Emblem Detection by Tracking Facial Features

Atul Kanaujia, CBIM, Rutgers University
Yuchi Huang, CBIM, Rutgers University
Dimitris Metaxas, CBIM, Rutgers University
Understanding Non-verbal Communication

► Emblems – An event or movement that symbolizes an Idea
  ▪ Head Nodding, Shaking and Head Tilting
  ▪ “Thieves” use fewer head movements, gestures and more self-touching.
Understanding Non-verbal Communication

- Facial Expressions Analysis and Synthesis
  - 6 Universal Facial Expressions i.e. Anger, Joy, Disgust, Surprise, Sadness and Fear.

- Eye Gestures
  - Blinking (Drowsiness)
  - Gaze (Interest in Conversation)
Detecting Emblems by Tracking Facial Features

- Extend Active Shape Models to handle Localized shape deformation.
  - Expressions and AU detection
- Track facial Features across large head movement.
- Use only 2D Shape tracking.
- Demonstrate algorithm on Emblem Detection – Head Nodding, Head Shaking, eye blinking.
Why Active Shape Models?

Active Shape Models
+ Larger Capture Range Compared to AAM (Search along profile)
+ Faster convergence
+ Less affected by Illumination variations.

- Does not use grey level information
- Needs finely located landmarks
- Poor Performance with textured background

Active Appearance Model
+ Lesser Land marks needed.
+ More Robust to textured Images

- Takes Longer to converge
- Large influence due to changes in Appearance and Illumination

Courtesy: Tim Cootes et al. Comparing Active Shape Models with Active Appearance Models
Shape Subspace for Localized Deformation

- Accurate tracking of Facial expressions largely depends on the characteristics of the shape basis vector.
- Localized Shape Representation using Local Factor Analysis (LFA), Local ICA, Non-negative Matrix Factorization (NMF).
- NMF – Factorizes the observed shapes into Non-Negative Linear combination of basis vectors.
NMF Basis Vector

\[ f_i \approx \sum_{j=1}^{M} w_j h_{ji}^T = Wh_i^T \]

\[ D(F||WH) = \sum_{i=1}^{N} \sum_{j=1}^{M} [F_{ij} \log(WH)_{ij} - (WH)_{ij}] \]

\[
\begin{align*}
    h_{kj} & \leftarrow h_{kj} \frac{\sum_{i=1}^{N} w_{ik} f_{ij}}{\sum_{r=1}^{R} w_{ir} h_{rj}} \\
    w_{ik} & \leftarrow w_{ik} \frac{\sum_{j=1}^{M} f_{ij} h_{kj}}{\sum_{r=1}^{R} w_{ir} h_{rj}}
\end{align*}
\]
NMF vs PCA
(Qualitative Comparison)
NMF vs PCA (Quantitative Comparison)

- Prediction accuracy on Cohn-Kanade database on Facial Expressions.

- PCA vectors ranged from 35 – 45 (captured 98% variance)
- Trained on specific emotion
- NMF with 40 basis vectors gave consistently better results.
Handling Large Head Movement

- Aspect changes cause Shapes to lie on a Non-Linear Manifold.
- Linear Subspaces cannot model non-linearities.
- Use Multiple ASM models to learn shapes from different viewpoints.
• Train a Multi-Category Classifier to recognize head pose
  - Trained on ~2000 Images, ~400 images of each pose.
  - Images aligned by nose tip.
  - ~93% accuracy
Tracking the Features

- Active Shape Model search cannot be used at every frame.
- Tracking using SSID point tracker
  - Distorts the facial feature shapes
- Constrain the shape to lie within the NMF subspace at every frame.
Tracking Head Rotations (25 FPS)
Detecting Head Nodding

Detect Nodding and Shaking by tracking Y and X co-ordinates of the Nose Tip
Detecting Eye Blinking using Template Matching.

ROC curves obtained by varying the thresholds in the detection algorithm
- Head Nodding ROC area – 96.24%
- Head Shaking ROC area – 94.7%
- Eye Blinking ROC area – 95.46%
Tracking with Glasses and Varying Illumination
Tracking Facial Expressions

Joy

Surprise
Real time Online demonstration

Day: Wednesday
Timings: 8:30 AM to 4:30 PM
Location: Kimmel Center, Room 906