



Hochschule für Angewandte Wissenschaften Hamburg
Hamburg University of Applied Sciences

Ausarbeitung: Anwendung 1

Jan-Uriel Lorbeer

Gestenerkennung mit Partikelfiltern

Inhaltsverzeichnis

1	Einleitung	1
2	Partikelfilter und Gestenerkennung	2
2.1	Partikelfilter	2
2.2	Nutzung von Partikelfiltern bei der Gestenerkennung	5
3	Forschungsumfeld und Focus	8
3.1	Gesetzter Focus	8
3.2	Forschungsumfeld	9
	Literaturverzeichnis	10

Abbildungsverzeichnis

2.1	Prozessschritte eines Partikelfilter [SWTO04]	3
2.2	Lokalisierung eines Roboters mit Hilfe eines Partikelfilters [Uda]	4
2.3	Hirarisches Handmodell [BLL02]	6
2.4	Partikelfilter Handerkennung mit hirarischem Handmodell [BLL02]	6
2.5	Hautfarbenfilter [BLL02]	7

Kapitel 1

Einleitung

Mit wachsender Anzahl und Komplexität von Maschinen im alltäglichen Umfeld, wird dessen Steuerung eine zunehmend interessante Herausforderung. Die Erstellung und Verbesserung geeigneter Schnittstellen zwischen Mensch und Maschine, dem *Human-machine interface* (HMI), spielt dabei eine große Rolle. Ein Teil dieser Überlegungen ist die Gestensteuerung, hierbei sollen Gesten erkannt werden und in Befehle für den Computer umgesetzt werden.

Gesten sind Bewegungen, der Arme, des Kopfes und vor allem der Hände. Bei denen eine Information übertragen werden soll. Als Teil der nonverbalen Kommunikation sind die Gesten Teil des alltäglichen Lebens und könnten eine intuitive Alternative zur Steuerung von Elektronik bieten.

Die Entwicklung einer effektiven Gestensteuerung kann in vielen Bereichen Anwendung finden.[MA07] Dazu zählen zum Beispiel Hilfen für Hörgeschädigte, das Erkennen von Zeichensprache [MPB⁺10], die Möglichkeit für Kleinkinder mit Computern zu interagieren, die Überwachung von emotionalen Zuständen und Stressleveln von Personen (Patienten, Autofahrer, etc.) oder die Bedienung von Virtuellen Umgebungen.

Schon heute unterstützen einige Maschinen die Bewegungssteuerung. Hierbei ist zum Beispiel die Spiele Industrie zu nennen. Die in den vergangenen Jahren erschienen neuen Spielekonsolen unterstützen alle eine gewisse Art von Bewegungsteuerung. Zum einen kamerabasierte Systeme wie beider Xboxone von Microsoft mit der stereo kinect Kamera die Spieler eigenständig erfasst oder die Sony Play Station 4 dessen Kamera sich an einer am Spielekontroller angebrachten Leuchte orientiert. Zum anderen die hardwarebasierte WiiU, die über Bewegungssensoren in den Spielekontrollern funktioniert.

Einige Versuche haben gezeigt, dass sich Partikelfilter auch für die Gestenerkennung als nützlich erweisen können.[BLL02] In der Robotik wurden Partikelfilter bereits erfolgreich getestet. Zum Beispiel haben bei der DARPA (*Defense Advanced Research Projects Agency*) Challenge Partikelfilter in Autonomen PKW Verwendung gefunden, wo sie unter anderem für das Verfolgen von Objekten oder die selbst Lokalisierung des Fahrzeuges eingesetzt wurden.[TMD⁺06]

In Kapitel 2 wird sowohl die Funktionsweise von Partikelfiltern, als auch deren Einsatzmöglichkeiten für die Gestenerkennung, erläutert. Im darauffolgenden Kapitel 3 wird auf das weitere Vorgehen für die Veranstaltungen Anwendung 2 und Projekt 1 Bezug genommen und ein Überblick über das Forschungsumfeld gewährt.

Kapitel 2

Partikelfilter und Gestenerkennung

In diesem Kapitel wird die Verwendung von Partikelfiltern für die Gestenerkennung erläutert. Zunächst werden die Funktionsweise von Partikelfiltern in Abschnitt 2.1 dargestellt und anhand eines Beispiels aus der Robotik veranschaulicht. Anschließend wird in Abschnitt 2.2 beschrieben wie Partikelfilter für die Gestenerkennung genutzt werden können.

2.1 Partikelfilter

Bei dem Partikelfilter handelt es sich um einen stochastischen Filter der auf dem Satz der bedingten Wahrscheinlichkeit von Bayes aufbaut.

Der Partikelfilter dient dazu, mithilfe von Messungen, den Zustand eines dynamischen Systems zu schätzen. Der Zustand wird hierbei durch einen Vektor von Attributen beschrieben, dabei dürfen die Attribute auch kontinuierlicher Natur sein.

Zur Veranschaulichung wird die Lokalisierung eines Roboters herangezogen. Der Zustandsvektor $s = (x, y, a)$ besteht hierbei aus der Position des Roboters (x, y) und einem Ausrichtungswinkel a .

Zu Beginn erzeugt der Partikelfilter eine Vielzahl von Instanzen (Partikeln) des Zustandsvektors. Jede dieser Instanzen wird mit zufälligen Werten (innerhalb der Systemgrenzen) für die Attribute initialisiert. Auf diese Weise bildet sich ein Partikelschwarm, wobei jeder Partikel eine Hypothese für den tatsächlichen Systemzustand darstellt.

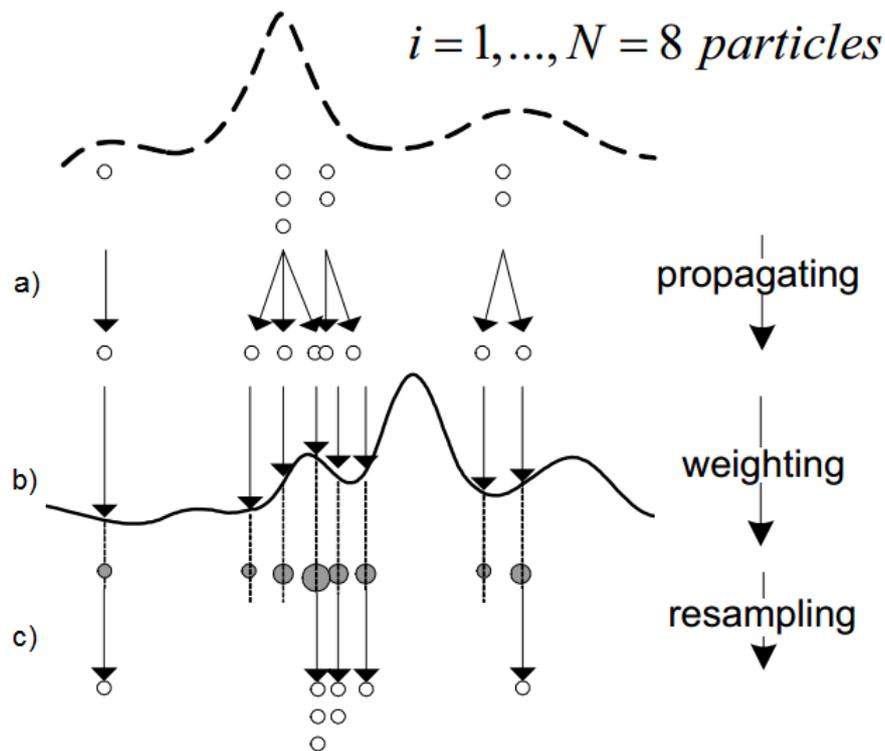


Abbildung 2.1: Prozessschritte eines Partikelfilter [SWTO04]

Durch den iterativen Prozess des Partikelfilters (Abb. 2.1) soll eine Approximation des tatsächlichen Systemzustandes erreicht werden.

Bei diesem Prozess wird zunächst von jedem Partikel die Aktion des Systems durchgeführt. Dieser Schritt nennt sich *propagation* oder auch *prediction* (Abb. 2.1 a)). Bei der Lokalisierung wäre dies zum Beispiel die Bewegung des Roboters, was eine Veränderung des Zustandsvektors zufolge hat.

Nach einer Messung (*measurement*) im System wird jeder der Partikel gewichtet. Das Gewicht repräsentiert die Wahrscheinlichkeit mit der die Messung zu dem Zustandsvektor des Partikels passt. Hierdurch erhalten plausible Partikel ein höheres Gewicht als weniger plausible. (Abb. 2.1 b)) Für die Lokalisierung kann nun der Partikel mit der höchsten Gewichtung, als Position des Roboters angenommen werden.

Im nächsten Schritt, dem *resampling*, werden N neue Partikel erstellt (für einen Partikelfilter mit N Partikeln). Jeder dieser Partikel ist eine Kopie eines bereits existierenden Partikels. Hierbei werden Partikel mit geringerer Gewichtung weniger häufig oder gar nicht kopiert. Partikel mit höherer Gewichtung werden häufiger kopiert. (Abb. 2.1 c))

Dieser Prozess wird immer wieder durchgeführt. Dies führt dazu, dass sich der Partikelschwarm immer weiter um den tatsächlichen Systemzustand konzentriert.

In Abbildung 2.2 ist die Lokalisierung eines Roboters mithilfe eines Partikelfilters visualisiert. Der Roboter verfügt über eine Karte der Umgebung und mehrere Abstandssensoren. Die vom Roboter angenommene Position ist in Grün dargestellt, weicht die tatsächliche Position von der angenommenen Position ab so ist sie in Schwarz verzeichnet. Jeder rote Punkt stellt ein Partikel da, welcher sich aus einer Position (x,y) und einem Ausrichtungswinkel a zusammen setzt.

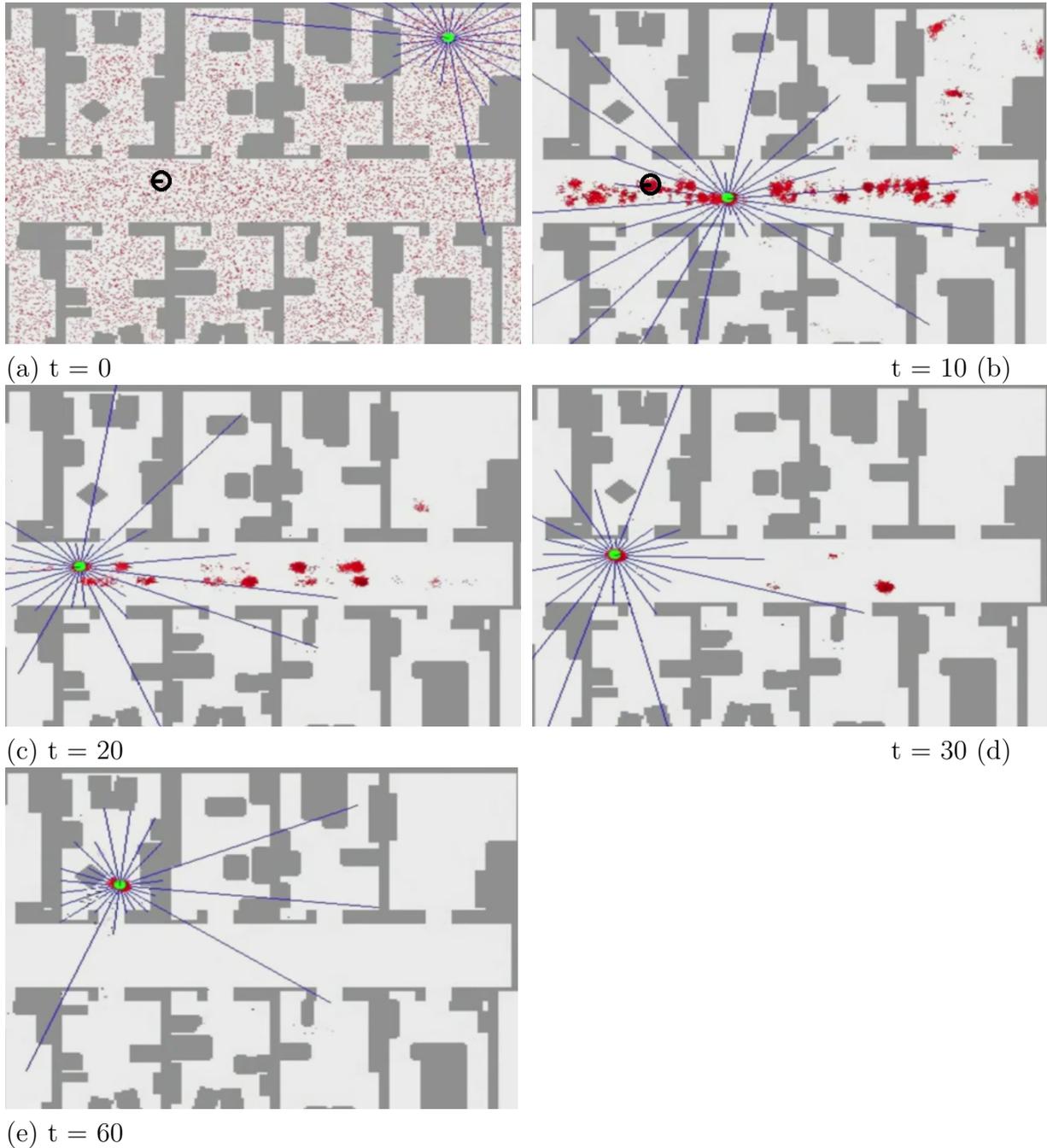


Abbildung 2.2: Lokalisierung eines Roboters mit Hilfe eines Partikelfilters [Uda]

Bei (a) lässt sich die initiale zufällige Verteilung der Partikel erkennen. Die vom Roboter angenommene Position ist ebenso zufällig.

Nach einigen Schritten in denen der Roboter sich vorwärts bewegt hat, hat der Roboter erkannt, dass er sich mit einer hohen Wahrscheinlichkeit auf dem Gang befindet. In der Abbildung 2.2 bei (b) lässt sich ebenso erkennen, dass sich der Partikelschwarm auf kleinere Bereiche zusammen gezogen hat, wobei er jedoch seine tatsächliche Position noch nicht erkannt hat.

Der Roboter hat bei (c) seine tatsächliche Position gefunden. Es existieren jedoch noch weitere Ballungszonen, was darauf schließen lässt, dass die gefundene Position noch nicht sicher ist. Der Partikelfilter verfolgt auch die anderen Hypothesen weiterhin. In diesem Multi-modalen Ansatz liegt, neben der geringen Komplexität von Partikelfiltern, eine der größten Stärken von Partikelfiltern, wodurch sie sich gegenüber Fehlannahmen robust verhalten (Fehlannahmen wie in Abb. 2.2 (b)) .

Erst nachdem der Roboter bei (e) einen Raum betreten hat, konnte er seine Position sicher bestimmen und hat alle anderen Hypothesen ausgeschlossen. Aufgrund der Symmetrie des Ganges gab es bei (d) immer noch einen anderen möglichen Aufenthaltsort des Roboters.

Für die initiale Findung des Systemzustandes ist eine hohe Anzahl von Partikel notwendig, um den Systembereich großflächig abzudecken. Diese hohe Anzahl an Partikel ist jedoch auf Dauer nicht immer notwendig, daher existieren ebenfalls Partikelfilter mit einer variablen Anzahl von Partikel.

Eine Erläuterung des Partikelfilters bis ins (mathematische) Detail kann auch in [AMG02] gefunden werden.

2.2 Nutzung von Partikelfiltern bei der Gestenerkennung

Um Partikelfilter für die Handerkennung nutzbar zu machen, muss ein geeignetes Modell für die Partikel gewählt werden. Bei dem Lokalisierungsbeispiel aus Abschnitt 2.1 bestand das Modell für den Roboter aus seiner Position und seiner Orientierung. Für die 2D Handerkennung kann ein dem Modell des Roboters gar nicht so unähnliches Modell gewählt werden. Ein hierarchisches Handmodell, wie in Abbildung 2.3 dargestellt, verfügt in diesem Fall über eine Position im Bild (x,y) , eine Ausrichtung der Hand a , also die Richtung in die die Finger zeigen, und einen großen skalierungs Faktor s . Der Faktor s ist nötig, da die Hand, je nach Entfernung, unterschiedlich groß sein kann. Eine Veränderung der Größe über die Zeit erlaubt Rückschlüsse auf die Entfernung zur Kamera. Das Modells setzt sich zusammen aus einem Handballen, fünf Fingern und einer Fingerspitze je Finger.

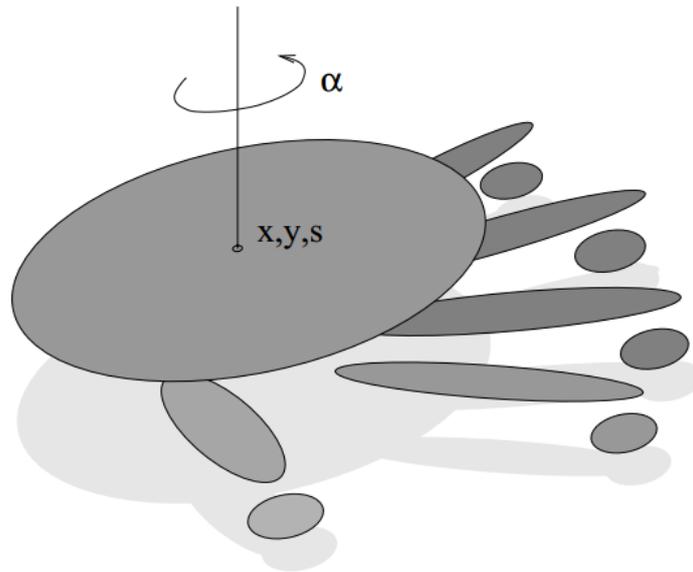


Abbildung 2.3: Hierarisches Handmodell [BLL02]

Die Gewichtung der Partikel hängt in diesem Zusammenhang davon ab, wie gut ein Partikel dem darunter liegenden Objekt im Bild entspricht. Für die Gewichtung werden ebenfalls Metadaten eines Partikel herangezogen. Dies sind zum Beispiel die Geschwindigkeit oder Beschleunigung eines Partikels, welche aus der Positionsdivergenz gewonnen werden.

In Abbildung 2.4 ist ein Partikelfilter mit hierarchischem Handmodell in Aktion zu sehen. Die Partikel sind dabei in Geld abgebildet. Der Partikel welcher am wahrscheinlichsten die Hand abbildet ist in Rot gekennzeichnet.

Mit einer ausreichend großen Anzahl an Partikel wird so die Problemstellung, die Hand zu finden, auf ein Validierungsproblem reduziert.



Abbildung 2.4: Partikelfilter Handerkennung mit hierarischem Handmodell [BLL02]

Der Ansatz mit Partikelfilter lässt sich mit anderen Ansätzen kombinieren, um die Effektivität der Handerkennung zu erhöhen. Auf Bildebene kann dies zum Beispiel durch einen Hautfarbenfilter geschehen, wie in Abbildung 2.5 dargestellt. Videomaterial kann mit einem Bewegungsfiler vorbehandelt werden.

Auf Partikelebene, lässt sich der Partikelfilter zum Beispiel mit unimodalen Filtern, wie dem Kalman Filter, kombinieren. In dem Hinzufügen eines weiteren Schrittes im iterativen Prozess wie in [SWTO04] vorgeschlagen besteht ebenfalls eine Verbesserungsmöglichkeit.

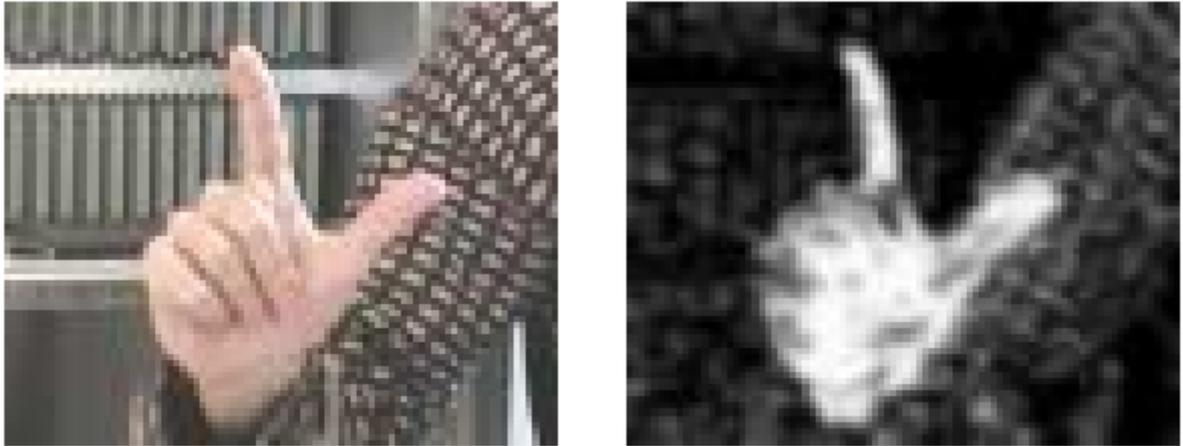


Abbildung 2.5: Hautfarbenfilter [BLL02]

Kapitel 3

Forschungsumfeld und Focus

3.1 Gesetzter Focus

Der gesetzte Focus liegt auf einer robusten Gestenerkennung. Robust bedeutet in diesem Zusammenhang, dass die Gestenerkennung auch funktionieren soll wenn erschwere Umstände hinzu kommen wie:

- ▷ schlechtes Bildmaterial, wie z.B. Webcams es liefern
- ▷ schwer von den Händen unterscheidbare Hintergründe
- ▷ die Geste füllt nur einen geringen Teil des Bildes aus, dies ist z.B. bei Überwachungskameras der Fall

Die handbasierte Gestenerkennung kann Grob in drei Schritte Aufgeteilt werden:

1. die Hände finden
2. die Hände verfolgen
3. interpretieren der Geste

An der HAW-Hamburg beschäftigte sich in diesem Semester Peter Oltmann, im Rahmen von Anwendung 2, mit einem Verfahren, dass aktive Konturen und Partikelfilter nutzt. Dieses Verfahren eignet sich gut für die Verfolgung von Objekten (Schritt 2). Auf Grund seiner Rechenkomplexität kann dieses Verfahren jedoch bei der initialen Objektfindung auf Probleme stoßen. Für die kommenden Veranstaltungen Anwendung 2 und Projekt 1 soll für mich daher zunächst Schritt 1 im Focus stehen.

Das vorerst gesetzte Ziel soll sein, verschiedene Methoden für die Handmustererkennung zu Untersuchen.

3.2 Forschungsumfeld

Im Bereich der Gestenerkennung existieren einige interessante Konferenzen. Dazu zählt die *IFIP Conference on Human-Computer Interaction (INTERACT)* welche sich mit der akademischen und industriellen Sicht auf HCI beschäftigt. Eine weitere Konferenz ist die *HCI International*, die versucht alle von HCI betroffenen Bereiche, sowohl technische als auch nicht technische, miteinander zu verbinden. Die *IEEE FG : Automatic Face and Gesture Recognition* ist spezieller auf Gesten- und Gesichtserkennung ausgerichtet.

Einer der führenden Geister in Sache HMI ist Thomas S. Huang von der University of Illinois at Urbana-Champaign. Im Bereich der Partikelfilter ist Sebastian Thrun von der Stanford University einer der führenden Forscher, wobei er unter anderem bei der Erstellung des autonomen PKW „Stanley“ mitgewirkt hat.[TMD⁺06]

Für die angestrebte (Hand-) Mustererkennung bieten die *Convolutional Networks* einen Ansatz, einer der führenden Köpfe auf diesem Gebiet ist Yann LeCun von der New York University.

Des weiteren bieten Ansätze, die auch in der Gesichtserkennung angewandt werden, Möglichkeiten für die Handerkennung. In [SAS12] ist einer dieser Ansätze dargelegt, wobei es sich dabei um einen Hand tracker auf der Basis von Eigenvektoren (*Eigenhand*) handelt, ähnlich den *Eigenfaces* aus der Gesichtserkennung [TPS].

Ebenfalls erwähnenswert sind die *deformable Templates* [WSFZ]. Diese werden für die Erkennung von verschiedensten Objekten verwendet und bieten ebenfalls einen untersuchenswerten Ansatz im gesetzten Focus.

Literaturverzeichnis

- [AMG02] ARULAMPALAM, M. S. ; MASKELL, Simon ; GORDON, Neil: A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. In: *IEEE TRANSACTIONS ON SIGNAL PROCESSING* 50 (2002), S. 174–188
- [BLL02] BRETZNER, Lars ; LAPTEV, Ivan ; LINDBERG, Tony: Hand Gesture Recognition using Multi-Scale Colour Features, Hierarchical Models and Particle Filtering. In: *Shortened version in Proc. Face and Gesture 2002*, 2002, S. 423–428
- [MA07] MITRA, S. ; ACHARYA, T.: Gesture Recognition: A Survey. In: *Trans. Sys. Man Cyber Part C* 37 (2007), Mai, Nr. 3, 311–324. <http://dx.doi.org/10.1109/TSMCC.2007.893280>. – DOI 10.1109/TSMCC.2007.893280. – ISSN 1094–6977
- [MPB⁺10] MADEO, Renata C. B. ; PERES, Sarajane M. ; BÍSCARO, Helton H. ; DIAS, Daniel B. ; BOSCARIOLI, Clodis: A Committee Machine Implementing the Pattern Recognition Module for Fingerspelling Applications. In: *Proceedings of the 2010 ACM Symposium on Applied Computing*. New York, NY, USA : ACM, 2010 (SAC '10). – ISBN 978–1–60558–639–7, 954–958
- [SAS12] SHAHRIMIE, Mohd ; ASAARI, Mohd ; SUANDI, Shahrel A.: Real Time Hand Tracking System Using Predictive Eigenhand Tracker. In: *Australian Journal of Basic and Applied Sciences* 6 (2012), Nr. 2, S. 190–198. – ISSN 1991–8178
- [SWTO04] SHAN, Caifeng ; WEI, Yucheng ; TAN, Tieniu ; OJARDIAS, Frederic: Real Time Hand Tracking by Combining Particle Filtering and Mean Shift. In: *Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition (FGR '04)*, 2004 (IEEE '04). – ISBN 0–7695–2122–3
- [TMD⁺06] THRUN, Sebastian ; MONTEMERLO, Mike ; DAHLKAMP, Hendrik ; STAVENS, David ; ARON, Andrei ; DIEBEL, James ; FONG, Philip ; GALE, John ; HALPENNY, Morgan ; HOFFMANN, Gabriel ; LAU, Kenny ; OAKLEY, Celia ; PALATUCCI, Mark ; PRATT, Vaughan ; STANG, Pascal ; STROHBAND, Sven ; DUPONT, Cedric ; JENDROSSEK, Lars-Erik ; KOELEN, Christian ; MARKEY, Charles ; RUMMEL, Carlo ; NIEKERK, Joe van ; JENSEN, Eric ; ALESSANDRINI, Philippe ; BRADSKI, Gary ; DAVIES, Bob ; ETTINGER, Scott ; KAEHLER, Adrian ; NEFIAN, Ara ; MAHONEY, Pamela: Stanley: The Robot That Won the DARPA Grand Challenge:

Research Articles. In: *J. Robot. Syst.* 23 (2006), September, Nr. 9, 661–692. <http://dx.doi.org/10.1002/rob.v23:9>. – DOI 10.1002/rob.v23:9. – ISSN 0741–2223

[TPS] TAYAL, Yogesh ; P, Pramod K. ; SINGH, D. B. V.: *Face Recognition using Eigenface*

[Uda] <https://www.udacity.com/course/viewer#!/c-cs373/1-48704330>

[WSFZ] WU, Ying N. ; SI, Zhangzhang ; FLEMING, Chuck ; ZHU, Song chun: *Deformable Template As Active Basis*