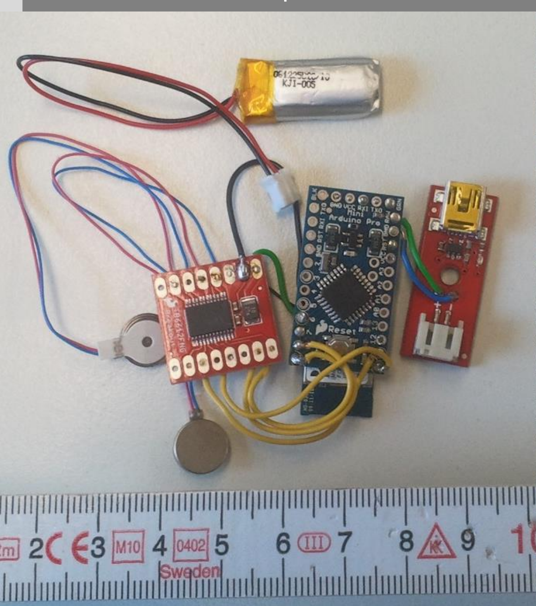
Computer Vision for HCI
Towards a mobility aid for seeing impaired users

Rainer Stiefelhagen
Institute for Anthropomatics and Robotics & Study Center for Visually Impaired Students, Karlsruhe Institute of Technology



My Background: Computer Vision for HCI
(& 3 years Study Center for Visually Impaired Students)

Karlsruher Institut für Technologie

- Visual Perception of Humans… (Computer Vision)
  - Person tracking & Identfication, age, gender, facial expression, body pose, action recognition, saliency & attention, video-to-text-alignment
- … for Human-Computer Interaction (HCI)
  - Robots & Interactive Environments
  - Surveillance & Multimedia-Analysis
  - <u>Assistive Technology for Blind & seeing impaired /</u> health-care
- Since 11/ 2011: Professor for „IT-Systems for Seeing Impaired Students" at Karlsruhe Institute of Technology
  - Director of the Study Center for Seeing Impaired at KIT
  - Director of Computer Vision for HCI Lab – Inst. f. Anthropomatics & Robotics
  - Close collab. with Perceptual User Interfaces Lab at Fraunhofer IOSB

Karlsruher Institut für Technologie

Human Robot Interaction

- Gaze estimation
- Action recognition in the kitchen
- Robot interaction

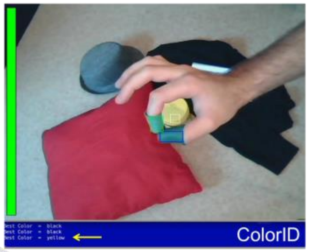
Multimedia Analysis and Retrieval

Facial Analysis

- Face recognition
- Expression analysis
- Facial features

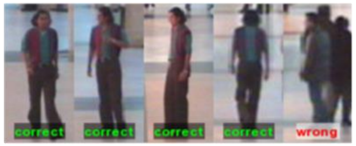
Computer Vision for the Blind

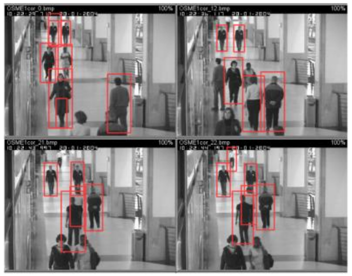
- Interfaces: Haptic / audio / etc.
- Navigation
- Object recognition



Security and Safety

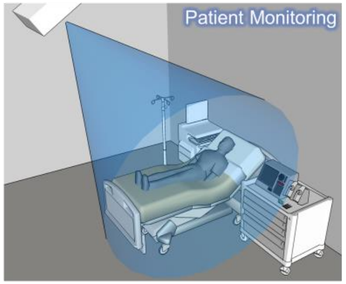
- Camera networks
- Person retrieval





Health-care scenario

- Dangerous event detection
- Monitoring vital signs
- Multimodal sensor



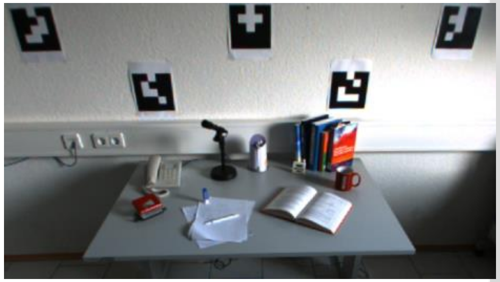
Smart Environments

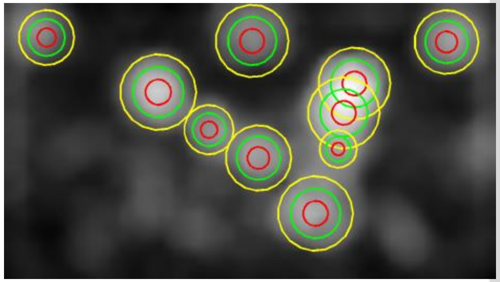
in collaboration with Fraunhofer IOSB

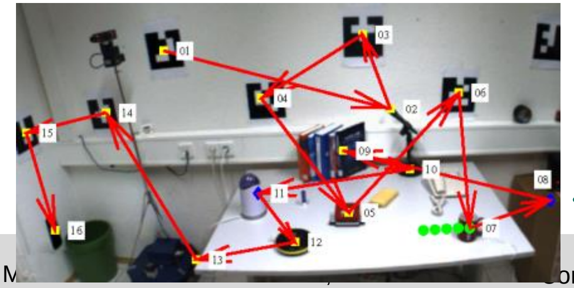
- Multimodal interaction
- Real-time situation monitoring

Multimodal Human-Robot Interaction

- Using speech, gesture, gaze for interaction
- With which object is the user interacting?
- Combination of bottom-up and top-down cues
  - Bottom-up: saliency
  - Top-down: gestures, speech
- Also work on action and acitivity recognition

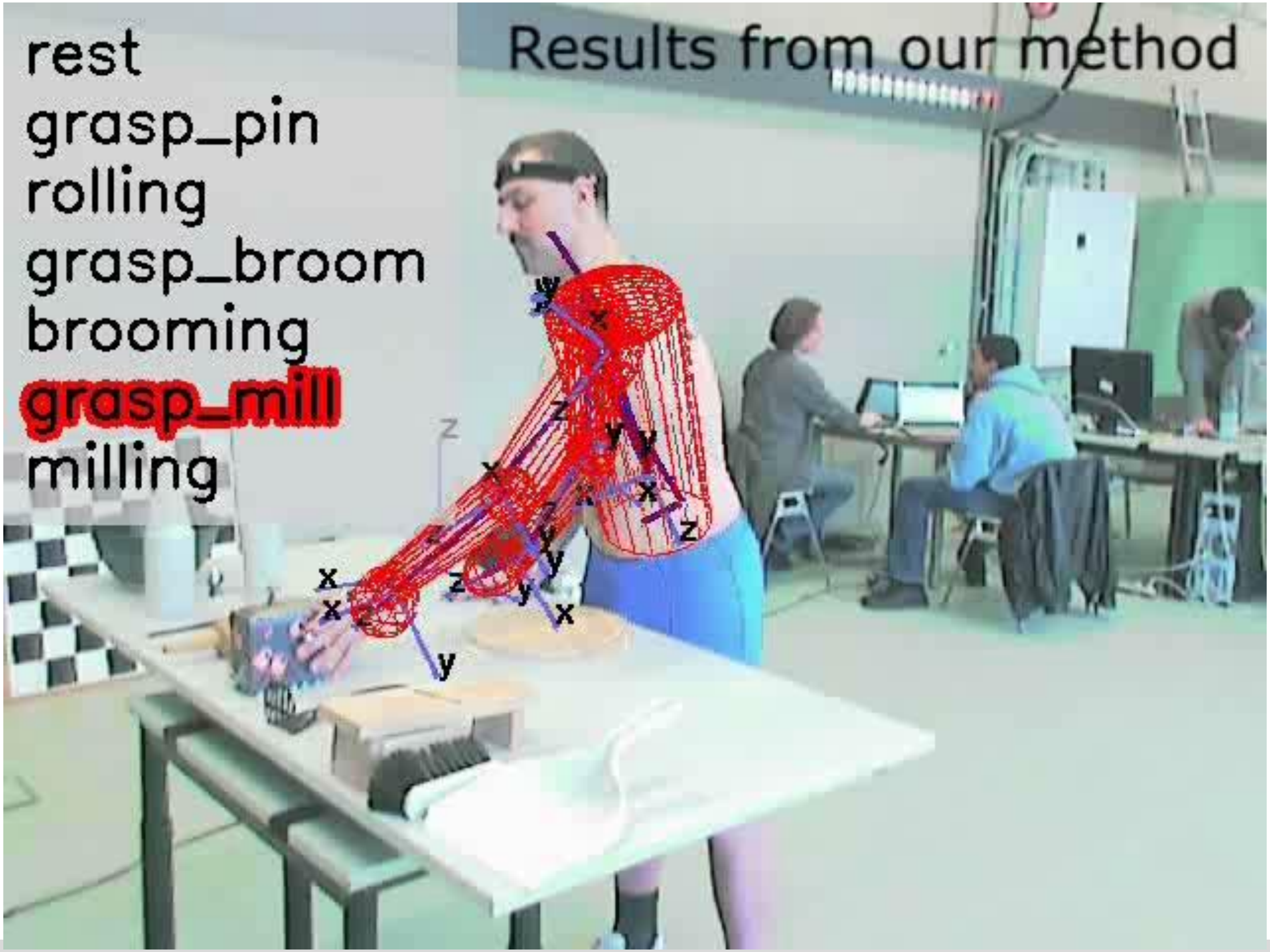




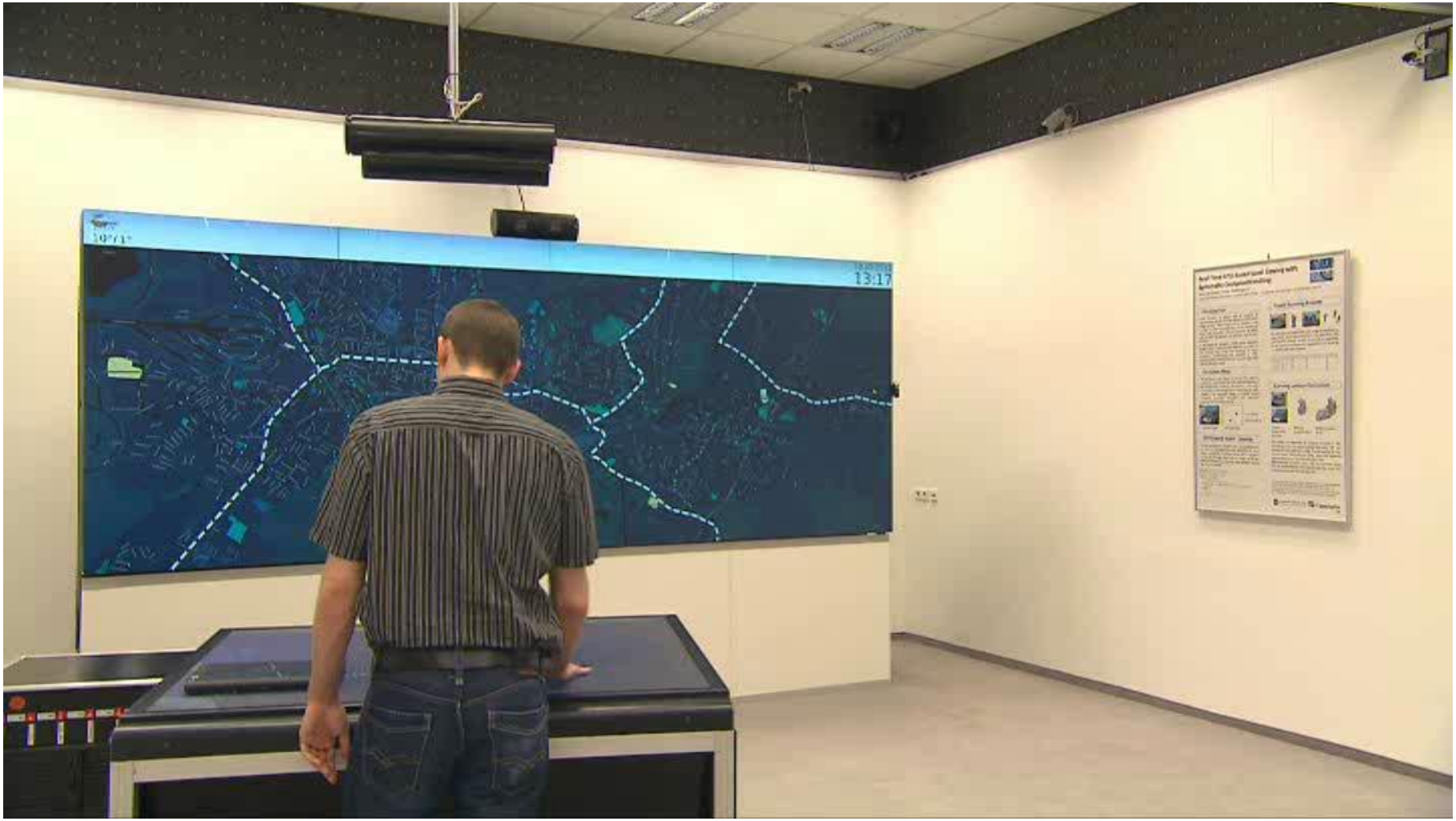


[Schauerte & Stiefelhagen, IROS 2014]

Actions and Activities



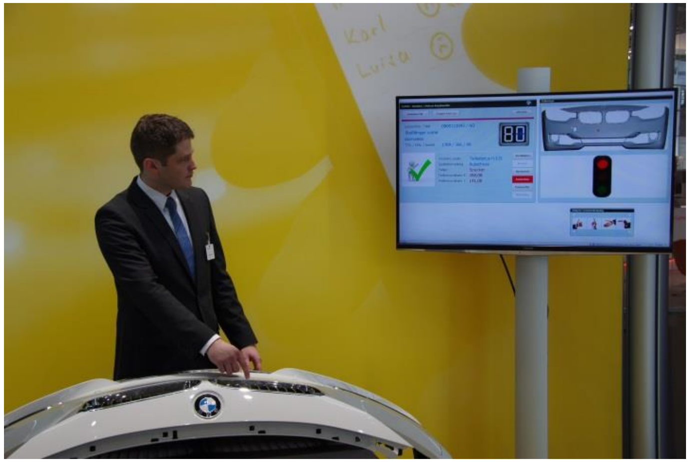
Interactive Environments

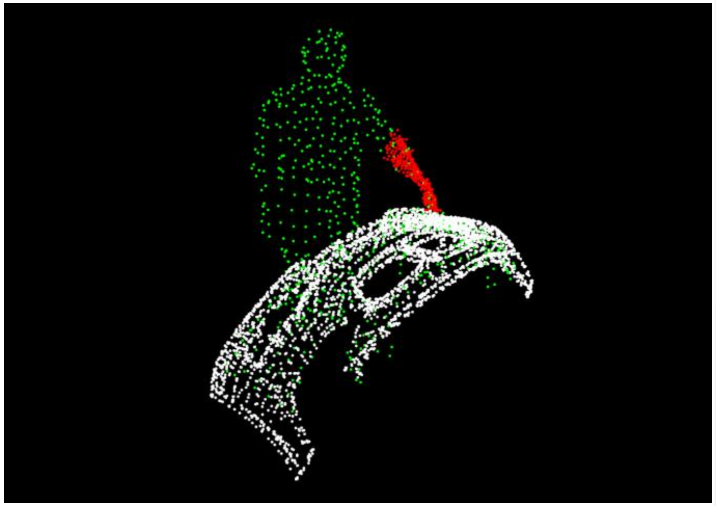


The „Smart Control Room“ at Fraunhofer IOSB

Interactive Rooms in Industry / Production

- Gesture-based interaction in quality control
  - Industry cooperation
- Worker and bumper tracked in real-time



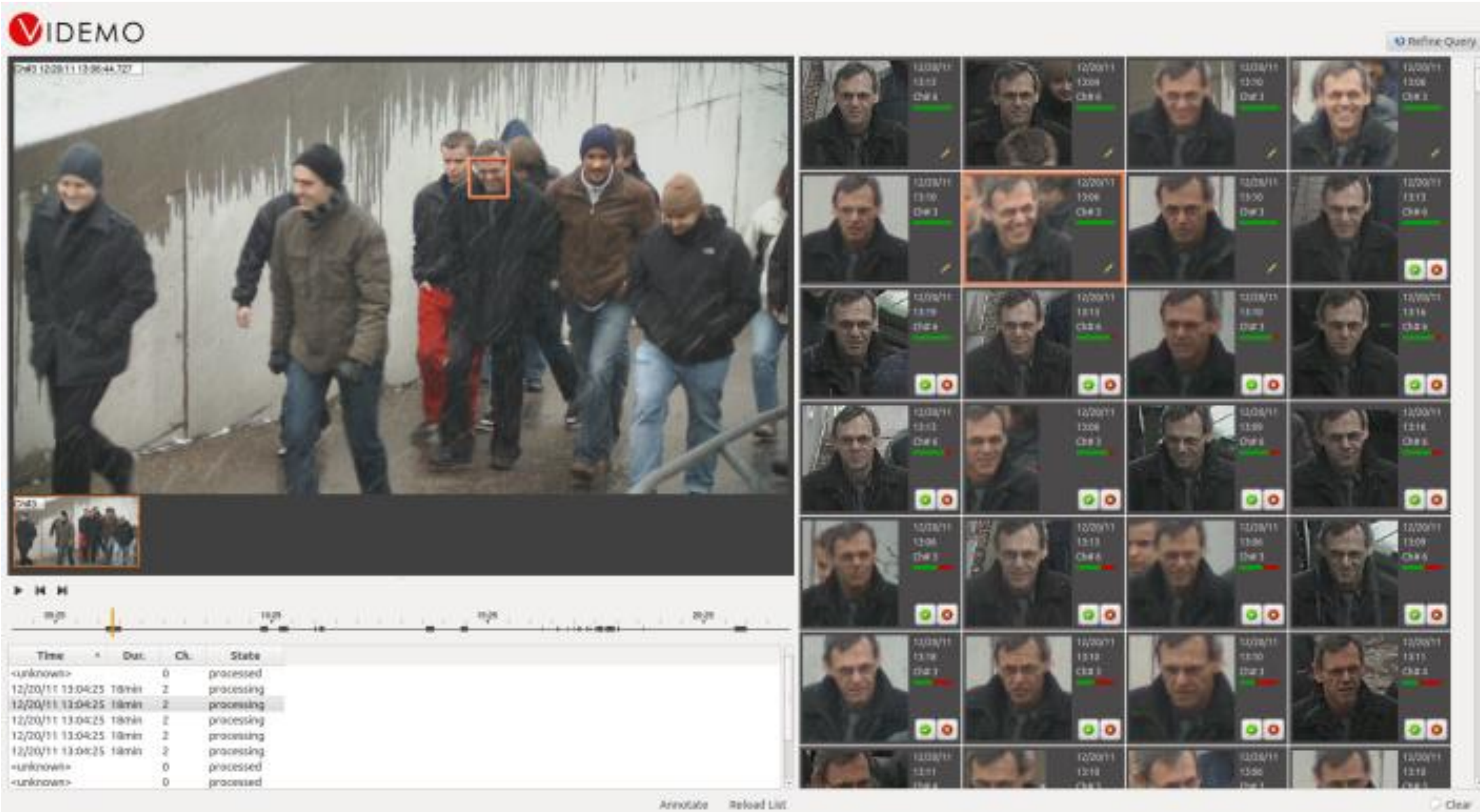


Work with A.Schick, F. v. d. Camp, M. Voit, Frauhunhofer IOSB

Indexing and Search in Multimedia

(Tapswi et al. CVPR 2012, CVPR 2014)

Interactive Search in Surveillance videos



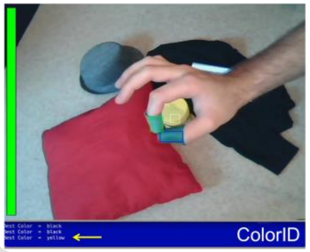
Bäuml et al, CVPR 2013 / BMBF Projects Pagevi & Mispel

Human Robot Interaction

- Gaze estimation
- Action recognition in the kitchen
- Robot interaction

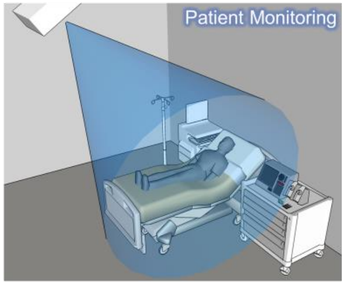

Computer Vision for the Blind

- Interfaces: Haptic / audio / etc.
- Navigation
- Object recognition



Health-care scenario

- Dangerous event detection
- Monitoring vital signs
- Multimodal sensor

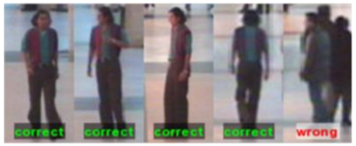
Multimedia Analysis and Retrieval

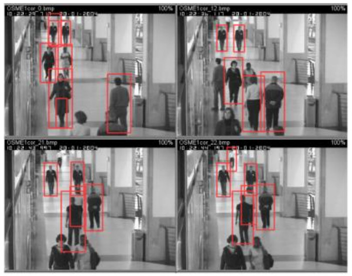
Facial Analysis

- Face recognition
- Expression analysis
- Facial features

Security and Safety

- Camera networks
- Person retrieval





Smart Environments

in collaboration with Fraunhofer IOSB

- Multimodal interaction
- Real-time situation monitoring


KIT Karlsruher Institut für Technologie

Computer Vision 4 the Visually Impaired

- 295 million visually impaired worldwide / 39 million blind
- Loss of sight leads to a number of problems

- Mobility
- Activities of daily living
- Information access
- Communication

- Education, Job, Recreation

Mobility

- Three out of ten visually impaired never go out without company (Clark-Carter et al 1981); elderly visually impaired: two out of three (Yerasimou 2002)
  - Fear of collisions and difficulties in orientation
- Long Cane is the most common assistive device
  - Detect obstacles and uneven ground
- Guide dogs
  - Recognizes and signals
    - obstacles, traffic signs, parked cars, pedestrians
    - Doors, steps, road crossing, mailboxes, free seating …
  - Very expensive (~ 20K Euros, plus food etc.)
  - Only 1-2% of blind have a guide dog

Existing Electronic Devices

- Based on Ultrasound, Laser or Infrared
- Feedback either haptic or audio

- Technical limitations
  - Ultrasound: limited resolution & precision, problems with reflections, not possible to detect stairs or holes
  - Laser: does not detect glass surfaces
  - Infrared: problematic outside

- Systems are not widely used

“Ultracane”

“K-Sonar”

“Miniguide”

“Sonic Pathfinder”

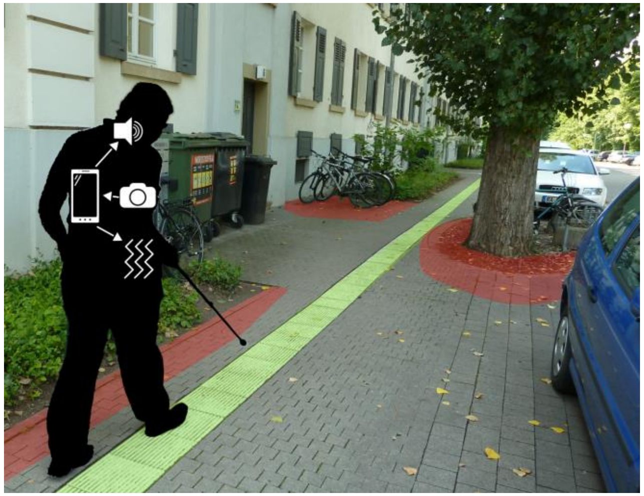
Computer Vision for the Visually Impaired

- Computer Vision is making great advances
  - Persons, ID, Age, Gender, head pose, …
  - Objects, Landmarks, Text
  - Semantic Scene Analysis
- Applications for VI
  - Urban mobility assistance
  - Human-Human-Interaction support
  - Activities of Daily Living
- How to design the interface ?
  - Audio: Speech / Sounds
  - Haptic
  - Braille

Towards a mobility aid that detects obstacles

- improve “urban mobility and autonomy”
- supplement the white cane (“extended range”)
- smoothly guide people around obstacles

- but, almost any physical object can be an obstacle, if it lies on the ground — detectors?



Obstacle-free Areas





- Basic assumptions and ideas
  - the area in front of the user is obstacle free, because he/she uses the system continuously
  - human-in-the-loop: the user can recover from mistakes (using his/her white cane)

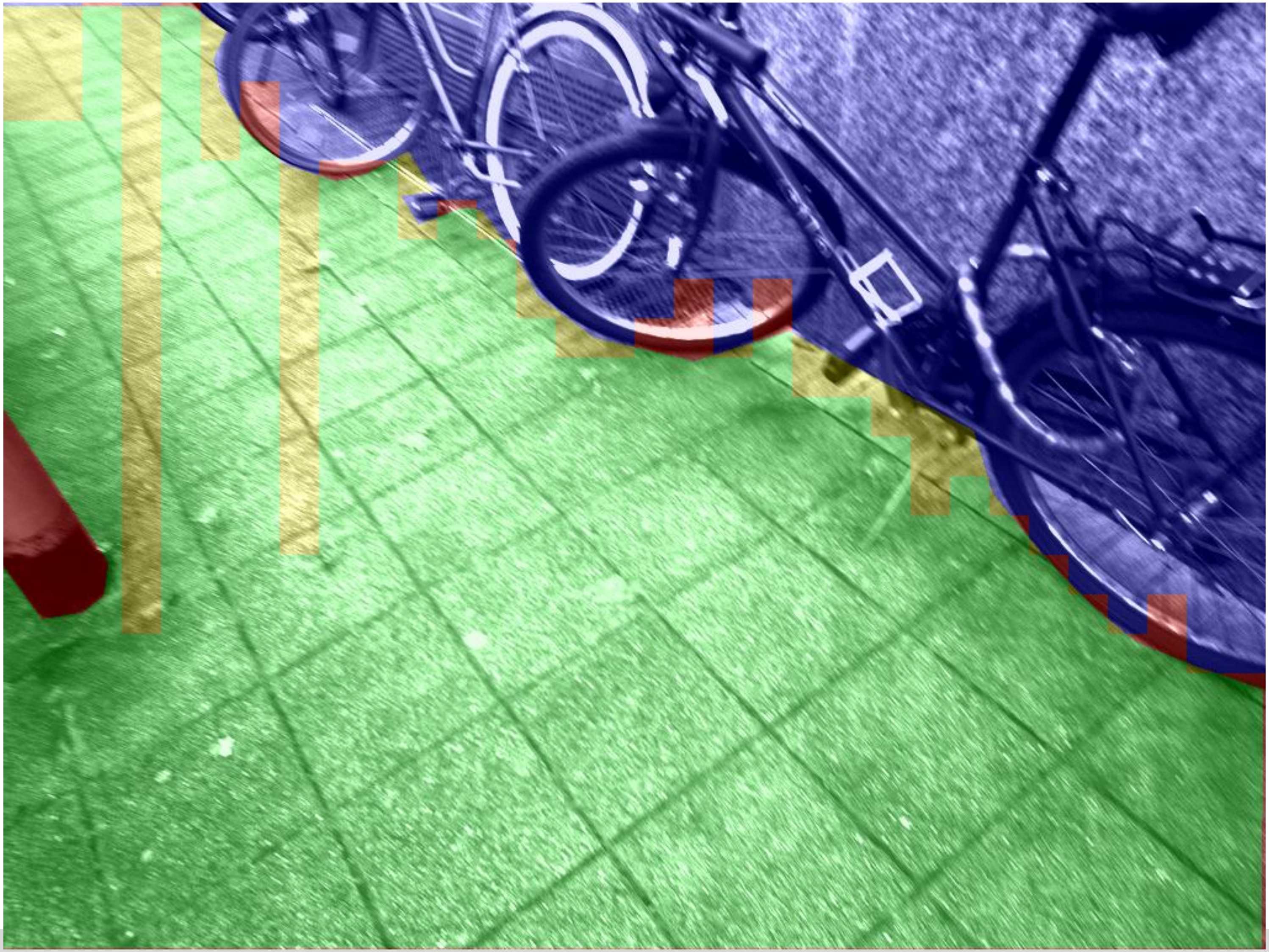
Detecting free space

- First system [Koester et al 2013]:
  - Estimate the surface angle based on depth map
  - Start from lower image border and search for non-plane angles/discontinuities
  - Collect (upwards) all blocks that fit criteria

- Second system [Schauerte et al. 2014]
  - Use CRF to model free space
  - Works with monocular images, too
  - Significantly better thatn depth-based heuristic
  - Adding depth information improves results

D. Koester et al., MAP4VIP 2013
B. Schauerte et al., ECCV Workshops 2014

Example Video



A small user study: Audio vs. Haptics to signal obstacles

- Experimental Setup
  - Obstacle course: 8 obstacles form a maze (20m x 5m)
  - We assume a working system that detects those obstacles and guides the user around the maze
  - One hour per test
  - user to familiarize with the test

[Martinez et al., ICCHP 2014]

Color Finder System

- System to detect objects and colors [Martinez et al. 2012]
- Detects obstacles (colors) at 30Hz feedback with 5-20ms latency
- We manually signaled obstacles in cases of illumination or communication problems → Wizard of Oz

Audio Interface

- Open headphones (others possible)
- 20ms beeps at 800Hz
- Horizontal image coordinate → sound panorama (pitch change removed)
- Up to 4 items could be differentiated by focused testers

Haptic Interface

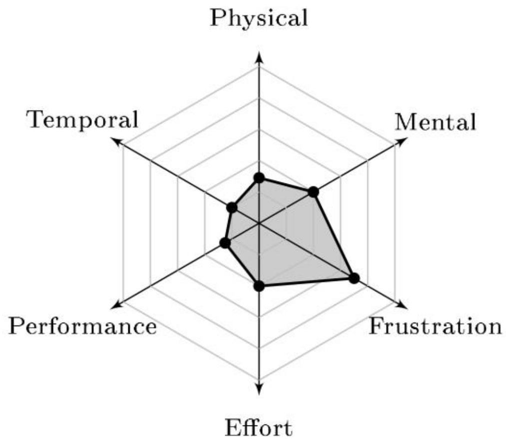
- Custom lightweight and small electronics
- Vibration motors and bluetooth module
- Mounted to white cane, vibration bursts signal obstacle in front of user (left/center/right)

Evaluation

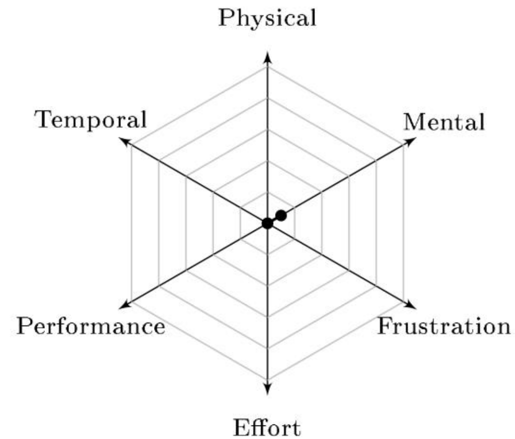
- Metrics such as time of completion and success rate not very useful
- Measure workload instead
  - → NASA TLX methods / questionaires
- Preliminary user study with N=6 subjects
  - 3 white cane users
  - 3 blind folded users

Workload Results

74.7%



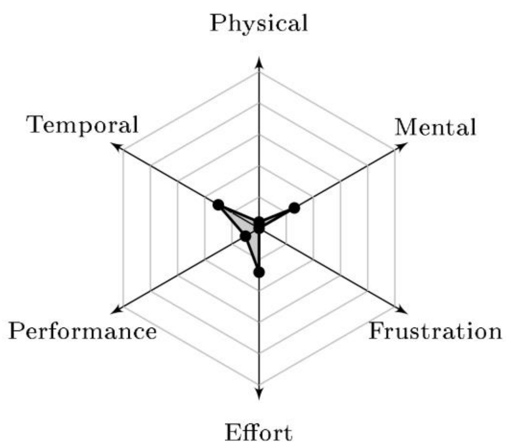
(a) White Cane Users - Audio



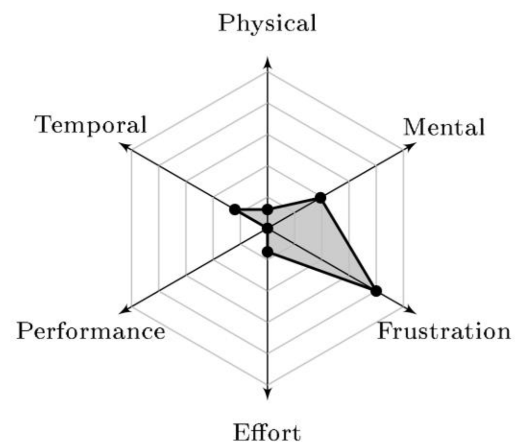
(b) White Cane Users - Haptic

3.3%

32.6%



(c) Non White Cane Users - Audio



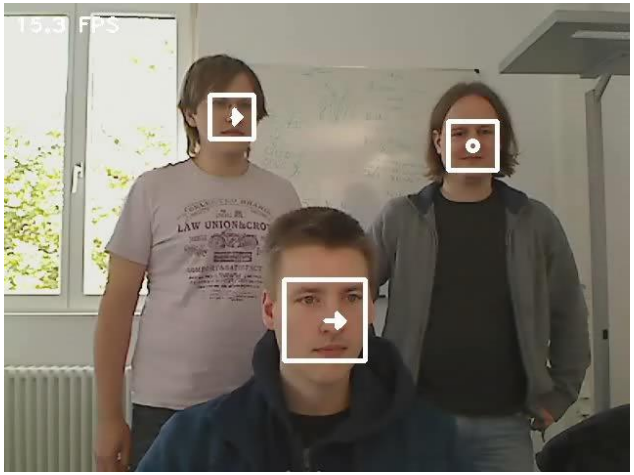
(d) Non White Cane Users - Haptic

56.0%

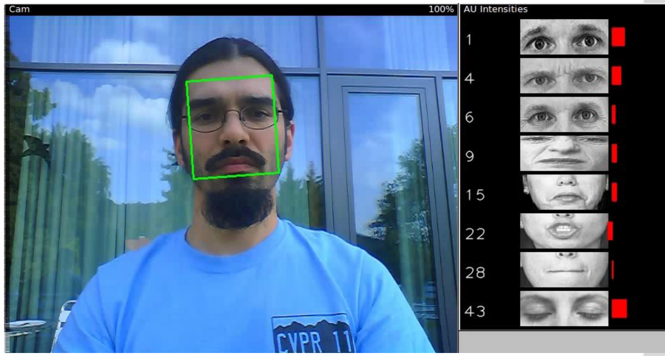
Outlook: Supporting Human Communication

(Funded by BMBF & Tubitak)

- Detect and signal
  - Who is there (face recognition, age, gender)
  - Who is looking at me (head pose / eye gaze)
  - Facial Expressions



- Initial user study
  - 19 out of 25 would like to have visual information about others in formal situations
  - 24 out of 25 would like to have it in informal situations
  - 24 out of 25 would like to know when they are looked at
  - 24 out of 25 interested in facial expressions



- CV-Components ready, have to work on interfaces / user studies

Conclusions

- Computer Vision methods have huge potential to build assistive technology for visually impaired
  - Mobility & Orientation
  - Communication
  - Activities of Daily Living
- Designing the user interface is crucial
  - User-centred design necessary
  - Blind-folded users are different from blind users
  - Have to work with visually impaired users!
  - → Collaboration with Study Center for Seeing Impaired
- Great topic for multimodal interfaces!

Thank you

Prof. Dr.-Ing. Rainer Stiefelhagen
Karlsruher Institut für Technologie
Institut für Anthropomatik und Robotik
Vincenz-Priessnitz-Str. 3
76131 Karlsruhe

Tel.: +49 721 608-46385
rainer.stiefelhagen@kit.edu

https://cvhci.anthropomatik.kit.edu/
http://www.szs.kit.edu/