Simply use the force
Implementation of RF-based gesture interaction

Christoph Rauterberg, Mathias Velten, Stephan Sigg, Xiaoming Fu
University of Göttingen

Motivation

Recent approaches were able to achieve gesture interaction using *specialised hardware*. We were aiming for *online detection of gestures* based in the RF-signal strength.

Related Work

Sigg et al. observed that 10 RSSI packets per second are sufficient to distinguish between simple classes and also hand gestures in proximity of the receiver

A showcase application

We developed two showcase applications, where the user is able to *train* certain gestures in a given scenario and is then able to *control a slideshow* with these gestures or *take a snapshot* with the camera of the mobile phone when a certain gesture is detected.

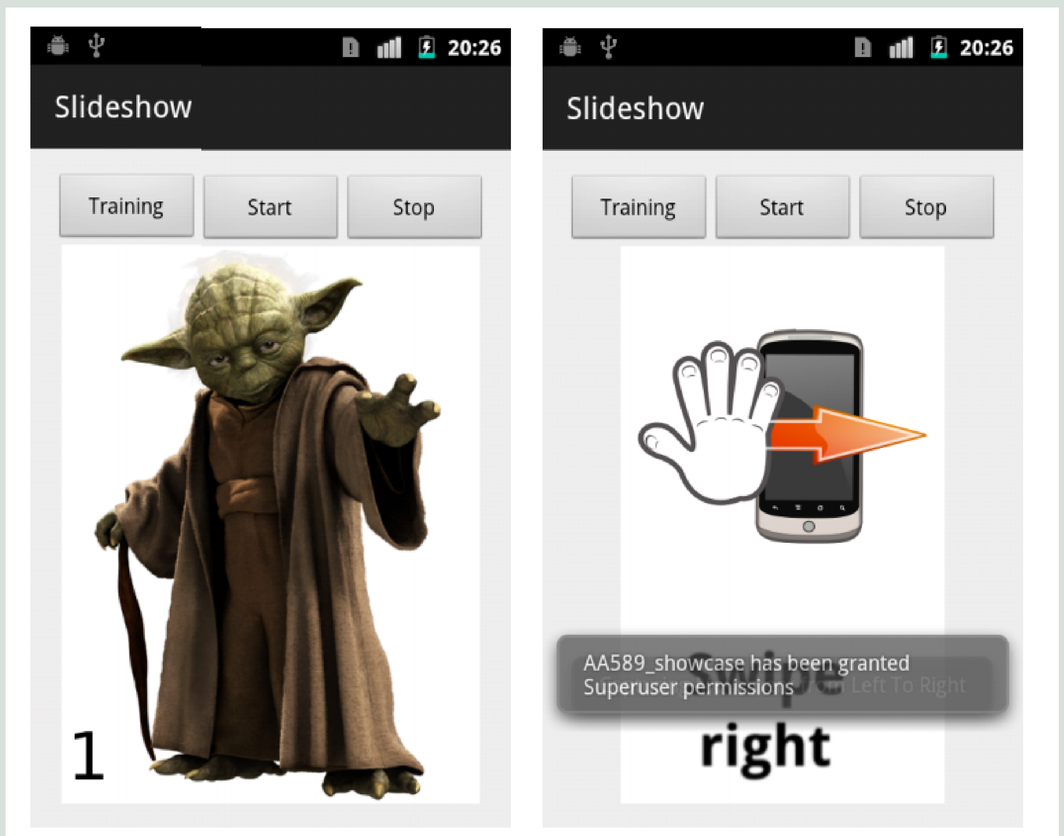

Figure 1: The UI of the showcase application. The user can swipe over the phone to display the next picture in the slideshow

Technical Details

In order to improve recognition rates compared to previous work in the literature, we employ further features from the data:

1. The user will be asked to perform movements for a period of time t. This data will be recorded and stored into a file.
2. The recorded data will be divided into n slots. Each packet is sorted into a slot according to its *time stamp*.
3. We then use statistical analysis to compute the *features* for Machine Learning. Those features are for one timeslot the *statistical mean*, *standard derivation*, *highest* recorded RSSI value and *lowest* recorded RSSI value, *total number* of packets as well as *numerical representation* of the source MAC-Addresses.
4. With these features, we use a *k-nearest-neighbour*-Classifier to classify the data.

Gesture Control

We used the following clases within the slide show:

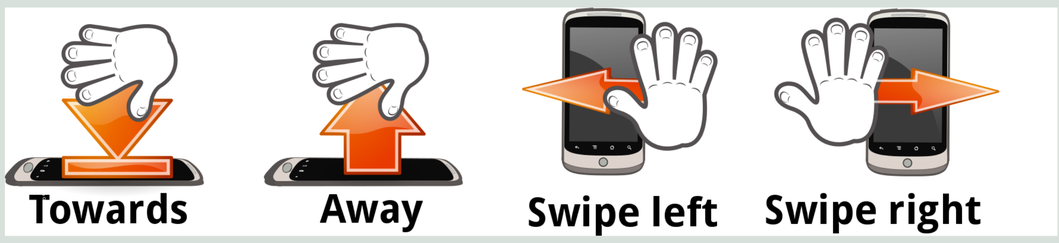

Figure 2: All gestures considered. The Showcase Slideshow uses swipes over the phone to navigate within the slideshow.

Evaluation

To evaluate the performance of the proposed method we evaluate the performance in the following scenarios:

Standard: After a training phase, while holding the phone in the hand, we record test samples to evaluate the accuracy of the classification. We will refer to this scenario as the *standard scenario*.

Training: We train the classifier whilst holding the phone *on a certain spot in a certain room* and then *walk about the room* and evaluate the classifier on different positions in the room.

Ad-hoc: We train the classifier whilst holding the phone *on a certain spot in a certain room* and then *completely leave the context* and go to other spaces within the city and evaluate the performance there.

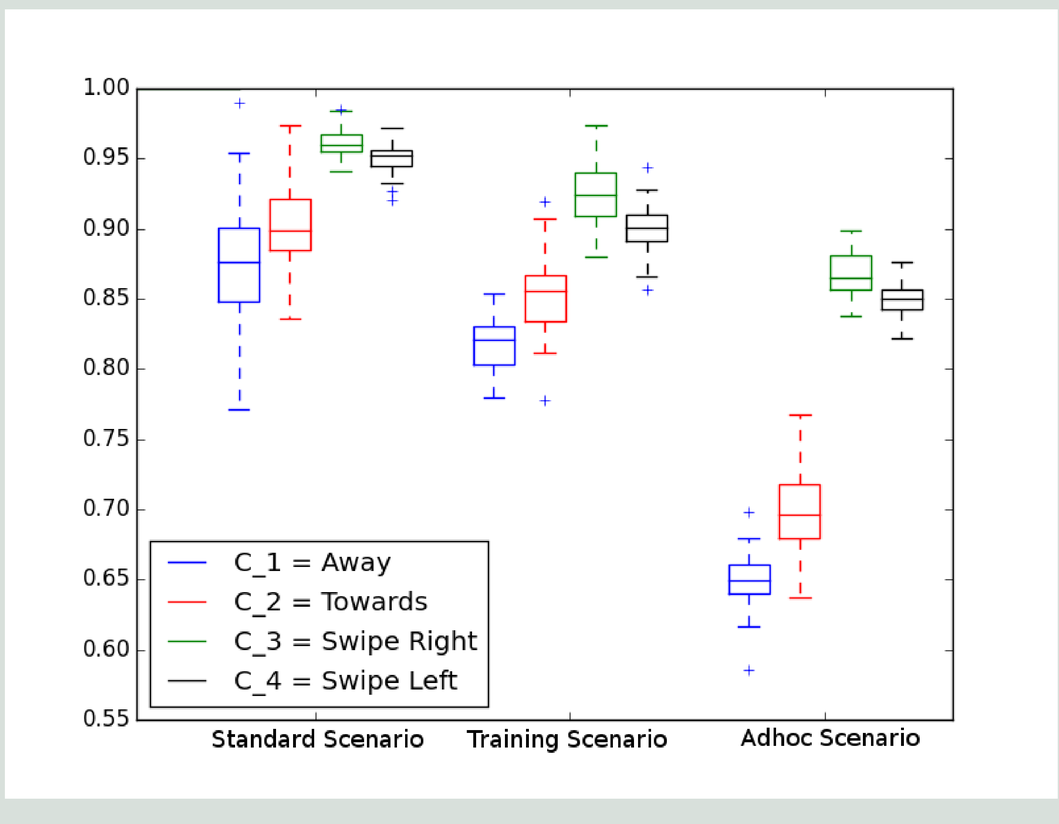

Figure 3: The results of the evaluation in the described scenarios.

Conclusion

1. We are able to detect the used gestures with an accuracy of more than **90%**.
2. The proposed method is not robust to *changes in the location*.
3. The test, that have been described above, where all contucted in an environment with *more than 10 packets per second*, as described by Sigg et al.