

Erkennung dynamischer Gesten im Living Place Hamburg

Arne Bernin
Wintersemester 2009/2010
20.01.2010

Inhalt

- ▶ Einleitung
- ▶ Vision
- ▶ Vorarbeiten
- ▶ Prototyp
- ▶ Herausforderungen
- ▶ Nächste Schritte

Einleitung

- ▶ Stellen wir uns vor:
 - ▶ Ich sitze vor dem Fernseher und würde gern den Kanal wechseln...
 - ▶ Ich sitze auf dem Sofa, höre entspannt Musik und möchte die Lautstärke ändern...
 - ▶ Ich sitze auf der Toilette und betrachte den Comic auf dem Bildschirm an der Wand. Es wird Zeit, die Seite umzublättern...

... Die Frage ist nur ...

Einleitung



Wo ist nun wieder die Fernbedienung ?



Einleitung

- ▶ Fragen über Fragen:
 - ▶ Wo ist die Fernbedienung ?
 - ▶ Wo ist die passende Fernbedienung ?
 - ▶ Warum sind die Batterien schon wieder leer ?

- ▶ Die passende Antwort(en):
 - ▶ Steuerung von Multimediageräten mit Hilfe einfacher dynamischer Gesten

Vision

Vision

- ▶ Erkennung einfacher dynamischer Gesten
- ▶ Zur Steuerung von Geräten über den iROS Eventheap
- ▶ Ohne Hardware/Marker am Körper
- ▶ Unabhängig von der Beleuchtung
- ▶ Günstige und verfügbare Hardware
- ▶ Einfaches, effizientes Verfahren (embedded geeignet)
- ▶ Opensource Software
- ▶ Unter Beachtung des Schutzes der Privatsphäre

Vision



[Quelle: Youtube]

Vorarbeiten

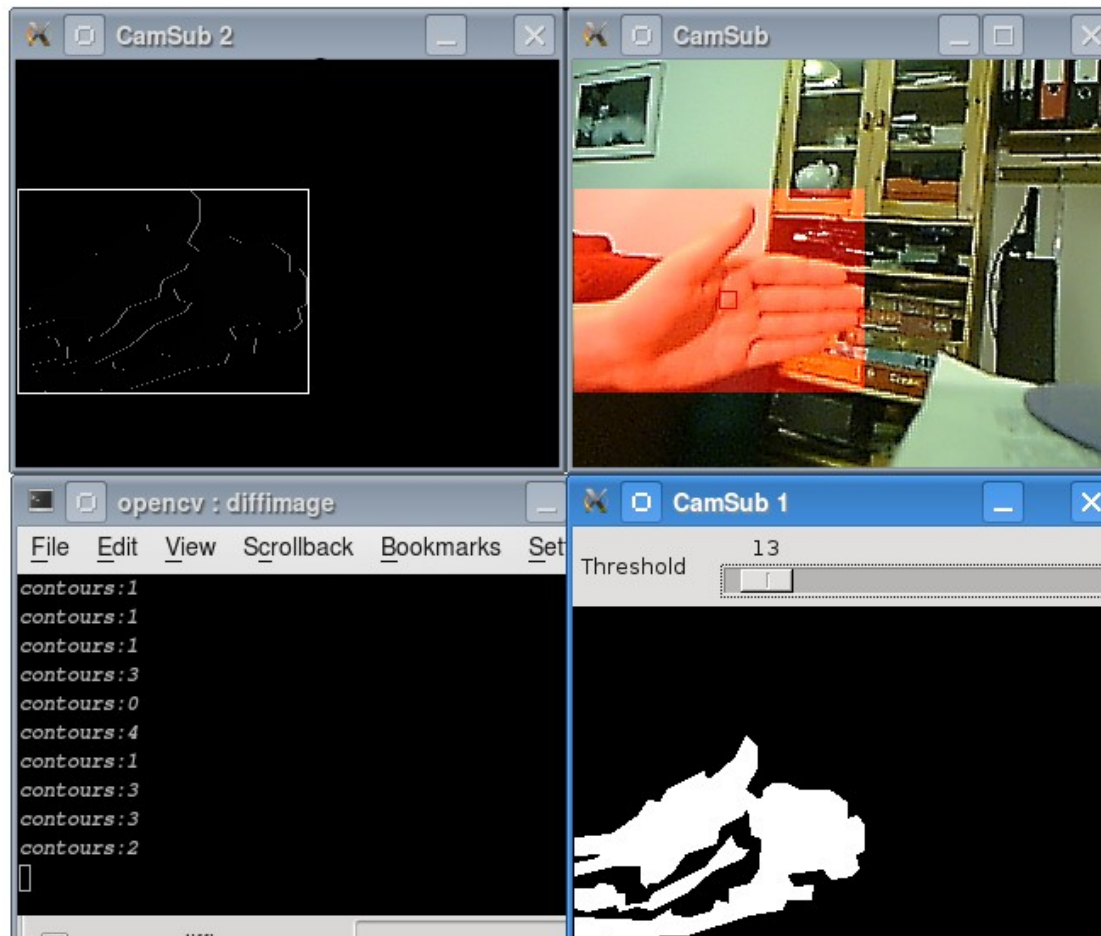
Vorarbeiten

- ▶ Bewegungserkennung als Grundlage für dynamische Gesten
- ▶ Bewegungserkennung mit Differenzbildern
- ▶ Rein optisches System mit USB-Kameras
- ▶ Basierend auf der OpenCV Bibliothek
- ▶ Einzelheiten im Bericht der Masterprojektes „Räumliche Segmentierung mit Differenzbildern“ vom Sommer 2009



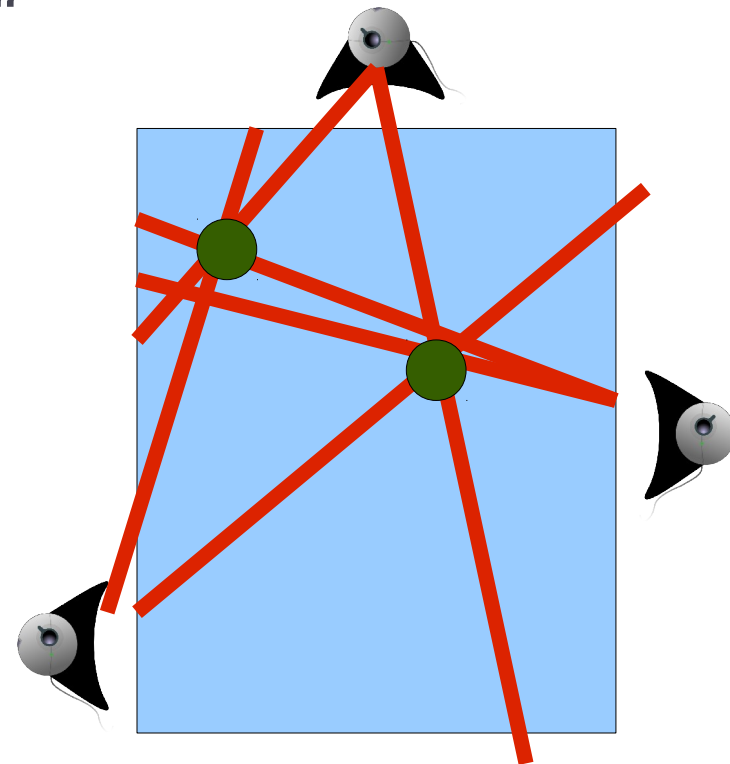
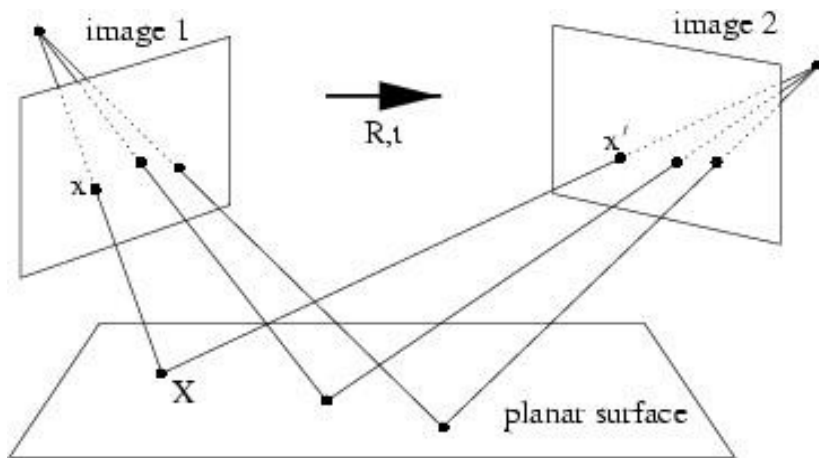
<http://users.informatik.haw-hamburg.de/~ubicomp/projekte/master2009-proj/bernin.pdf>

Vorarbeiten



Vorarbeiten

Diffcam: Position im „Raum“



[Quelle: Hartley2000]

Vorarbeiten: Fazit

- ▶ Differenzbilder-Verfahren grundsätzlich geeignet
 - ▶ Passt sich gut wechselnden Lichtverhältnissen an
 - ▶ Erkennt Bewegungen zuverlässig
 - ▶ Besonders gut bei bekannter Konfiguration
 - ▶ Schnell und günstig in Bezug auf Rechenzeit
 - ▶ Es wird schwierig bei
 - ▶ Mehreren Personen
 - ▶ Reflektionen/Spiegelungen
 - ▶ Zu großer Entfernung /geringer Auflösung
 - ▶ **Das Verfahren alleine reicht nicht aus !**
-

Vorarbeiten: Lösungsansatz

- ▶ Mehr Informations-Kanäle benutzen
 - ▶ erweiterte Bildverarbeitung
 - ▶ Zusätzliche Sensoren
- ▶ Mögliche Kandidaten sind
 - ▶ Hautfarbe
 - ▶ 2D-Position
 - ▶ 3D-Position
 - ▶ Tiefeninformation durch spezielle Hardware
 - ▶ Modell des Körpers / der möglichen Bewegungen

Prototyp

Prototyp

- ▶ Erster Prototyp :
 - ▶ Mehr als einer nötig
 - ▶ Aufbau der Infrastruktur
 - ▶ Verbindung zum Eventheap
 - ▶ Schaffen von Vergleichsmöglichkeiten
 - ▶ Verwirklichung von Kai`s Traum ;-)

Prototyp: Verfahren

- ▶ Erkennen von Bewegung durch Differenzbilder
- ▶ Erkennen von Hautpartien (Transformation in HSV)
- ▶ Erstellen von Masken aus beiden Verfahren
- ▶ Maskenabgleich
- ▶ Bestimmen des Bewegungsmittelpunktes
- ▶ Erkennen von Gesten durch Beobachtung des Mittelpunktes
- ▶ Wie wird das System aktiviert ?

Prototyp: Kommandomodus

- ▶ Zur Differenzierung von Kommandos und Gesten im Rahmen der zwischenmenschlichen Kommunikation
- ▶ Einleiten des Kommandomodus
- ▶ Möglichkeiten:
 - ▶ Dynamische Geste zur Einleitung
 - ▶ Statische Geste
 - ▶ Aktivierung per Sprache
- ▶ Dynamische Geste bietet sich an

Prototyp: Gesten

- ▶ Welche Gesten sollen erkannt werden ?
- ▶ Beispiele:
 - ▶ Hoch ↔ runter
 - ▶ Lautstärke
 - ▶ Links ↔ rechts
 - ▶ Fernseh-Kanal, Musiktitel, Folie
 - ▶ Winken
 - ▶ Aufmerksamkeit
 - ▶ Kreis
 - ▶ Noch offen für Ideen ;-)

Prototyp: Technik

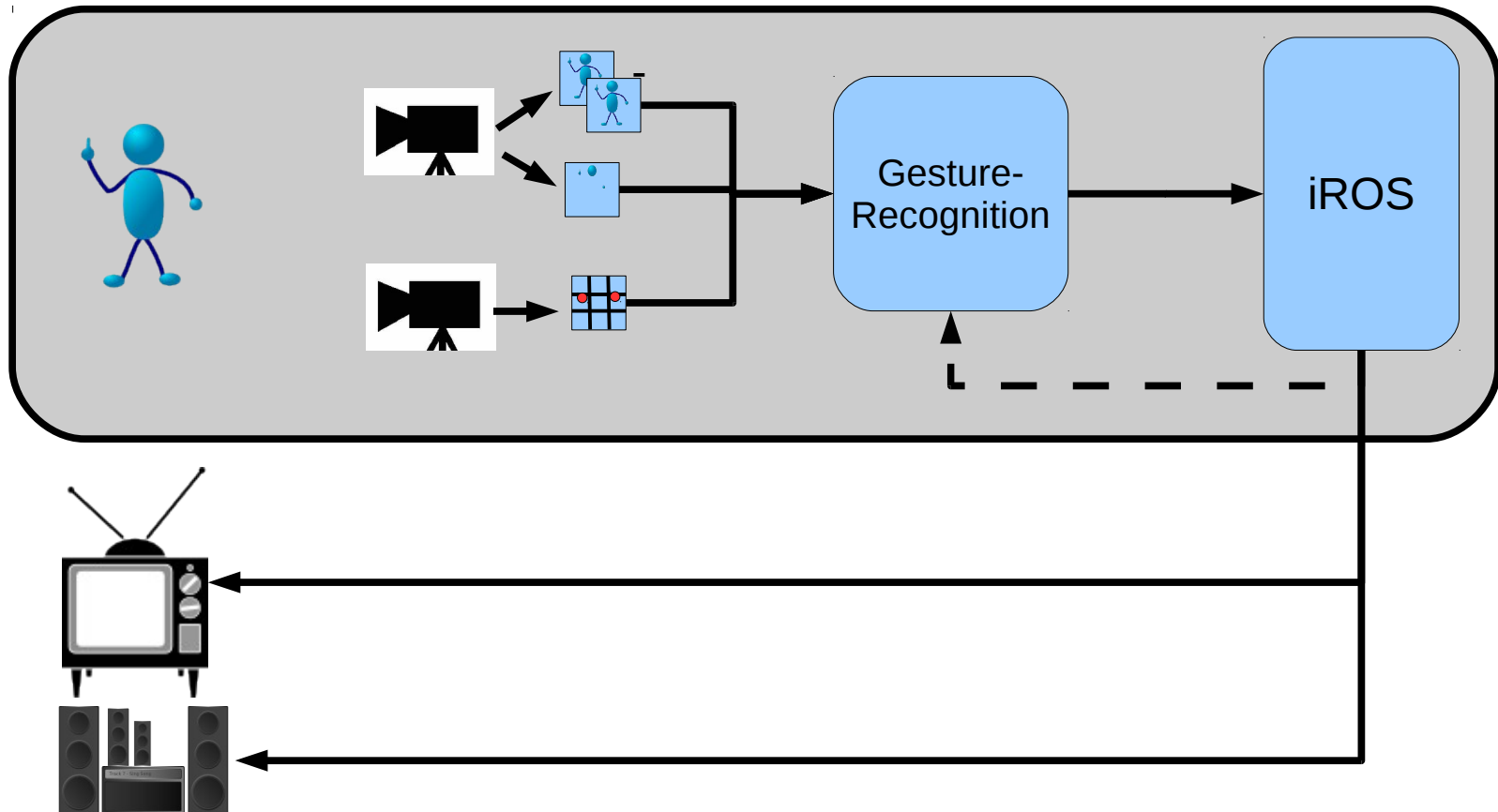
Quelle: [mobotix]



- ▶ Kamera Mobotix Q24
 - ▶ 3 Megapixel
 - ▶ 30 Bilder pro Sekunde
 - ▶ Fisheye Objektiv (Hemispherik)
 - ▶ Videostrom über Netz
 - ▶ Anbindung an OpenCV möglich
 - ▶ Embedded Linux



Prototyp: Aufbau



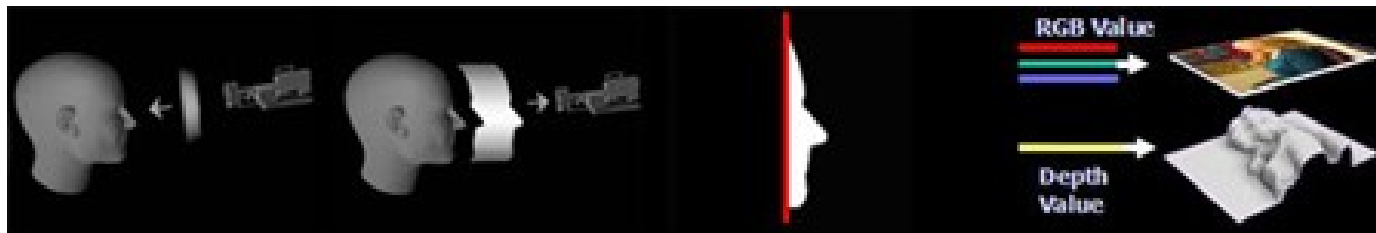
Evaluierung

Evaluierung

- ▶ Evaluierung von zusätzlicher Hard-/Software für den Prototyp 2
- ▶ Fragestellungen
 - ▶ Verfügbarkeit
 - ▶ Verfügbarkeit von Sourcecode
 - ▶ Lizenz
 - ▶ Ansteuerung / Schnittstellen
 - ▶ Maximaler Sichtbereich
 - ▶ Benötigte Lichtbedingungen und Voraussetzungen

Evaluierung

- ▶ Time-of-Flight Kameras
 - ▶ Relativ neue Entwicklung (2003 ESPROS)
 - ▶ Infrarotpuls und Laufzeitmessung
 - ▶ Hochgeschwindigkeits-shutter
 - ▶ Teilweise kombiniert mit Farbkamera
 - ▶ Nutzbarkeit bei Tageslicht unklar



[Quelle:3dvsystems]

Evaluierung: Hardware

- ▶ **Projekt Natal von Microsoft**
 - ▶ preislich günstig
 - ▶ Derzeit leider nicht kommerziell verfügbar
- ▶ **[Vision]S3 von PMD Technologies**
 - ▶ Preislich nicht günstig
 - ▶ Verfügbar
- ▶ **SR4000 von Mesa Imaging**
 - ▶ Auch nicht wirklich günstig
 - ▶ Verfügbar



Quellen: [natal],[pmdsystems],[sr4000]

Evaluierung: Software

- ▶ **Gesturetek**
 - ▶ Webcam basiert
 - ▶ Middleware
 - ▶ Kommerziell
 - ▶ Seit Jahren im Einsatz (xbox, eye toy, hitachi)
 - ▶ Schnittstellen unklar

Herausforderungen

Herausforderungen

- ▶ Problemklassifizierung bei der Gestenerkennung
 - ▶ Nicht Erkennen von Gesten
 - ▶ Korrektes Erkennen der Geste aber zur falschen Zeit
 - ▶ Falsch-positive Erkennung von Gesten

Herausforderungen

- ▶ Nicht Erkennen von Gesten
 - ▶ Beleuchtungsprobleme
 - ▶ Außerhalb des Sichtbereiches
 - ▶ Verdeckung
 - ▶ Fehlinterpretation der Bewegung

Herausforderungen

- ▶ Korrektes Erkennen der Geste aber zur falschen Zeit
 - ▶ Gestik in normaler zwischenmenschlicher Kommunikation
 - ▶ Falsche Person (bei mehreren im Raum)
 - ▶ Lösung: Kommandomodus

Herausforderungen

- ▶ Falsch-positive Erkennung von Gesten

- ▶ Falsch erkannte Bewegungen durch
 - ▶ Haustiere
 - ▶ Fenster
 - ▶ Reflektionen
 - ▶ Bildschirme
 - ▶ Helligkeitswechsel

Herausforderungen: Datenschutz

My Home is my castle
or
Big brother ist watching me ?

Herausforderungen: Datenschutz

- ▶ Es geht um Kameras in einem sensiblen Bereich
- ▶ Warum ist das ein Thema ?
 - ▶ Akzeptanz/Vertrauen
 - ▶ Begehrlichkeiten
 - ▶ Peinlichkeiten
 - ▶ Einblick in intimste Bereiche
 - ▶ Verhaltensänderungen
- ▶ Datenschutz sollte beim Design berücksichtigt werden!

Herausforderungen: Datenschutz

- ▶ Lösungsansätze
 - ▶ Datenvermeidung
 - ▶ Geschlossenes System
 - ▶ Intimsphäre beachten
 - ▶ Kontrolle durch Benutzer
 - ▶ Transparenz

Weitere Schritte

- ▶ Entwicklung eines Prototyps mit einer Kamera
- ▶ Evaluierung von TOF-Kameras
- ▶ Entscheidung treffen, ob zusätzliche Hardware verwendet werden soll
- ▶ Aufbau eines erweiterten Prototyps
- ▶ Test des Systems
- ▶ Zusammenfassen der Ergebnisse

.... und fertig ist die Masterarbeit ;-)

Risiken

- ▶ Das System funktioniert nicht
- ▶ ... oder nicht wie gewünscht
- ▶ Jemand anders ist schneller
- ▶ Patente ?

Fragen?

Arne Bernin
Wintersemester 2009/10
20.1. 2010

Literatur

- ▶ **[3dvsystems 2009]** 3Dvsystems. Verifiziert am 18.1.2010.
<http://www.3dvsystems.com>
 - ▶ **[artrack 2009]** Advanced Realtime Tracking GmbH. Verifiziert am 16.1.2010. <http://ar-tracking.eu>
 - [Bianchi-Berthouze u. a. 2007]** B IANCHI -B ERTHOuze, Nadia ; K IM, Whan W. ; PATEL, Darshak: Does Body Movement Engage You More in Digital Game Play? and Why? In: ACII '07: Proceedings of the 2nd international conference on Affective Computing and Intelligent Interaction. Berlin, Heidelberg : Springer-Verlag, 2007. – ISBN 978– 3–540–74888–5, S. 102–113
 - ▶ **[Natal 2009]** Projekt Natal. Verifiziert am 18.6.2009. <http://www.xbox.com/en-US/live/projectnatal/>.
Version: 2009 **[OpenCV 2009]** Welcome - OpenCV Wiki. Verifiziert am 18.6.2009.
<http://opencv.willowgarage.com/wiki/>
 - ▶ **[pmds3]** PMD Technologies S3 Vision Kamera Verifiziert am 16.1.2010
<http://www.pmdtec.com/products-services/pmdvisionr-cameras/pmdvisionr-s3/>
 - ▶ **[sr4000]** Mesa Imaging, Swiss-Ranger 4000 Verifiziert am 16.1.2010
<http://mesa.imaging.ch>
 - ▶ **[ab2009]** „Räumliche Segmentierung mit Differenzbildern“ , **verifiziert am 16.1.2010**
<http://users.informatik.haw-hamburg.de/~ubicomp/projekte/master2009-proj/bernin.pdf>
-

Literatur

- ▶ **[Corporation 2009]** CORPORATION, Surveyor: Surveyor Stereo Vision System. Verifiziert am 18.6.2009. <http://surveyor-corporation.stores.yahoo.net/srblstca.html>
- ▶ **[E3 2009]** Project Natal: Microsoft und die Revolution im Wohnzimmer. Verifiziert am 16.6.2009. <http://www.golem.de/0906/67480.html>. Version: 2009
- ▶ **[Farrugia u. a. 2006]** FARRUGIA, J.-P. ; H ORAIN, P. ; G UEHENNEUX, E. ; A LUSSE, Y.:GPU CV: A Framework for Image Processing Acceleration with Graphics Processors. In: Multimedia and Expo, 2006 IEEE International Conference on, 2006, S. 585–588
- ▶ **[GpuCV 2009]** GpuCV: GPU-accelerated Computer Vision. Verifiziert am 18.6.2009. <https://picoforge.int-evry.fr/cgi-bin/twiki/view/Gpucv/Web/>
- ▶ **[C. Niclass, E. Charbon]** A Single Photon Detector Array with 64x64 Resolution and Millimetric Depth Accuracy for 3D Imaging. In: IEEE International Solid-State Circuits Conference (ISSCC). 2005, S. 364-366.
- ▶ **[Andrew Wilson]:** Video range camera provides 3-D data. In: Technology Trends - Image Processing (Hrsg.): Vision Systems Design. Nr. 11, 2006, ISSN 1089-3709, S. 19-20.

Literatur

- ▶ **[Garbe u. a. 2003]** GARBE, C. S. ; SPIES, H. ; JAEHNE, B.: Estimation of complex motion from thermographic image sequences. In: RAMER, K. E. (Hrsg.) ; MALDAGUE, X. P. (Hrsg.): Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Bd. 5073, 2003 (Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference), S. 303–317
- ▶ **[Hartley u. Zisserman 2000]** HARTLEY, R. I. ; ZISSERMAN, A.: Multiple View Geometry in Computer Vision. Cambridge University Press, ISBN: 0521623049, 2000
- ▶ **[Lochmatter u. a. 2008]** LOCHMATTER, Thomas ; ODUIT, Pierre ; IANCI, Chris ; ORRELL, Nikolaus ; ACOT, Jacques ; MARTINOLI, Alcherio: SwisTrack - A Flexible Open Source Tracking Software for Multi-Agent Systems. In: IEEE/RSJ 2008 International Conference on Intelligent Robots and Systems (IROS 2008), IEEE, 2008, 4004–4010
- ▶ **[Schwarte u.a.]** Schwarte, R., Xu, Z., Heinol, H., Olk, J., Buxbaum, B.: New optical four-quadrant phase-detector integrated into a photogate array for small and precise 3D-cameras. In: Proc. SPIE. Nr. 3023, 1997, S. 119-128.
- ▶ **[mobotix]** The HiRes Video Company Verifiziert am 16.1.2010
http://www.mobotix.com/ger_DE