

Modellbasierte Handposenerkennung

Hannes Dieck
Studiengang Technische Informatik

**Betreuender Prüfer: Prof. Dr.-Ing.
Birgit Wendholt**
**Zweitgutachter : Prof. Dr.-Ing.
Andreas Meisel**

Dezember 2013



GLIEDERUNG

- ▶ INTRO
- ▶ ANALYSE
- ▶ LÖSUNG
- ▶ FAZIT & AUSBLICK



GLIEDERUNG

- ▶ **INTRO**
- ▶ ANALYSE
- ▶ LÖSUNG
- ▶ FAZIT & AUSBLICK



Motivation

Warum Posenerkennung ?

- **Erweiterte Grasping-HCI**
 - **Robotersteuerung und -training**
 - **Interaktion mit virtuellen und realen**
 - **Objekten**
 - **Aktivitätserkennung**

*Warum visuelle
Handposenerkennung ?*

- **Kein *Cyber Glove***

**Zielsetzung: Entwicklung eines Framework zur
modellbasierten Handposenerkennung**



GLIEDERUNG

- ▶ INTRO
- ▶ **ANALYSE**
 - ▶ PROBLEMBEREICH
 - ▶ VORVERARBEITUNG
 - ▶ VIEW-BASED ESTIMATION
 - ▶ MODEL-BASED ESTIMATION
- ▶ LÖSUNG
- ▶ FAZIT & AUSBLICK



Definition Posenerkennung

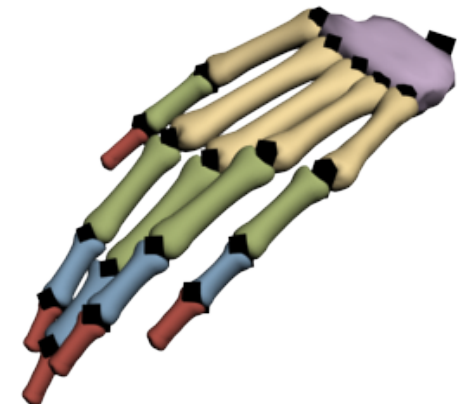
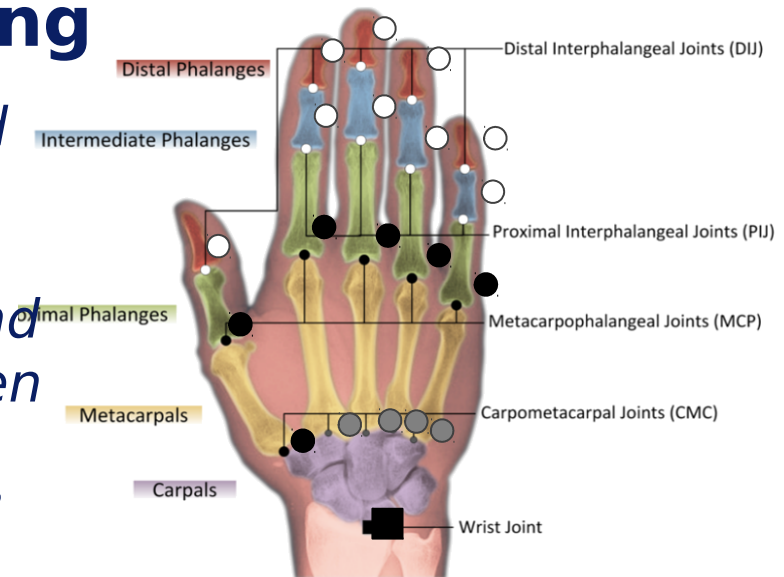
*Pose = aktueller Bewegungszustand
der artikulierten Hand*

Zustand ergibt sich aus

*Degrees of Freedom (DOFs) der Hand
Bewegung um die Verbindungsstellen
der Knochen (Joints)*

Ziel visuelle Handposenerkennung ?

- **Wertmäßige Bestimmung
(Messung) der DOFs aus einem
Kamerabild**



Schwierigkeiten

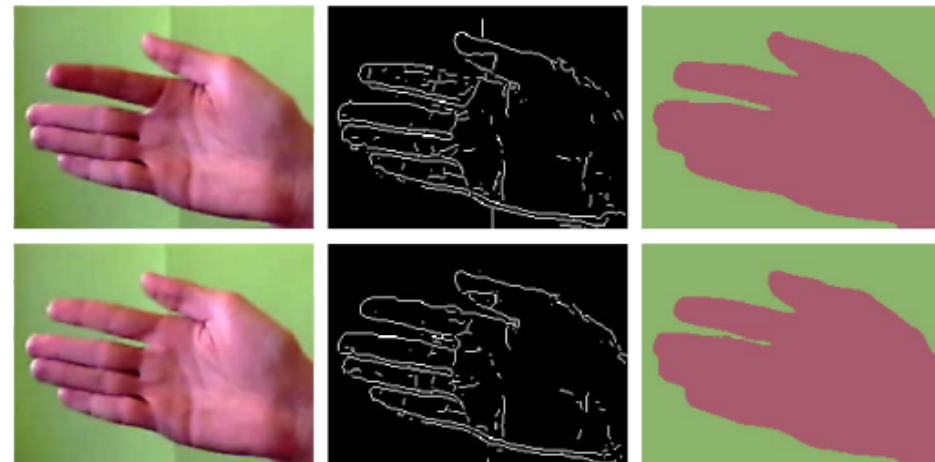
*Allgemeine Schwierigkeit der
Objekterkennung*

- **Unkontrollierte Umgebung**
- **Große Anzahl zu
verarbeitender Daten**

Selbstverdeckung

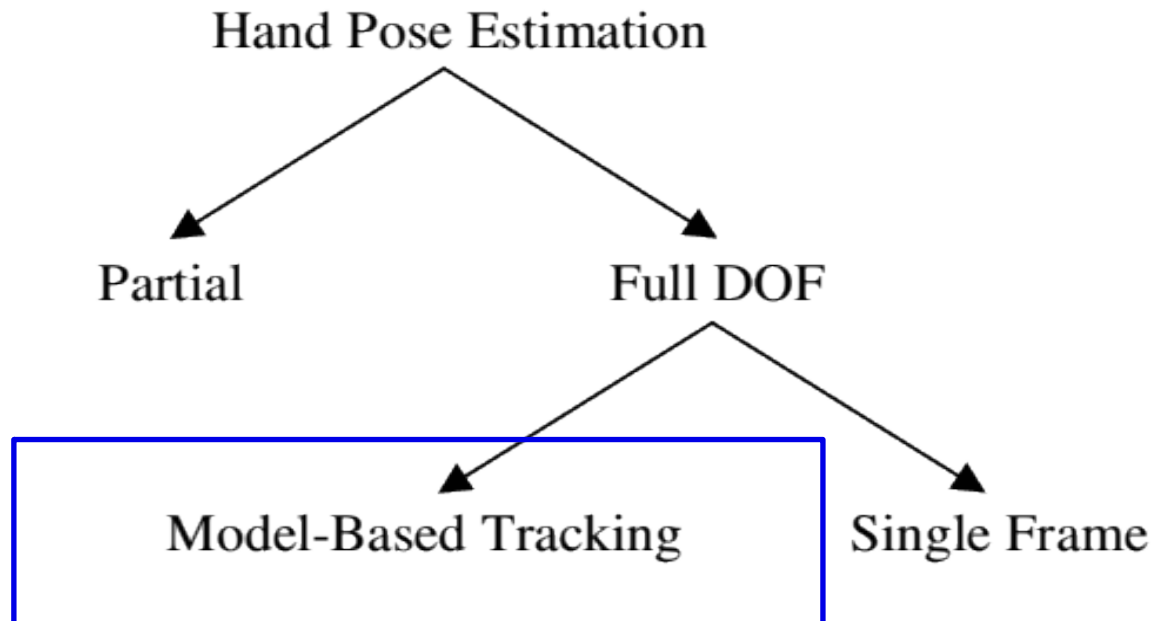
Nichtlineare Abbildung

*⇔ Mehrdeutigkeit
(Ambiguity)*



Abgrenzung und Zielsetzung

Ersten Ansatzpunkt liefern Erol u.a. (2007)



GLIEDERUNG

- ▶ INTRO
- ▶ **ANALYSE**
 - ▶ PROBLEMBEREICH
 - ▶ **VORVERARBEITUNG**
 - ▶ VIEW-BASED ESTIMATION
 - ▶ MODEL-BASED ESTIMATION
- ▶ LÖSUNG
- ▶ FAZIT & AUSBLICK



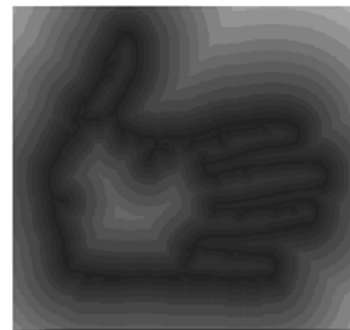
Segmentierung & Merkmalsextraktion

Segmentierungsmethoden:

- **Hautfarbenmodell**
- **Depth Thresholding**

Basis-Merkmalstypen:

- **Keypoints (z.B. Ecken o. Wendepunkte)**
- **Kantenzüge \Rightarrow Distanztransformation**
- **Tiefenmerkmale**



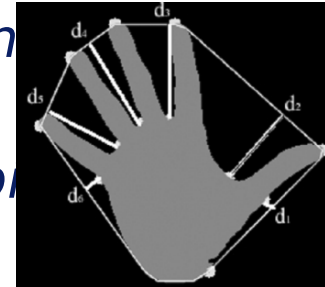
GLIEDERUNG

- ▶ INTRO
- ▶ **ANALYSE**
 - ▶ PROBLEMBEREICH
 - ▶ VORVERARBEITUNG
 - ▶ **VIEW-BASED ESTIMATION**
 - ▶ MODEL-BASED ESTIMATION
- ▶ LÖSUNG
- ▶ FAZIT & AUSBLICK



Partielle Handposenerkennung (Appearance Based)

*Ziel: Bestimmung von Merkmalen mit hohem Informationsgehalt (High-Level Features)
= Detektion von Handteilen (z.B. Fingerspitzen)
Verarbeitung:*



- **direkt die Merkmalsposition**
- **dynamische Gestenklassifikation**
- **Klassifizierung ausgewählter statischer**
- **Posen**

Meist sehr performant in Echtzeit (> 60 Hz)

+umsetzbar

+Keine allgemeine Garantie für die

–Messbarkeit der Merkmale



Single View Pose Estimation

Ziel: Allgemeine Bestimmung der kompletten 3D-Handpose zu jeder Zeit aus einem einzelnen Aufnahmebild

**Exempel-basierte Beschreibung der Abbildung zwischen Handmerkmalsraum
⇒ und DOF-Zustandsraum**

1. *Template Matching* = Direkter Vergleich mit Exempeln

2. *Stetige Abbildung erlernen*

3. Klassifizierung einzelner Handteile und anschließende

• **explizite Zusammenführung [Yuan Yao u.a., 2012]**

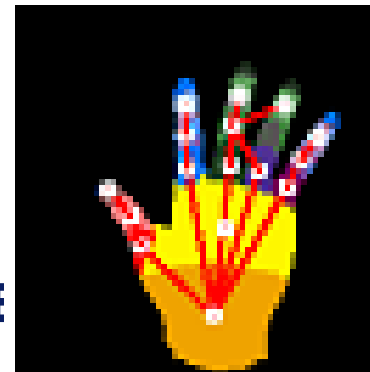
Nachteile:

Direkte Abhängigkeit von

– **dem Exempel-Datensatz**

– **Stark von Mehrdeutigkeiten**

– **betroffen ⇒ weniger robuste E**



GLIEDERUNG

- ▶ INTRO
- ▶ **ANALYSE**
 - ▶ PROBLEMBEREICH
 - ▶ VORVERARBEITUNG
 - ▶ VIEW-BASED ESTIMATION
 - ▶ **MODEL-BASED ESTIMATION**
- ▶ LÖSUNG
- ▶ FAZIT & AUSBLICK

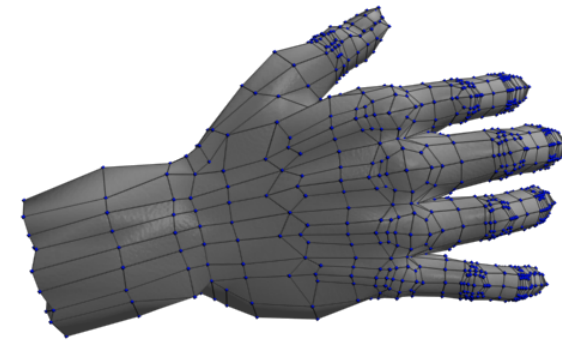
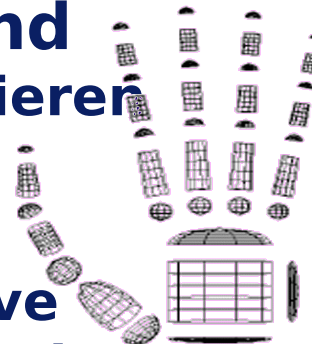


Modellbasierter Ansatz

- ▶ **Ziel: Online-Anpassung und Vergleich eines parametrisierten Handmodells mit der aufgenommenen Hand**
- ▶ **Modellparameter reflektieren**
 - ▶ **Hand-DOF**

1. Strukturmodell

- **Unabhängige 3D-Primitive**
- **Einheitliches Polygon Mesh**

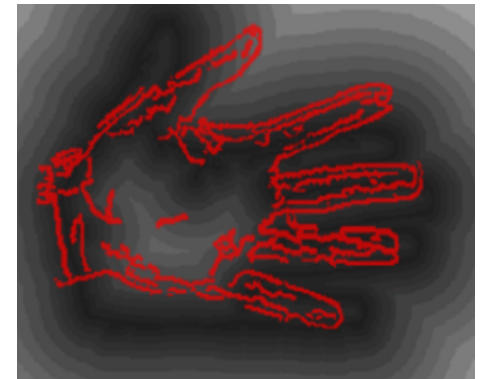


2. Kinematikmodell zur Modelltransformation

- **z.B. Skeleton Model**
- **inklusive Constraints**

3. Observationsmodell

- **Messung des Anpassungsfehlers**
- **Distanztransformation oder**
- **direkter Vergleich von Tiefenpixeln**



Parametersuche und Tracking

*Suche Parameterkonstellation, die
minimalen Fehler verursacht =*

Optimierungsproblem

*Problem: Mehrdeutigkeiten verursachen
mehrere Extrema*

*Lösung: Ausnutzen der temporalen
Kontinuität*

⇒ Rekursives Tracking

Anfangswertproblem ⇒ Initialisierung:

- **Iterative Closed Point (ICP)**
Algorithmus [Breuer 2005]
- **Single View Pose Estimation**
Verfahren



Trackingverfahren

*Single Hypothesis Tracking (SHT)
mit lokaler Optimierung*

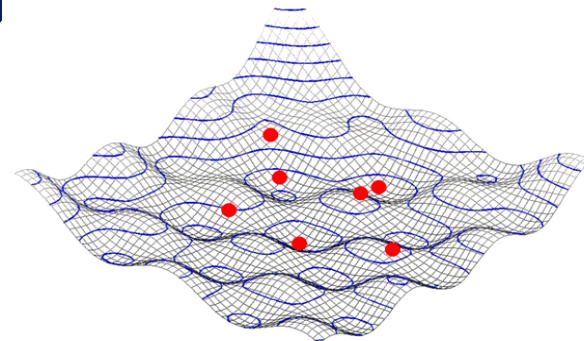
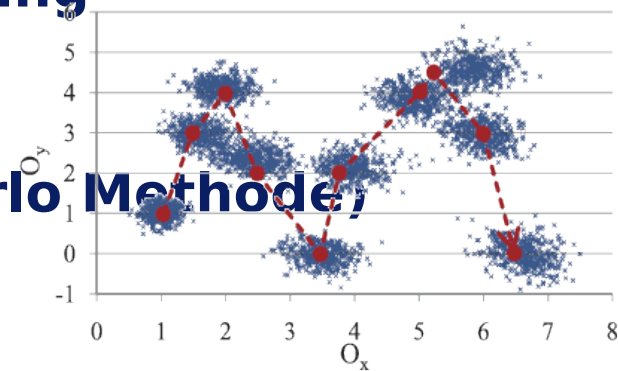
- **Gradientenbasierte Verfahren**
- **Stochastische Verfahren ohne Ableitung**
- **z.B. Unscented Kalman Filter (UKF)**

Multiple Hypothesis Tracking (MHT):

- **Partikelfilter (Sequenzielle Monte Carlo Methode)**
- **Particle Swarm Optimization (PSO)**

[Oikonomidis u.a. 2011]

- **SHT mit globaler Optimierung**



GLIEDERUNG

- ▶ INTRO
- ▶ ANALYSE
- ▶ **LÖSUNG**
 - ▶ ROADMAP
 - ▶ OFFLINE PHASE
 - ▶ PSO TRACKING
 - ▶ ARCHITEKTUR
 - ▶ ONLINE MODELLE
- ▶ FAZIT & AUSBLICK



Entwicklungsziele

1.Offline

- **Kamerakalibrierung**
- **Modellierung**
 - **Strukturmodell = Polygon-Mesh**
 - **Offline Kinematikmodell = Skeleton**
 - **Statistisches Hautfarbenmodell**

2.Online

- **Framework-Implementierung**
 - **Hautfarben-Segmentierungsmodul**
 - **Online Kinematikmodell**
 - **Observationsmodell** ⇒ Tiefenbildvergleich mit GPU-Beschleunigung + Distanztransformation
 - **Tracking = Particle Swarm Optimization**
 - **Initialisierung durch einfache Startpose**



Plattform, Hardware, Tools und Bibliotheken

C/C++

Windows 7 auf Quad-Core

Microsoft Kinect

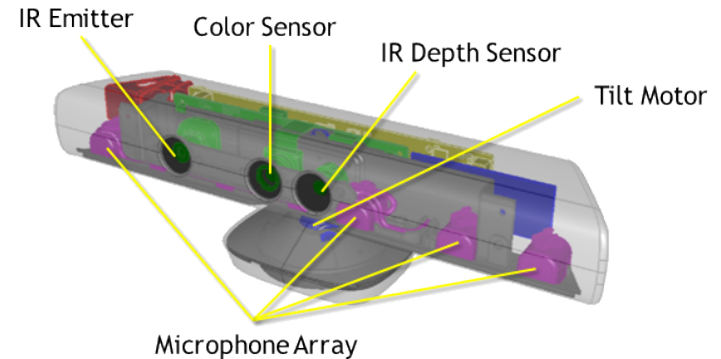
AMD Radeon HD 6870

Tools:

- **Autodesk 3DS MAX**

Bibliotheken:

- **Kinect for Windows SDK**
- **DirectX 11 SDK (jetzt Windows SDK)**
- **OpenCV**
- **Collada DOM und boost c++ Libraries**



GLIEDERUNG

- ▶ INTRO
- ▶ ANALYSE
- ▶ **LÖSUNG**
 - ▶ ROADMAP
 - ▶ **OFFLINE PHASE**
 - ▶ PSO TRACKING
 - ▶ ARCHITEKTUR
 - ▶ ONLINE MODELLE
- ▶ FAZIT & AUSBLICK



Offline Handmodellierung

Erstellung in 3DSMax:

Modellierungsvorlagen aus

1.ausgewählten Tiefenbildern

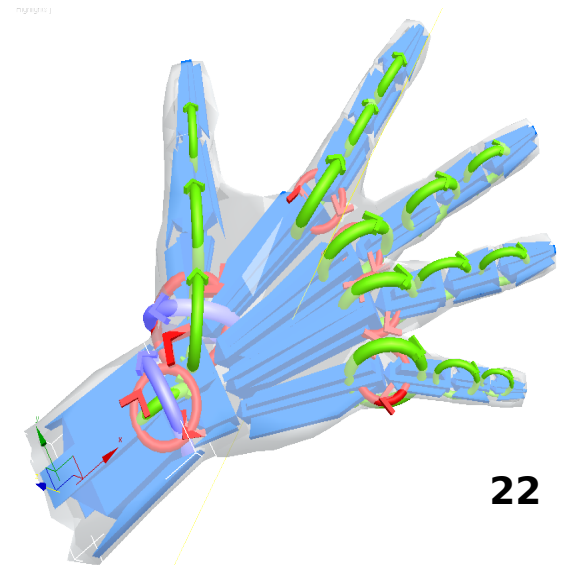
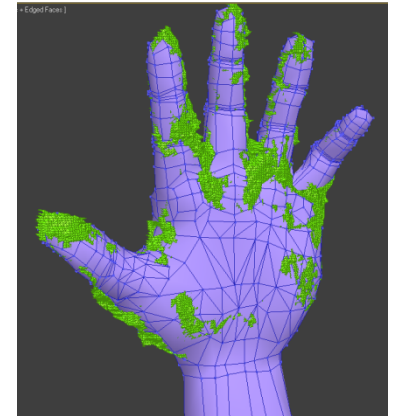
2.Strukturmodellierung

3.Handskelett = Bones System

- **Bone = Transformationsmatrix**
- *Skinning = Zuordnung*
- *Struktur-Vertices zu*

4.transformierbaren Bones

Ergebnis: Modell mit 27 DOFs



Statistisches Hautfarbenmodell

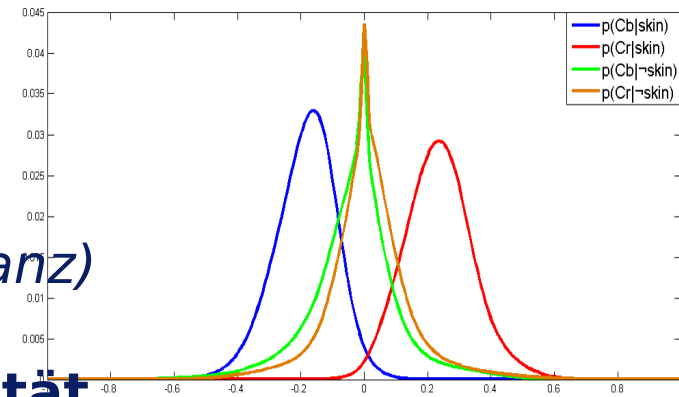
*Aufbauend auf Gaussian Mixture Models
von Jones und Rehg (1999)*

GMM approximiert bedingte

Wahrscheinlichkeitsverteilungen

Konvertierung RGB \rightarrow CbCr (ohne Luminanz)

- **Komprimierte Lookuptable**
- **Unabhängigkeit von Farbintensität**



GLIEDERUNG

- ▶ INTRO
- ▶ ANALYSE
- ▶ **LÖSUNG**
 - ▶ ROADMAP
 - ▶ OFFLINE PHASE
 - ▶ **PSO TRACKING**
 - ▶ ARCHITEKTUR
 - ▶ ONLINE MODELLE
- ▶ FAZIT & AUSBLICK



Particle Swarm Optimization

Algorithmus

Iterativer Optimierungsalgorithmus von Kennedy und Eberhart (1995)

simuliert Sozialverhalten schwarmorientierter Organismen

Ausgangslage: über den Parameterraum zufällig erzeugte

Partikelmenge

Ziel: Partikelpositionen ändern, bis Minimum gefunden ist oder Abbruch

Speicherung von Bestwerten:

- **Lokaler Bestwert = minimal verursachter Fehler eines Partikels**
- **Globaler Bestwert = minimaler Fehler aller Partikel**
- *Anpassung der Partikel*
- *bewegungsrichtung durch soziales Modell*

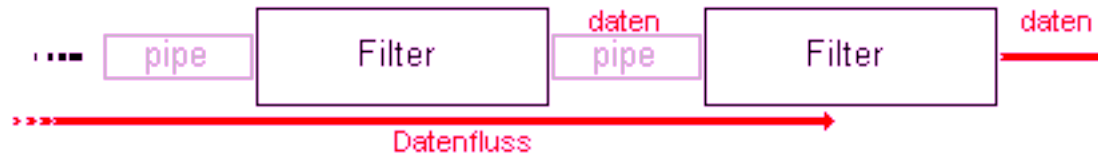


GLIEDERUNG

- ▶ INTRO
- ▶ ANALYSE
- ▶ **LÖSUNG**
 - ▶ ROADMAP
 - ▶ OFFLINE PHASE
 - ▶ PSO TRACKING
 - ▶ **ARCHITEKTUR**
 - ▶ ONLINE MODELLE
- ▶ FAZIT & AUSBLICK



Erkennungspipeline



Mögliches Architektur-Pattern ?!?

z.B. Pipes and Filters



GLIEDERUNG

- ▶ INTRO
- ▶ ANALYSE
- ▶ **LÖSUNG**
 - ▶ ROADMAP
 - ▶ OFFLINE PHASE
 - ▶ PSO TRACKING
 - ▶ ARCHITEKTUR
 - ▶ **ONLINE MODELLE**
- ▶ FAZIT & AUSBLICK



Segmentierung

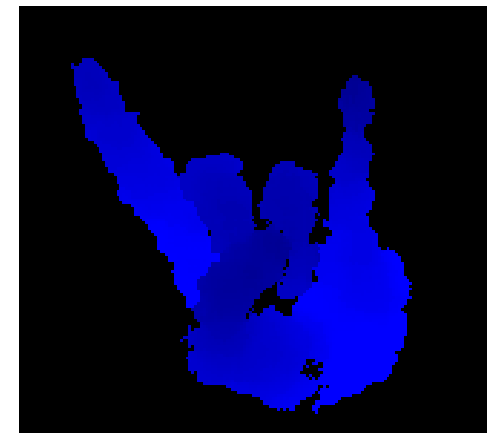
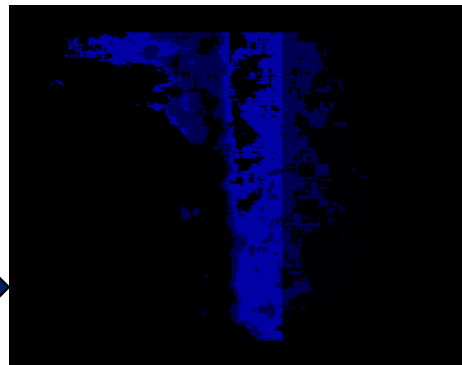
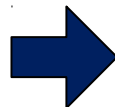
Problem: Hautfarbenfilterung liefert keine zusammenhängende Pixelregion

Lösung:

- 1. Suche eines besten Hautfarbenpixels**
- 2. Pixel ist Seed-Point für**
- 3. Tiefensegmentierung**
- 4. Tiefensegmentierung durch Flood-Fill**



keine Hand im Bild



Kinematikmodell und Rendering

- 1. Constraining der Modellparameter = statische Wertebereichsgrenzen*
- 2. Handmodell-Transformation in zwei Teilen*
- 3. CPU-basierte Anpassung der Bone-Transformation**
 - 1. durch Modellparameter**
 - 2. GPU-basiertes lineares Skinning im**
 - 3. Vertex Shader**
- 4. Implizite Projektion der transformierten homogenen Koordinaten*
- 4. ⇒ Bedarf Projektionsmatrix aus Camera Intrinsics*



Observationsmodell (1)

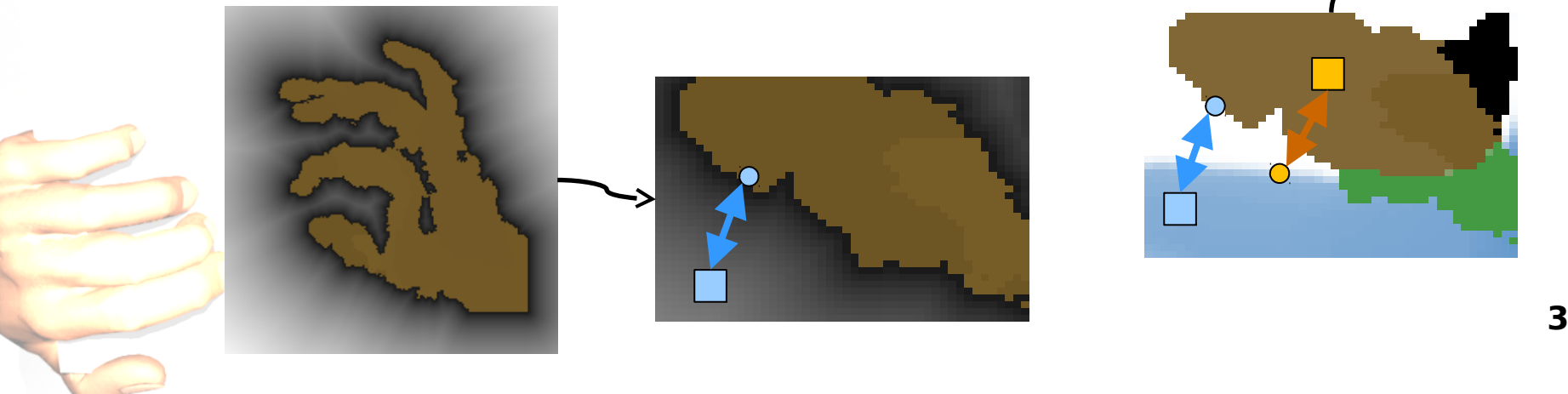
Ziel: GPU-basierte Fehlermessung durch direkten Pixelvergleich (Tiefenbild vs. Model-Rendering)

- **Per-Pixel-Fehler = | segmentierte Tiefenregion – Rendering |**

- *Problem: Was ist, wenn kein korrespondierender Pixel zur Messung vorhanden ist ?*

Lösung: 3D-Euklidische Distanz zum nächsten

- **Bild/Modell-Pixel [Hamer u.a. 2009]**
- **z.B. durch Distanztransformierte bestimmen**



Observationsmodell (2)

*Zusammenfassung der Per-Pixel-Fehler di
ähnlich wie Oikonomidis u.a. 2011*

1. Berechnung durchschnittlicher Per-Pixel-Fehler

- Nivellierung von Ausreißern durch Maximalwert δ

2. Normalisieren durch Koeffizient σ

3. Variabler Penalty-Term zur Gewichtung von

4. nicht übereinstimmenden Pixeln

5. Problem: Viele unabhängige parallele GPU-Threads

⇒ Wie den Fehler GPU-basiert zusammenfassen?

z.B. durch Multiskalenansatz (hier Quadtree)

- im Geometry Shader



GLIEDERUNG

- ▶ INTRO
- ▶ ANALYSE
- ▶ LÖSUNG
- ▶ FAZIT & AUSBLICK






Was wurde erreicht ?

1.Offline

- **Kamerakalibrierung** ✓
- **Modellierung**
 - **Strukturmodell = Polygon-Mesh** ✓
 - **Offline Kinematikmodell = Skeleton** ✓
 - **Statistisches Hautfarbenmodell** ✓

2.Online

- **Framework-Implementierung** ✓
 - **Hautfarben-Segmentierungsmodul** ✓
 - **Online Kinematikmodell** ✓
 - **Observationsmodell** ⇒ Tiefenbildvergleich mit GPU-Beschleunigung + Distanztransformation 
 - **Tracking = Particle Swarm Optimization** 
 - **Initialisierung durch einfache Startpose** 



Fazit

Modellbasiertes Tracking ist hilfreiches Werkzeug zur allgemeingültigen Posenerkennung

- **Laufzeit vs. Generalität**
- *Lösung mit direkter Hardware-Unterstützung ohne GPGPU-Toolkit-Overhead*

Direkte Verfahrensverbesserungen:

- **Initialisierung/Kombination mit Single View Pose Estimation**
- **Integrierte Segmentierung mit effektivem Hautfarben-Objekt-Tracking**
- **Komplexeres Constraint-Modell (z.B. erlernt)**
- **Tile-basierte Fehlerauswertung im Observationsmodell**



Ausblick

Hardware:

- **Schnellere Tiefensensoren mit höherer Auflösung**
- **Weitere General Purpose SIMD-Hardware (nicht nur GPGPU)**

Modellerweiterungen:

Dynamische Modellkalibrierung = Rekonstruktion

- **der Handstruktur aus Aufnahme (z.B. Kinect Fusion)**
- **Dynamisches Level of Detail**
- **Physik- und Kollisionsmodelle**

Erweiterte Perzeption:

- **2 oder N Hände im Eingabebild [Oikonomidis u.a. 2012]**
- **Interaktion mit Objekt**
- **[Hamer u.a. 2009; Oikonomidis u.a. 2011]**

Anwendungsspezifisch:

- **Mobile Geräte z.B. Google Glass 2.0**
- **Augmented- bzw. Mixed Reality**



Danke
!!!

Fragen
?!?



Bildquellen

[Folie 4] Spielberg, Steven (2002): Minority Report.

[Folie 4] AnyTouch (2013): Digitas Labs Paris. Online verfügbar unter <http://vimeo.com/43108191>.

[Folie 4] Simulations- und Softwaretechnik (Hg.) (2013): Interaction Techniques. DLR. Online verfügbar unter http://www.dlr.de/sc/Portaldata/15/Resources/images/abteilung/scsrv/vroos/vroos_direct_interaction_250.jpg.

[Folie 4] Karlsruhe Institute of Technology (2013): Grasping and Manipulation. Online verfügbar unter <http://www.sfb588.uni-karlsruhe.de/userfiles/image/grasping.jpg>.

[Folie 4] Kawasaki & Mouri Laboratory (Hg.) (2013): The Research of Force Feedback Glove. Gifu University. Online verfügbar unter <http://robo.mech.gifu-u.ac.jp/content/research/VR/FFG.jpg>.

[Folie 7] La Gorce u.a. (2011): Ambiguity. In: [La Gorce u.a 2011], S. 1794.

[Folie 8] Erol u.a. (2007): Hand Pose Estimation Taxonomie. In: [Erol u.a. 2007], S. 54.

[Folie 10] Zhigeng Pan u.a. (2010): Skin Color Segmentation; Fingertip Filter. In: [Zhigeng Pan u.a 2010], S. 220.

[Folie 10] Gang Hu; Gao, Q. (2011): Perceptual feature keypoints; CPP-GET. In: [Gang Hu u.a. 2011], S. 206.

[Folie 10,15] Ben Henia, Ouissem; Hariti, Mohamed; Bouakaz, Saida (2010): Distance Transform. In: [Ben Henia u.a. 2010], S. 4.



Bildquellen

- [Folie 10] Farfield Technology (Hg.) (2013): Pointcloud. Online verfügbar unter http://www.farfieldtechnology.com/products/toolbox/pointcloud/hand_dots_up.jpg.
- [Folie 12] Tofighi, G.; Monadjemi, S.A; Ghasem-Aghaee, N. (2010): convexity defects. In: [Tofighi u.a. 2010], S. 3.
- [Folie 12] Hackenberg, G.; McCall, R.; Broll, Wolfgang (2011): Finger Extraction; Application. In: [Hackenberg u.a. 2011], S. 22, 25.
- [Folie 12] Alon, Jonathan; Athitsos, V.; Quan Yuan; Sclaroff, S. (2009): Number Gestures. In: [Alon u.a. 2009], S. 1686.
- [Folie 12] Chan Wah Ng; Surendra Ranganath (2002): Pose Set. In: [Chan Wah Ng u.a. 2002], S. 994.
- [Folie 13] Athitsos, Vassilis; Sclaroff, Stan (2003): Edge Database. In: [Athitsos u.a. 2003], S. 2.
- [Folie 13] Keskin, C.; Kirac, F.; Kara, Y.E; Akarun, L. (2011): Hand Labeling. In: [Keskin u.a. 2011], S. 1231.
- [Folie 15,16] Stenger, Björn Dietmar Rafael (2004): Quadrics Model; Error Function. In: [Stenger 2004], S. 33, 58, 85.
- [Folie 17] angermayer.de (Hg.) (2013) : Gradientenverfahren. Online verfügbar unter <http://www.angermayer.de/padochp/pics/papic/Image104.gif>
- [Folie 17] Brown, J.A; Capson, D.W (2012): Particle Filter Convergence. In: [Brown u.a. 2012], S. 71.

Bildquellen

- [Folie 17] Universität Stuttgart (Hg.) (2013): PSO Quickoverview. Institut für Technische und Numerische Mechanik. Online verfügbar unter http://www.itm.uni-stuttgart.de/research/pso_opt/bilder/pso.gif.
- [Folie 20] Microsoft (Hg.) (2013): Kinect Sensor Specification. Online verfügbar unter <http://i.msdn.microsoft.com/dynimg/IC584396.png>.
- [Folie 26] Wagner, Stefan (2010): PSO Soziales Modell. In: [Wagner 2010].
- [Folie 28] de.wikipedia.org (Hg.) (2013): Pipes and Filters. Online verfügbar unter <http://upload.wikimedia.org/wikipedia/de/3/33/PipesFilters.png>.
- [Folie 38] Oikonomidis, I.; Kyriazis, N.; Argyros, A.A (2012): Two Hand Perception. In: [Oikonomidis u.a. 2012], S. 1862.
- [Folie 38] Hamer, H.; Schindler, K.; Koller-Meier, E.; van Gool, L. (2009): Hand-Object Perception. In: [Hamer u.a. 2009], S. 1481.
- [Folie 38] Kastrenakes, Jacob (2013): Google Glass. Hg. v. theverge.com. Online verfügbar unter <http://www.theverge.com/2013/8/12/4614966/google-glass-update-evernote-path-voice-commands-video-controls>.



Literatur

- [Alon u.a. 2009] Alon, Jonathan; Athitsos, V.; Quan Yuan; Sclaroff, S. (2009): A Unified Framework for Gesture Recognition and Spatiotemporal Gesture Segmentation. In: *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 31 (9), S. 1685-1699. DOI: 10.1109/TPAMI.2008.203.
- [Albrecht u.a. 2003] Albrecht, Irene; Haber, Jörg; Seidel, Hans-Peter (2003): Construction and animation of anatomically based human hand models. In: Proceedings of the 2003 ACM SIGGRAPH/Eurographics symposium on Computer animation. Aire-la-Ville, Switzerland, Switzerland: Eurographics Association (SCA '03), S. 98-109. Online verfügbar unter <http://dl.acm.org/citation.cfm?id=846276.846290>.
- [Argyros u.a. 2004] Argyros, Antonis A.; Lourakis, Manolis I. A. (2004): Real-Time Tracking of Multiple Skin-Colored Objects with a Possibly Moving Camera. In: ECCV, S. 368-379.
- [Athitsos u.a. 2003] Athitsos, Vassilis; Sclaroff, Stan (2003): Estimating 3D Hand Pose from a Cluttered Image. In: *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on* 2, S. 432-432-9 vol.2. DOI: 10.1109/cvpr.2003.1211500.
- [Autodesk 2013] Autodesk (2013): 3DS MAX. Online verfügbar unter <http://www.autodesk.de/products/autodesk-3ds-max/overview>, zuletzt geprüft am 16.09.2013.
- [Barnes u.a. 2008] Barnes, Mark; Finch, Ellen Levy (2008): COLLADA 1.5.0 Specification. Online verfügbar unter http://www.khronos.org/files/collada_spec_1_5.pdf, zuletzt geprüft am 26.09.2013.
- [Ben Henia u.a. 2010] Ben Henia, Ouissem; Hariti, Mohamed; Bouakaz, Saida (2010): A two-step minimization algorithm for model-based hand tracking. In: 18th International Conference on Computer Graphics, Visualization and Computer Vision (WSCG). Online verfügbar unter <http://liris.cnrs.fr/publis/?id=4574>.
- [Bender u.a. 2003] Bender, Michael; Brill, Manfred (2003): Computergrafik. Ein anwendungsorientiertes Lehrbuch ; [Website mit Quellcodes, Lösungen zu den Aufgaben, Einführungen in die Tools und mehr]. München, Wien: Hanser.

Literatur

- [boost.org 2013] boost.org (Hg.) (2013): Boost C++ Libraries. Online verfügbar unter <http://www.boost.org/>, zuletzt geprüft am 16.09.2013.
- [Bradski u.a. 2008] Bradski, Gary R.; Kaehler, Adrian (2008): Learning OpenCV. Computer vision with the OpenCV library. 1. Aufl. Sebastopol, CA: O'Reilly.
- [Bray u.a. 2004] Bray, M.; Koller-Meier, E.; Muller, P.; van Gool, L.; Schraudolph, N.N (2004): 3D hand tracking by rapid stochastic gradient descent using a skinning model. In: Visual Media Production, 2004. (CVMP). 1st European Conference on, S. 59-68.
- [Breuer 2005] Breuer, Pia (2005): Entwicklung einer prototypischen Gestenerkennung in Echtzeit unter Verwendung einer IR Tiefenkamera. Diplomarbeit. Universität Koblenz Landau, Koblenz, GE. Institut für Computervisualistik.
- [Brown u.a. 2012] Brown, J.A; Capson, D.W (2012): A Framework for 3D Model-Based Visual Tracking Using a GPU-Accelerated Particle Filter. In: *Visualization and Computer Graphics, IEEE Transactions on*, title=A Framework for 3D Model-Based Visual Tracking Using a GPU-Accelerated Particle Filter 18 (1), S. 68-80. DOI: 10.1109/TVCG.2011.34.
- [Causo u.a. 2008] Causo, A.; Ueda, E.; Kurita, Y.; Matsumoto, Y.; Ogasawara, T. (2008): Model-based hand pose estimation using multiple viewpoint silhouette images and Unscented Kalman Filter. In: Robot and Human Interactive Communication, 2008. RO-MAN 2008. The 17th IEEE International Symposium on, S. 291-296.
- [Chan Wah Ng u.a. 2002] Chan Wah Ng; Surendra Ranganath (2002): Real-time gesture recognition system and application. In: *Image and Vision Computing* 20 (13-14), S. 993-1007. DOI: 10.1016/S0262-8856(02)00113-0.



Literatur

- [Clerc u.a. 2002] Clerc, M.; Kennedy, J. (2002): The particle swarm - explosion, stability, and convergence in a multidimensional complex space. In: *Evolutionary Computation, IEEE Transactions on* 6 (1), S. 58-73. DOI: 10.1109/4235.985692.
- [collada-dom.sourceforge.net 2013] collada-dom.sourceforge.net (Hg.) (2013): COLLADA Document Object Model. Online verfügbar unter <http://collada-dom.sourceforge.net/>, zuletzt geprüft am 16.09.2013.
- [Eberhart u.a. 1995] Eberhart, R.; Kennedy, J. (1995): A new optimizer using particle swarm theory. In: *Micro Machine and Human Science, 1995. MHS '95., Proceedings of the Sixth International Symposium on*, S. 39-43.
- [Erol u.a. 2007] Erol, Ali; Bebis, George; Nicolescu, Mircea; Boyle, Richard D.; Twombly, Xander (2007): Vision-based hand pose estimation: A review. In: *Comput. Vis. Image Underst.* 108 (1-2), S. 52-73. DOI: 10.1016/j.cviu.2006.10.012.
- [Gang Hu u.a. 2011] Gang Hu; Gao, Q. (2011): Gesture Analysis Using 3D Camera, Shape Features and Particle Filters. In: *Computer and Robot Vision (CRV), 2011 Canadian Conference on*, S. 204-211.
- [Hackenberg u.a. 2011] Hackenberg, G.; McCall, R.; Broll, Wolfgang (2011): Lightweight palm and finger tracking for real-time 3D gesture control. In: *Virtual Reality Conference (VR), 2011 IEEE*, S. 19-26.
- [Hamer u.a. 2009] Hamer, H.; Schindler, K.; Koller-Meier, E.; van Gool, L. (2009): Tracking a hand manipulating an object. In: *Computer Vision, 2009 IEEE 12th International Conference on*, S. 1475-1482.
- [Hui Liang u.a. 2012] Hui Liang; Junsong Yuan; Thalmann, Daniel (2012): 3D fingertip and palm tracking in depth image sequences. In: *Proceedings of the 20th ACM international conference on Multimedia*. New York and NY and USA: ACM (MM '12), S. 785-788. Online unter <http://doi.acm.org/10.1145/2393347.2396312>.



Literatur

- [Jones u.a. 1999] Jones, M.J; Rehg, J.M (1999): Statistical color models with application to skin detection. In: Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on, Bd. 1, S. -280 Vol. 1.
- [Kennedy u.a. 1995] Kennedy, J.; Eberhart, R. (1995): Particle swarm optimization. In: Neural Networks, 1995. Proceedings., IEEE International Conference on, Bd. 4, S. 1942-1948 vol.4.
- [Keskin u.a. 2011] Keskin, C.; Kirac, F.; Kara, Y.E; Akarun, L. (2011): Real time hand pose estimation using depth sensors. In: Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on, S. 1228-1234.
- [La Gorce u.a. 2011] La Gorce, M. de; Fleet, D.J; Paragios, N. (2011): Model-Based 3D Hand Pose Estimation from Monocular Video. In: *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, title=Model-Based 3D Hand Pose Estimation from Monocular Video 33 (9), S. 1793-1805. DOI: 10.1109/TPAMI.2011.33.
- [Lepetit u.a. 2005] Lepetit, V.; Lagger, P.; Fua, P. (2005): Randomized trees for real-time keypoint recognition. In: Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, Bd. 2, S. 775-781 vol. 2.
- [Microsoft 2013] Microsoft (2013): DirectX Graphics and Gaming. Hg. v. Microsoft Developer Network (MSDN). Online verfügbar unter <http://msdn.microsoft.com/en-us/library/ee663274%28v=vs.85%29.aspx>, zuletzt geprüft am 16.09.2013.
- [Microsoft 2013] Microsoft (2013): Kinect for Windows. Hg. v. Microsoft Corporation. Online verfügbar unter <http://www.microsoft.com/en-us/kinectforwindows/>, zuletzt geprüft am 15.09.2013.
- [Microsoft 2013] Microsoft (Hg.) (2013): Kinect for Windows SDK. Online verfügbar unter <http://msdn.microsoft.com/en-us/library/hh855347.aspx>, zuletzt geprüft am 17.09.2013.

Literatur

- [Oikonomidis u.a. 2011] Oikonomidis, I.; Kyriazis, N.; Argyros, A.A (2011): Full DOF tracking of a hand interacting with an object by modeling occlusions and physical constraints. In: Computer Vision (ICCV), 2011 IEEE International Conference on, S. 2088-2095.
- [Oikonomidis u.a. 2012] Oikonomidis, I.; Kyriazis, N.; Argyros, A.A (2012): Tracking the articulated motion of two strongly interacting hands. In: Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on, S. 1862-1869.
- [Oikonomidis u.a. 2011] Oikonomidis, Iason; Kyriazis, Nikolaos; Argyros, Antonis A. (2011): Efficient model-based 3D tracking of hand articulations using Kinect. In: Proceedings of the British Machine Vision Conference: BMVA Press, S. 101.1-101.11.
- [opencv.org 2013] opencv.org (Hg.) (2013): OpenCV. Online verfügbar unter <http://opencv.org/>, zuletzt geprüft am 16.09.2013.
- [Petersen 2013] Petersen, Iwer (2013): Using object tracking for dynamic video projection mapping. Bachelorthesis. Hochschule für Angewandte Wissenschaften, Hamburg. Fakultät Technik und Informatik.
- [Romero u.a. 2009] Romero, J.; Kjellstrom, H.; Kragic, D. (2009): Monocular real-time 3D articulated hand pose estimation. In: Humanoid Robots, 2009. Humanoids 2009. 9th IEEE-RAS International Conference on, S. 87-92.
- [Rosales u.a. 2001] Rosales, Romer; Athitsos, Vassilis; Sigal, Leonid; Sclaroff, Stan (2001): 3D Hand Pose Reconstruction Using Specialized Mappings. In: In ICCV, S. 378-385.
- [Rosenfeld u.a. 1966] Rosenfeld, Azriel; Pfaltz, John L. (1966): Sequential Operations in Digital Picture Processing. In: *J. ACM* 13 (4), S. 471-494. DOI: 10.1145/321356.321357.



Literatur

- [Sharp 2008] Sharp, Toby (2008): Implementing Decision Trees and Forests on a GPU. In: David Forsyth, Philip Torr und Andrew Zisserman (Hg.): Computer Vision - ECCV 2008, Bd. 5305: Springer Berlin Heidelberg (Lecture Notes in Computer Science), S. 595-608. Online verfügbar unter http://dx.doi.org/10.1007/978-3-540-88693-8_44.
- [Shimada u.a. 2001] Shimada, N.; Kimura, K.; Shirai, Y. (2001): Real-time 3D hand posture estimation based on 2D appearance retrieval using monocular camera. In: Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems, 2001. Proceedings. IEEE ICCV Workshop on, S. 23-30.
- [Shotton u.a. 2011] Shotton, J.; Fitzgibbon, A.; Cook, M.; Sharp, T.; Finocchio, M.; Moore, R. et al. (2011): Real-time human pose recognition in parts from single depth images. In: Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition. Washington, DC, USA: IEEE Computer Society (CVPR '11), S. 1297-1304. Online verfügbar unter <http://dx.doi.org/10.1109/CVPR.2011.5995316>.
- [Stenger 2004] Stenger, Björn Dietmar Rafael (2004): Model-Based Hand Tracking Using A Hierarchical Bayesian Filter. Dissertation. University of Cambridge, Cambridge, UK.
- [Tofighi u.a. 2010] Tofighi, G.; Monadjemi, S.A; Ghasem-Aghaee, N. (2010): Rapid hand posture recognition using Adaptive Histogram Template of Skin and hand edge contour. In: Machine Vision and Image Processing (MVIP), 2010 6th Iranian, S. 1-5.
- [Viola u.a. 2001] Viola, P.; Jones, M. (2001): Rapid object detection using a boosted cascade of simple features. In: Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on, Bd. 1, S. I-511-I-518 vol.1.



Literatur

- [Wagner 2010]** Wagner, Stefan (2010): Partikelschwarmoptimierung. Technische Universität Chemnitz,. Online verfügbar unter <http://www-user.tu-chemnitz.de/~wags/about/particle-swarm-optimization.pdf>, zuletzt geprüft am 17.10.2013.
- [Yuan Yao u.a. 2012]** Yuan Yao; Yun Fu (2012): Real-Time Hand Pose Estimation from RGB-D Sensor. In: Multimedia and Expo (ICME), 2012 IEEE International Conference on, S. 705-710.
- [Zhigeng Pan u.a. 2010]** Zhigeng Pan; Yang Li; Mingmin Zhang; Chao Sun; Kangde Guo; Xing Tang; Zhou, S.Z (2010): A real-time multi-cue hand tracking algorithm based on computer vision. In: Virtual Reality Conference (VR), 2010 IEEE, S. 219-222.

