

Computer Vision für Mensch-Maschine Schnittstellen

Vorlesung WS 2013 / 14

Prof. Dr.-Ing. Rainer Stiefelhagen

Maschinensehen für MMI, Institut für Anthropomatik

Studienzentrum für Sehgeschädigte

Institut für Anthropomatik, Fakultät für Informatik



Heutige Vorlesung

- Warum ist Computer-Vision interessant?
- Wie funktioniert das?
- Anwendungsbeispiele
 - → viele Videos
 - Themen werden dann im Verlauf der Vorlesung besprochen
- Organisatorisches
 - Themenüberblick / Termine
 - Prüfbarkeit, Website
 - Programmierprojekte

Mensch-Computer Interaktion



1982 Audi V4, 2002
 Saatchi & Saatchi
 Mercedes DICE, 2012

Bolt, Put that there, 1982
 HBO, EyeToy, 2003

Ein- / Ausgabe

- Schalter, Lochkarten, LED
- Tastatur, Bildschirm (`77)
- Maus, Graphische Benutzeroberfläche (`83)
- Touchdisplays (`83)
- **Gesten (´74)**
- **Gesten + Sprache (`83)**
- **Körperbewegungen (2002)**

- Spiele
- TV
- Automobil
- ... ?

In naher Zukunft ...



- Interaktive Roboter, Robot Companions
 - Intelligente Umgebungen
 - → die uns unterstützen
 - → die mit uns interagieren können
-
- Bei der Arbeit, zu Hause, Freizeit, im Fahrzeug, in Medizin & Pflege, ...

Diese Systeme sollten Sehen können!

- Wahrnehmung des Menschen durch das System ist Voraussetzung für sinnvolle Interaktion und für angemessenes Handeln
 - Wo sind Personen?
 - Wer ist die Person?
 - Was macht die Person?
 - Wohin schaut die Person?
 - Mit wem oder was interagiert sie?
 - Worauf richtet sie ihre Aufmerksamkeit?
 - Wie geht es der Person?

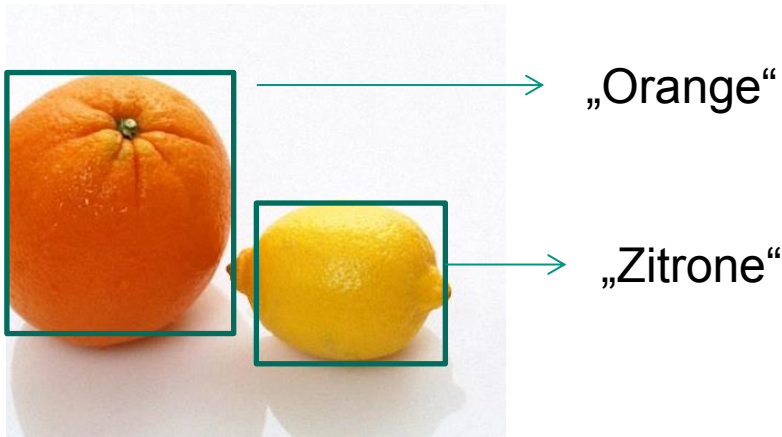
- Kontext ist wichtig:
 - Umgebung, Objekte



Wie funktioniert das?

Bildverarbeitung / Maschinelles Sehen

- Automatische Analyse von Bildern und Videos
 - Was ist zu sehen?



- Teilgebiet der Mustererkennung
- Teilgebiet der Künstlichen Intelligenz
- Teilgebiet der Informatik (und Elektrotechnik)

Wie funktioniert das?

■ Auge → Kamera

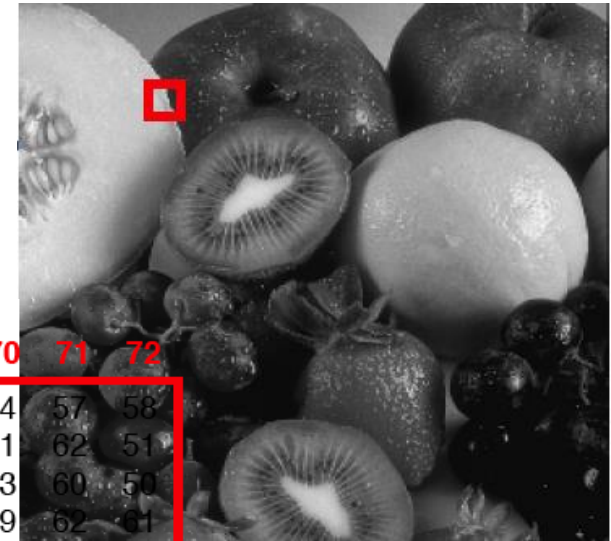


Bilder bestehen aus vielen Bildpunkten (Pixel)

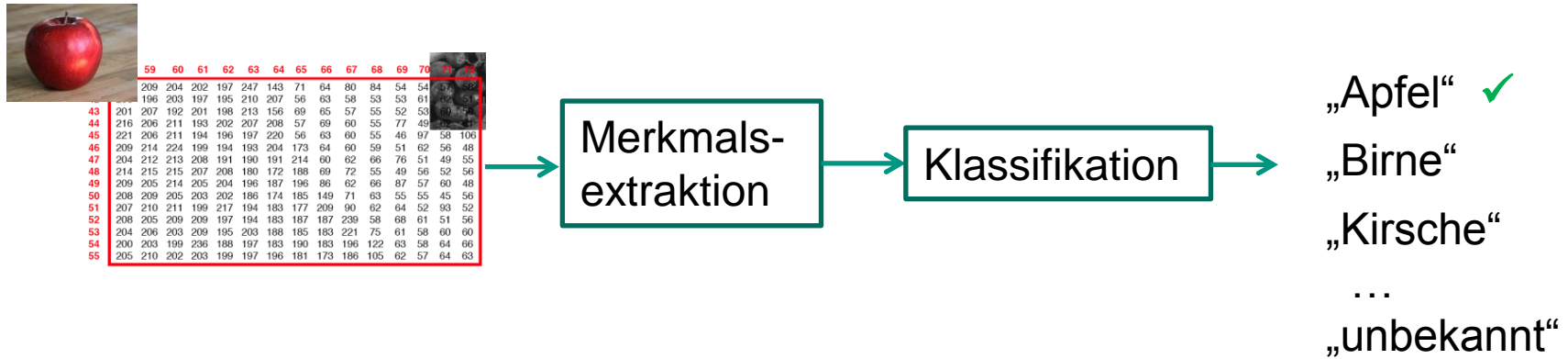


Pixel sind als Zahlen repräsentiert

		x =														
		58	59	60	61	62	63	64	65	66	67	68	69	70	71	72
y =	41	210	209	204	202	197	247	143	71	64	80	84	54	54	57	58
	42	206	196	203	197	195	210	207	56	63	58	53	53	61	62	51
	43	201	207	192	201	198	213	156	69	65	57	55	52	53	60	50
	44	216	206	211	193	202	207	208	57	69	60	55	77	49	62	61
	45	221	206	211	194	196	197	220	56	63	60	55	46	97	58	106
	46	209	214	224	199	194	193	204	173	64	60	59	51	62	56	48
	47	204	212	213	208	191	190	191	214	60	62	66	76	51	49	55
	48	214	215	215	207	208	180	172	188	69	72	55	49	56	52	56
	49	209	205	214	205	204	196	187	196	86	62	66	87	57	60	48
	50	208	209	205	203	202	186	174	185	149	71	63	55	55	45	56
	51	207	210	211	199	217	194	183	177	209	90	62	64	52	93	52
	52	208	205	209	209	197	194	183	187	187	239	58	68	61	51	56
	53	204	206	203	209	195	203	188	185	183	221	75	61	58	60	60
	54	200	203	199	236	188	197	183	190	183	196	122	63	58	64	66
	55	205	210	202	203	199	197	196	181	173	186	105	62	57	64	63



Bildverarbeitung – Typischer Ablauf



■ Beispiele für „Merkmale“

- Farbe – Farbhistogramme, Farbmomente, Correlogramme, ...
- Form & Textur – Kanten, Gabor-filter, Frequenzanalyse, Segmentierung, ..

■ Klassifikation → Maschinelle Lernverfahren

- Meistens viele Lernbeispiele notwendig

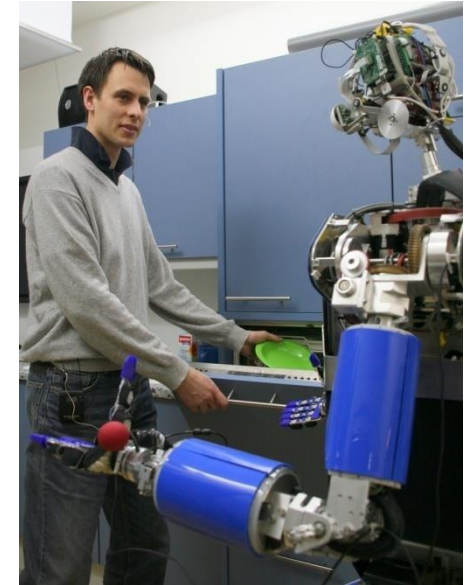
Eigene Forschung

- Ziel:
 - Entwicklung von Methoden zur Wahrnehmung von Menschen
 - Entwicklung wahrnehmender unterstützender Systeme
- Algorithmen zur Erfassung von Menschen:
 - Personendetektion und -tracking → Wo?
 - Personenidentifikation / Gesichtserkennung → Wer?
 - Erkennung von Gesten und Handlungen → Was ?
 - Erkennung von Blickrichtung / Aufmerksamkeit → Wohin?
 - Fusion und Interpretation
- Nutzerfreundliche Assistenzsysteme
 - Roboter und Intelligente Umgebungen
 - Suche nach Personen und Bildinhalten
 - Anwendungen im Bereich Gesundheit
 - Assistive Technologie für Sehgeschädigte

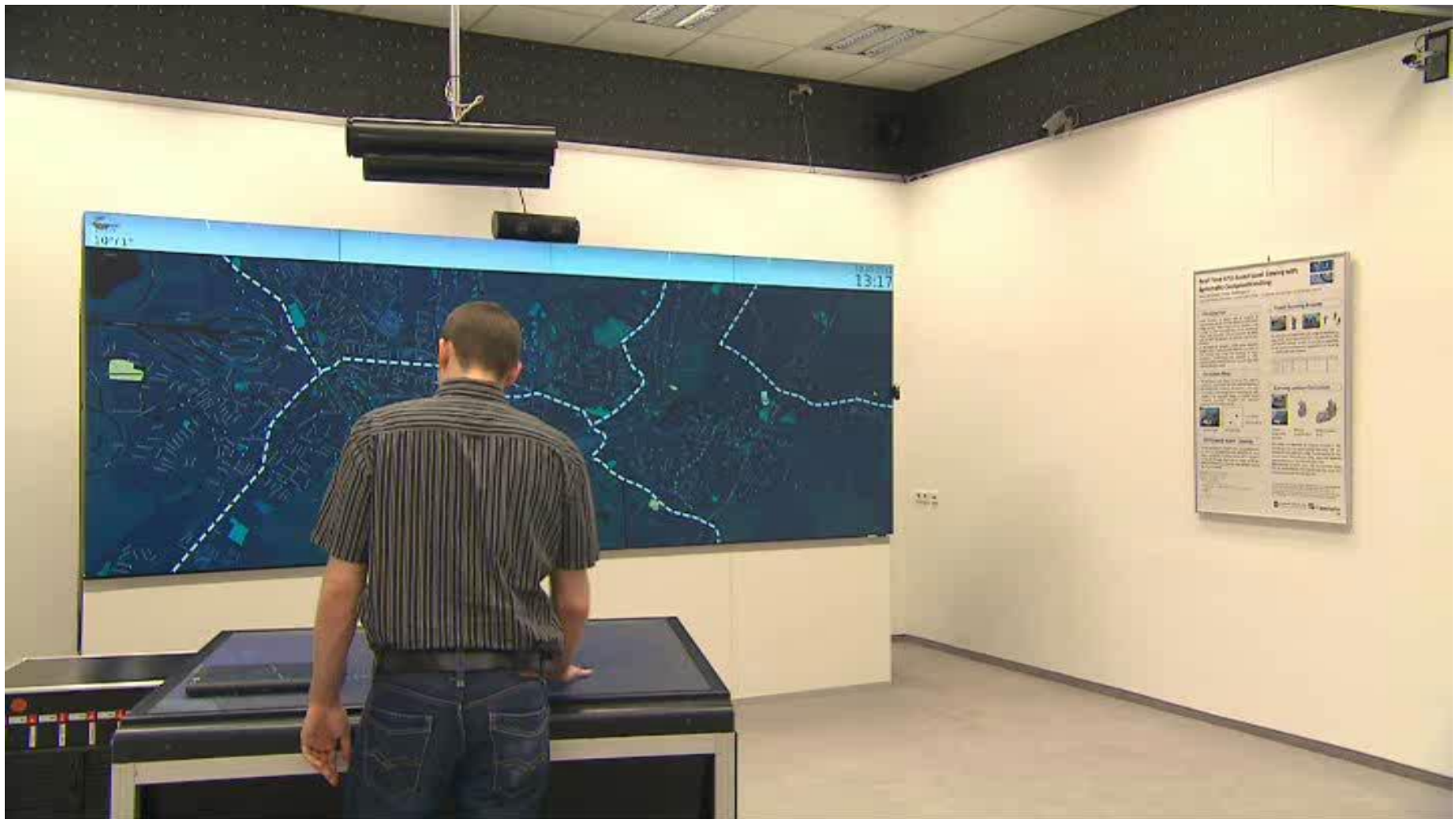


Mensch-Roboter Interaktion

- Vom Industrieroboter zum interaktiven Helfer
 - Im Haushalt
 - In der Pflege
 - In der Produktion
- Diese Roboter müssen den Mensch wahrnehmen
 - Wo ist jemand?
 - Wer ist da?
 - Wohin schaut die Person?
 - Mit wem oder was interagiert die Person?
- Erfassung der nicht-verbalen Ausdrucksmittel sind wichtig für die Interaktion
 - Gesten, Blickrichtung, Mimik

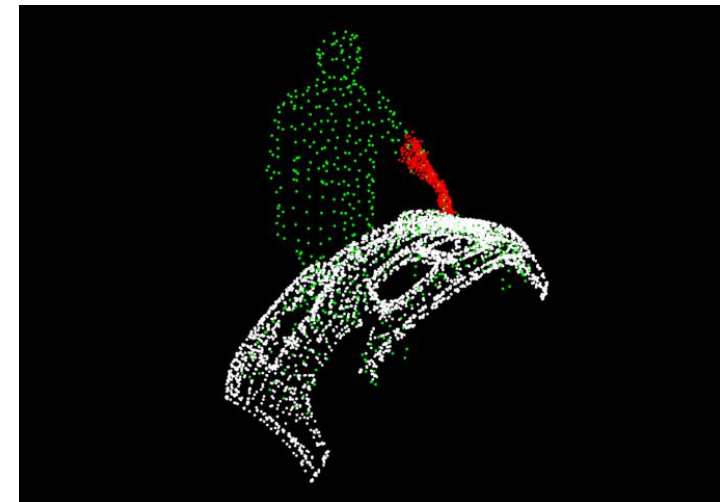


Interaktive Räume



Interaktive Räume in der Produktion

- Gestenbasierte Interaktion bei der Qualitätsprüfung
 - Qualitätskontrolle von Stoßfängern
 - Ergebnis muss protokolliert werden
- Interaktion über Wisch- oder Zeigegeste
- Werker und Bauteil werden in Echtzeit erfasst



Automatische Indexierung von Schauspielern



Interaktive Suche nach Schauspielern



18 23.10.2013

Monitoring von Patienten

Gezieltes Monitoring von Patienten nach der OP in der Intensivstation

1. Automatisches Erkennen von gefährlichen Situationen:

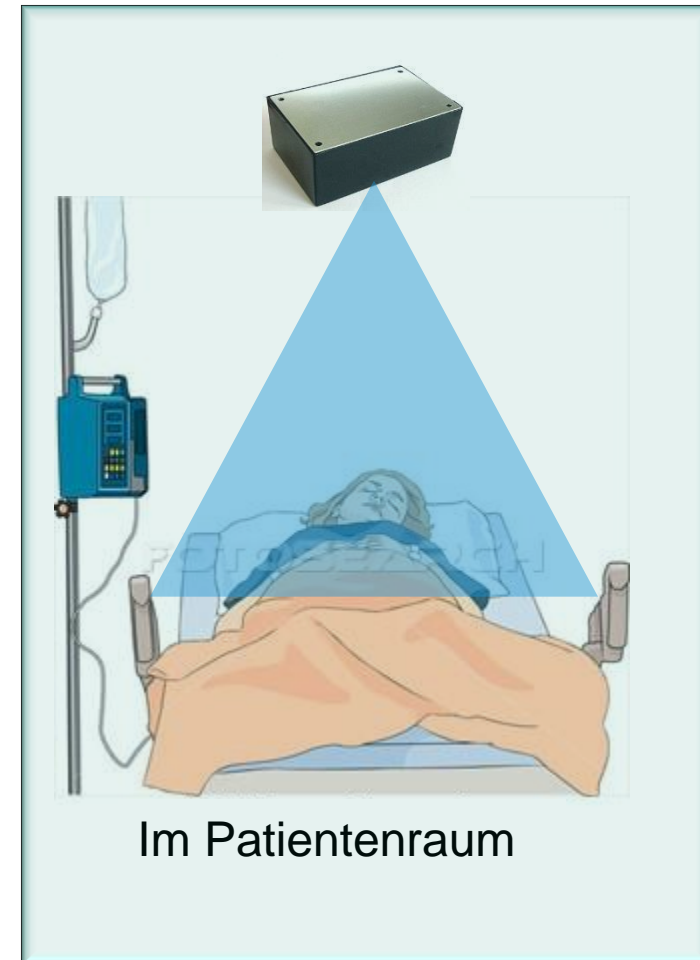
- Unruhige/plötzliche Bewegungen
- Entfernen von Schläuchen, Beatmungsgeräten, usw.
- Fall aus dem Bett

2. Statistische Analyse von Umweltfaktoren und Patientenaktivität → Einfluss auf das Patientenbefinden (z.B. Delirium)

Partner:

Städtisches Klinikum Karlsruhe

Videmo Intelligente Videoanalyse GmbH & Co. KG

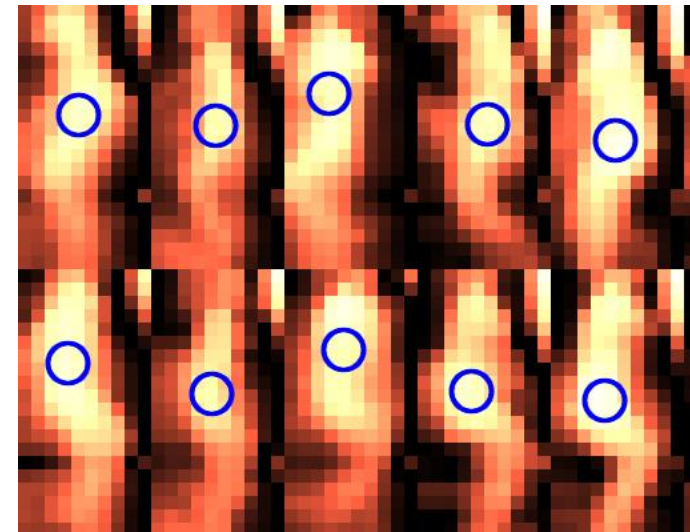
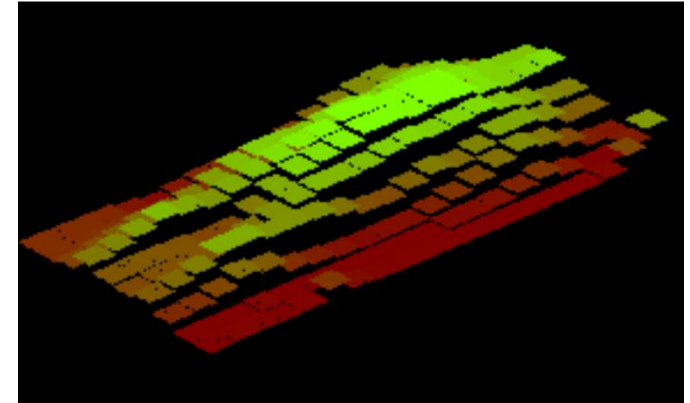


Monitoring von Patienten

- Lage der Person
 - Auf dem Rücken liegend, seitlich liegend
 - Position im Bett

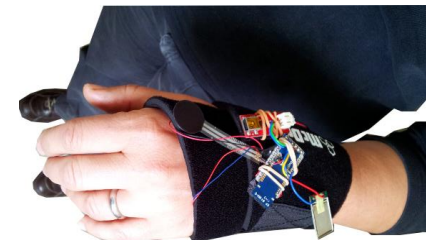
- Unruhiges Verhalten
 - Bewegungen
 - Grimassen
 - → Schmerz? Unruhe?

- Atemfrequenz



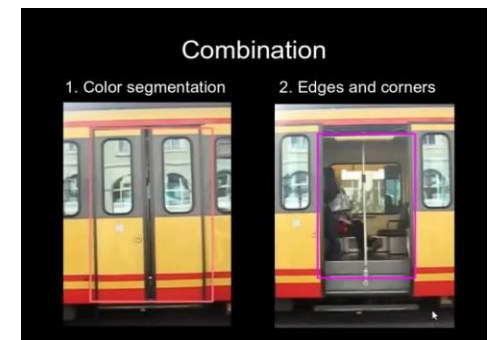
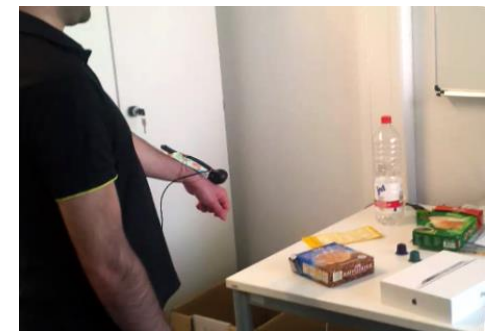
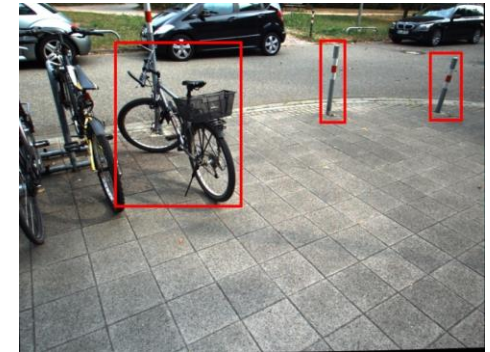
Assistive Technologie für Blinde Menschen

- Ziel: Entwicklung tragbarer Assistenzsysteme
 - Zur Mobilitätsunterstützung und Navigation
 - Zur Erkennung von Objekten, Farben, Personen
- Erkennung von
 - Personen (Geschlecht, Alter, Identität)
 - Objekten und Hindernissen
 - Gebäuden
 - Schildern und Text
- Ein- / Ausgabe: Haptisch, akustisch, Text, Braille, ...
- Diskrete Einarbeitung von Sensorik und Computer in die Kleidung

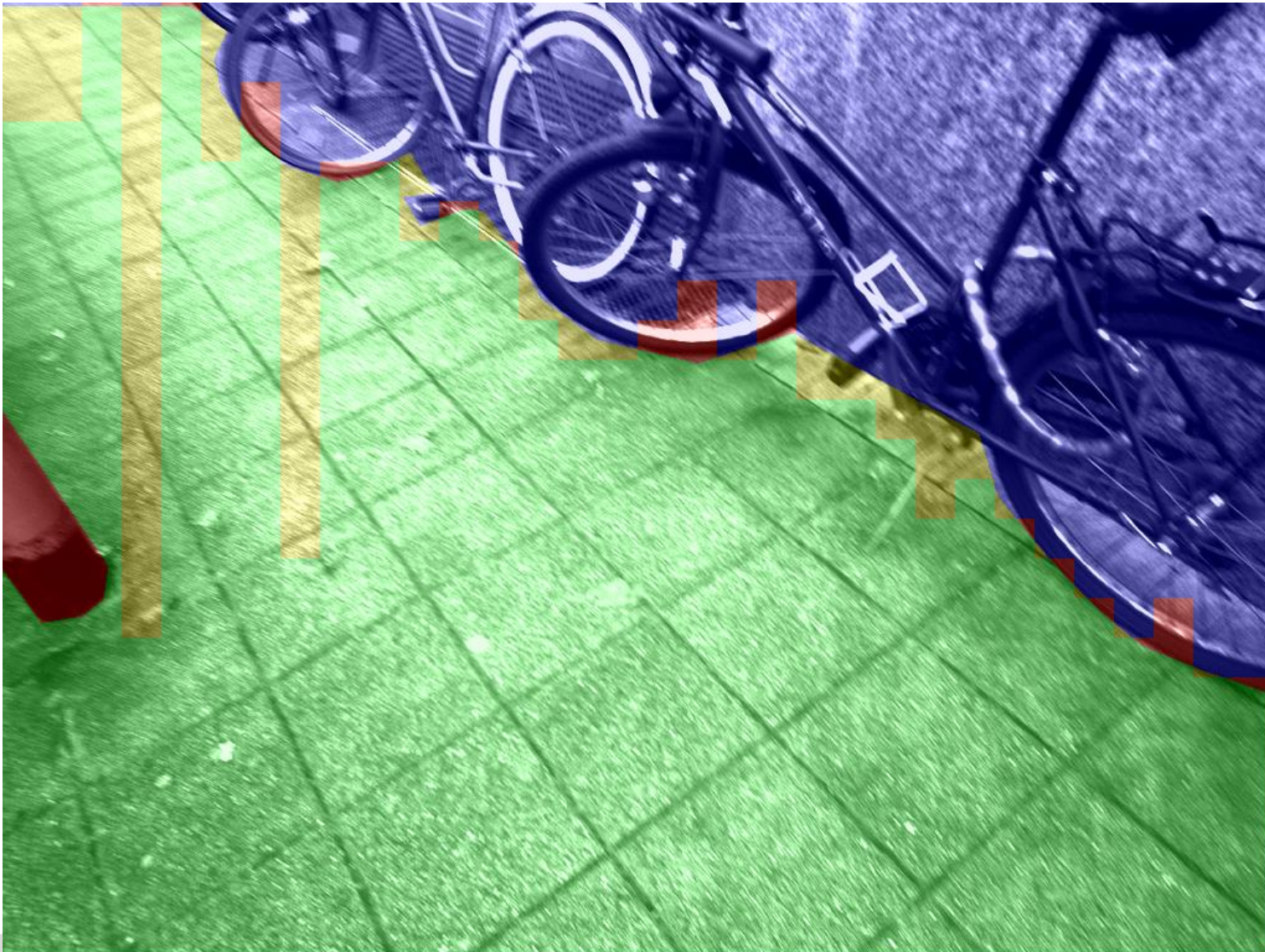


Assistive Technologie für Blinde (2)

- Assistenzsystem zur Unterstützung der Mobilität
 - Aufbau einer modularen Softwareplattform ✓
 - Hinderniserkennung in Arbeit
 - Experimente zu audio-haptischen Schnittstellen
 - Als nächstes: Erkennung von Straßenübergängen, Ampeln und Landmarken
- „Object-Finder“ ✓
 - Suche nach eingelernten Objekten oder Farben
 - Ausgabe über 3D-Audio
- Erkennung von Straßenbahntüren
- Benutzerstudien und Umfragen



Erkennung freier Flächen



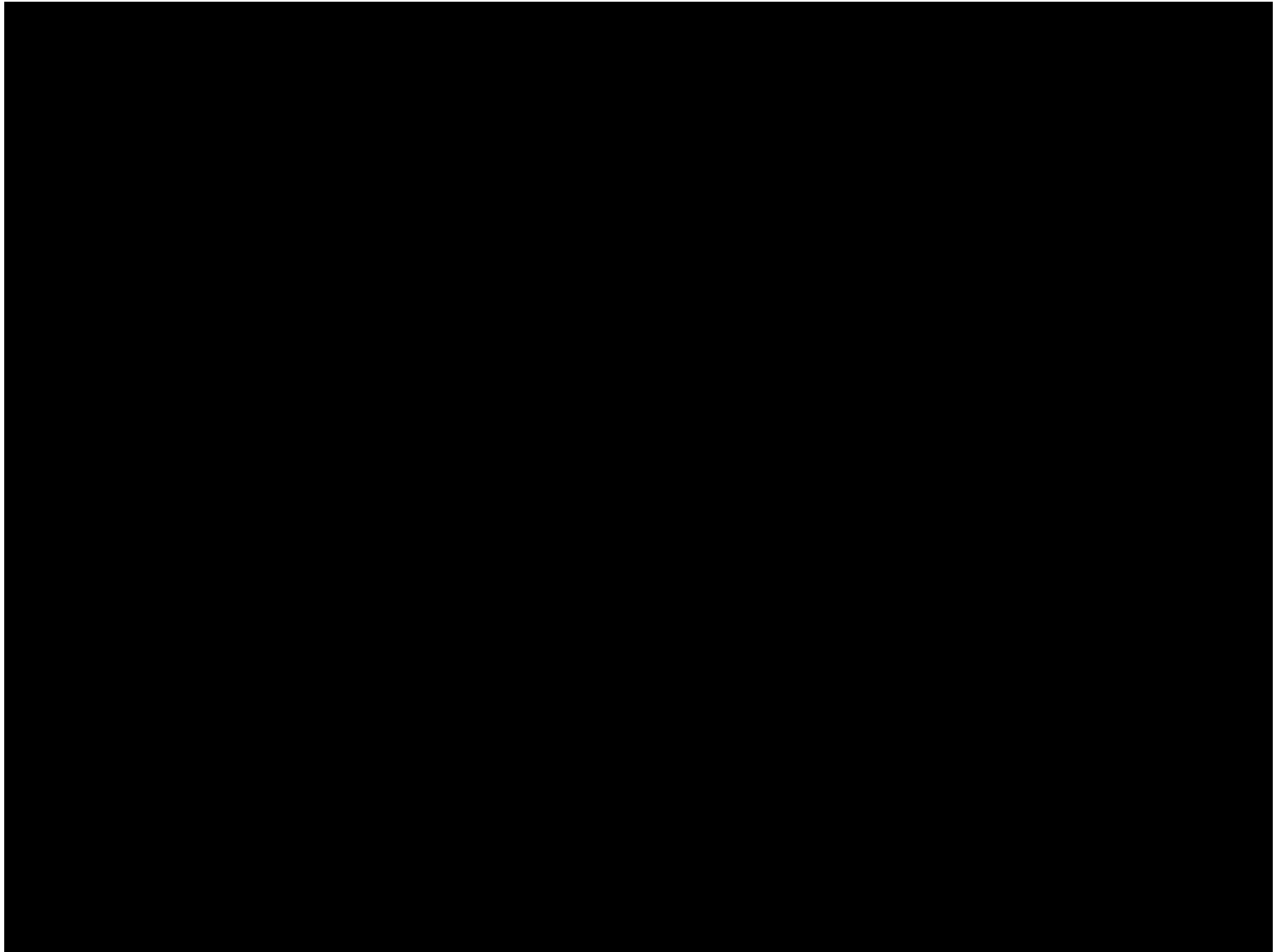
Videobasierte Erfassung von Menschen

- Position → Wo sind Personen / wo gehen sie entlang ?
- Körperhaltung → Handlungen

- Gesichtserkennung → Wer ist das ?
- Blickrichtung → Aufmerksamkeit / Intention
- Mimik → Emotion, Schmerz, Belastung, Stress, ..
- Alter & Geschlecht

- Weitere Merkmale möglich: Bart, Brille, ...

Personen finden



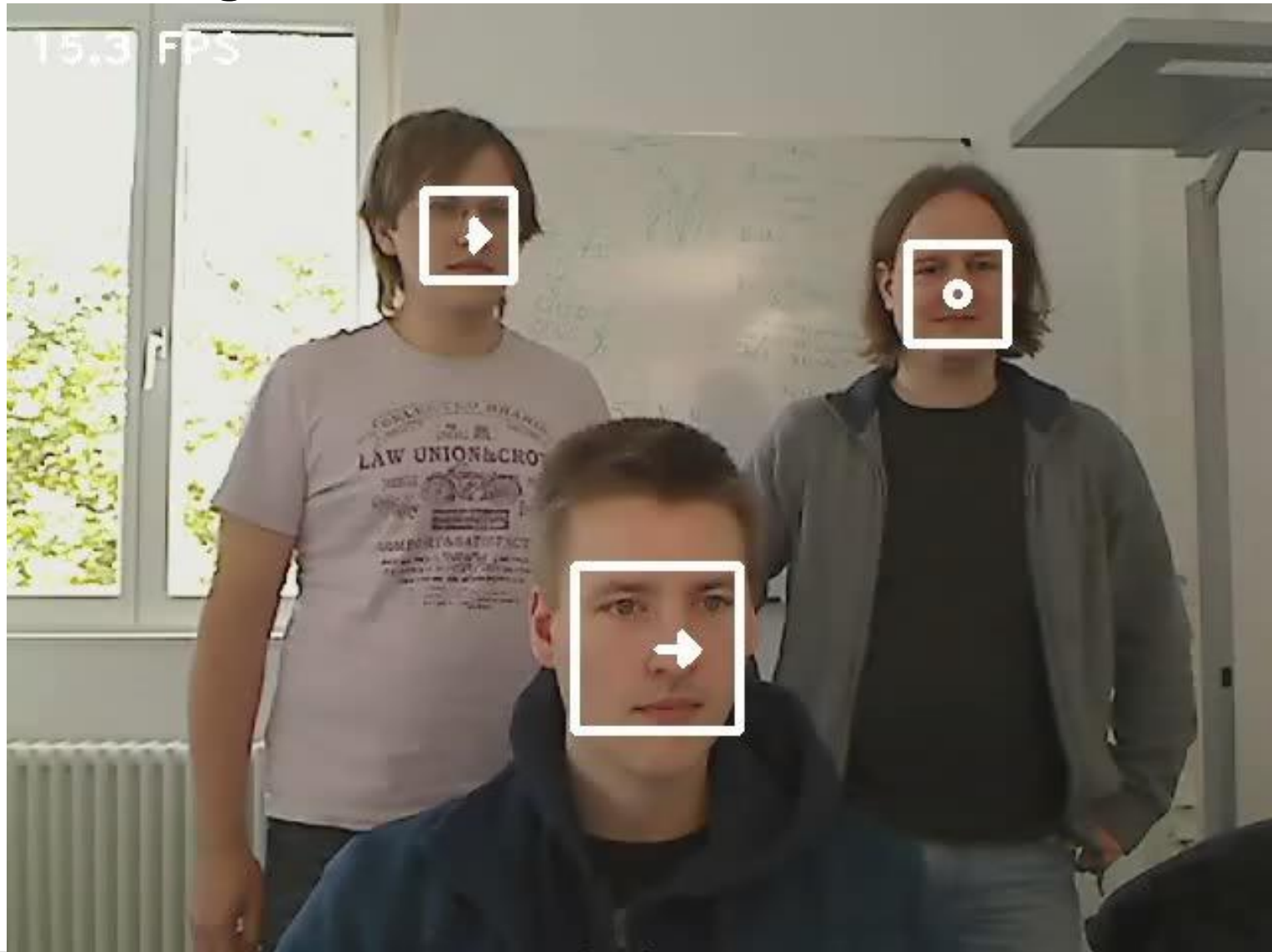
Personen erkennen – Wer ist das?



Personen erkennen – Wer ist das?



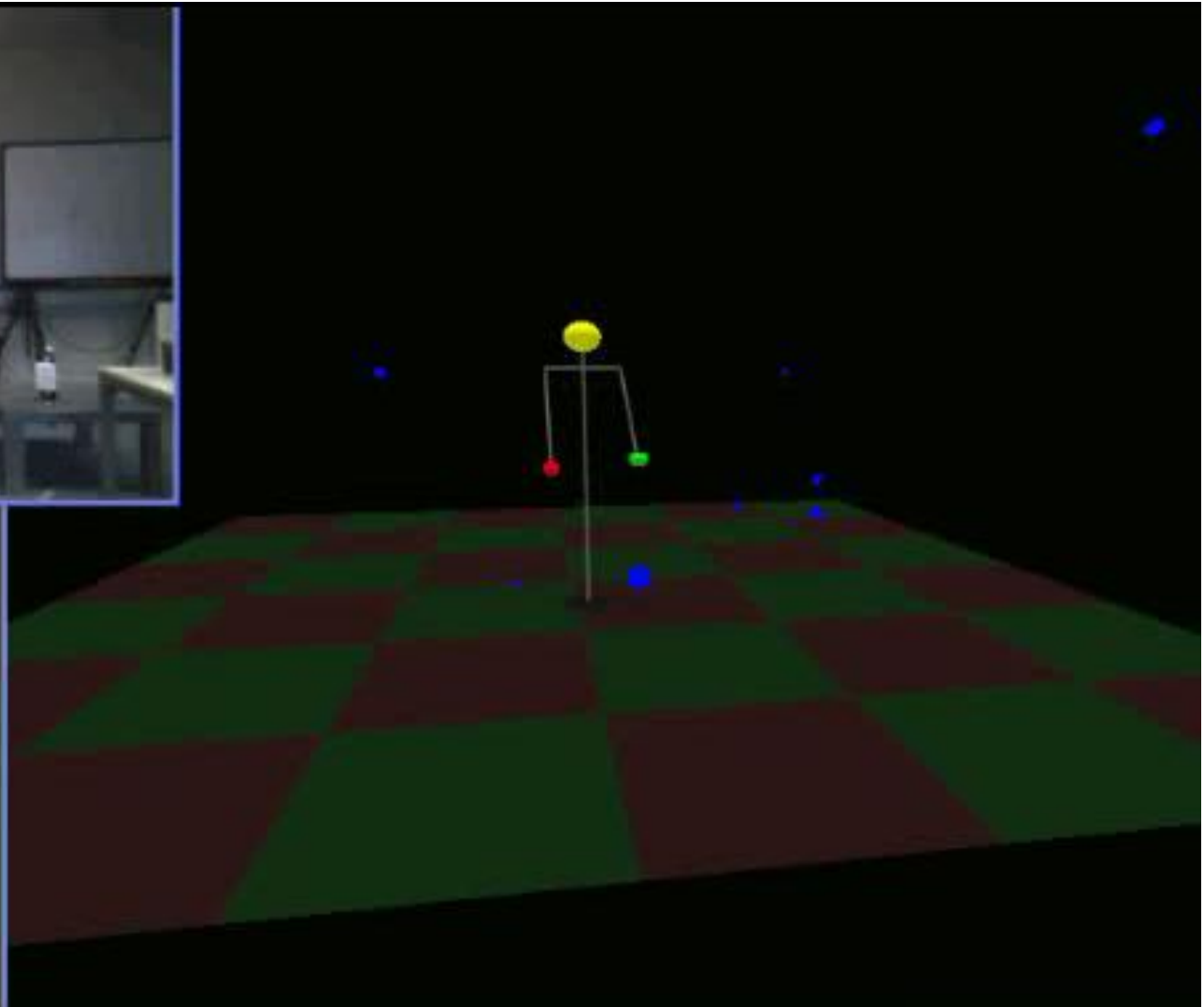
Blickrichtung und Aufmerksamkeit



Blickrichtung und Aufmerksamkeit (2)

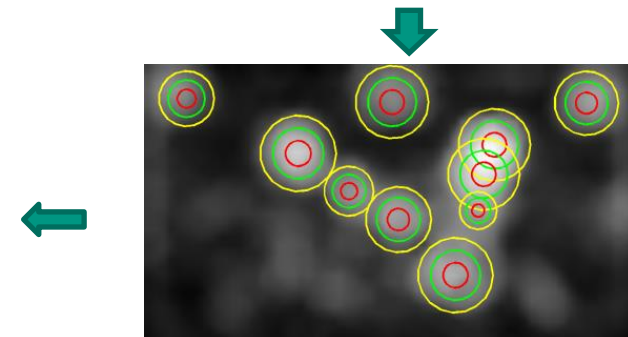
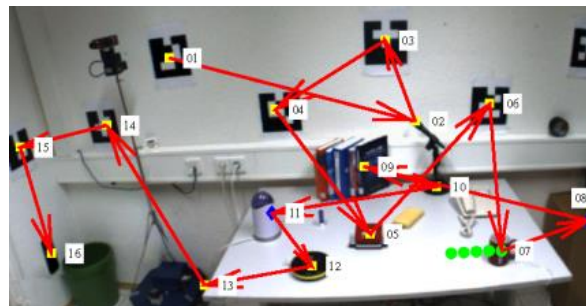
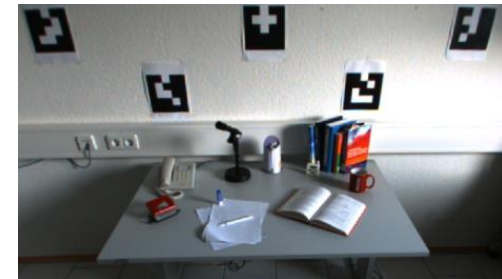


Zeigegesten

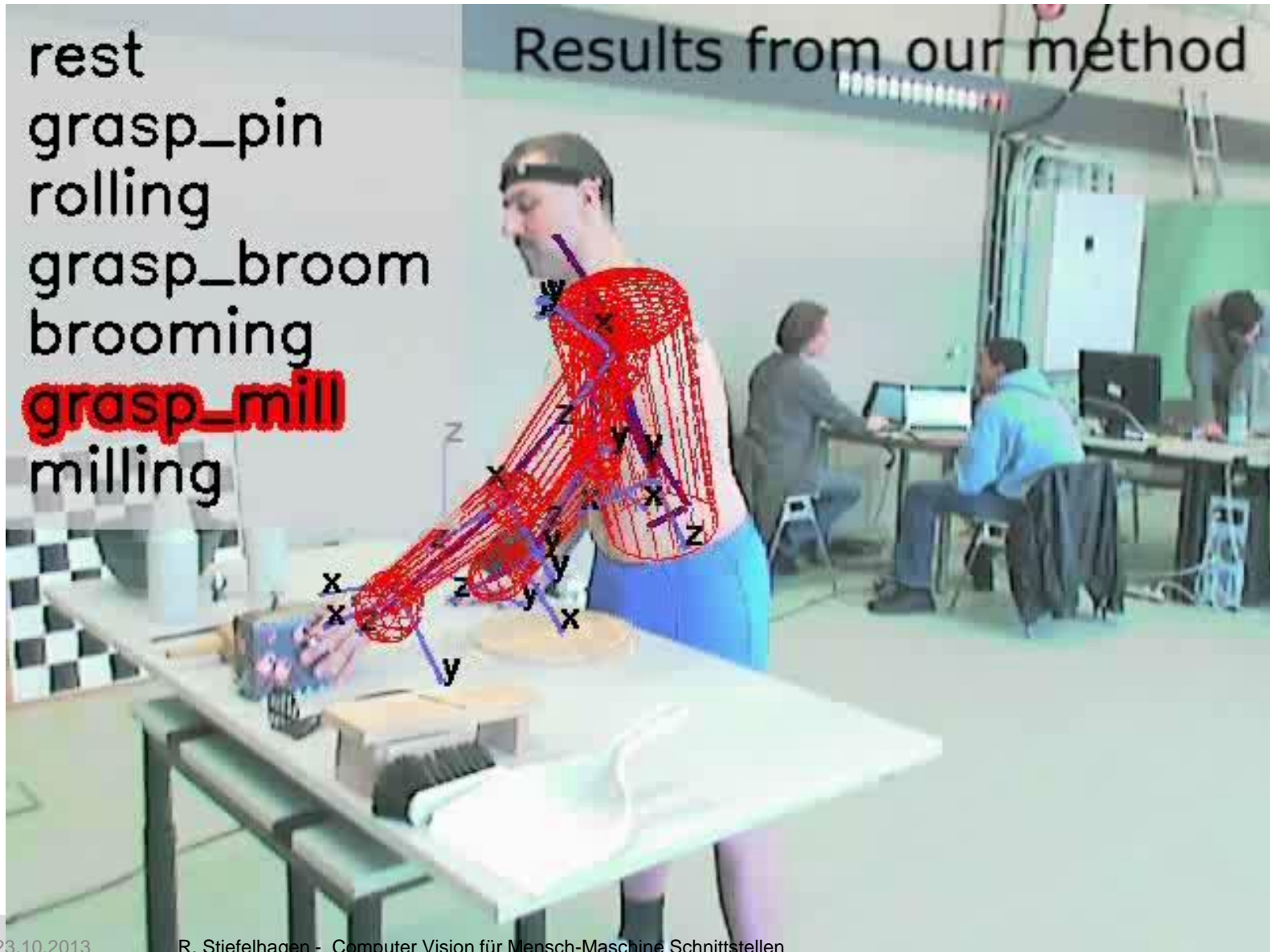


Multimodale Interaktion

- Nutzung von Sprache, Gestik, Blickrichtung zur Interaktion
- Welches Objekt referenziert der Nutzer?
 - Wichtig für Interaktion & Lernen
- Entwicklung eines neuen Salienzmodells
 - Ermöglicht Segmentierung von Objektregionen
 - Erlaubt eine Prädiktion der Blickrichtung)



Bewegungen und Handlungen

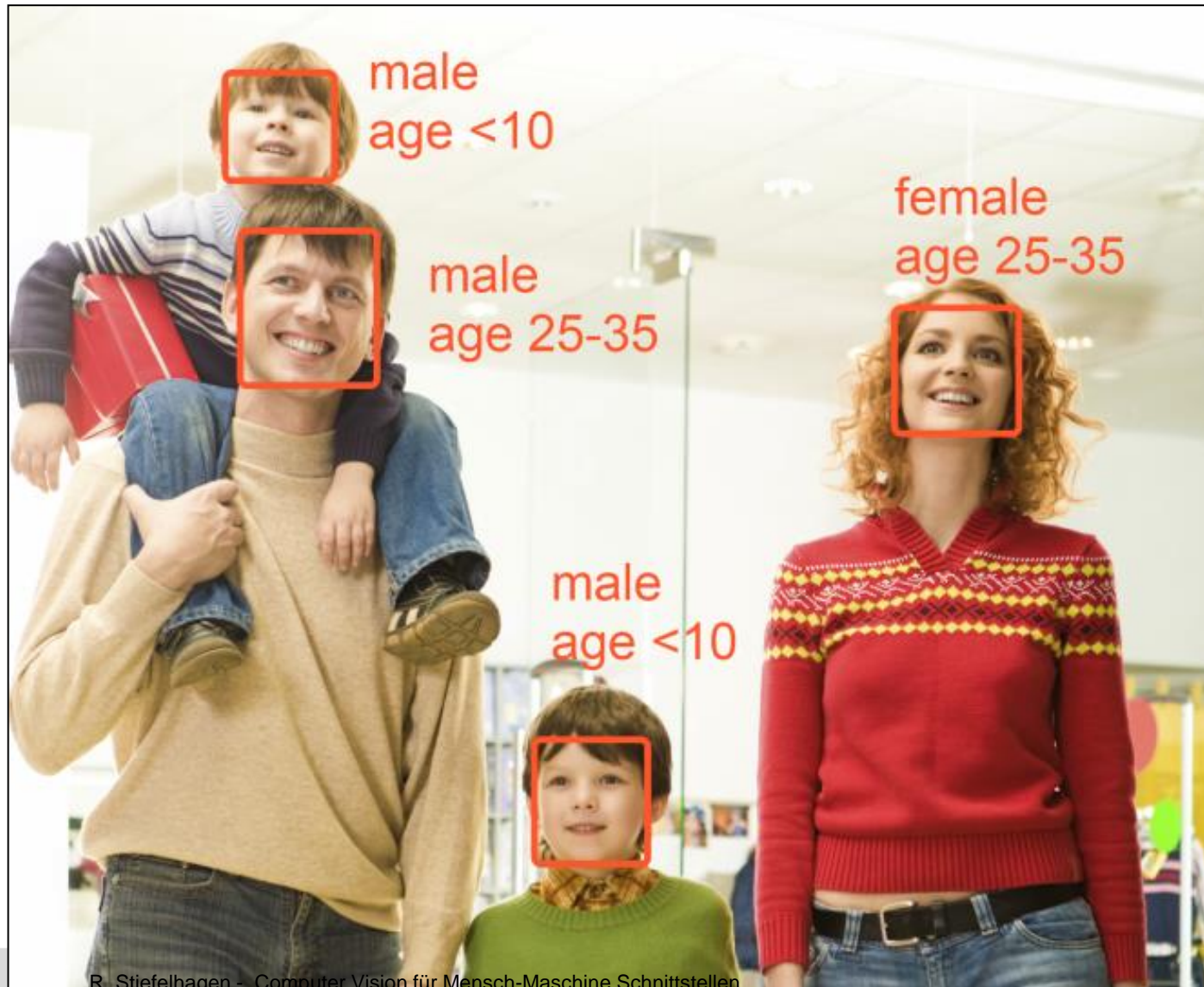




■ Erkennung von Mimik („Facial Action Units“)

- Emotionen / Schmerz / Belastung, Stress
- Kognitive, psychische Störungen

Geschlecht und Alter



Zusammenfassung

- Maschinelles Sehen ist ein spannendes Forschungsthema!
- Vielzahl von Anwendungsmöglichkeiten
 - Intelligente interaktive Roboter
 - Intelligente Umgebungen, Häuser, Fahrzeuge
 - Monitoring / Sicherheit von Patienten
 - Suche in und nach Bild- und Videoinhalten
 - Sicherheit: Suche nach Personen / Zugangskontrollen
 - Marketing: Kundenerfassung
 - Assistive Systeme für blinde und sehbehinderte Menschen
- Verfahren halten bereits Einzug in viele Bereiche
 - Spiele, Automotive, Sicherheit ...
- Diese Entwicklung wird sich fortsetzen
 - weitere Bereiche, weitere Funktionalität, bessere Verfahren

Organisatorisches

Organisatorisches

■ VL-Website:

- <http://cvhci.anthropomatik.kit.edu/teaching/visionhci>

■ Termine

- Montags, 14:00 - 15:30 Uhr, SR -120 / Geb. 50.34
- Freitags, 9:45 - 11:15 Uhr, SR -120 / Geb. 50.34

■ Vorlesung ist prüfbar mit 4 SWS

- Vertiefungsgebiet 13 „Anthropomatik“
- Vertiefungsgebiet 14 „Kognitive Systeme“
- Modul „Multimodale Benutzerschnittstellen“
- Modul „Maschinelle Visuelle Wahrnehmung“

Organisatorisches (2)

■ Vorbereitung

- Vorlesungsfolien auf Website
- ergänzend Papers lesen (werden online bereitgestellt)
 - L: coursemember
 - PW: 321meins
- Projektarbeit (siehe später)!
- Fragen, falls irgendwas unklar ist!
 - In der Vorlesung und/oder Termin
- Ergänzend: → Seminar „Bildverarbeitung für MMI“

Überblick über die Vorlesung

- Einleitung, Motivation, Überblick (heute)
 - Anwendungsbeispiele, einzelne Technologien

- Wiederholung von Grundlagen
 - Maschinensehen: Definitionen, Aufgaben, Challenges, Metriken,
 - Bilderverarbeitung, Bildtransformationen
 - Mustererkennung, Bayes-Klassifikation, Gauss-Mischverteilung,

Überblick über die Vorlesung (2)

■ Maschinensehen von Menschen:

- Face Detection
- Face Recognition
- Facial Analysis, Facial Feature Detection & Tracking
- Estimating Head Pose & Focus of Attention
- Gesture Recognition
- People Detection
- Person Tracking: Kalman-filter, Particle-filter, Applications; Articulated Body Tracking
- Activity Analysis

Termine (1)

Termine	Thema
21.10.13	Introduction, Applications
25.10.13	Basics: Computer Vision & Image Understanding
28.10.13	Basics: Image Preprocessing
1.11.13	Feiertag
4.11.13	Basics: Pattern recognition
8.11.13	Machine Learning
11.11.13	Mid-Level Representations (Dense Features and their encoding)
15.11.13	Project 1: Intro + Programming tips
18.11.13	Face Detection I: Color, Edges
22.11.13	Face Detection II: ANNs, Viola & Jones
25.11.13	Face Recognition I: Traditional Approaches(ASM/AAM)
29.11.13	Face Recognition 2: State-of-the-art
2.12.13	Project 1: Questions
6.12.13	Facial Expression Recognition
9.12.13	People Detection I - Holistic models
13.12.13	People Detection II - Part-based models
16.12.13	People Detection III (Part-Based Models)
20.12.13	Project 1: Student Presentations, Project 2: Intro

Termine (2)

10.1.14	Head Pose Estimation: Model-based, NN, Texture Mapping, Focus of Attention
13.1.14	Tracking I : Kalman, Particle Filter & Applications (AV-Tracking, Body pose)
17.1.14	Tracking II: Articulated Body Tracking
20.1.14	Gesture Recognition: Taxonomy, Neill, Starner (ASL), Nickel, ...
24.1.14	Activity Analysis / Action recognition I
27.1.14	Activity Analysis / Action recognition II
31.1.14	MS Kinect, Kinect-based body part detection & Programming
3.2.14	keine VL (Zeit zum Programmieren)
7.2.14	Project 2: Student Presentations & Lessons Learned
10.2.14	Laborbesuch

Programmieraufgaben

- There will be 2 programming assignments
 - Organized by Manel Martinez, manel@ira.uka.de
 - 15.11. - 20.12. Skin Color Detection
 - 20.12. – 7.2. People Classification
- Framework: C++, Qt
- A small intro will be given on the 15.11.
 - *Details soon!*

Assignments: It's a challenge

- Results have to be quantitatively evaluated
 - We compare the results of the different groups
- Results have to be presented in front of the class
 - Short presentation with approach and examples
- Good results will have a positive influence on the final grade
- To make it a little easier: Teams of 3 students

Assignments: Goal

- Have fun!
 - One of the rare occasions, where you as a student get hands-on experience
 - Caution: It can also be frustrating sometimes ;-)
- See what computer vision research is like
 - Best practice for your diploma or student thesis

Seminar im WS 2013/2014

Computer Vision für Mensch-Maschine-Interaktion

■ Themen in den Bereichen

- Face/Object Detection
- Face Identification
- Person Tracking
- Person (Re-)Identification
- Articulated Body Tracking
- Head-pose & Gaze Estimation
- Activity Recognition
- Facial Expressions
- Gesture Recognition

■ Organisatorisches

- Termin: Mittwochs, 15:45 – 17:45
- Ort: **Vincenz-Prießnitz-Str. 3, Geb. 07.08, R003)**
- **erster Termin: Mittwoch, 23.10.2013, 15:45**
(Vorgesprechung, Themenvergabe)
- ECTS-Punkte: 3
- Kontakt: Rainer Stiefelhagen <rainer.stiefelhagen@kit.edu>
Saquib Sarfraz <saquib.sarfraz@kit.edu>
- Webseite: <http://cvhci.anthropomatik.kit.edu/teaching>



Ort

- Institut für Anthropomatik
Karlsruher Institut für Technologie
Vincenz-Prießnitz-Str. 3
76131 Karlsruhe



Danke und bis Mittwoch (Seminar) / Freitag (VL)

