IGKRI+VBTO

Identifikation von Gesichtern im Kontext von Kind-Roboter-Interaktion + Vision-based Tracking of Humans in Arbitrary Orientations

Oliver Ast & Jan Philip Göpfert & Katharina Klein

Supervisors: Philipp Cimiano, Katharina Rohlfing, Malte Schilling, Johannes Wienke, Britta Wrede, Sebastian Wrede

Bielefeld University, Faculty of Technology

Abstract

The automatic detection and recognition of objects, particularly human faces, inside images is a problem that has long reached the public mainstream. However, when analyzing videos, it may be desirable to not only identify objects inside individual frames, but to understand information about their presence, specifically their position in space, over time, i.e. over a series of frames. In the scope of this project, we gathered a multitude of techniques in order to tackle this problem, using the RSB[2] framework and OpenCV[1] as a basis.

Main Objectives

- Investigate how well children can be detected with existing applications.
- Detect humans in frames from videos.
- Combine the results of per-frame detection over the course of a video, thereby tracking the humans present.
- Implement this tracker using the RSX framework.
- Compare the results to existing solutions in the RSX framework.

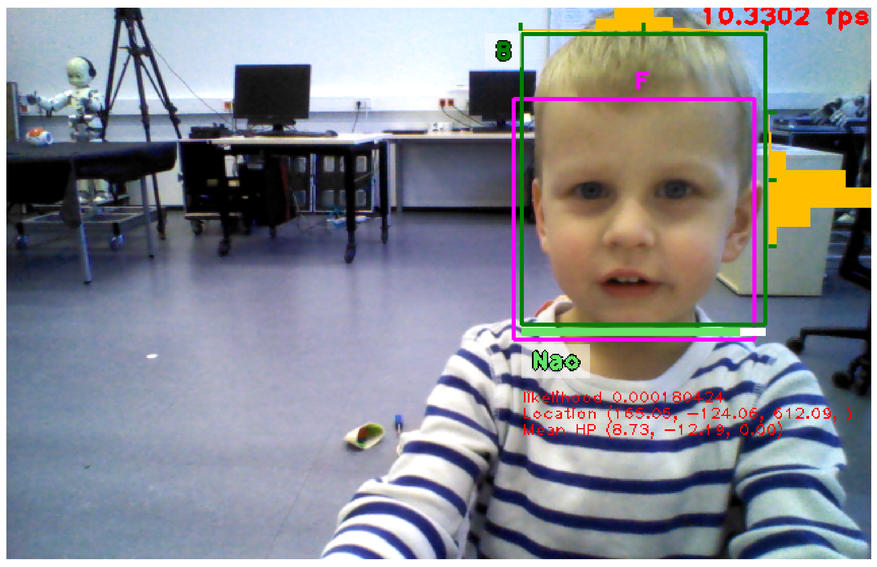


Figure 1: A 4-year-old's face, properly detected by the RSB face tracker.

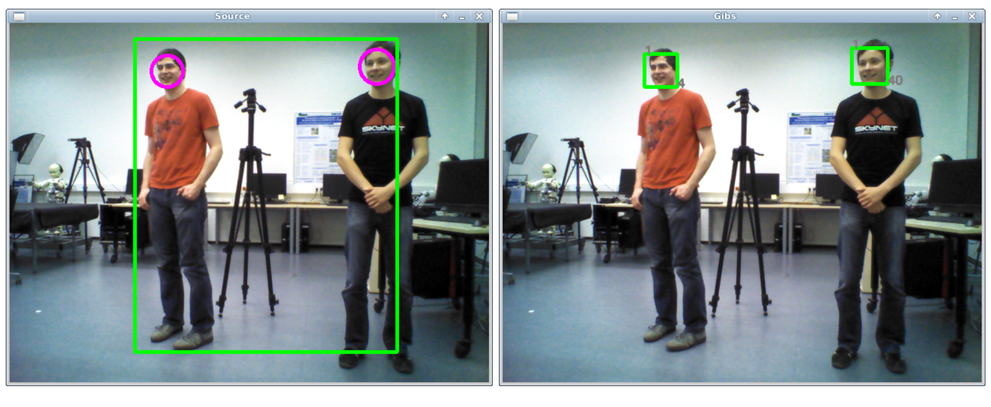
Issues

- Detectors do not always recognize faces.
- False positives.
- Detecting faces is not enough, as people do not face the camera at all times.
- Available human shape detectors (whole body or body parts, e.g. torso or back of the head) are not very advanced.
- Detectors usually work on individual frames only.
- People leave and reenter the image or disappear behind other objects.

Methods

To overcome the problems that individual detectors face, we utilize a collection of different detectors and aggregate their results. Additionally, we compare the result of any detector for a given frame to the detector's results for a number of preceding frames, in order to track the detected objects over time.

We build upon the RSB framework and use detectors supplied by the RSB framework and OpenCV.



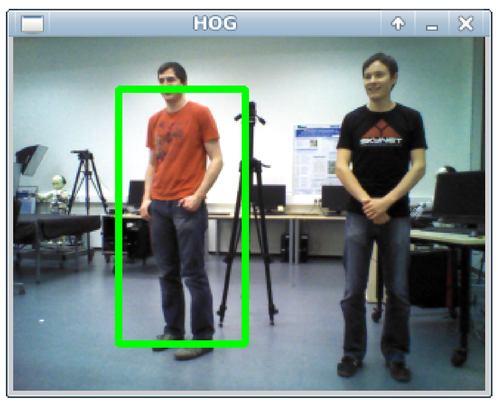


Figure 2: top left: ISY-FaceDetection, top right: RSBFacedetection, bottom: ISY-HOG (Histogram of Gradients)

Results

From OpenCV, we currently use background subtraction, Cascade Classifier[4] and the Histogram of Gradients object detector[5]. To keep in line with RSB's philosophy, we implement each detector as an individual application which then broadcasts its findings on individual RSB scopes. The aggregator is yet another RSB application and can function independently from the number of detectors running.

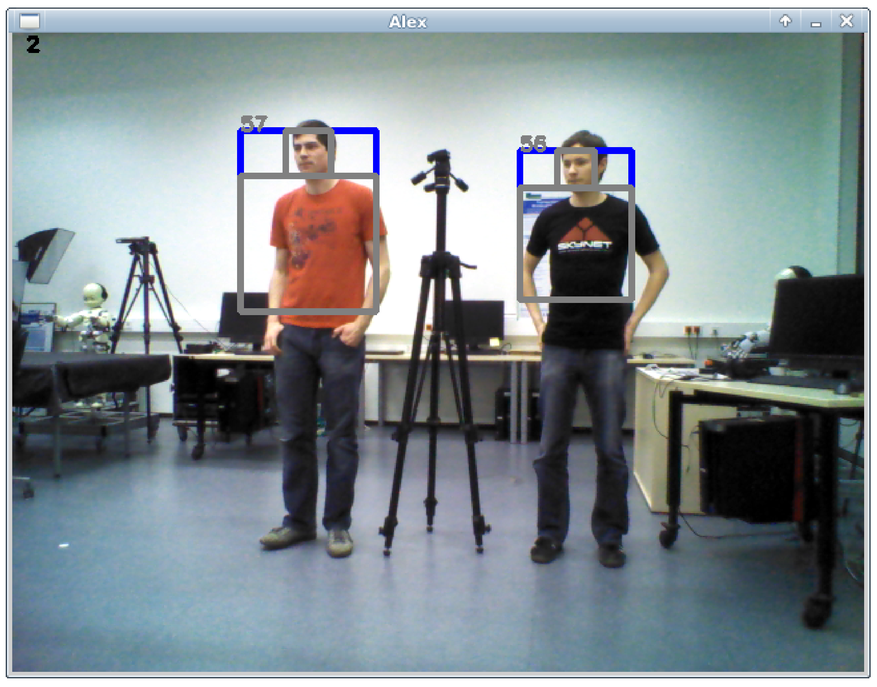


Figure 3: ISY-Tracking, combining outputs of multiple detectors.

Outlook

We want to build upon our current results and will further look into the following items:

- Activate / deactivate detectors depending on available resources.
- Assign individual trust levels and optionally respect certainties.
- Gather, combine and compare additional data, e.g. name, gender, hair color or eye color.
- Learn a person's features on the fly, thereby enabling the use of further detection methods.

Evaluation

In order to be able to properly judge the quality and compare the results of different trackers and detectors, we are looking to manually annotate video clips and then automatically determine numeric values to express said characteristics.

References

[1] OpenCV. `http://opencv.org/`, February 2014.

[2] RSB. `https://code.cor-lab.de/projects/rsb`, February 2014.

[3] Bastian Leibe, Edgar Seemann, and Bernt Schiele. Pedestrian detection in crowded scenes. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 1, pages 878–885. IEEE, 2005.

[4] OpenCV. Cascade Classification. `http://docs.opencv.org/modules/objdetect/doc/cascade_classification.html`, February 2014.

[5] OpenCV. Object Detection. `http://docs.opencv.org/modules/gpu/doc/object_detection.html`, February 2014.