

# Intuitive und markerlose Interaktion in einer mobilen Virtual Reality Anwendung auf Basis von RGBD-Daten

Masterstudium:  
Medieninformatik

Daniel Fritz

Technische Universität Wien  
Institut für Softwaretechnik und Interaktive Systeme  
Arbeitsbereich: Interactive Media Systems  
Betreuer: Priv.-Doz. Mag. Dr. Hannes Kaufmann  
Mitwirkung: Dipl.-Ing. Annette Mossel

## Einleitung

In dieser Arbeit wurden zur 3D-Interaktion mit einer mobilen Virtual Reality (VR) Anwendung marker- und berührungslose Interaktionstechniken verwendet. Die Evaluierung untersucht die implementierten Ansätze und zeigt die Vor- und Nachteile.

**Ziel:** Robuste und intuitive 3D-Interaktion per Hand und Kopf mit VR-Szene auf Tablet und Tiefenkamera als Erweiterung (RGBD)

### Herausforderungen

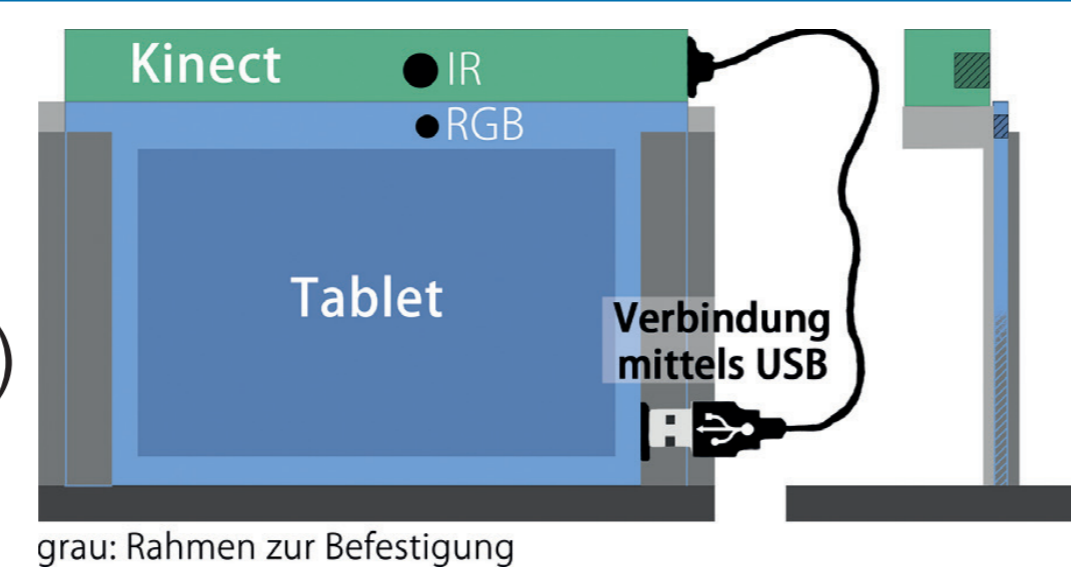
- ▶ Hardware Prototyp (Kameras fixieren)
- ▶ Tiefendaten auslesen (USB, Treiber)
- ▶ RGBD-Daten (Kalibrierung, Mapping)
- ▶ 3D-Interaktion mit VR-Szene
- ▶ Training von Klassifikatoren
- ▶ Robuste Hand- und Kopferkennung



## Methodik

### Hardware Prototyp

- ▶ Android Tablet Asus TF201 (Recheneinheit, RGB-Kamera, VR-Szene)
- ▶ Kinect for Windows (Tiefendaten)

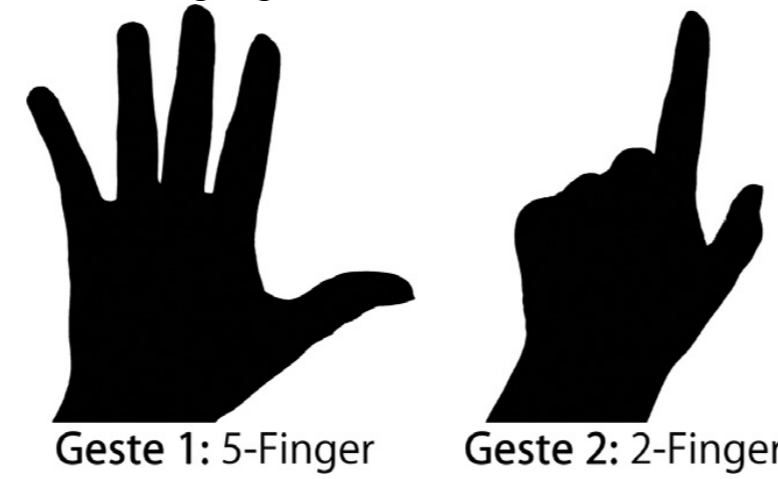


### Interaktion

Geste 1: Selektion von Objekten

Geste 2: Manipulation von Objekten

Kopf: Ändern des Betrachtungswinkels

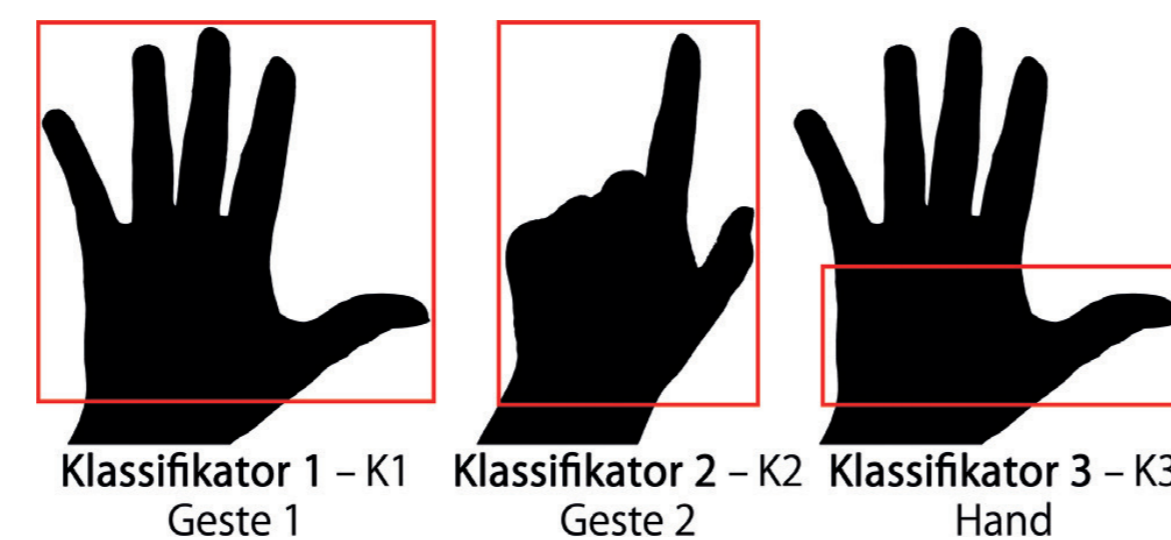


### Klassifikatoren

Kopf: Bereits trainiert

Fingergeste und Hand: Training

- ▶ Haar-Merkmale [1]
- ▶ Boosting Lernalgorithmus [2]



### Ansätze zur Fingergeste- und Kopferkennung, 3D-Position

#### Erster Ansatz - A1 (RGB)

1. Fingergeste- und Kopferkennung: Klassifikatoren
2. 3D-Position: Relative Tiefenschätzung (Objektgröße, max. Grauwert)

#### Zweiter Ansatz - A2 (RGB)

1. Hand- und Kopferkennung: Klassifikatoren
2. 3D-Position: Relative Tiefenschätzung (Objektgröße, max. Grauwert)
3. Fingerspitzenerk.: Computer Vision Algorithmen (A2-C, A2-H, A2-T, A2-D)
4. Bestimmung der Fingergeste: Anzahl von ermittelten Fingerspitzen

#### Erweiterung (RGBD)

5. RGBD-Daten: Absolute Tiefenschätzung, Fingerspitzenerkennung

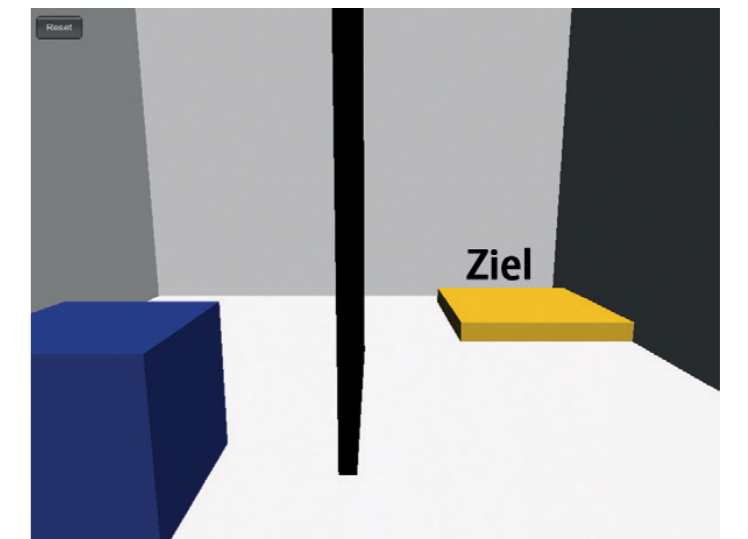
## Evaluierung

### Fingergestenerkennung

- ▶ Performance der Applikation (Framerate)
- ▶ Robustheit ermitteln (F-Score, Accuracy)
  - ▶ F-Score = Trefferquote und Genauigkeit als harmonisches Mittel
  - ▶ Accuracy = Anteil aller richtig erkannten Bilder
- ▶ Aufbau - Bezugsdaten
  - ▶ 5 Testsets mit je mehr als 100 Bilder (Hintergrund, Beleuchtung variiert)
  - ▶ Handmittelpunkt annotieren (Bezugsbilder)
- ▶ Auswertung
  - ▶ Ergebnis und Bezugsbilder vergleichen (Handmittelpunkt)
  - ▶ Richtig Positiv: Abstand < 15 Pixel

### 3D-Interaktion

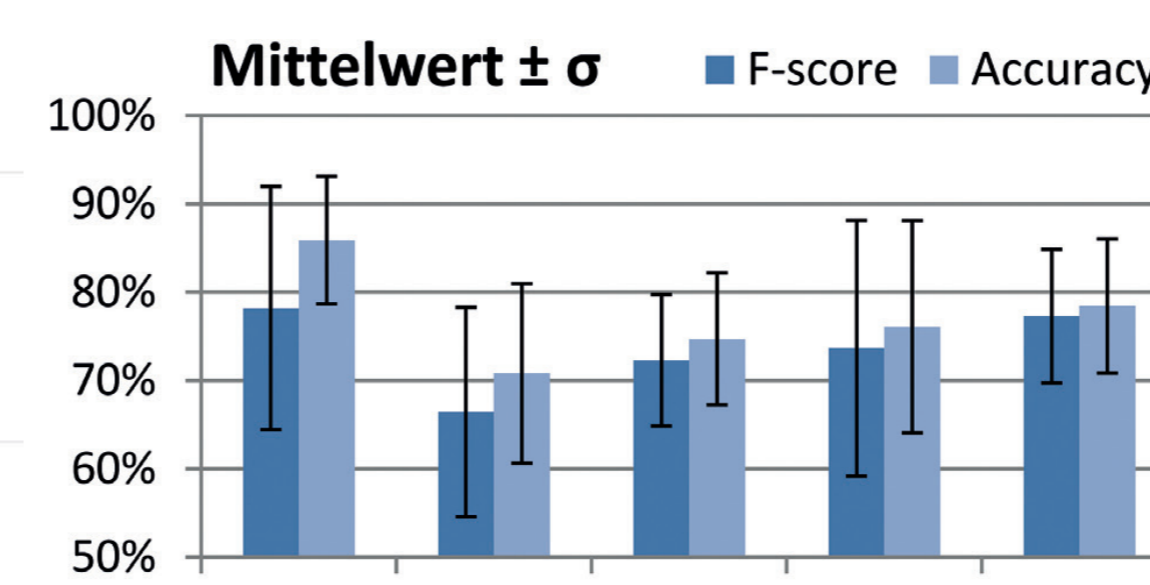
- ▶ Praktische Anwendung (Aufgabe lösen, Fingergeste)
- ▶ Methoden zur Tiefenschätzung (3D-Position)
- ▶ Testszenario
  - Aufgabe 1: Selektion des Würfels (Geste 1)
  - Aufgabe 2: Positionierung im Ziel (Geste 2)
- ▶ Aufbau
  - ▶ Experimentelle Studie mit 1 Teilnehmer
- ▶ Auswertung
  - ▶ 3D-Position während Interaktion als Funktion der Zeit



## Resultate

### Fingergestenerkennung

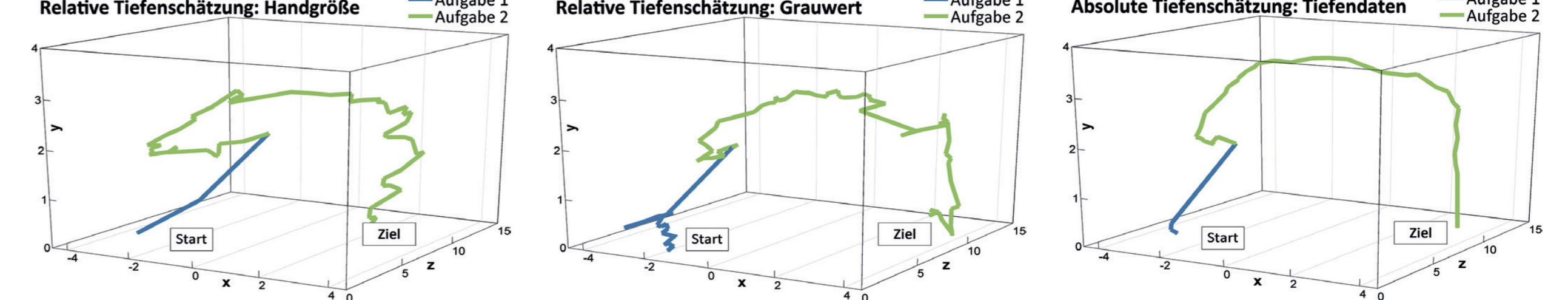
- ▶ Performance: Alle Ansätze erreichen interaktive Framerate (8,1 - 17 fps)
- ▶ Robustheit: A1 und A2-D am besten in allen Sets
- ▶ Empfehlungen



- A1: Robust, sofern Klassifikatoren gut trainiert; geringere Framerate
- A2-C: Nicht geeignet; Hand von anderen Konturen nicht unterscheidbar
- A2-H: Schnell, automatische Anpassung
- A2-T: Vor weißem Hintergrund geeignet; einfache Implementierung, schnell
- A2-D: Robuster Ansatz in allen Situationen; geringere Framerate

### 3D-Interaktion

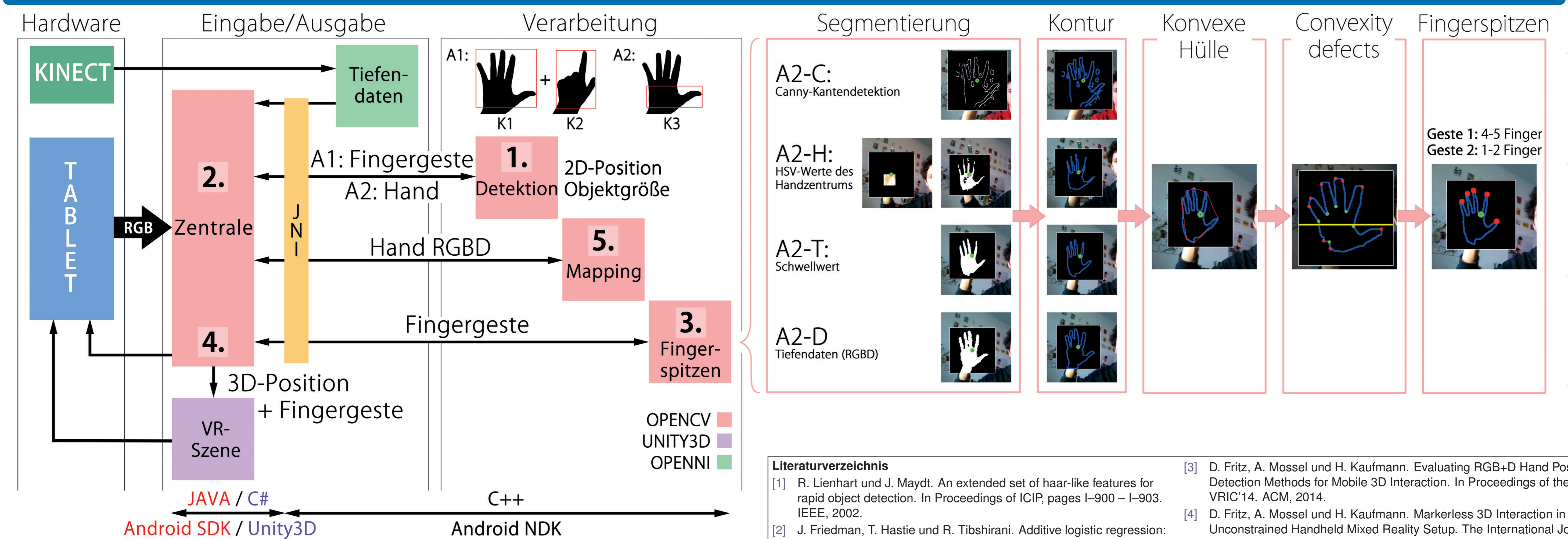
- ▶ 3D-Position während der Interaktion:



- ▶ z-Position unruhig (Klassifikator Performance)
- ▶ Abweichungen in z (Handbeleuchtung)
- ▶ Robust, kontinuierlich, keine Abweichungen

Publikationen: Teile der Ergebnisse wurden in [3, 4] veröffentlicht

## Softwareaufbau



### Literaturverzeichnis

- [1] R. Lienhart und J. Maydt. An extended set of haar-like features for rapid object detection. In Proceedings of ICIP, pages 1-900 - 1-903. IEEE, 2002.
- [2] J. Friedman, T. Hastie und R. Tibshirani. Additive logistic regression: A statistical view of boosting. Annals of Statistics, 28(2):337-407, 2000.
- [3] D. Fritz, A. Mossel und H. Kaufmann. Evaluating RGB+D Hand Posture Detection Methods for Mobile 3D Interaction. In Proceedings of the VRIC'14. ACM, 2014.
- [4] D. Fritz, A. Mossel und H. Kaufmann. Markerless 3D Interaction in an Unconstrained Handheld Mixed Reality Setup. The International Journal of Virtual Reality, 2014.

Kontakt: mail@danielfritz.at