

Philipps-Universität Marburg

Fachbereich 12 - Mathematik und Informatik



Masterarbeit

Analyse von Videodaten aus SmartHome Systeme zur Erkennung von außergewöhnlichen Situationen

von
Anh Duc Nguyen
April 2018

Betreuer:
Prof. Dr. Thorsten Thormählen
Dr. Tobias Kuhl
Dr. Marcel Schneider

Arbeitsgruppe Grafik und Multimedia Programmierung

Erklärung

Ich, Anh Duc Nguyen (Informatikstudent an der Philipps-Universität Marburg, Matrikelnummer xxxxxx), versichere an Eides statt, dass ich die vorliegende Masterarbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe. Die hier vorliegende Masterarbeit wurde weder in ihrer jetzigen noch in einer ähnlichen Form einer Prüfungskommission vorgelegt.

Marburg, 01. April 2018

Anh Duc Nguyen

Zusammenfassung

Bei meiner Masterarbeit geht es um eine real-time Anwendung, wo man zur Erkennung von außergewöhnlichen Situationen für allein wohnende Alten verwenden kann. Die Eingaben des Programms sind Frames von einer Kamera, und im nächsten Schritt werden Informationen erzeugt, die uns sagen ob da eine Bewegung gibt. Aus die Bewegungsmeldungen wird ein Vordergrund von dem Hintergrund mit Hilfe von Background Subtraktion und Update extrahieren. Nach dem Extrahieren des Vordergrundes wird eine Silhouette von Schwarzweiß erstellt. Mit Hilfe von einer Histogrammanalyse auf Silhouette wird ein Körperhaltung vorhersagen. Das Programm kann sagen ob das bewegende Objekt steht, sitzt, beugt oder liegt. Anhand von Positionen wo das Objekt oft bewegt und wann das Objekt steht, sitzt oder liegt... kann außergewöhnlichen Situationen erkannt werden.

Abstract

... (exakte englische Übersetzung der deutschen Kurzfassung)

Inhaltsverzeichnis

Inhaltsverzeichnis	I
1. Einleitung	1
1.1. Motivation	1
1.2. Ziele	1
1.3. Aufbau der Arbeit	2
1.4. Verwandte Arbeiten	2
1.4.1. Background Subtraktion	2
2. Grundlagen	5
2.1. Background Subtraktion	5
2.1.1. subsection	5
2.2. Histogrammanalyse	5
2.3. OpenCV Framework	5
3. Eigenes Verfahren	7
3.1. LaTeX-Editoren	7
3.2. Beispiel für eine Tabelle	7
3.3. Formel in Latex und Konventionen zur Verwendung mathematischer Symbole	7
3.4. Beispiel für eine Vektorgrafik	8
3.5. Beispiel für eine Vektorgrafik mit mathematischen Symbolen	9
3.6. Beispiel für eine Rastergrafik	9
3.7. Beispiel für die Darstellung von Algorithmen	10
3.8. Beispiel für die Darstellung von Quellcode-Auszügen	10
3.9. Doppelseitiger Druck	10
3.10. Zitieren und Quellenangaben	10
4. Ergebnisse und Evaluation	13
5. Zusammenfassung und Ausblick	15
Literaturverzeichnis	16
Abkürzungsverzeichnis	19
Abbildungsverzeichnis	21

Tabellenverzeichnis	22
Liste der Algorithmen	24
Listings	26
A. Anhang	28
Danksagung	29

1. Einleitung

Diese Masterarbeit beschäftigt sich mit den Background Subtraktion um eine Bewegung zu erkennen, Segmentieren eine ROI um eine Silhouette zu verbessern, Histogrammanalyse um Körperhaltung vorherzusagen, Wahrscheinlichkeitsberechnungen um die abnormale Situationen zu schätzen. OpenCV ist hier angewendet, das Framework vereinfacht den Weg zur Implementierung sowie erhöht die Geschwindigkeit des Programms.

1.1. Motivation

In Deutschland werden die Menschen immer älter und die Meisten wohnen oft allein in einer Wohnung oder Haus. Manche Familien haben nur ein Kind und nicht selten beide Elternteile arbeiten gehen. Außerdem verlangt die Arbeitswelt immer mehr Flexibilität, deshalb wohnen die Kinder oder Verwandten von Senioren oft in anderen Städten und sind nur schwer zu erreichen. Im Besonderen je älter die Senioren sind, desto größere Angst haben die vor Altenheimen. Deutschland stellt vor neue Herausforderung. Heutzutage leben immer mehr und mehr Senioren allein. Beispielsweise wohnt jede zweite Seniorin allein und 20 Prozent sind über 65 Jahre alt. Wenn die Alte einsam sind, erhöht auch das Risiko, dass die unter Depressionen, Herzerkrankung oder Unfall leiden. Das ist aber kein Ende des Tunnels, mit der Entwicklung von Technologien kann das Problem mit Assisted Ambient Living ALL schon gelöscht werden. Derzeit springt der IT-Bereich besonders im Computer Vision einen großartigen Schritt. Durch Anwendung einer Kamera können die Verwandten oder Kinder von Senioren einfach im real-time sehen. Aber nicht jeder hat Zeit den ganzen Tag vor Bildschirm beobachten, deshalb das Programm hier ist eine Hilfe zur Erkennung der Situationen, wenn irgendwas mit allein wohnende Alte passiert.

1.2. Ziele

Das Ziel dieser Arbeit ist die Entwicklung eines Programms, das mit Hilfe von Kamera Bewegungen erkennt, und dadurch Vordergrund von Hintergrund extrahieren analysiert. Am Ende soll das Programm die Situationen abschätzen, ob alles noch gut ist oder irgendwas schief läuft.

Echtzeit

Da es um eine Unfallmeldung geht, deshalb muss die Anwendung schon in Echtzeit laufen. Die durchschnittliche Laufzeit des Programms soll höchstens 3 Frames pro Sekunde benötigen, um die Situation in Echtzeit zu erkennen.

Kosten

Das Programm sollte mit der normalen, kostengünstigen Privatanwender-Hardware laufen, sodass für die Ausführung keine Spezialrechner benötigt werden. Die einzige Voraussetzung an das System ist eine Kamera. Diese ist aber in den meisten Desktoprechnern vorhanden oder kann kostengünstig nachgerüstet werden.

1.3. Aufbau der Arbeit

Im ersten Abschnitt des Kapitels 2 soll zunächst die angewendete Hintergrund Subtraktionen verdeutlicht werden. Dann soll im zweiten Abschnitt die Histogrammanalyse beschrieben werden, bevor im Abschnitt 2.3 ein Überblick über das OpenCV Framework geben wird. An dieser Stelle werden auch die Vorteile der Anwendung von OpenCV bei dieser Arbeit genauer erläutern. Kapitel 3 beschäftigt sich mit der Theorie des im Rahmen dieser Masterarbeit entwickelten Algorithmus. Dabei sollen die einzelnen Schritte des Algorithmus genauer von erläutert werden. Die Implementierung des Algorithmus wird im diesen Kapitel besprochen, dabei sollen die benutzten Bibliotheken sowie die Umsetzung des Programms beschrieben werden. Anschließend werden im Kapitel 4 die durchgeführten Tests und deren Ergebnisse dokumentiert und diskutiert, sowie im Kapitel 5 ein Ausblick auf weiterführende Arbeiten gegeben.

1.4. Verwandte Arbeiten

1.4.1. Background Subtraktion

Bewegungserkennung spielt heutzutage eine wichtige Rolle in Überwachungssystem. Durch die Subtraktion von Hintergrund wird die Bewegung von dem Hintergrund extrahiert. Es gibt verschiedene wissenschaftliche Arbeiten für diese Methode. Ein der Verfahren ist “Adaptive background mixture models” [SG99], das ist eine gängige Methode zur Echtzeit-Segmentierung bewegter Regionen in Bildsequenz. Die Methode beinhaltet eine “Hintergrundsubtraktion” oder eine Schwellenwertbindung des Fehlers zwischen einer Schätzung des Bild ohne sich bewegende Objekte und dem aktuellen Bild. In [SG99] wird das Modellieren jedes Pixels als eine Mischung von Gauß-Werten und das Verwenden einer Online-Approximation zum Aktualisieren des Modells erläutert. Die Gaußschen Verteilungen des adaptiven Mischungsmodells werden dann ausgewertet, um zu bestimmen, welche am wahrscheinlichsten aus einem Hintergrundprozess resultieren.

Eine verbesserte Version von Gauß’schen Mischungsmodells ist in [ZVDH06] als “KNN” Methode beschrieben. Rekursive Gleichungen werden präsentiert, die dazu dienen, die Parameter eines Gauß’schen Mischungsmodells ständig zu aktualisieren und gleichzeitig die geeignete Anzahl von Komponenten für jedes Pixel auszuwählen.

In [EHD00] wird eine Methode, die “kernel density estimation” (KDE) heißt, besprochen. KDE ist ein neuartiges nicht-parametrisches Hintergrundsubtraktion. Das Modell kann mit Situationen umgehen, in denen der Hintergrund der Szene überladen und nicht vollständig statisch ist, sondern kleine Bewegungen wie Äste und Büsche enthält.

Das Modell schätzt die Wahrscheinlichkeit der Intensitäten von Pixel auf der Grundlage einer Stichprobe von Intensitäten für jedes Pixel. Das Modell passt sich schnell an Veränderungen in der Szene an, was eine sehr empfindliche Erkennung bewegter Ziele ermöglicht.

In [BVD09] und [BVD11] wird “Vibe”, eine leistungsfähige Methode zur Hintergrundextraktion vorgestellt. “Vibe” verbessert die Genauigkeit und reduziert die Rechenlast. Die Hauptinnovation betrifft die Verwendung einer zufälligen Strategie zur Auswahl von Werten, um eine stichprobenartige Schätzung des Hintergrunds zu erstellen. Es war das erste Mal, dass eine zufällige Wahrscheinlichkeit im Bereich der Hintergrundextraktion verwendet wird.

Für diese Arbeit werden alle oben genannten Methoden verglichen und die Vergleichen aller oben genannten Methoden werden im nächsten Kapitel genauer beschrieben.

Körperhaltungserkennung basiert auf Histogrammanalyse

Histogrammanalyse ist eine einfache und schnelle Methode zur Erkennung der Körperhaltung. Es gibt viele wissenschaftliche Arbeiten, die über Histogrammanalyse studiert und präsentiert werden.

In [GM06] wird das Projektionshistogramm verwendet, um unterschiedliche Körperhaltungen zu unterscheiden. Ein neues Verfahren wird vorgeschlagen, das ein Projektionshistogramm zur statischen Erkennung der menschlichen Haltung verwendet. Diese Arbeit in [GM06] besteht aus drei Schlüsselmodulen: Hintergrundsubtraktion, Projektionshistogrammberechnung und Vorlagenabgleich.

Eine andere Methode ist in [HHD98] beschrieben, die als “Ghost” System benannt. “Ghost” ist ein Echtzeitsystem zum Schätzen der menschlichen Körperhaltung und zum Detektieren von Körperteilen. Es erstellt ein silhouetten-basiertes Körpermodell, um die Position der Körperteile zu bestimmen, während sich Personen in generischen Positionen befinden. Es kombiniert eine hierarchische Schätzung der Körperposen mit Histogrammanalyse, eine konvexe Rumpfanalyse der Silhouette und eine teilweise Abbildung von den Körperteilen zu den Silhouetten-Segmenten. Die Histogrammanalyse in dieser Masterarbeit basiert auf die Methode, die in “Ghost” beschrieben wird.

Fußgänger Erkennung

2. Grundlagen

text

2.1. Background Subtraktion

2.1.1. subsection

text

2.2. Histogrammanalyse

Wie schon in Kapitel 2.1 beschrieben ...

2.3. OpenCV Framework

text

3. Eigenes Verfahren

In diesem Kapitel soll das eigene Verfahren beschrieben werden. Es geht dabei nicht nur darum zu beschreiben was gemacht wurde, sondern ebenfalls darum zu begründen, weshalb bestimmte Design-Entscheidungen getroffen wurden.

3.1. LaTeX-Editoren

Ein guter Cross-Plattform (Windows/Linux/Mac) Latex-Editor mit englischer und deutscher Rechtschreibkorrektur ist z.B. TexStudio (<http://texstudio.sourceforge.net/>). Unter Windows verwende ich diesen Editor gerne zusammen mit dem Sumatra PDF Viewer (<http://blog.kowalczyk.info/software/sumatrapdf/free-pdf-reader.html>), da dieser ein automatisches Neuladen unterstützt.

3.2. Beispiel für eine Tabelle

In Tabelle 3.1 sind die verwendeten Bibliotheken aufgelistet.

Bibliothek	Version
CUDA SDK	2.3
CUDA Toolkit	2.3
OpenGL	3.2
GLUT	3.7
GLEW	1.5.1
CUDPP	1.1

Tabelle 3.1.: Die verwendeten Bibliotheken.

3.3. Formel in Latex und Konventionen zur Verwendung mathematischer Symbole

Mathematische Symbole können in Latex sehr leicht erzeugt werden. Beispiel für Symbole im Text: Gegeben sei ein Skalar $a \in \mathbb{R}$. Tabelle 3.2 liste einige Konventionen zur Verwendung mathematischer Symbole. Abgesetzte Formel lassen sich ebenfalls leicht erzeugen:

$$f(x) = x^2 + 3 \tag{3.1}$$

Außerdem kann leicht auf abgesetzte Formel verwiesen werden (siehe Gleichung 3.1). Für mehrere ausgerichtete Formeln bietet sich die Umgebung `eqnarray` an:

$$f(x) = x^2 + 3 \quad (3.2)$$

$$g(\theta) = \cos(2\theta) = \cos^2 \theta - \sin^2 \theta \quad (3.3)$$

$$h(x) = \int_0^\infty e^{-x} dx \quad (3.4)$$

Typ	Schriftart	Beispiele
Variablen (Skalare)	kursiv	a, b, x, y
Funktionen	aufrecht	$f, g(x), \max(x)$
Vektoren	fett, Elemente zeilenweise	$\mathbf{a}, \mathbf{b} = \begin{pmatrix} x \\ y \end{pmatrix} = (x, y)^\top$
Matrizen	Schreibmaschine	$\mathbf{A}, \mathbf{B} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$
Mengen	kalligrafisch	$\mathcal{A}, \{a, b\} \in \mathcal{B}$
Zahlenbereiche	doppelt gestrichen	$\mathbb{N}, \mathbb{Z}, \mathbb{R}^2, \mathbb{R}^3$

Tabelle 3.2.: Konventionen zur Verwendung mathematischer Symbole

3.4. Beispiel für eine Vektorgrafik

Wenn möglich sollten immer Vektorgrafiken verwendet werden. Rastergrafiken sollten nur dann eingesetzt werden, wenn die Original-Quelle ebenfalls eine Rastergrafik ist. Ein Cross-Plattform Editor zur Erstellung von Vektorgrafiken ist z.B. Inkscape: <http://inkscape.org/download/>. Nach der Erstellung in Inkscape sollte die Grafik zum einen zur späteren Weiterverarbeitung als SVG gespeichert werden. Zum anderen zwecks Import in Latex als PDF. Abbildung 3.1 zeigt ein Beispiel.

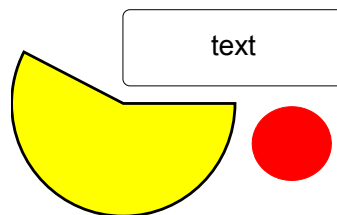


Abbildung 3.1.: Eine Vektorgrafik

3.5. Beispiel für eine Vektorgrafik mit mathematischen Symbolen

Um beliebigen Latex-Code in eine Vektorgrafik einzufügen (z.B. um mathematische Symbole zu setzen) kann in Inkscape beim Speichern der Datei als PDF die Option „Pdf+Latex: Text in PDF weglassen und Latex Datei erstellen“ angewählt werden (siehe Abb. 3.2)

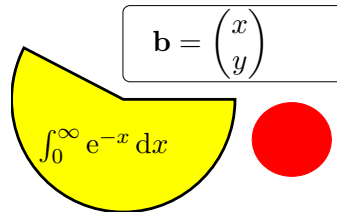


Abbildung 3.2.: Eine Vektorgrafik mit mathematischen Symbolen

3.6. Beispiel für eine Rastergrafik

Abbildung 3.3 zeigt, wie eine Rastergrafik eingebunden werden kann.

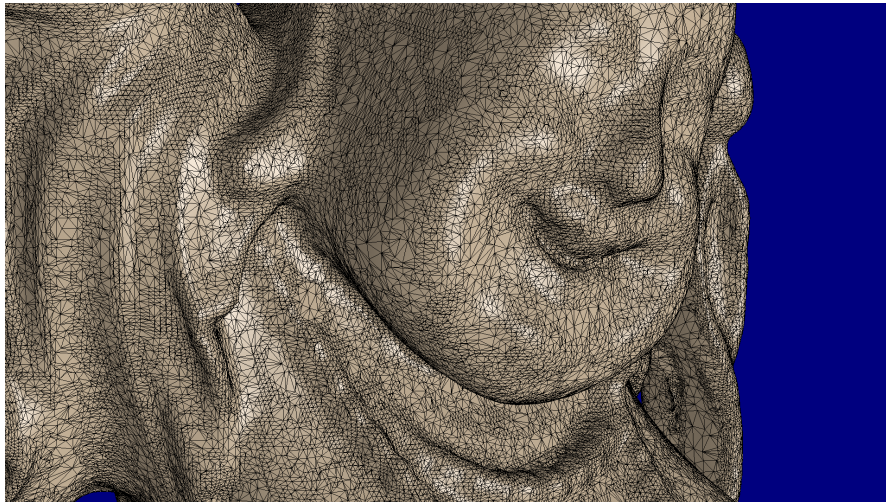


Abbildung 3.3.: Eine Rastergrafik

Bild- bzw. Tabellen-Beschriftungen sollten möglichst informativ sein. Aus der Beschreibung sollte die Bedeutung der Abbildung vollständig hervorgehen, so dass der Haupttext zum Verständnis nicht notwendigerweise gelesen werden muss.

3.7. Beispiel für die Darstellung von Algorithmen

Algorithmus 1 zeigt ...

Phase 1: Reduktion

```
for  $d := 0$  to  $\log_2 n - 1$  in parallel do
  for  $k := 0$  to  $n - 1$  by  $2^{d+1}$  in parallel do
     $x[k + 2^{d+1} - 1] := x[k + 2^d - 1] + x[k + 2^{d+1} - 1]$ 
```

Phase 2: Propagation

```
for  $d := \log_2 n$  to 0 in parallel do
  for  $k := 0$  to  $n - 1$  by  $2^{d+1}$  in parallel do
     $t := x[k + 2^d - 1]$ 
     $x[k + 2^d - 1] := x[k + 2^{d+1} - 1]$ 
     $x[k + 2^{d+1} - 1] := t + x[k + 2^{d+1} - 1]$ 
```

Algorithm 1: Pseudocode der zwei Phasen vom SCAN-Algorithmus [?].

3.8. Beispiel für die Darstellung von Quellcode-Auszügen

Listing 3.1 zeigt ...

```
1 void Adapt() {
2   cudaGLMapBufferObject(..., vertVbo);
3   cudaGLMapBufferObject(..., indVbo);
4   ...
5   cudaGLUnmapBufferObject(vertVbo);
6   cudaGLUnmapBufferObject(indVbo);
7 }
```

Listing 3.1: Pseudocode für die Kontrolle der VBOs.

3.9. Doppelseitiger Druck

Diese Vorlage ist für den doppelseitigen Druck konzipiert und enthält daher am Anfang des Dokuments mehrere leere Rückseiten. Falls ein einseitiger Ausdruck erfolgen soll, können diese zusätzlichen `\newpage`-Befehle entfernt werden.

3.10. Zitieren und Quellenangaben

Richtiges Zitieren ist eine Kernkompetenz für wissenschaftliches Arbeiten. Es dienen u. a. dazu zwischen dem eigenen Beitrag und der Arbeit anderer zu differenzieren. Nur so wird der eigene Beitrag nachvollziehbar und bewertbar.

Wörtliche Zitate (auch wörtliche Übersetzungen) müssen in Anführungszeichen gesetzt werden. Beispiel:

“I love deadlines. I like the whooshing sound they make as they fly by.”
-*Douglas Adams*

Auf das Kopieren ganzer Textblöcke sollte verzichtet werden. Nicht in Anführungszeichen gesetzte wörtliche Textübernahmen werden als Täuschungsversuch gewertet.

Wird etwas nur sinngemäß in eigenen Worten dargestellt (indirektes Zitat), reicht eine Referenz auf die Quelle durch den Befehl `\cite`. Beispiel: [?]. Auch Quellen von Bildern, Tabellen oder Code müssen referenziert werden.

Referenzen auf Internetseiten müssen immer mit einem Datum versehen werden. Da die Seriosität von Internetseiten oft als niedriger angesehen wird als die von vollständig referenzierbaren Veröffentlichungen, sollten Internetquellen nur in Ausnahmefällen verwendet werden (z. B. wenn die Informationen sehr aktuell sind und keine anderweitig referenzierbaren Werke existieren). Um den Unterschied der Quellen hervorzuheben, ist es sinnvoll, bei einer Internetseite oder andere Quellen mit problematischer Zitierwürdigkeit statt `\cite` den Befehl `\footnote` zu verwenden¹.

Wird etwas als gedankliches Allgemeingut angesehen (z. B. Grundrechenarten, Grundkonzepte objektorientierter Programmierung, etc.) muss hierfür keine Quelle angegeben werden.

¹Quelle: <http://www.uni-marburg.de/fb12/> am 17. Dezember 2009

4. Ergebnisse und Evaluation

In diesem Kapitel sollen die Ergebnisse dieser Diplomarbeit diskutiert werden.

5. Zusammenfassung und Ausblick

In diesem Kapitel sollen zunächst die erreichten Ziele diskutiert und abschließend ein Ausblick auf mögliche, weiterführende Arbeiten gegeben werden.

Literaturverzeichnis

- [BP12] Bhole C., Pal C.: Automated person segmentation in videos. In *Pattern Recognition (ICPR), 2012 21st International Conference on* (2012), IEEE, pp. 3672–3675.
- [BTVG06] Bay H., Tuytelaars T., Van Gool L.: Surf: Speeded up robust features. *Computer vision—ECCV 2006* (2006), 404–417.
- [BVD09] Barnich O., Van Droogenbroeck M.: Vibe: a powerful random technique to estimate the background in video sequences. In *Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on* (2009), IEEE, pp. 945–948.
- [BVD11] Barnich O., Van Droogenbroeck M.: Vibe: A universal background subtraction algorithm for video sequences. *IEEE Transactions on Image processing* 20, 6 (2011), 1709–1724.
- [DG13] Duffner S., Garcia C.: Pixeltrack: a fast adaptive algorithm for tracking non-rigid objects. In *Proceedings of the IEEE international conference on computer vision* (2013), pp. 2480–2487.
- [EHD00] Elgammal A., Harwood D., Davis L.: Non-parametric model for background subtraction. *Computer Vision—ECCV 2000* (2000), 751–767.
- [FW06] Fleet D., Weiss Y.: Optical flow estimation. In *Handbook of mathematical models in computer vision*. Springer, 2006, pp. 237–257.
- [GM06] Guo P., Miao Z.: Projection histogram based human posture recognition. In *Signal Processing, 2006 8th International Conference on* (2006), vol. 2, IEEE.
- [GRB13] Godec M., Roth P. M., Bischof H.: Hough-based tracking of non-rigid objects. *Computer Vision and Image Understanding* 117, 10 (2013), 1245–1256.
- [HHD98] Haritaoglu I., Harwood D., Davis L. S.: Ghost: A human body part labeling system using silhouettes. In *Pattern Recognition, 1998. Proceedings. Fourteenth International Conference on* (1998), vol. 1, IEEE, pp. 77–82.

- [KHD05] KIM K., HARWOOD D., DAVIS L. S.: Background updating for visual surveillance. In *International Symposium on Visual Computing* (2005), Springer, pp. 337–346.
- [MPWSH16] MÄRKI N., PERAZZI F., WANG O., SORKINE-HORNUNG A.: Bilateral space video segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2016), pp. 743–751.
- [Pic04] PICCARDI M.: Background subtraction techniques: a review. In *Systems, man and cybernetics, 2004 IEEE international conference on* (2004), vol. 4, IEEE, pp. 3099–3104.
- [PS02] POWER P. W., SCHOONEES J. A.: Understanding background mixture models for foreground segmentation. In *Proceedings image and vision computing New Zealand* (2002), vol. 2002, pp. 10–11.
- [SG99] STAUFFER C., GRIMSON W. E. L.: Adaptive background mixture models for real-time tracking. In *Proceedings. 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No PR00149)* (1999), vol. 2, p. 252 Vol. 2.
- [ZVDH06] ZIVKOVIC Z., VAN DER HEIJDEN F.: Efficient adaptive density estimation per image pixel for the task of background subtraction. *Pattern recognition letters* 27, 7 (2006), 773–780.

Abkürzungsverzeichnis

KDE kernel density estimation

ROI region of interest

ALL assisted ambient living

Abbildungsverzeichnis

3.1. Eine Vektorgrafik	8
3.2. Eine Vektorgrafik mit mathematischen Symbolen	9
3.3. Eine Rastergrafik	9

Tabellenverzeichnis

3.1. Die verwendeten Bibliotheken.	7
3.2. Konventionen zur Verwendung mathematischer Symbole	8

List of Algorithms

1. Pseudocode der zwei Phasen vom SCAN-Algorithmus [?]. 10

Listings

3.1. Pseudocode für die Kontrolle der VBOs.	10
---	----

A. Anhang

Thema 1

Beispiel für einen Anhang

Thema 2

Danksagung

Hiermit möchte ich mich besonders bei Prof. Dr. XXXX, Prof. Dr. xxxx, Dipl. Inf. xxxx und Dipl. Inf. xxx für die Betreuung meiner Arbeit, hilfreiche Diskussionen und viel Geduld bei zahlreichen Fragen bedanken.