



Hochschule für Angewandte Wissenschaften Hamburg  
*Hamburg University of Applied Sciences*

## Projektbericht PO2

Hosnia Najem

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

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

Projektbericht PO2 eingereicht im Rahmen der Masterprojekt  
im Studiengang Master 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 31.August.2011

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>4</b>
<b>2</b>	<b>Projekt Umsetzung</b>	<b>6</b>
2.1	Anwendungsszenarien . . . . .	6
2.1.1	Kameragestützte Suche . . . . .	6
2.1.2	Sprachliche Lokalisierung des Objektes . . . . .	7
2.2	Entwurf . . . . .	9
2.3	Projekt Ergebnisse . . . . .	11
2.3.1	Teaching . . . . .	11
2.3.2	Searching . . . . .	12
2.3.3	3-D-Raumposition . . . . .	14
2.3.4	Lokalisation . . . . .	14
<b>3</b>	<b>Fazit</b>	<b>20</b>
3.1	Ausblick . . . . .	20
	<b>Literaturverzeichnis</b>	<b>22</b>

# 1 Einleitung

Das Projekt „Modellbasiertes Suchen von Objekten“ hat den Anspruch, ein computergestütztes Kamerasystem zum Auffinden von Objekten in intelligenten Wohnungen (Smart Homes) zu entwickeln. Der Living Place Hamburg (vgl. Abb 1.1) integriert bereits verschiedenste Anwendungssysteme und Sensor-Technologien, auf die das Projekt zugreift und mit deren Hilfe sich das Living Place Hamburg zu einer intelligenten Wohnung entwickelt. Eines dieser Anwendungssysteme, das 3-D-Modell des Living Place Hamburg ermöglicht den Zugriff auf eine virtuelle Abbildung der Wohnung. Über eine Schnittstelle zur UBISENSE<sup>1</sup> wird diese 3-D-Modell mit Echtzeitdaten über beweglichen Gegenständen angereichert. Das Projekt „Modellbasiertes Suchen von Objekten“ verbindet die Daten des 3-D-Modelles mit den live Bildern der Pen-Ten-Zoom-Domekamera (PTZ-Domekamera<sup>2</sup>) um diese auswerten zu können

Dementsprechend liegt der Fokus des Projekts auf der Entwicklung einer Schnittstelle, die Daten der Kameras und des 3-D Modelles vereinigt um die Lokalisierung von beweglichen Objekten im Living Place zu unterstützen.

Die Grundlagen und theoretischen Vorüberlegungen des Projektes wurden bereits im Rahmen des Masterstudiums erarbeitet (Vgl. Najem, Feb 2011), (Vgl. Najem, 2011). Dieser Projektbericht gibt einen kurzen Überblick über die bislang in der Umsetzung erzielten Ergebnisse.

---

<sup>1</sup>RFID basiertes Sensornetzwerk zur Lokalisierung, mit Tag versehenen Gegenständen

<sup>2</sup><http://www.dallmeier-electronic.com/de/cctv-ip-videoueberwachung/kameras/modelle/ptz-domekamas/ddz4010-yyhshd-domerar.html>

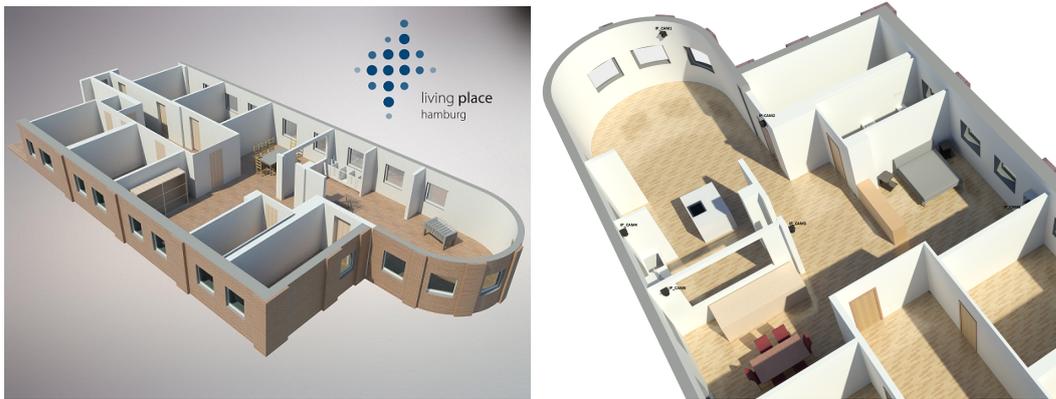


Abbildung 1.1: 3-DModell Living Place Hamburg, erstellt von Bastian Karstaedt (Vgl. [Karstaedt, 2011a](#))

Im weiteren Verlauf ist dieser Projektbericht wie folgt gegliedert. Das Kapitel Projekt Umsetzung [2] beschreibt ausgehend von einer Analyse der Anwendungsszenarien den Entwurf des Projektes um abschließend die erreichten Ergebnisse in Abschnitt 2.3.1 zu diskutieren. Das abschließende Kapitel Fazit [3] gibt schließlich einen Ausblick auf zu erwartende Entwicklungen im weiterem Projektverlauf.

## 2 Projekt Umsetzung

### 2.1 Anwendungsszenarien

Dem für das Projekt erstellten Entwurf der in dem Abschnitt 2.2 gezeigt wird, liegen folgende Szenarien zugrunde.

#### 2.1.1 Kameragestützte Suche



Abbildung 2.1: Szenario: Suche des Objektes "Buch" im Living Place Hamburg [Quelle:Hosnia Najem]

Die Abbildung 2.1 zeigt das Szenario, in dem ein Bewohner des Living Place Hamburg, das in diesem Projekt entwickelte Kamerasystem für die Suche nach einem Buch nutzt.

Zur Erfassung von Objekten bietet das Kamerasystem die Möglichkeit, ein Objekt vor die Kamera zu halten, um daraus das Referenzmodell zu erstellen und dieses im Gesamtsystem zu hinterlegen.

Für die Suche nach einem bestimmten Objekt wird aus den im System hinterlegten Referenzmodellen eines, durch den Nutzer selektiert.

Im oben dargestellten Szenario wurde das gesuchte Objekte „Buch“ vom Kamerasystem auf der Kommode im Schlafbereich lokalisiert.

## 2.1.2 Sprachliche Lokalisierung des Objektes

Um dem Bewohner auf die Serviceanfrage „Wo ist das Buch?“ mit umgangssprachlichen Lokalisationsäußerungen antworten zu können, wurde in diesem Projekt eine Schnittstelle zum Masterprojekt „Indoor Spatial Information Service“ (ISIS) (Vgl. [Karstaedt, 2011b](#)), (Vgl. [Karstaedt, 2011a](#)) hergestellt.

„Das dreidimensionale auf Basis der IFC-Klassen generierte Modell des Living Place Hamburg erfasst alle im Living Place Hamburg vorhandenen Möbel und deren Positionen.“ [zitat: Bastian]

Mit dem 33-D-Modell ist der Indoor Spatial Information Service (ISIS) implementiert, der Anfragen über räumliche Relationen von Objekten erlaubt.

Das Szenario in Abbildung 2.2 zeigt die Interaktion zwischen dem Kamerasystem und dem 3-D-Modell im Living Place Hamburg.

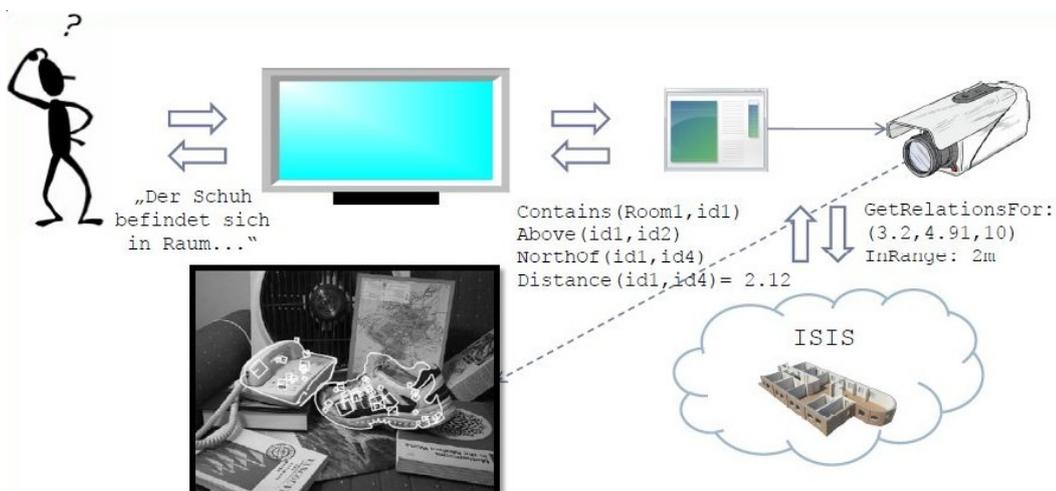


Abbildung 2.2: Interaktion zwischen Anwender, Kamerasystem und ISIS (erstellt von. [Karstaedt, 2011a](#))

Nachdem das gesuchte Objekt vom Kamerasystem des Living Place Hamburg gefunden wurde, wird mit den ermittelten 3-D-Koordinaten des Objektes die Serviceanfrage „Welche Objekte befinden sich in unmittelbarer Umgebung des gefundenen Objektes?“ an den ISIS gestellt

Der ISIS liefert Informationen darüber, in welchem Raum sich das gesuchte Objekt bzw. seine Koordinaten befinden und welche Nachbarobjekte in unmittelbarer Umgebung der 3-D-Koordinaten des gesuchten Objektes im 3-D-Modell vorhanden sind.

Darüber hinaus liefert er Informationen über räumliche Relationen der Nachbarobjekte untereinander und in Bezug auf das gesuchte Objekt.

Die räumlichen Relationen werden auf Basis der im Projekt „Indoor Spatial Information Service“ definierten Prädikate wie „contains“, „above“, „below“, „northOf“, usw. angegeben.

Um dem Bewohner die Position des gesuchten Objektes und seine relative Lage zu den anderen Gegenständen in der Wohnung umgangssprachlich mitzuteilen, werden diese Prädikate vom Kamerasystem in lokalisierende Ausdrücke wie „im“, „auf“, „unter“ usw. übersetzt.

Die Kooperation mit dem Projekt „Indoor Spatial Information Service“ ermöglicht dem Kamerasystem, dem Bewohner die Position des gesuchten Objektes wie folgt zu formulieren:

„DAS BUCH IST IM SCHLAFBEREICH AUF DER KOMMODE.“

## 2.2 Entwurf

Für die Umsetzung der Szenarien wurden im Rahmen des Projektes die im folgenden abgebildeten Entwürfe erstellt.

Anhand dieser Entwürfe soll ein Einblick in das Zusammenspiel der einzelnen Funktionalitäten des Kamerasystemes und die Integration des externen Services ISIS gegeben werden.

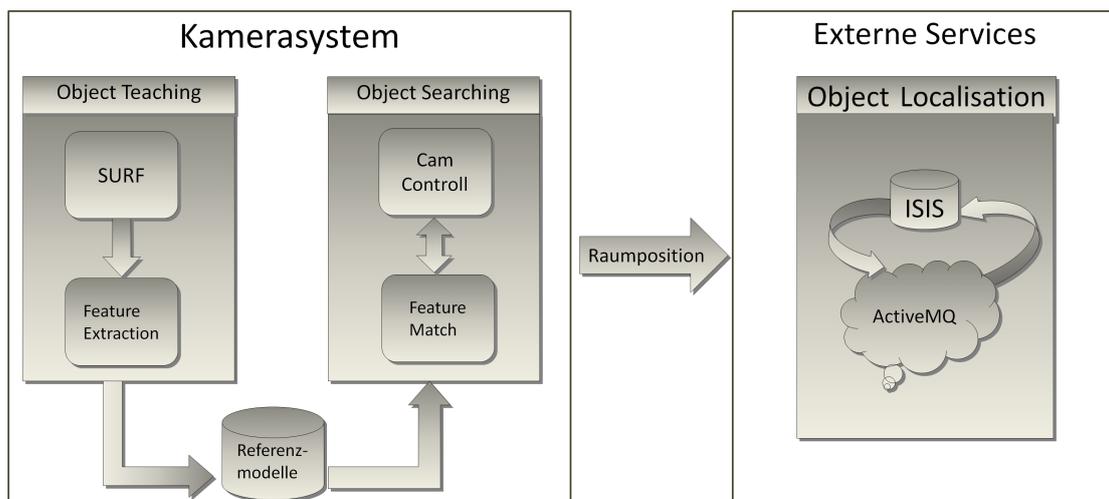


Abbildung 2.3: Projekt-Entwurf

Im Laufe des Projektes haben sich aus der Zielsetzung und den Szenarien die Phasen "Teaching", "Searching" und „Localisation“ entwickelt, die im Folgenden näher erläutert werden.

**Teaching** Die erste Phase ist das Erfassen und Erstellen von Referenzmodellen.

Das Teaching befasst sich unter anderem mit den Fragen:

- Was soll gesucht werden?
- Wie können Serviceanfragen generiert werden?
- Wie können Referenzmodelle aus den Kamerabildern erstellt werden?
- Welche Auflösung müssen die Bilder haben, um robuste und markante Bildpunkte aus den Bildern zu extrahieren?

**Searching** Die zweite Phase umfasst die Objektsuche anhand der Referenzmodelle in den Bildsequenzen der Kamera.

Das Searching integriert unter anderem die Lösung für die folgenden Fragestellungen:

- Wo befindet sich das Objekt in den Bildsequenzen?
- Wo befindet sich das Objekt in der realen Umgebung, d. h. wo im Living Place Hamburg befindet es sich?

**Localisation** Die letzte Phase integriert die sprachliche Lokalisierung des gefundenen Objektes im Raum. Die Localisation thematisiert unter anderem die Fragestellungen:

- Wie können die räumlichen Informationen, die das ISIS liefert, semantisch interpretiert werden?
- Wie kann das Kamerasystem lokalisierende Aussagen in einem Antwortsatz definieren, sodass der Bewohner die Position des gefundenen Objektes im Raum erkennt.

Wie in Abb. 2.3 dargestellt, bilden Teaching- und Searching-Module die Kernkomponenten des Kamerasystems in diesem Projekt.

Über die zentrale Kommunikationsschnittstelle im Living Place Hamburg, das ActiveMQ<sup>1</sup> (Vgl. [Otto und Voskuhl, 2011](#)), wird die Interaktion mit weiteren Applikationen und bereits integrierten Projekten (externe Services) in das in diesem Projekt erstellte Kamerasystem eingebunden.

---

<sup>1</sup><http://activemq.apache.org/>

## 2.3 Projekt Ergebnisse

Im folgenden Abschnitt sind die Ergebnisse der Implementierung der einzelnen Phasen aufgeführt.

### 2.3.1 Teaching

Der in der Grafik [2.3](#) dargestellte Teilbereich „Object Teaching“ umfasst die automatische Erkennung und Speicherung von Referenzmodellen.

Die Umsetzung der Funktionalitäten basiert auf dem SURF-Verfahren (Vgl. [Lowe, 1999](#)), (Vgl. [Najem, 2009](#)). Die erste Umsetzung wurde in einem früheren Projektbericht (Vgl. [Najem, Feb 2011](#)) vorgestellt.

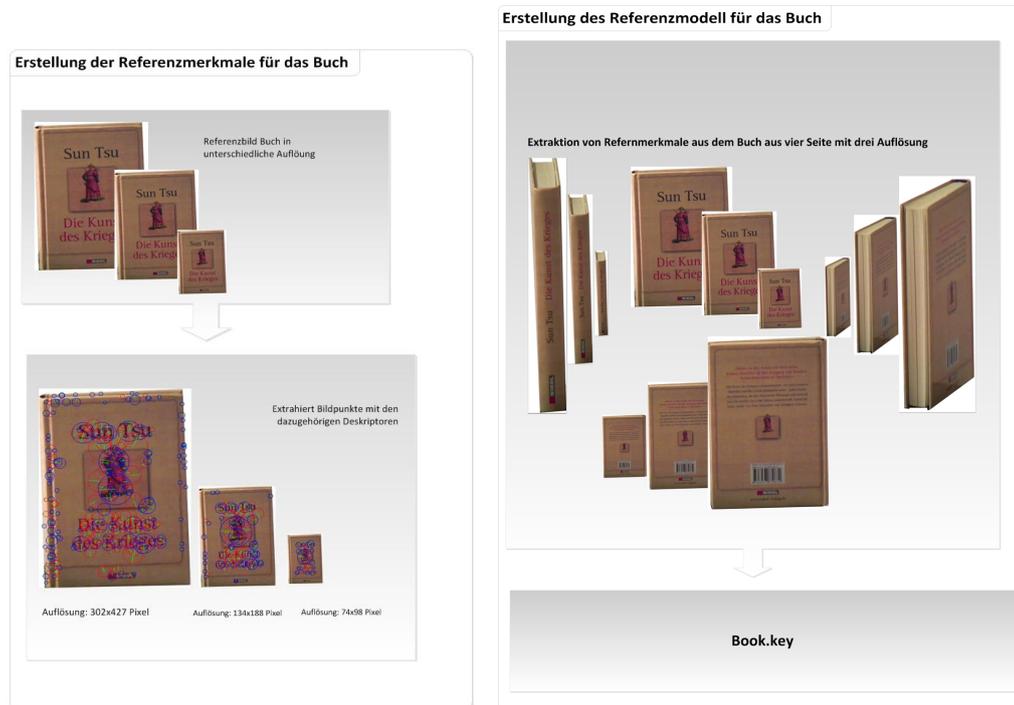
Funktionale Tests mit der SURF-Implementierung haben gezeigt, dass bei der Erstellung des Referenzmodells für ein Objekt, die Bilder sehr hohen Kontrast aufweisen müssen, um robuste und markante Bildpunkte aus den Bildern zu extrahieren (Vgl. [Najem, Feb 2011](#)).

Um das Objekt in der Wohnung mit unterschiedlichen Zoomeinstellungen und somit auch mit unterschiedlichen Auflösungen der Kameras wiedererkennen zu können, ist es von Vorteil, wenn ein Modell des gesuchten Objektes von allen Seiten und in unterschiedlichen Auflösungen besteht, wie in [Abbildung 2.4 \(b\)](#) dargestellt.

Aus den Referenzbildern werden durch den SURF-Algorithmus Referenzen extrahiert und aus diesen Merkmale generiert. Diese Merkmale werden in einer "\*.key"Datei pro Referenzmodell gespeichert. Zusätzlich wird zu jedem Objekt eine Reihe von beschreibenden Informationen bzw. Eigenschaften, der Tag abgelegt.

Diese Eigenschaften werden von dem Benutzer beim Erstellen des Referenzmodells manuell eingeben und im Gesamtsystem hinterlegt.

Die Speicherung von Tags soll im weiteren Projektverlauf zur Analyse der Objekte in einer Szene bzw. in einem Kontext dienen. Diese eröffnet weitere Möglichkeiten im Rahmen des Projektes.



(a) Extraktion von Referenzpunkten mit ihren Merkmalen für das Buch mit jeweils drei unterschiedlichen Auflösungen aus einer Perspektive.

(b) Die Book.key-Datei speichert die extrahierten Referenzpunkte und ihre Merkmale für das Buch aus vier Perspektiven mit jeweils drei unterschiedlichen Auflösungen

Abbildung 2.4: Erstellung des Referenzmodells „Buch“ aus vier Perspektiven mit jeweils drei unterschiedlichen Auflösungen

### 2.3.2 Searching

Das in Abbildung 2.3 dargestellte Modul „Object Searching“ integriert zum einen die Funktionalitäten zur Steuerung der im Living Place Hamburg eingebauten PTZ-Domekamera<sup>2</sup> und zum anderen das sogenannte Matching-Verfahren.

Beim Matching-Verfahren werden Korrespondenzen zwischen dem gewählten Referenzmodelle und den Bildmerkmalen der vorkommenden Objekte in den Bildsequenzen der Kamera gesucht.

<sup>2</sup><http://www.dallmeier-electronic.com/de/cctv-ip-videoueberwachung/kameras/modelle/ptz-domekameras/ddz4010-yyhshd-domerar.html>

Für die erste Umsetzung des Szenarios in Abbildung 2.1 wurde zunächst mit nur einer Kamera im Living Place Hamburg gearbeitet, die manuell über die von der PTZ-Domekameran zur Verfügung gestellte Web-Oberfläche gesteuert wurde.

In der Praxis hat sich gezeigt, dass die Suche mit den Kameras beschleunigt wird, wenn das Referenzmodell Informationen über die Objekte in verschiedenen Zoomstufen enthält. 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-Domekameran beim Matching-Verfahren keine HD-Auflösungen aufweisen müssen.

Die Ergebnisse des „Object Searching“ sind in der Abbildung 2.5 dargestellt. Diese zeigt die Wiedererkennung des Objektes „Buch“ anhand des zuvor erstellten Referenzmodells ("book.key") mit den PTZ-Domekameran in zwei Szenen im Living Place Hamburg .

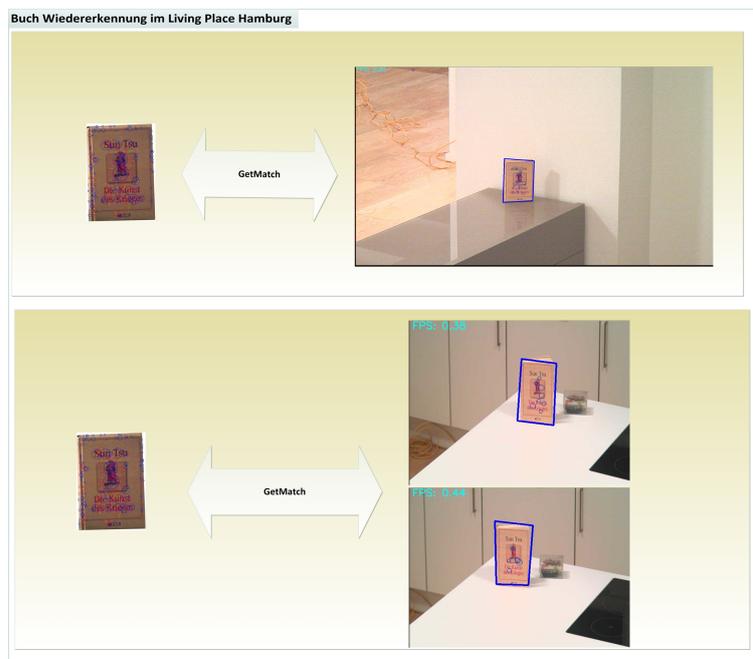


Abbildung 2.5: Wiederfinden des Buches anhand des Referenzmodells im Living Place Hamburg

### 2.3.3 3-D-Raumposition

Das SURF-Verfahren liefert die Position des gefundenen Objektes im Bild (z. B. Bildkoordinaten). Um die Position des Objektes im Raum zu bestimmen, müssen außer den Bildpunkten die dazugehörigen Raumpunkte bestimmt werden.

Für die Kommunikation zwischen dem Kamerasystem und dem 3-D-Gebäudemodell ist die Umrechnung der 2-D-Bildkoordinaten in die 3-D-Raumkoordinaten eine grundlegende Voraussetzung.

Für die ersten Tests wurden die 3-D-Koordinaten des Objektes in der Wohnung manuell ermittelt. Hierfür wurde die PView-Software<sup>3</sup>

Eine weitere Möglichkeit, um für eine Position im Living Place Hamburg die genauen Raumkoordinaten manuell zu ermitteln, bietet das Tool „Ubisense Simulator“<sup>4</sup>. Dabei wird ein Objekt auf einem 2-D-Modell des Living Place Hamburg bewegt und dieses liefert die äquivalenten 3-D-Raumkoordinaten.

### 2.3.4 Lokalisation

Das Modul „Object Localisation“ aus der Abbildung 2.3 stellt die Schnittstelle zum 3-D-Modell des Living Place Hamburg dar, entwickelt von Bastian Karstaedt in Rahmen seines Masterprojektes (Vgl. [Karstaedt, 2011a](#)), (Vgl. [Karstaedt, 2011b](#)).

Die zentrale Kommunikationsschnittstelle zum Austausch von Nachrichten und Daten zwischen den einzelnen Applikationen im Living Place Hamburg bildet der ActiveMQ.

Unter Verwendung des ActiveMQ wurde eine Blackboard-Architektur als Zentrale Kommunikationsplattform im Living Place Hamburg realisiert. Die Umsetzung beinhaltet die Anforderung an einer intelligenter Umgebung.(Vgl. [Otto und Voskuhl, 2011](#)).

Für Serviceanfragen an ISIS wurde für den Zugriff auf den ActiveMQ eine Client-Server Applikation in der Sprache C# implementiert, welche alle Anfragen (Request) und Antworten (Response) an ISIS über einen „ActiveMQ-PushPullClient“ abgewickelt.

ISIS bietet für die Kommunikation mit dem Kamerasystem zwei Schnittstellen an: eine ActiveMQQueue namens „LP3D.REQUEST.SPATIALRELATIONSHIPS“, um Anfragen entgegenzunehmen, und ein ActiveMQTopic namens „LP3D.RESPONSE.SPATIALRELATIONSHIPS“, um die angeforderten Informationen zu einer Raumkoordinate zu publizieren.

---

<sup>3</sup>PView ist die zentrale Steuerelement der Dallmeier Video Management Systeme , der den Anwender Steuer Funktionalitäten der Kameras anbietet. Weiteres: . [PView-Softwar von Dallmeier](#).

<sup>4</sup>Ubisense: <http://www.ubisense.net/en/resources/factsheets/ubisense-developer.html>

Eine ActiveMQ-Message im Living Place Hamburg ist definiert als eine Textnachricht im JSON-Format (Vgl. [Crockford, 2006](#)), die mindestens *eine Versionsnummer* und eine eindeutige ID der Applikation enthält. Für die Umsetzung des Szenarios in der Abbildung [2.2](#) wurde die ActiveMQ-Message um ein weiteres Feld namens „*ReferencPoint*“ erweitert, das die 3-D-Koordinaten des gesuchten Objektes enthält.

Beispielhaft ist in Listing [2.1](#) die JASON-Nachricht dargestellt, der die Serviceanfrage „*Welche Objekte befinden sich in unmittelbarer Umgebung der Koordinaten(x,y,z)?*“ an den ISIS auslöst.

Listing 2.1: Beispiel: Serviceanfrage von dem Kamerasystem an den ISIS

```
{ "id":1235,
  "Version":"1.0",
  "ReferencPoint":{"x,y,z}"
}
```

Die angeforderten räumlichen Informationen und die Relationen werden durch ISIS auf der ActiveMQTopic namens „*LP3D.RESPONSE.SPATIALRELATIONSHIPS*“ publiziert und von hier auf das Kamerasystem, welches mit der ActiveMQTopic verbunden ist, weitergeleitet.

Für die Kommunikation wurden für die erfassten Möbel im 3-D-Modell des Living Place Hamburg globale Bezeichnungen, sogenannte Objekt-IDs definiert. Mit deren Hilfe wird gewährleistet, dass Teilnehmer, die eine Serviceanfrage an das 3-D-Modell des Living Place Hamburg stellen, immer über dieselben Möbel sprechen.

Die bislang definierten IDs im Living Place Hamburg sind in der Abbildung [2.6](#) dargestellt. Auf Basis dieser definierten Objekt-IDs wird die Antwortnachricht von ISIS generiert.

Die Antwortnachricht ist eine JASON-Massage mit ID und Versionsnummer der Anfrage. Die angeforderten semantischen Informationen zu den angegebenen 3-D-Koordinaten werden im Feld namens „*SpatialRelations*“ gespeichert.

Die Informationen über die Relationen der gefundenen Objekte werden auf Basis definierter Prädikate, die Beispielhaft in der Listing [2.2](#) dargestellt sind, aufgebaut.

Listing 2.2: Beispiel JASON-Nachricht: ISIS Service Antwort

```
{ "id":11235,
  "Version":"1.0",
  "SpatialRelations":{"Contains(Room1;ref _Point),Above(ref
    _Point;commode),Distance(commode;bed) = CLOSE}"
}
```

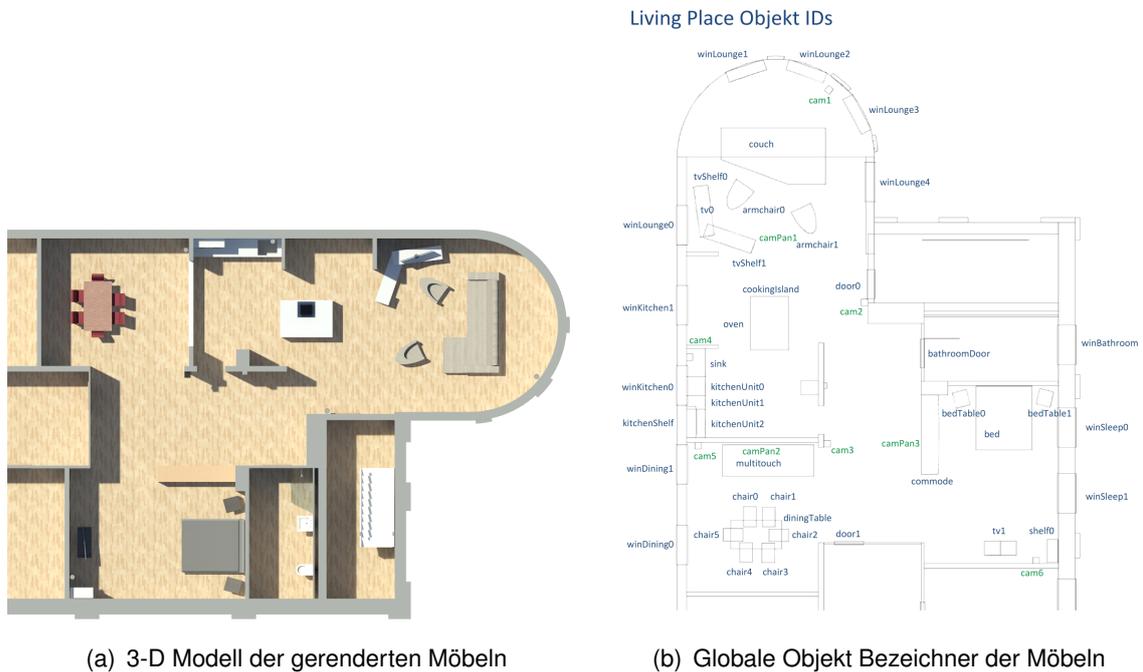


Abbildung 2.6: 3-D Modell Living Place Hamburg (erstellt von. [Karstaedt, 2011a](#))

Um den Prädikaten weitere semantische Informationen zuordnen zu können, wurden ihnen Kategorien wie topologische, direktionale und metrische Relationen zugeordnet.

**Lokalisationsäußerungen** Damit der Bewohner solche Lokalisationsäußerungen verstehen kann, muss das Kamerasystem (wie bereits erwähnt) die gewonnenen Informationen analysieren und interpretieren, um diesen semantischen Informationen Bedeutungen zuzuweisen zu können.

Für die sprachliche Lokalisierung eines Objektes im Raum muss der Zusammenhang zwischen den räumlichen und sprachlichen Strukturen betrachtet werden (Vgl. [Najem, 2010](#)).

Um einen Satz zu formulieren, muss das Kamerasystem auf Basis der gewonnenen Informationen noch Entscheidungen treffen wie:

- Welche Objekte kann das ISIS wählen, um eine aussagekräftige Lokalisierung formulieren zu können?
- Werden die Lokalisationsäußerung aus seiner eigenen Sicht, aus der Sicht des Bewohners oder bezogen auf ein weiteres Objekt in der Nähe gewählt?

Es muss somit Entscheidungen zwischen verschiedenen Perspektiven, Bezugssystemen und den gelieferten Nachbarobjekte treffen.

In welchen aussagelogischen Ausdrücken Prädikate von dem Kamerasystem interpretiert werden können, wird im Folgenden dargestellt.

**Topologische Relationen** Das Prädikat „*Contains*“ spezifiziert die Informationen über die Topologie. Es liefert die Information, in welchem Raum sich die angefragten Koordinaten((x,y,z) (hier *ref\_Point*) befinden.

„*Contains*“ wird interpretiert als Präposition wie „IM“ oder das Prädikat „ENTHÄLT“.

Das Aussage hinter folgender gelieferter Information *Contains(Roomid; ref\_Point)* lautet:

„Der Raum enthält die Koordinate(x,y,z).“

Die Aussage hinter folgender gelieferter Information *Contains(ref\_Point; Roomid)* lautet:

„Die Koordinate(x,y,z); ist im Raum “

Mit der Äußerung "DAS BUCH LIEGT AUF DER KOMMODE" würde es dem Bewohner schwerfallen, die physikalischen Gegebenheiten der räumlichen Konstellation intern zu rekonstruieren. Denn diese Aussage enthält keine Information darüber, in welchem Raum sich die Kommode befindet. Um eine Lösung für diese Problemstellung zu finden, bzw. um diese Problematik zu unterbinden, wurden im Rahmen des Projektes folgende Regelungen für die Formulierung bzw. den Satzbau getroffen.

Das Kamerasystem bildet aus den gelieferten topologischen Relationen einen Satzanfang, der unabhängig vom Standpunkt des Betrachters und von der Kameraperspektive ist. Die gewählte Formulierung des Satzes geht immer vom Prädikat „*Contains*“ aus, das als Lokalisationsäußerung „IM“ interpretiert wird.

Somit wird jeder Antwortsatz wie folgt anfangen.

„Das gesuchte Objekt BEFINDET SICH IM :. „  
bzw. „DAS „BUCH“ BEFINDET SICH IM Schlafbereich.“

Darauf aufbauend werden mithilfe der gelieferten direktionalen Relationen die weiteren Objekte als Bezugspunkte genommen, welche dann die weitere Struktur des Satzes bestimmen.

**Direktionale Relationen** Das Prädikat „Above“ wird von dem Kamerasystem übersetzt in das Prädikat „AUF“

Die Aussage hinter dem gelieferten Information  $above(commode; ref\_point)$  lautet: Die Koordinate (x,y,z) befindet sich /ist „auf“ der *commode*.

Das Kamerasystem interpretiert anhand der Objekt-ID *commode* als die Kommode und *ref\_point* entspricht den 3-D-Koordinaten des gesuchten Objektes, hier am Beispiel ist es das „Buch“.

Der von dem Kamerasystem formulierte Satz würde mithilfe der gelieferten Informationen über die direktionalen Relationen wie folgt anfangen:

„AUF“ bzw. hier im Beispiel „AUF DER KOMMODE BEFINDET SICH DAS BUCH.“

Das Prädikat „Below“ wird von dem Kamerasystem übersetzt in das Prädikat „UNTER“.

Die Aussage hinter dieser gelieferten Information  $below(commode; ref\_point)$  lautet:

Die Koordinate (x,y,z) befindet sich /ist „unter“ der *commode*.

Der vom Kamerasystem formuliert Satz würde lauten: „UNTER DER KOMMODE .“

Im Rahmen dieses Projektes werden für den Aufbau einer Satzstruktur zunächst nur die Prädikate „above“ und „below“ betrachte.

Aufbauend auf dem ersten definierten Satz

DAS GESUCHTE „BUCH“ BEFINDET SICH IM SCHLAFBEREICH

würden mit den gelieferten direktionalen Relationen folgende Sätze von dem Kamerasystem formuliert werden können:

Bei geliefertem direktionalen Relationen  $above(commode; ref\_point)$  würde der Satz mit dem Prädikat „AUF“ ergänzt werden. z.B. wie folgt:

DAS GESUCHTE OBJEKT, „BUCH“ BEFINDET SICH IM SCHLAFBEREICH  
„AUF DER KOMMODE.“

Das Kamerasystem ordnet jedes Referenzobjekt, an Hand deren Form und Größe nur bestimmt Prädikat zu. Dieser dient der Entscheidung, einen Satz zu formulieren, der den Bewohner einen Mehrwert an Informationen liefert und einen Sinn ergibt.

Bei geliefertem direktionalen Relationen  $below(commode; ref\_point)$  würde von dem Kamerasystem keine übersetzt statt finden, da für das gelieferte Referenzobjekt „Kommode“ keine Zuordnung mit den Prädikat „UNTER“ besteht. Denn die Aussage DAS GESUCHTE OBJEKT, „BUCH“ BEFINDET SICH IM SCHLAFBEREICH.UNTER DER KOMMODE würde für den Bewohner keinen Sinn ergeben.

Weitere definierte Prädikate, die Informationen über direktionale Relationen beinhalten sind „northOf“, „southOf“, „eastOf“ und „westOf“. Die genaue Interpretation dieser semantischen Informationen in lokalisierende Ausdrücke soll im weiteren Projektverlauf stattfinden.

**Metrische Relationen** Die metrischen Informationen werden mit dem Prädikat „*Distance*“ angegeben. Diese liefern Angaben über den Abstand eines Objektes zu anderen Objekten. Als Bezugspunkt der metrischen Informationen wird entweder ein weiteres Objekt oder die angegebene 3-D-Koordinate gesehen.

Diese Angaben werden nicht quantitativ, sondern qualitativ mit den Prädikaten „VERY\_CLOSE, CLOSE, FAR, VERY\_FAR“ gemacht.

Die Aussage hinter der gelieferten Information

$Distance(commode; bed) = CLOSE$  lautet: Die *commode* ist in der Nähe vom *bed*.

Diese Aussage wird von dem Kamerasystem interpretiert als:

„DIE KOMMODE' BEFINDET SICH IN DER NÄHE VOM BETT.“.

Der oben formulierten Satz würde den gelieferten metrischen Informationen  $Distance(commode; bed) = CLOSE$  wie Folgt ergänzt werden:

DAS GESUCHTE OBJEKT, „BUCH“ BEFINDET SICH IM SCHLAFBEREICH.

„AUF DER KOMMODE.“

„NEBEN DAS BETT.“

Die Interpretation der weiteren metrischen Informationen wurden im Projektverlauf bislang nicht in Betracht gezogen. Denn dies erfordert eine weitere Analyse der räumlichen Beschreibung, der räumlichen Informationen und der räumlichen Strukturen. Deren Umsetzungen werden im weiteren Projektverlauf stattfinden.

## 3 Fazit

Im Rahmen dieses Projektes wurden die ersten Experimente für die Interaktion der Kameras mit dem Masterprojekt „ISIS“ durchgeführt, um so einen Austausch von Informationen aus den beiden Projekten im Rahmen des Living Place Hamburg zu ermöglichen. Die Einbindung der „ISIS“ als ein Serviceanbieter für die Anfragen über räumliche Informationen und die Relationen von Objekten, ermöglicht es dem Kamerasystem eine sprachliche Lokalisierung zu formulieren.

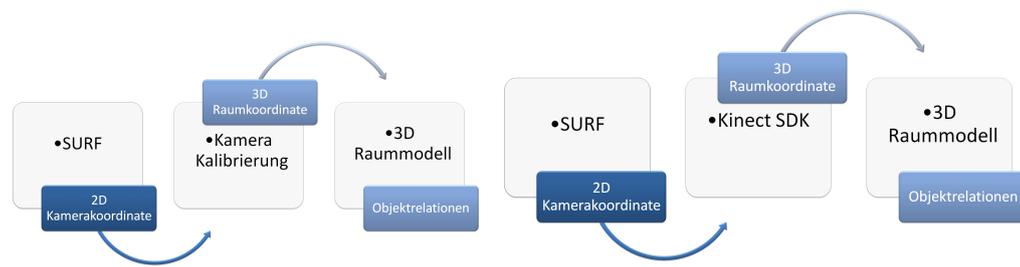
Hierfür wurde in dem Kapitel „Projekte Experimente“ gezeigt, wie die Referenzmodelle im Kamerasystem hinterlegt werden, um diese in den Bildersequenzen der Kameras im Living Place Hamburg wiederzuerkennen und deren Position in den Bildern zu lokalisieren. In dem Abschnitt 3.2.3 wurde experimentell gezeigt, wie die gelieferten Prädikate in die lokalisierenden Ausdrücke „IM“, „AUF“ und „UNTER“ übersetzt werden, damit das Kamerasystem daraus einen Satz wie „DAS BUCH BEFINDET SICH IM SCHLAFBEREICH AUF DER KOMMODE“ strukturieren kann.

### 3.1 Ausblick

Hauptthemen, die als weitere Schritte in dem Projekt gesehen werden, sind die automatische Ermittlung der 3-D-Koordinaten und die automatische Steuerung der PTZ-Domekamas im Living Place Hamburg.

**Automatische Ermittlung der Raumkoordinaten** Bislang wurden die Raumkoordinaten des Objektes manuell ermittelt. Zu den gelieferten 2-D-Bildkoordinaten aus dem SURF-Verfahren müssen für die Serviceanfragen an den „Indoor Spatial Information Service“ die zugehörigen Raumkoordinaten im Living Place Hamburg automatisch errechnet werden, so wie in Abb. 3.1 dargestellt.

Die Kamerakalibrierung ist die Grundlage dafür, dass aus Bildkoordinaten (Bildpunkte) Raumkoordinaten (Raumpunkte) mathematisch errechnet werden können



(a) Kamerakalibrierung zur Ermittlung Raumkoordinaten zu einem Bildkoordinaten (b) Einsatz von Kinect Technologie zur Ermittlung der Raumkoordinate

Abbildung 3.1: Verfahren zur Ermittlung der Raumkoordinaten

Die Kalibrierung der PTZ-Domekamas stellt im Vergleich zu statischen Kameras eine weitere Herausforderung in diesem Projekt dar. Der erste Lösungsansatz ist, für die PTZ-Domekamas eine Grundposition zu definieren, in der die Kamera kalibriert wird. Nach dem das Kamerasystem die Bildkoordinate des Objektes ermittelt hat, werden für die Berechnung der 3-D-Raumkoordinaten die Kameras in diese Grundposition eingefahren.

Ein weiterer Ansatz zur Bestimmung der 3-D-Raumkoordinaten eines Objektes ist der Einsatz einer Kinect-Kamera <sup>1</sup>[Kinect (2011)] mit deren Hilfe die Tiefeninformation eines Objektes bestimmt wird.

Im weiteren Verlauf des Projektes soll herausgefunden werden, welcher Lösungsansatz sich in der Praxis bewährt.

**Die Kamerasteuerung** Diese Suche von Objekten, wie in Abschnitt 2.3.1 dargestellt, wurde bislang mit nur einer Kamera und durch manuelle Steuerung der Kamera erzielt. Für die Steuerung der PTZ-Domekamas soll eine Applikationen erstellt werden, die über die Open Network Video Interface Forum<sup>2</sup> (ONVIF)-Schnittstelle die Steuerungsfunktionen Zoomen und Schwenken (360-Grad-Neigung) automatisch einstellt und steuert.

Zusätzlich zu der Steuerung wird eine Planungs- und Koordinationseinheit für die Kameras benötigt, die entscheidet, wann, welche und wie viele PTZ-Domekamas im Living Place Hamburg bei eine Suchanfrage beauftragt werden.

Mit der Umsetzung dieser Funktionalitäten wäre für die Integration der Kameras in der Smart-Home-Umgebung der Grundstein gelegt. Somit würde die Interaktion der Kameras, in weitere Projekte im Rahmen des Living Place Hamburg, nichts mehr im Wege stehen.

<sup>1</sup>Xbox: <http://www.xbox.com/de-DE/kinect/>

<sup>2</sup><http://www.dallmeier-electronic.com/de/cctv-ip-videoueberwachung/integration/onvif.html>

# Literaturverzeichnis

- [Crockford 2006] CROCKFORD, D.: *The application/json Media Type for JavaScript Object Notation (JSON)*. RFC 4627 (Informational). Juli 2006. – URL <http://www.ietf.org/rfc/rfc4627.txt>
- [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)
- [Kinect 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
- [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 objekt-suche im Living Place Hamburg*, Hochschule für Angewandte Wissenschaften hamburg, Diplomarbeit, Februar 2011
- [Najem Feb 2011] NAJEM, Hosnia: Aufbau der Infrastruktur für das Model-basierte objekt-suche im Living Place Hamburg. In: *MA-INF1 Projektbericht PO1* (Feb 2011)
- [Otto und Voskuhl 2011] OTTO, Kjell ; VOSKUHL, Sören: Weiterentwicklung der Architektur des Living Place Hamburg. In: *Berichte Masterprojekt 2 Master Informatik HAW Hamburg Winter 2010 / 11* (2011)