eines in C# implementierten Interfaces für die automatische Erstellung und Speicherung von Referenzmodellen.

Die Implementierung basiert auf dem SURF-Algorithmus von David Lowe ([Lowe, 1999](#)), einem in der Bildverarbeitung bestbewertetem rotations- und skalierungsinvariantem feature-based Verfahren zur Objekterkennung und Objektverfolgung in Videosequenzen.

Um das Objekt in der Wohnung aus jeder Perspektive sowie auch mit unterschiedlichen Zoomeinstellungen der Kameras wiedererkennen zu können, wurde der SURF-Algorithmus dahingehend modifiziert, dass das Referenzmodell eines Objektes von allen Seiten und in unterschiedlichen Auflösungen, wie in Abbildung [2] dargestellt, erstellt wird.

Funktionale Tests mit der SURF-Implementierung haben gezeigt: Je mehr Informationen das Modell enthält, desto weniger muss die Kamera gesteuert werden. Das Matching-Verfahren wird auf diese Weise beschleunigt. Ein weiterer Performance-Vorteil, der dabei entsteht, ist, dass die Bildsequenzen der PTZ-Kameras beim Matching-Verfahren keine HD-Auflösungen aufweisen müssen. (vgl. [Najem, Feb 2011](#))

Der Bewohner kann unabhängig von der Suchanfrage jederzeit Referenzmodelle im Gesamtsystem hinterlegen und diese über den Namen des entsprechenden Objektes für die Suche selektieren und die Serviceanfrage an das System starten.



Object Searching Object Searching integriert die Funktionalitäten zur Steuerung der eingebauten PTZ-Kameras im Living Place Hamburg (Cam-Controll in der Abb.[2]) und zum anderen das sogenannte Matching-Verfahren (Future -Macht in der Abb.[2]).

Beim Matching-Verfahren werden mithilfe des SURF-Algorithmus Korrespondenzen zwischen den gewählten Referenzmodellen und den Bildmerkmalen der vorkommenden Objekte in den Bildsequenzen der Kamera gesucht.(vgl. [Najem, 2009](#))

Cam-Control Das Module cam-control integriert die Funktionalitäten zur Steuerung der eingebauten Pan- Tilt -Zoom Dome Kameras der Firma Dallmeier [PTZ-Kameras](#).

Welche Eigenschaften die PTZ-Kameras besitzen und welche Funktionalitäten diese integrieren, wurde im PO1-Bericht [Najem \(Feb 2011\)](#) beschrieben. Darüber hinaus wurde eine einfache Kamerasteuerung, mit dem Dallmeier Camera Communication Protocol (DCCP) realisiert.

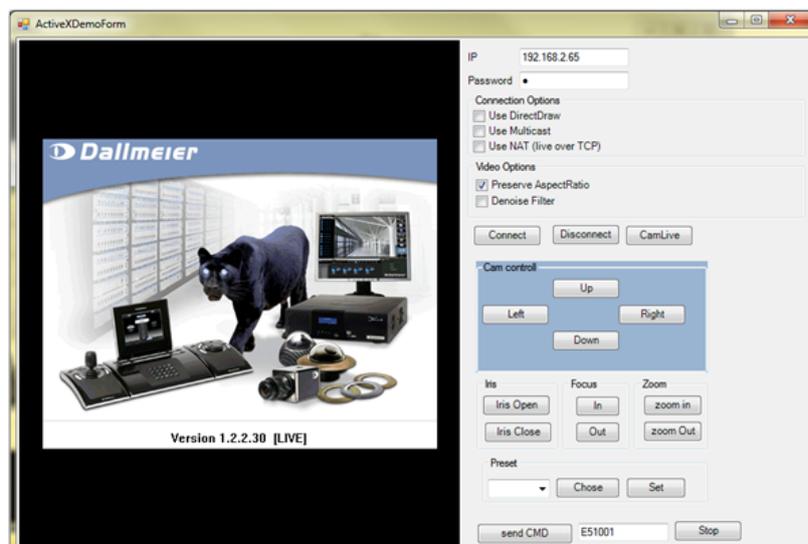


Abbildung 2.1: Applikation zur Steuerung der PTZ-Kamera [Quelle:Hosnia Najem]

Wie in der Abbildung 2.1 dargestellt ist, wurde für die Steuerung eine erste Testanwendung in C# implementiert, die ermöglicht, Steuerbefehle über das Netzwerk an die einzelnen Kameras zu senden.

Das DCC-Protokoll ermöglicht, jedes Steuersignal (rechts, oben, unten, links) über das Netzwerk an die Kamera zu senden. Jedoch muss jeder Steuerbefehl mit einem Stopp-Signal abgeschlossen werden, um die Aktion zu beenden. Auch der aktuelle Status der Kameras

(Zoomeinstellung, Drehung, genaue Position der Kameras) kann nicht über das Protokoll abgefragt werden.

Somit ist Protokolle als eine Schnitte zu den PTZ-Kameras zum Durchsuchen der Wohnung sowie erstellen eine Planungs- und Koordinationseinheiten der Kameras ungeeignet. Welcher Ansatz in der Masterthesis verfolgt wird, ist in dem Abschnitt 3.2 dargestellt.

Raumposition Das SURF-Verfahren liefert die Position des gefundenen Objektes im Bild. Um die Position des Objektes im Raum zu bestimmen, müssen zu den 2D Bildkoordinate die dazugehörigen 3-D-Koordinaten bestimmt werden.

Für die ersten Tests wurden die 3-D-Koordinaten des Objektes in der Wohnung mithilfe der integrierten UbiSense-Sensoren, der die Lokalisierung eines Takte in der Wohnung ermöglicht, ermittelt.

Welcher Lösungsansatz in der Masterthesis verfolgt wird, ist in dem Abschnitt 3.1 dargestellt.

2.2 AW2 und Projekt 2

Eine weitere Herausforderung, die sich aus dem Szenario ergibt, ist die sprachliche Ausgabe des Suchergebnisses. Die verbale Übermittlung der Lage des Objektes an den Bewohner stellt die letzte Teilkomponente der Masterthesis dar [siehe Object Localisation in der Abb. 2].

Für die sprachliche Lokalisierung eines Objektes im Raum muss der Zusammenhang zwischen den räumlichen und sprachlichen Strukturen betrachtet werden.

Das Thema „Sprache und Raum“ ist Gegenstand der gegenwärtig stattfindenden interdisziplinären Diskussionen des Sonderforschungsprojekts SFB/TR 8 Spatial Cognition¹: Reasoning, Action, Interaction.

In der Raumkognition beschäftigen sich die Wissenschaftler u.a. mit den Fragen:

Welche kognitiven Prozesse finden statt, wenn wir uns in einem Raum bewegen, uns im Raum orientieren, jemandem einen Weg oder die Lage eines Gegenstandes sprachlich beschreiben?

Das Thema Raumkognition wurde in der Anwendung-2-Hausarbeit Najem (2010) thematisiert, und soll hier an Hand eines Beispiels nur einige Aspekte eine sprachliche Ausgabe verdeutlicht werden.

¹[Quelle: <http://www.sfbtr8.uni-bremen.de/index.html>]

Szenen Beschreibung: Ein Buch liegt auf dem Tisch. Links von dem Buch befindet sich eine Tasse.

Die Position des Buch in dieser Szene soll dem Bewohner sprachlich beschrieben werden.

Auf die Frage „Wo ist das Buch?“ wäre folgende Beschreibungen in Abhängigkeit von weiteren Nachbarobjekten möglich. Dabei wären der Tisch und die Tasse die weiteren Objekte in der Nähe des Buches.

Das Buch befindet neben die Tasse.

Das Buch befindet links von der Tasse.

Das Buch befindet auf dem Tisch.

Das Buch befindet sich auf dem Tisch, neben die Tasse.

Abhängig von dem Nachbarobjekt wurden die entsprechenden lokalisierenden Ausdrücke wie „neben“, „links“, oder „auf“ gewählt.

Hieraus wird deutlich, dass für ein computergestütztes Sprachausgabesystem viele Hürden zu überwinden sind, um einen den Menschen gebräuchlichen Satz zu formulieren.

Die Anforderungen an das kameragestützte System ist aus der [Najem \(2010\)](#) zu entnehmen und hier kurz zusammengefasst dargestellt, sind:

Es müssen weitere geeignete Objekte in der Nähe des gesuchten Objektes erkannt und bestimmt sowie deren Lage in Beziehung zueinander gesetzt werden. Es muss ein Bezugssystem bestimmt und für die Beschreibung eine Perspektive gewählt werden.

Die Einbindung der „Indoor Spatial Information Service“(ISIS) von Bastian Karstaedt [[Karstaedt \(2011b\)](#) [Karstaedt \(2011a\)](#)] als ein Serviceanbieter für die Anfragen über räumliche Informationen und die Relationen von Objekten im Living Place Hamburg unterstützt das kameragestützte System bei der Formulierung einer sprachlichen Lokalisierung.

Die räumlichen Relationen sind topologische, relationale und direktionale Informationen und werden auf Basis der im Projekt „Indoor Spatial Information Service“ definierten Prädikate , z.B. „above“, „below“ und usw. angegeben.

Die Anforderung an die Sprachausgabe des kameragestürzten Systems ist, auf Basis der gewonnenen räumlichen Informationen einen kurzen und aussagekräftigen Satz zu formulieren, der die Lage des Objektes im Raum eindeutig wiedergibt.

Hierfür wurde im Projekt 2 Konvention vereinbart, wie und wann die gelieferte Prädikate von ISIS in lokalisierende Ausdrücke wie „im“, „auf“, „unter“ usw. interpretiert.

Die Experimente in der Living Place Hamburg haben gezeigt, dass die gelieferte topologischen und direktionalen Relationen ausreichen, um einen beschreibenden Satz wie folgt zu formulieren:

„Das Buch ist im Schlafbereich auf der Kommode.“

Dieser Satz könnte optional mit den metrischen Informationen ergänzt werden. Sie werden mit dem Prädikat wie „near by“ angegeben und liefern Angaben über den Abstand eines Objektes zu anderen Objekten. Als Bezugspunkt der metrischen Informationen wird entweder ein weiteres Objekt (z. B. „Die Kommode ist in der Nähe des Bettes“) oder die angegebene 3-D-Koordinate („Das Buch ist in der Nähe der Tasse“) gesehen.

In welchen Fällen es Sinn ergibt, den Satz mit den metrischen Informationen für die sprachliche Aussage zu ergänzen, ist abhängig vom Informationsgehalt des vorherigen Satzes und bedarf weiter Untersuchungen in der Masterthesis.

Annahmen und Präzision bei der Sprachausgabe Bislang wurde davon ausgegangen, dass jede von ISIS gelieferte räumliche Relationen ein sinnvolles Ergebnis liefert. Was passiert, wenn ISIS eine räumliche Information mit folgenden Aussagen liefert:

Die angegebene Koordinate liegt zwischen einem Stuhl und das Bett.

Die angefragte Koordinate liegt unter der Kommode.

Die angegebene Koordinate befindet sich einen halben Meter über der Kommode.

Um einen kurzen Satz zu formulieren, der die Lage des Objektes im Raum eindeutig wiedergibt, muss das kameragestützte System auf Basis der gewonnenen räumlichen Informationen Entscheidungen zwischen verschiedenen Perspektiven, Bezugssystemen und den gelieferten Nachbarobjekten treffen. Hierfür wurden Annahmen getroffen und im Gesamtsystem implementiert.

Für die Wahl eines Nachbarobjektes für die sprachliche Ausgabe wird voraus gesetzt, dass in der Applikation definiert wird, welche Gegenstände in der Wohnung als prominente Objekt und welche Gegenstände als sekundäre Objekte betrachtet werden. Diese Kategorisierung basiert auf der Analyse der Form und Größe jedes Gegenstandes in der Wohnung. Ein prominentes Objekt, wird definiert als ein Gegenstand in der Wohnung, welche eine eindeutig einen Bereich in der Wohnung definiert (z.B. das Bett im Schlafbereich, das Sofa im Wohnbereich, das Herd in der Küche usw). Ein prominentes Objekt in der Wohnung kann als eine Landmarke (vgl. [Najem, 2010](#)) in einem Bereich der Wohnung gesehen werden.

Bislang wurde die Landmarken in der Implentierung nicht mit berücksichtigt, doch soll im Rahmen der Masterthesis umgesetzt werden.

Eine sichere Annahme, von der ausgegangen werden kann, ist, dass ein Buch nicht in der Luft schweben kann, da während der Suche, die Objekte in Ruhepositionen betrachtet werden.

Eine weitere Annahme, die implementiert wurde, ist das bestimmte Gegenstände aufgrund ihrer Form, Größe und Position im Living Place Hamburg keine weiteren Gegenstände unter sich verbergen können. So kann sich z. B. ein Buch nicht unter einer Kommode befinden. Hierfür wird die Auswertung der geliefert räumliche Relationen von ISIS, mit dem Prädikat „below“, im Bezug zu bestimmte Gegenstand (z.B.Kommode,Küchentressen,Sofa usw.) ignoriert bzw. explizite unterbunden.

Neben diese Annahmen müssen in der Masterthesis noch einige weitere Annahmen getroffen und in Impelentierung berücksichtigt werden, um einen Satz zu formulieren, der für den Bewohner auch einen Sinn ergibt.

3 Zukünftiges Vorhaben in der Masterthesis

Die Hauptthemen, die als weitere Schritte in der Masterthesis gesehen werden, sind die automatische Ermittlung der 3-D-Koordinaten, welche in Abschnitt 3.1 dargestellt wird. Erstellung einer Planungs- und Koordinationseinheit zum automatischen Durchsuchen der Wohnung. Damit verbunden ist die Steuerung der PTZ-Kamera, der Lösungsansätzen werden in dem Abschnitt 3.2 erläutert.

3.1 Bestimmung der Raumposition

Für die sprachliche Beschreibung eines Objektes im Raum und somit auch für die Kommunikation mit dem 3-D-Gebäudemodell ist die Umrechnung der 2-D-Bildkoordinaten in die 3-D-Raumkoordinaten eine grundlegende Voraussetzung. Angelehnt an das menschliche räumliche Sehen, wird in der digitalen Bildverarbeitung das Stereoanalyseverfahren verwendet, um die Position des Objektes im Raum anhand geometrischer Beziehungen zu bestimmen.

3.1.1 Stereoanalyse

In der Masterthesis soll nur statische Stereoanalyseverfahren betrachtet werden, da hier von der Annahme ausgegangen wird, dass keine Objekt- oder Kamerabewegung während der Aufnahme der Bilder stattfinden.

Im Folgenden werden die Verarbeitungsschritten des Verfahren und der Lösungsweg erläutert.

1. Bildaufnahme:

In einer Szene wird das Objekt von zwei unterschiedlichen Standorten bzw. von zwei verschiedenen Kameras aufgenommen.

Der Lösungsansatz in der Masterthesis ist wie folgt: Sobald das Objekt von einer Kamera (CAM-1) in einem Bereich der Wohnung erkannt wird, soll eine zweite Kamera

beauftragt werden, in den Bereich CAM-1 zu durchsuchen und das selbe Objekt wieder erkennen.

2. Merkmalextraktion:

Es werden signifikante Bildmerkmale (hier SURF-Feature-Points) in den Bildern bestimmt.

3. Korrespondenzanalyse:

Bestimmung von korrespondierenden Elementen in beiden Bildern. Diese wird mit Hilfe des SURF-Matching-Verfahrens realisiert.

4. Tiefenwertbestimmung:

Die Berechnung der Disparität, der Differenzvektor zweier korrespondierenden Merkmalen liefert die Tiefeninformation eines Objektes in der Szene. Hierfür werden zwei eindeutige Matching-points gewählt und daraus die Disparität berechnet.

Die Grundvoraussetzung der Stereoanalyseverfahren ist die Kamerakalibrierung und die parallele Anordnung der Kameras in der Szene.

Annahmen und Einschränkung der Stereo Vision Im Allgemeinen stellte die Korrespondenzanalyse das schwierigste Teilproblem im Stereoverfahren dar. Da beim Bestimmen der korrespondierenden Bildmerkmale in Stereobildern Mehrdeutigkeit bei der Zuordnung in einer Szene auftauchen können.

Das Problem wird durch das eingesetzte SURF-Matching-Verfahren reduziert bzw. teilweise auch behoben. Das SURF-Matching-Verfahren ist eine merkmalsbasierte Korrespondenzanalyse (vgl. [Reinhard Klette, 1996](#), Seite 156-165). Hierdurch sind die Mehrdeutigkeit bei Korrespondenzen geringer und die Bestimmung der Disparitäten erfolgt genauer, denn die Position der Merkmale kann subpixelgenau berechnet werden.

Dennoch bietet das statische Stereoanalyseverfahren nicht in allen Bereichen der Wohnung eine Lösung zur Gewinnung der Tiefeninformation an. Zum einen sind nicht alle Bereiche der Wohnung mit zwei Kameras ausgestattet, zum anderen kann nur mit aufwendiger Steuerung (Justierung) der Kameras nur in einigen Bereichen der Wohnung realisiert werden. Da nicht überall die optischen Achsen (Cl und Cr) der Kameras parallel verlaufen und die Basislinie (X- und Y- Achse) richtungsmäßig mit der Bildebene, wie in der Abbildung [??vgl.\]\[fig:KameraParalle](#) dargestellt ist, übereinstimmt.

Sind die Kameras beliebig zueinander ausgerichtet wie in der [Abb.3.2](#) dargestellt, müssen die Epipolarlinien und Epipolarebenen bestimmt werden, was einen höheren Rechenaufwand mit sich bringt. (vgl. [3D-Bildverarbeitung Prof. Andreas Meisel, 2012](#))

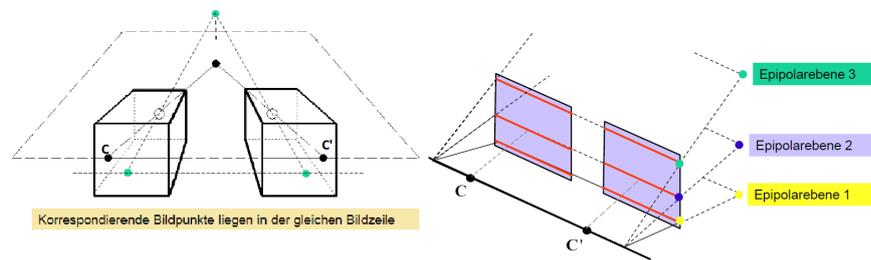
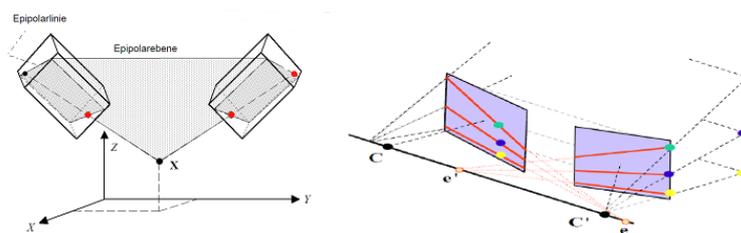


Abbildung 3.1: Achsen-parallele Anordnung der Kameras

Abbildung 3.2: Epipolarebenen und Epipolarlinien Quelle[(vgl. [3D-Bildverarbeitung Prof. Andreas Meisel](#) , 2012)]

In der Masterarbeit ist zu untersuchen, welche weiteren Verfahren für die Bestimmung des 3-D-Punktes im Raum auch mit einer Kamera und evtl. mithilfe eines aktiven Lichtstrahls realisierbar sind und sich als effizient für die Umgebung im Living Place Hamburg erweisen. Dies kann z. B. durch die Verfahren trinokulares Stereo, aktives Stereo und axiales Stereo (vgl. [Jiang und Bunke, 1997](#)) realisiert werden.

Die Implementierung der oben genannten mathematischen Stereoanalyseverfahren zur Bestimmung von 3-D-Koordinaten bietet die OpenCV-Bibliothek mit ihren Methoden an. Diese Bibliothek soll später auch für die Kamerakalibrierung verwendet werden.

Die Kamerakalibrierung Die Kamerakalibrierung ist ein Verfahren anhand dessen die Geometrie der Kamera (Bestimmung der externen und internen Parameter der Kamera) und der Relation zwischen den Kamerakoordinaten und den Weltkoordinaten bestimmt wird.

Dieses stellt bei den PTZ-Kameras aufgrund ihrer Eigenschaft, im Vergleich zu statischen Kameras eine weitere Herausforderung in dieser Masterarbeit dar. Gegenüber statischen Kameras können die Dome PTZ-Kameras geschwenkt, gezoomt und gedreht werden. Hierbei sind Lösungswege zu finden, um mit möglichst wenigen Parameterneinstellung die Kamerakalibrierung durchzuführen.

3.2 Die Kamerasteuerung

Die Integrierung der Dome PTZ-Kamera im Living Place Hamburg stellt den Entwickler vor eine weitere Herausforderung.

Bevor die Raumpunkte über die oben genannten Verfahren ermittelt werden könnten, muss die Voraussetzung geschaffen werden, die Kameras in jeder beliebigen Position zu positionieren.

Für die Steuerung der PTZ-Domekameras soll eine Anwendung erstellt werden, die über eine Schnittstelle die Steuerungsfunktionen Zoomen, Schwenken und Lenken automatisch einstellt und steuert. Laut Hersteller unterstützen die Kameras das standardisierte Protokoll Open Network Video Interface (ONVIF) für den Zugriff auf eine IP-Kamera über eine SOAP Schnittstelle. Dies ermöglicht unter anderem, den Status der Kameras jeder Zeit abzufragen und jede einzelne Kamera in variabler Geschwindigkeit in jede Position zu schwenken, zu zoomen und zu lenken.

Für die technische Umsetzung ist in der Masterarbeit herauszufinden, wie die Open Network Video Interface (ONVIF) in eine C#-Anwendung eingebunden werden kann.

Diese Steuerung der Kamera ist die Grundlage für die Erstellung einer Planungs- und Koordinationsalgorithmus für das Durchsuchen der Wohnung, und wird im Folgenden Abschnitt dargestellt.

Planungs- und Koordinationseinheit Die Planungs- und Koordinationseinheit der Kameras soll entscheiden, wann, welche und wie viele PTZ-Kameras im Living Place Hamburg bei einer Suchanfrage beauftragt werden. Der Entscheidungsalgorithmus sollte adaptiv sein, d. h. sich an unterschiedliche Bedingungen und Kontexte anpassen können.

Der hohe Anspruch für einer Planungs- und Koordinationseinheit liegt darin, dass möglichst wenige Video-Sequenzen und Kameras für das Durchsuchen der Wohnung genutzt werden.

Sie muss unterbinden, dass ein Bereich der Wohnung von zwei Kameras gleichzeitig durchsucht wird, um so effizient und schnell ein Ergebnis zu erzielen.

Nach der Implementierung und Einbindungen solch einer Planungs- und Koordinationseinheit im Gesamtsystem, sind weitere Analysen zu erarbeiten um Lösungswege für folgende Fragestellungen zu finden:

Inwieweit müssen die Kameraeinstellungen angepasst werden, um herauszufinden, ob sich das gesuchte Objekt im Suchbereich der jeweiliger Kamera befindet? Ab welchem Zeitpunkt wird die Suche einer Kameras als Misserfolg gewertet?

Mit welchen Einstellungen erzielt die Kamera den besten Sucherfolg?

Wie sollen allen Kameras benachrichtigt werden, dass ein Objekt im Raum gefunden wurde?

Neben diesen Fragestellungen, welche nur einige Aspekte der Komplexität beinhalten, werden im Rahmen der Masterthesis noch weitere Versuche in der Praxis nötig sein, um einen vollautomatischen Algorithmus zum Durchsuchen der Wohnung zur realisieren.

4 Fazit

Thesis Outline stellt die weitere Vorgehensweise der Masterthesis „Modellbasiertes Suchen von Objekten in einer Smart-Home-Umgebung“ vor.

Die Grundlagen zu den theoretischen Vorüberlegungen wurden im Rahmen des Masterstudiums in verschiedenen Hausarbeiten und Projekten erarbeitet. Die Ergebnisse der bisherigen Ausarbeitungen sind im Abschnitt 2 [Vorarbeiten zur Masterthesis] dargestellt.

Im Rahmen der vorgestellten Projekte wurden die ersten Implementierung und Anwendungssystem in der Praxis erprobt. Dabei wurden die ersten Schritte für die Realisierung eines kameragestützten Systems zur Auffindung von Objekten im Raum vorgenommen.

Ein Schwerpunkt der Masterthesis wird sich unter anderem mit der Fragestellung zu der Bestimmung der aktuellen Raumpositionen beschäftigen. Ein weiterer Schwerpunkt wird die Kamerasteuerung an sich sein. In dem Abschnitt 3 [Zukünftiges Vorhaben in der Masterthesis] wurden Lösungsansätze für diese Schwerpunkte erarbeitet.

Welcher der vorgestellten Lösungsansätze sich am Ende für die Realisierung eines „intelligenten Anwendungssystem“ eignet, wird sich durch die Erprobung in der realen Umgebung erst noch herausstellen.

Literaturverzeichnis

- [3D-Bildverarbeitung Prof. Andreas Meisel 2012] 3D-BILDVERARBEITUNG PROF. ANDREAS MEISEL , Prof. Dr.-Ing. Andreas M.: *Vortragsfolien zu Grundlagen der 3D-Bildverarbeitung*. 2012. – URL http://www.informatik.haw-hamburg.de/uploads/media/AW_3DBV_V03.pdf
- [Hardenack 2011] HARDENACK, Frank: Das intelligente Bett - Sensorbasierte Detektion von Schlafphasen. In: *Masterarbeit eingereicht im Rahmen der Masterprüfung im Studiengang Master Informatik am Department Informatik der Fakultät Technik und Informatik der Hochschule für Angewandte Wissenschaften Hamburg* (2011), Oktober
- [Jiang und Bunke 1997] JIANG, Xiaoyi ; BUNKE, Horst: *Dreidimensionales Computersehen: Gewinnung und Analyse von Tiefenbildern*. Springer-Verlag, Berlin und Heidelberg, 1997
- [Karstaedt 2011a] KARSTAEDT, Bastian: PO2 - Entwicklung und Integration der Indoor Spatial Information Services in das Living Place Hamburg. In: *Berichte Masterprojekt 1* (2011)
- [Karstaedt 2011b] KARSTAEDT, Bastian: Seminarring - Entwicklung eines Indoor Spatial Information Service für IFC-basierte Gebäudemodelle. In: *Ringvorlesung Seminarvorträgen* (2011)
- [Lowe 1999] LOWE, D.G.: Object recognition from local scale-invariant features. In: *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on* Bd. 2, 1999, S. 1150 –1157 vol.2
- [LP-HH] LP-HH: Living Place Hamburg.
- [Najem 2008] NAJEM, Hosnia: Realisierung einer Bibliothek für die schnelle Objektverfolgung und Objekterkennung in Bildsequenzen. In: *Bachelorarbeit* (2008), November
- [Najem 2009] NAJEM, Hosnia: AW1. (2009)
- [Najem 2010] NAJEM, Hosnia: *Hausarbeit - Anwendung 2 Raumkognition*, Studiengang Master Informatik am Studiendepartment Informatik der Fakultät Technik und Informatik an der Hochschule für Angewandte Wissenschaften Hamburg, Dissertation,

2010. – URL <http://users.informatik.haw-hamburg.de/~ubicomp/projekte/master2010-aw2/najem/bericht.pdf>

[Najem 2011] NAJEM, Hosnia: *Aufbau der Infrastruktur für das Model-basierte objektsuche im Living Place Hamburg MA-INF1 Projektbericht PO2*, Hochschule für Angewandte Wissenschaften hamburg, Diplomarbeit, Februar 2011

[Najem Feb 2011] NAJEM, Hosnia: *Aufbau der Infrastruktur für das Model-basierte objektsuche im Living Place Hamburg*. In: *MA-INF1 Projektbericht PO1* (Feb 2011)

[PTZ-Kameras] PTZ-KAMERAS: PTZ Domekameras Dallmeier.

[Reinhard Klette 1996] REINHARD KLETTE, Karsten S.: *Computer Vision Räumliche Information aus digitalen Bildern*. Vieweg Technik, 1996