Real-time Hand Gesture Recognition on Unmodified Wearable Devices

Jie Song, Gábor Sörös, Fabrizio Pece, Otmar Hilliges

INTRODUCTION
We present a novel algorithm that performs real-time hand gesture classification and hand depth regression, using only an RGB camera.

Our algorithm extends the interaction space from the touch screens to around the mobile devices.

ADVANTAGES

- runs in real time on unmodified smartphones, smartwatches, smartglasses
- has significantly lower memory footprint than traditional random forest approaches
- offers great flexibility in the gesture set
- is robust to segmentation errors

CLASSIFICATION-REGRESSION PIPELINE

Our technique implements a multi-layer random forest (RF) architecture, which consists of established image processing steps interwoven with a new, staged classification-regression pipeline.

Each stage of the pipeline is trained to perform a different task. As a result, the input to each pipeline is modified by the previous stage, reducing the variation of the data passed to the subsequent stages.

CLASSIFICATION

- Segmentations
- Coarse Depth Classification
- Shape Classification
- Part Classification

REGRESSION

- Pre-processing segments hands from background
- PCA is used to find in-plane rotation of hands
- A Depth Classification Forest (DCF) coarsely estimates the depth of the hand, greatly reducing the variation in depth of frames which are forward-facing.
- A Shape Classification Forest (SCF) classifies foreground pixel into 6 gesture classes + plus noise, passing down the pipeline only the ones containing pointing-like gestures.
- A Part Classification Forest (PCF) classifies location of fingertips for pointing-like gestures. Alternatively, a Depth Regression Forest (DRF) estimates the distance of the hand from the camera.

RESULTS

- Confusion matrix (20 users).
- Left: half-test/half-train cross-validation; avg. accuracy 98%
- Right: leave-one-subject-out; avg. per-frame accuracy 93%
- Classification accuracy as function of architecture when extreme gesture depth variations occur. DCF+SCF (orange) outperforms single SCF (blue) in terms of accuracy averaged over all classes.

For more information, please visit our homepage:
http://www.ai.t.tudelft.nl/projects/2014/InAirGesture/
http://www.ai.t.tudelft.nl/projects/2015/3Dfrom2D/