



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

## Ausarbeitung - WiSe 2009/2010

Arne Bernin

Erkennung minimaler dynamischer Gesten im  
Living Place Hamburg

## Inhaltsverzeichnis

<b>1 Einführung</b>	<b>3</b>
1.1 Einleitung . . . . .	3
1.2 Vision . . . . .	4
1.3 Vorarbeiten . . . . .	4
<b>2 Gestenerkennung</b>	<b>5</b>
2.1 Beispiele für Gesten . . . . .	5
2.2 Ablauf der Gestenerkennung . . . . .	5
<b>3 Technik</b>	<b>7</b>
3.1 Projekt Natal . . . . .	7
3.2 Prototyp . . . . .	7
3.3 Optische Kamera . . . . .	8
3.4 TOF-Kamera . . . . .	8
<b>4 Masterarbeit</b>	<b>9</b>
<b>5 Risiken</b>	<b>9</b>
5.1 Allgemeine Risiken . . . . .	9
5.2 Technische Risiken . . . . .	10
5.3 Clients . . . . .	11
5.4 Usability . . . . .	11
5.5 Akzeptanzprobleme . . . . .	12
<b>6 Fazit</b>	<b>13</b>
<b>Abbildungsverzeichnis</b>	<b>14</b>
<b>Literatur</b>	<b>14</b>

# 1 Einführung

## 1.1 Einleitung

Die Art der Bedienung von Maschinen unterliegt seit Anbeginn ihrer verbreiteten Nutzung einem Wandel. Wurden mit Beginn der Industrialisierung vor allem mechanische Maschinen mit mechanischen Reglern bedient, so wandelte sich dies mit dem Aufkommen elektrischer Apparaturen zu Schaltern und Knöpfen, im Bereich der Unterhaltungsgeräte dann schließlich zu tragbaren Schaltern und Knöpfen, den Fernbedienungen.

Auch bei den ersten Computern dienten Schalter und Knöpfe als direkte Schnittstelle zum Menschen, in späteren Versionen dann in der gleichen Anordnung wie bei einer Schreibmaschine, der Tastatur.

Im Zuge der weiteren Computerisierung der Unterhaltungselektronik verschwimmt die Grenze zwischen Computer und Fernseher, Musikanlage und Spielekonsole. Das gilt auch für die Bedienung dieser Gerätschaften. Neue Lösungen aus dem Bereich der Spiele, wie etwa die Steuerung durch Bewegung und Gesten, dringen auch in andere Bereiche als neue Möglichkeit der Mensch-Maschine-Interaktion (HCI) und Kommunikation.

Dies mag in naher Zukunft dazu führen, dass die leidige Suche nach der Fernbedienung beziehungsweise der richtigen Fernbedienung ( mit möglichst noch vollen Batterien ) der Vergangenheit angehört. Doch wie realisiert man ein solches System in einem realen oder realitätsnahen Kontext? Die geplante Masterarbeit des Autors soll darüber Aufschluss geben.

Die Entwicklung und Evaluierung des System findet im Kontext des „Living Place Hamburg“ (siehe auch [Gregor u. a. (2009)]) an der HAW Hamburg statt.



Abbildung 1: Quelle [tomshardware (2010)]

## 1.2 Vision

Ziel des zu realisierenden Systems ist die Erkennung einfacher (minimaler) dynamischer Gesten zur Steuerung von Geräten aus dem Multimediabereich über den iROS-Eventheap ([[Ponnekanti u. a. \(2003\)](#)]) im Living Place. Dazu soll weder Hardware noch passive Marker am Körper verwendet werden. Idealerweise ist ein solches System unabhängig von der verwendeten Beleuchtung, funktioniert also sowohl bei Tage als auch bei Nacht und abgedunkeltem Raum.

Realisiert werden soll dies möglichst durch günstige und derzeit verfügbare Hardware. Ist dies aufgrund mangelnder Verfügbarkeit nicht möglich, soll andere Hardware prototypisch zum Einsatz kommen.

Das Gesamtsystem soll geeignet sein als geschlossenes System auf einem Embedded-Computer zu funktionieren. Insbesondere für den Bereich Datenschutz ist dies wünschenswert, da so sichergestellt werden kann, dass nur akkumulierte Informationen weitergeben werden und nicht die Rohdaten (wie Videodaten). Eine weitere Nutzung der erstellten Software durch andere Projekte ist durch die Freigabe unter einer Opensource-Lizenz ([[oss \(2010\)](#)]) angedacht.

## 1.3 Vorarbeiten

Im Rahmen des Projektes im Sommersemester 2009 [[abprojekt \(2010\)](#)] wurden erste Erfahrungen mit der Bewegungserkennung mit Differenzbildern gesammelt. Dabei stellte sich heraus, dass Differenzbilder grundsätzlich ein geeignetes Verfahren darstellen, um Bewegungen einfach und effektiv zu erkennen. Allerdings hat dieses Verfahren seine Grenzen, insbesondere in Umgebungen mit vielen Reflektionen (Glasscheiben, Bildschirme) und Schattenwürfen werden falsch-positive Bewegungen erkannt.



Abbildung 2: Screenshot der Differenzbilder

## 2 Gestenerkennung

### 2.1 Beispiele für Gesten

Die folgende Tabelle enthält einige Beispiele für mögliche Gesten. Diese beziehen sich auf die Hand als beweglicher „Marker“, die Ausrichtung der Finger spielt dabei vorerst keine Rolle. Die Hand als Ganzes beziehungsweise ihr Mittelpunkt ist der Bezugspunkt für die Geste.

Bewegung	ausgelöste Aktion
Hand hoch bzw. runter	Verändern der Lautstärke
Hand links bzw. rechts	Wechseln des Fernseh-Kanals, Musiktitels
Winken	Auslösen und Beenden der Kommandomodus
Hand im Kreis führen	Auswahl aus mehreren Möglichkeiten (Auswahl-Rad)
Bewegen der Hand nach vorne	Bestätigen einer Auswahl

Tabelle 1: Beispielgesten zur Steuerung eines Gerätes

Wie der Tabelle zu entnehmen ist spielt auch die Position der Hand im Raum (vor und zurück) eine Rolle, somit sind alle 3 Achsen der möglichen Bewegung im Raum zu erfassen (3 Freiheitsgrade/Dimensions of Freedom).

### 2.2 Ablauf der Gestenerkennung

Die geplante Gestenerkennung beschränkt sich auf die Erkennung von Gesten der ganzen Hand beziehungsweise des damit verbundenen Armes. Gesten in diesem Sinne sind Bewegungen der Hand in eine definierte Richtung, mit einer Mindestbewegung und einer maximalen Dauer. Die Eingabe wird durch eine (Kommando-) Geste eingeleitet, dann erfolgt die eigentliche Funktionseingabe (Lautstärke, (Kanal) weiterschalten, etc). Und abschließend eine Bestätigung durch die erneute Verwendung der Kommando-Geste. Dazu ist es hilfreich, wenn über ein Display eine Rückmeldung an den Benutzer über den eingeleiteten Kommandomodus und die erkannte Geste erfolgt, so dass dieser diese Bestätigen oder wiederholen kann.

Dafür ist es notwendig die anfallenden Bilddaten auf die Daten zu reduzieren, die für eine Erkennung von Gesten notwendig ist (Vorverarbeitung). Entscheidend für die Gestenerkennung ist die Position der Hand des Benutzers, denn nur diese wird für die Eingabe verwendet.

Elementar hierfür ist die Identifikation der Hand. Durch die Kombination der Verfahren Segmentierung durch Differenzbilder, Hautfarberkennung und die Position im Raum sollte sich eine zuverlässige Erkennung von Händen erreichen lassen. Jedes dieser Verfahren hat seine

Stärken und Schwächen. Durch die Kombination von drei unterschiedlichen Verfahren sollen diese einzelnen Schwächen ausgeglichen und die Zuverlässigkeit des Gesamtsystems erhöht werden.

Da die Verfahren voneinander unabhängig sind (teilweise sogar über unterschiedliche Hardware erfolgen) kann die Verarbeitung parallelisiert werden.

Die Bewegungen der Hand müssen vom System ständig beobachtet werden, damit eine Kommandogeste zuverlässig erkannt wird.

Die Klassifikation der einzelnen Gesten erfolgt dann anhand eines Feature Vectors ([Cristianini u. Shawe-Taylor (2000)]). Der genaue Aufbau dieses Vectors wird noch Gegenstand praktischer Untersuchungen mit dem ersten Prototyp sein, als generelle Grundlage bietet sich eine Arbeit zur Gestenerkennung an der HAW ([Heitsch (2008)]) an, die zur Klassifizierung eine Support Vector Machine ([Schölkopf u. Alexander (2002)]) verwendet. Der Unterschied ist allerdings, dass in diesem Falle die dritte Dimension berücksichtigt werden soll.

Ziel der Arbeit wird es sein, die Liste der Gesten dynamisch zu halten, also auch das einfache Hinzufügen von Gesten ohne weiteren Programmieraufwand zu ermöglichen. Dazu ist eine dynamische Klassifizierung erforderlich. Der ausgewählte Algorithmus muss dies also unterstützen.

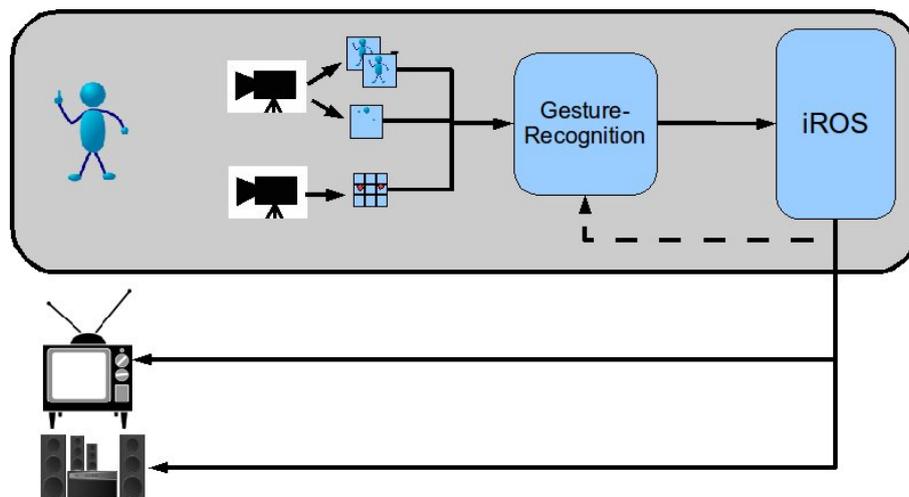


Abbildung 3: Übersicht über das System

## 3 Technik

Aus den Erfahrungen des Masterprojektes ergibt sich die Notwendigkeit, zusätzlich zu den Informationen einer Bewegungserkennung mit einer Kamera weitere Kanäle zu nutzen (siehe auch Abbildung 3) . Neben den klassischen optischen - aber leider rechenaufwendigen Verfahren - wie der Triangulation (3D Kamera, siehe auch [Yamamoto u. a. (2004)]) ,bietet es sich zugunsten der Performance an auf neuere Entwicklungen zurückzugreifen. Eine dieser Entwicklungen sind die sogenannten Time-of-Flight-Kameras ([Lange u. Seitz (2001)], [Yahav u. a. (2007)]), bei denen eine Laufzeitmessungen des verwendeten Lichtes (meist 850 nm, Infrarot) durchgeführt wird.

Da diese Kameras nur über eine recht geringe Auflösung von maximal 200x200 Pixel verfügen sind sie auf größere Distanz hin nicht geeignet als alleinige Kamera verwendet zu werden. Die Verbindung mit einer optischen Kamera ist sinnvoll, da diese die Möglichkeit bietet, die geplante weiteren Verfahren (Differenzbilder und Hauterkennung) durchzuführen.

### 3.1 Projekt Natal

Das „Projekt Natal“ von Microsoft (siehe [natal (2009)]) scheint dafür die ideale Hardware zu sein: günstig, performant, für Spiele entwickelt und somit schnell genug für eine Gestenerkennung. Leider sind diese Kameras noch nicht verfügbar und scheiden als nutzbare Variante aus, der „Nachbau“ mit verfügbaren Einzelteilen ist jedoch machbar.



Abbildung 4: Projekt Natal , Quelle [natalpic (2009)]

### 3.2 Prototyp

Die Hardware des Projektes Natal besteht aus einer TOF- und einer optischen Kamera, daher sollte auch ein Prototyp aus beiden Komponenten bestehen. Zur besseren Einbettung in die Infrastruktur des Living Place empfiehlt sich die Nutzung von Geräten mit Ethernet-Schnittstelle, die vorhandenen Kabel können so leicht genutzt werden. Geplant ist der Aufbau eines transportablen Prototyps aus beiden Kameras in einem Gehäuse.

### 3.3 Optische Kamera

Als geeignete Kamera im Normal-Optischen Bereich bietet sich die IP Cam 9070 CS von Aviosys an, diese verfügt über eine Auflösung im Megapixel Bereich (1280x720).



Abbildung 5: Aviosys IP Cam 9070cs, Quelle [[9070cs \(2010\)](#)]

### 3.4 TOF-Kamera

Time-of-flight Kameras stellen, mit Hilfe von Laufzeitmessungen des zur Beleuchtung verwendeten Lichtes, Tiefeninformationen zur Verfügung. Die vorgesehene Kamera SR4000 von Mesa Imaging tut dies über einen 16-bit Farbwert für jeden Pixel. Diese Farbe zeigt die Entfernung des Objektes an. Die Kamera hat eine Auflösung von etwa 1/4 Grad pro Pixel bei einem Gesamtwinkel von ca. 45 Grad in horizontaler und 40 Grad in vertikaler Richtung. Dadurch ergibt sich auf eine angenommene Entfernung von etwa 5m eine Breite des Sichtbereichs von ca. 3,5m.

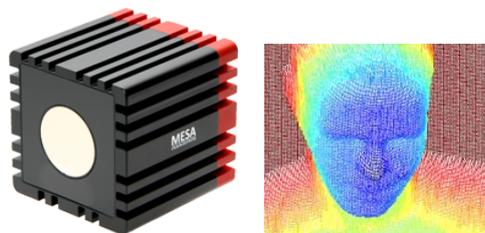


Abbildung 6: Swissranger SR4000, SR4000 Gesicht Quelle [[sr4000 \(2010\)](#), [mesaface \(2010\)](#)]

## 4 Masterarbeit

Für das weitere Vorgehen auf dem Weg zur Masterarbeit sind die folgenden Schritte geplant:

- Beschaffen der nötigen Technik
- Evaluierung und Test der Technik
- Aufbau eines ersten Prototyps zum Erstellen einer Datenbank mit Rohdaten, gewonnen aus Versuchen mit verschiedenen Probanden
- Analysieren der Daten und Erstellen des Feature Vectors / Extractors
- Entwickeln der nötigen Software zur Umsetzung der Gestenerkennung
- Weiterentwickeln des Prototyps zur kompletten Funktionalität
- Client für den iROS-Eventheap zur Anzeige der erkannten Gesten
- Erneute Testphase mit Probanden
- Optimierung/Feintuning
- Anbinden von realen Clients an den iROS-Eventheap
- Abschlusstests: Usability und Grenzen des Systems

## 5 Risiken

### 5.1 Allgemeine Risiken

Bei Arbeiten die die Entwicklung eines Prototyps enthalten sind einige Risiken immer in Betracht zu ziehen. Zum einen die Gefahr, dass das System nicht oder nur in Teilbereichen funktioniert, da beispielsweise die Komplexität der Materie unvorhergesehene Folgen verursacht. Eine gründliche Recherche und Zeit in der Designphase helfen dabei, dieses Risiko zu minimieren.

Zusätzlich dazu ist bei der geplanten Masterarbeit mit weiteren Risiken zu rechnen. In diesem Fall zusammengefasst in den folgenden Abschnitten.

## 5.2 Technische Risiken

### Einschränkungen durch die Umgebung

Es besteht die Möglichkeit, dass die verwendeten Kameras und Verfahren nicht oder nur bedingt für das konkrete Umfeld, also eine reale Wohnung mit Fenstern, Möbeln und anderen Gegenständen geeignet sind. Obwohl dies im Vorwege mit dem Hersteller ([mes (2009)]) der TOF-Kamera geklärt wurde, kann ein abschließendes Urteil erst im realen Betrieb gefällt werden. Hierbei besonders problematisch können einerseits Reflektionen und ein Zuviel an Tageslicht beziehungsweise direktem Sonnenlicht sein. Eine mögliche Lösung wäre das Einschränken auf ein bestimmtes „Setting“ wie das Verdunkeln durch Jalousien. Auf anderen der Seite ist eine zu schwache Beleuchtung problematisch. Die Erkennung der Hautfarbe bedingt die Verwendung von Farbbildern, Infrarotbilder liefern jedoch nur Graustufen. Sollte die Beleuchtung durch sichtbares Licht also nicht ausreichen, muss ein alternatives Verfahren anstatt der Hauterkennung verwendet werden. Eine Möglichkeit wäre die Verwendung eines Verfahrens zur Erkennung der Hand im Bild anhand ihrer Silhouette, dieses ist jedoch aufwendig.

### Gegenseitige Störung

Die Gestenerkennung soll auch bei schlechter beziehungsweise „dezentere“ Beleuchtung funktionieren. Sollte das sichtbare Licht dazu nicht ausreichen, muss auch mit der optischen Kamera im Infrarotbereich gearbeitet werden. Eine dafür notwendige Infrarot-Beleuchtung kann zu Interferenzen mit der Tiefenerkennung (TOF-Kamera) führen. Dies kann auch bei weiteren Geräten wie Infrarot basierten Fernsteuerungen auftreten.

### Datenübertragung

Das geplante System setzt auf den Transfer von Daten über Ethernet. Es ist, trotz des Einsatzes von entsprechender Technik (Switch), immer möglich, dass es zu Verzögerungen durch konkurrierenden Datentransfer kommt. Ebenso muss die Gesamtauslastung und -Beschränkung (maximale Datenrate) beachtet werden. Gegebenenfalls müssen auf dieser Ebene Maßnahmen, wie die Priorisierung des Datentransfers (QoS, [802.1q (2003)]) eingesetzt werden. Im Gegensatz zur Übertragung über Funk (WLAN) ist allerdings bei kabelgebundenen Systemen mit einer deutlich sichereren Übertragung zu rechnen.

### 5.3 Clients

Die externe Steuerung von echten Multimediageräten wie Fernseher oder Musikanlage ist nicht genormt. Die meisten Geräte verfügen nur über eine Infrarotschnittstelle zur Fernbedienung. Der Einsatz einer Universalsteuerung mit Ethernet-Schnittstelle (wie der IrTrans, [irtrans (2010)]) sollte daher erwogen werden.



Abbildung 7: LAN-IR-Konverter IrTrans, Quelle [irtrans (2010)]

### 5.4 Usability

In diesen Bereich entfallen Risiken, die zu einer eingeschränkten Benutzbarkeit des Systems führen. Im schlimmsten Fall zur Nichtbenutzbarkeit des Gesamtsystems.

#### Reaktionszeit

Im Bereich Usability von Benutzerschnittstellen gibt es verschiedene Klassen für die Reaktionszeit eines Systems ([Nielsen (1994)]). Dabei wird zwischen Systemen mit einer Reaktionszeit von weniger als 100 Millisekunden und 100 bis 1000 Millisekunden (sowie weiteren Einteilungen jenseits der 1 Sekunde) unterschieden. Im ersteren Fall gilt dies als Echtzeit, für den Benutzer erfolgt die Reaktion sofort. Im Bereich bis zu einer Sekunde ist eine Zuordnung Aktion zu Reaktion noch deutlich gegeben, darüber hinaus für den Benutzer nicht mehr klar ersichtlich. Eine Reaktionszeit größer als eine Sekunde ist somit nicht mehr tolerabel.

Eine Abschätzung der zu erwartenden Reaktionszeit (siehe Tabelle 2) führt zur Erkenntnis, dass ein Echtzeitsystem in dieser Konstellation nicht realisierbar ist, wohl aber der Bereich unterhalb einer Sekunde. Es gilt hierbei: Je schneller, je besser. Um unnötige Verzögerungen im Bereich der Bildverarbeitung zu vermeiden, sollte von vornherein eine Parallelisierung und die Verwendung entsprechender Hardware vorgesehen werden.

<sup>1</sup> Herstellerangabe, gilt sowohl für 9070CS als auch SR4000

<sup>2</sup> geschätzt auf Basis der Messung (21sec) mit diffimage[abprojekt (2010)]

<sup>3</sup> geschätzt

Aufgabe	Zeitliche Abschätzung
Erfassung Kamera Verarbeitung Kamera Übertragung ethernet + IP-Stack	kleiner 150ms <sup>1</sup>
Verarbeitung Bildverarbeitung	ca. 30ms <sup>2</sup>
Verarbeitung Gestenerkennung	ca. 30ms <sup>3</sup>
Nachricht an Eventheap	ca. 5 ms <sup>3</sup>
Übertragung ethernet + IP-Stack	<= 1ms
Verarbeitung Eventheap	<= 10ms
Übertragung ethernet + IP-Stack	<= 1ms
Reaktion des Zielsystems	100ms <sup>3</sup>
<b>Gesamt</b>	<b>max. 327ms</b>

Tabelle 2: Abschätzung der Reaktionszeit

### Fehlinterpretation

Eine Fehlinterpretation der Eingabe des Benutzers und damit verbunden eine falsche Reaktion führt, genauso wie eine ausbleibende Reaktion, zur Frustration des Benutzers und sollte möglichst vermieden werden. In diesen Bereich fällt auch die Interpretation von (legitimen) Gesten anderer Benutzer. Mit Hilfe einer räumlichen Zuordnung gilt der beabsichtigte Kommandomodus nur für eine Person mit einem (nach Auslösen des Kommandomodus) konstanten Aufenthaltsort.

### 5.5 Akzeptanzprobleme

Kameras in sensiblen Bereichen der privaten Lebensgestaltung (Privatsphäre) können zu Akzeptanzproblemen führen. Vor allem wenn nicht klar ersichtlich ist, wo diese Bilder landen. Im vorliegenden Fall handelt es sich allerdings um einen Prototypen. Dieser erhebt nicht den Anspruch alle Teilaspekte eines Systems für Endbenutzer voll abzubilden. Deshalb sollte diese Fragestellung zwar berücksichtigt werden, hat aber kaum praktische Relevanz.

## 6 Fazit

Der geplante Prototyp bietet die Gelegenheit schon jetzt einen Ausblick auf die kommenden Veränderungen der Mensch-Maschine-Schnittstelle im Consumerbereich zu werfen.

Es bietet sich die Chance, ein funktionierendes, robustes System zur Erkennung minimaler, dynamischer Gesten im Livingplace zu schaffen. Im günstigsten Falle wirklich unabhängig von Lichtverhältnissen und somit Tages- beziehungsweise Nachtzeit. Darauf aufbauend können weitere Arbeiten zur Optimierung der einzelnen Teilsysteme durchgeführt werden, beispielsweise der Berücksichtigung der Finger oder der Ausrichtung der Hand.

Entscheidend für ein weiteres Vorgehen - vor allem für die Entwicklung der Software - ist natürlich, ob und wie zuverlässig die geplante Hardware nutzbare Daten liefert und ob die vorgesehenen Verfahren die nötige Zuverlässigkeit in der realen Umgebung wirklich erreichen. Dies wird sich endgültig erst in praktischen Untersuchungen herausstellen.

## Abbildungsverzeichnis

1	Quelle [tomshardware (2010)] . . . . .	3
2	Screenshot der Differenzbilder . . . . .	4
3	Übersicht über das System . . . . .	6
4	Projekt Natal , Quelle [natalpic (2009)] . . . . .	7
5	Aviosys IP Cam 9070cs, Quelle [9070cs (2010)] . . . . .	8
6	Swissranger SR4000, SR4000 Gesicht Quelle [sr4000 (2010), mesaface (2010)]	8
7	LAN-IR-Konverter IrTrans, Quelle [irtrans (2010)] . . . . .	11

## Literatur

- [mes 2009] *Mesa Imaging SR4000*. Verifiziert am 16.2.2010. <http://www.mesa-imaging.ch>. Version:2009
- [802.1q 2003] *IEEE Standards for Local and metropolitan area networks, Virtual Bridged Local Area Networks*. Verifiziert am 26.2.2010. <http://standards.ieee.org/getieee802/download/802.1Q-2003.pdf>. Version:2003
- [9070cs 2010] *Aviosys IP Cam 9070cs*. Verifiziert am 18.2.2010. <http://www.aviosys.com>. Version:2010
- [abprojekt 2010] *Räumliche Segmentierung mit Differenzbildern*. Verifiziert am 18.2.2010. <http://users.informatik.haw-hamburg.de/~ubicomp/projekte/master2009-proj/bernin.pdf>. Version:2010
- [Cristianini u. Shawe-Taylor 2000] CRISTIANINI, Nello ; SHAWE-TAYLOR, John: *An introduction to support vector machines : and other kernel-based learning methods*. 1. Cambridge University Press, 2000 <http://www.cs.orst.edu/~bulatov/papers/CambridgeUniversityPress-Support.Vector.Machines.and.Oth.chm>. – ISBN 0521780195
- [Gregor u. a. 2009] GREGOR, S. ; RAHIMI, M. ; VOGT, M. ; SCHULZ, T. ; LUCK, K. v.: *Tangible Computing revisited: Anfassbare Computer in Intelligenten Umgebungen*. Verifiziert am 12.2.2010. <http://users.informatik.haw-hamburg.de/~ubicomp/arbeiten/papers/MMWismar2009.pdf>. Version:2009
- [Heitsch 2008] HEITSCH, Sebastian: *Ein Framework zur Erkennung von dreidimensionalen Gesten*. Verifiziert am 26.2.2010. <http://opus.haw-hamburg.de/volltexte/2008/650/>. Version:2008
- [irtrans 2010] *IRTrans*. Verifiziert am 26.2.2010. <http://www.connect-b.com/irtrans.htm>. Version:2010

- [Lange u. Seitz 2001] LANGE, R. ; SEITZ, P.: Solid-state time-of-flight range camera. In: *Quantum Electronics, IEEE Journal of*. Washington, DC, USA : IEEE Computer Society, March 2001. – ISSN 0018–9197, S. 390–397
- [mesaface 2010] *mesa imaging*. Verifiziert am 18.2.2010. <http://www.j-clavis.co.jp/Mesa-sr/APP.html>. Version: 2010
- [natal 2009] *Infos zum Projekt Natal*. Verifiziert am 26.2.2010. <http://www.xbox.com/de-DE/news-features/news/Project-Natal-in-detail-050609.htm>. Version: 2009
- [natalpic 2009] <http://xbox.pcmasters.de/attachments/xbox-360-hardware-zubehoer/98d1245589635-xbox-360-project-natal-natal1.jpg>
- [Nielsen 1994] *Kapitel 5*. In: NIELSEN, Jakob: *Usability Engineering*. Morgan Kaufman, 1994. – ISBN 0–12–518406–9
- [oss 2010] *Usability Engineering*. Verifiziert am 26.2.2010. [OpenSourceInitiative](http://opensourceinitiative.com). Version: 2010
- [Ponnekanti u. a. 2003] PONNEKANTI, Shankar R. ; JOHANSON, Brad ; KICIMAN, Emre ; FOX, Armando: Portability, Extensibility and Robustness in iROS. In: *PERCOM '03: Proceedings of the First IEEE International Conference on Pervasive Computing and Communications*. Washington, DC, USA : IEEE Computer Society, 2003. – ISBN 0–7695–1893–1, S. 11
- [Schölkopf u. Alexander 2002] SCHÖLKOPF, Bernhard ; ALEXANDER, J. S.: *Learning with Kernels. Supported Vector Machines, Regularization, Optimization, and beyond*. MIT Press, 2002. – ISBN 0–262–19475–9
- [sr4000 2010] *SR4000 ToF-Kamera*. Verifiziert am 18.2.2010. <http://www.mesa-imaging.ch>. Version: 2010
- [tomshardware 2010] *Universal Fernbedienungen*. Verifiziert am 18.2.2010. <http://www.tomshardware.com/>. Version: 2010
- [Yahav u. a. 2007] YAHAV, G. ; IDDAN, G.J. ; MANDELBAUM, D.: 3D imaging Camera for Gaming Application. In: *Consumer Electronics, 2007. ICCE 2007. Digest of Technical Papers. International Conference on (2007)*
- [Yamamoto u. a. 2004] YAMAMOTO, Y. ; YODA, I. ; SAKAUE, K.: Arm-pointing gesture interface using surrounded stereo cameras system, 2004. – ISSN 1051–4651, S. 965–970 Vol.4