A Real-time Hand Gesture Recognition System

Arun Ganesan and Caoxie (Michael) Zhang
EECS Department University of Michigan, Ann Arbor

**Introduction**

**Design goals:**
- Real-time, Just-hands, Accurate

**Methodology:**
- Machine learning as opposed to rule-based
- Train on a large amount of label data
- Use simple algorithm in prediction

**Contributions:**
- A system
- A rapid way to generate massive labeled data
- Extensive experiments
- Computational analysis of SVM vs random forest

**System Architecture**

**Per-pixel Classification**

**Feature extraction:**
- \(J(x) = d(x, y_1) - d(x, y_2)\)

Random forest for classification
- Ensemble of decision trees

**Prediction complexity:**
- \(O(d \times n_{tree})\) vs. linear SVM \(O(n_{class} \times n_{features})\)

**Use GPU for real-time prediction:**
- Massive parallelism: 307,200 threads per frame
- OpenCL: general purpose computing for GPU

**Generating Training Sets**

**Color glove:**
- An inexpensive approach to label gestures
- Map RGB pixel to depth pixel

**Experiments**

Use EC2 to train many data sets
- Largest data set > 25 GB, our demo takes 24 hours to train

**Pooling & Experience**

Propose gesture by pooling per-pixel classification

**Use clustering algorithms:**
- Kmeans
- Subject to outliers
- Fix number of clusters
- Density-based clustering
  - No need to specify the number of clusters
  - Resilient to noise

**Experience:**
- Late optimization
- Virtual wall
- Test accuracy is not enough