FACIAL EXPRESSION CLASSIFICATION USING VISUAL CUES AND LANGUAGE

Abhishek Kar
MOTIVATION

- Long standing problem
- Applications in HCI, indexing of videos, affective computing
- Availability of a large number of datasets
  - Extended Cohn-Kanade (CK+) Dataset
  - RU FACS Dataset
  - JAFFE
  - MMI Dataset
- Vast amount of literature available
THE PROBLEM

Image

- Angry
- Disgust
- Happy
- Fear
- Sadness
- Neutral
- Surprise
METHODOLOGY

Face detection (Viola Jones)

Feature Extraction using Gabor Filters

Dimensionality Reduction/Feature Selection

Classification
FEATURE EXTRACTION

- Face detection done on the CK+ dataset and face patches resized to 48x48
- Face patch converted into Gabor magnitude representation
- 72 Gabor filters used at 8 orientations and 9 frequencies
- Feature vector size for each image = 48x48x72 = 165888
FEATURE SELECTION/DIMENSIONALITY REDUCTION

- PCA
  - Feature vector was reduced to various dimensions between 10 and 359
  - Best dimensionality was found to be around 60.
  - Interesting to note that the Facial Action Coding System used to code various emotions has 64 action units.
  - PCA able to find rough mapping to the Action Unit intensities??
FEATURE SELECTION/DIMENSIONALITY REDUCTION

- Adaboost
  - Iterative algorithm combining a cascade of weak classifiers to classify a pattern
  - We select the best features (weak learners) obtained by Adaboost for every one versus rest classification task.
  - Final set of features – Union of all features obtained in the above step.
  - Used these set of features for further classification
CLASSIFICATION

SVM

- Used multiclass SVM (1 vs. 1) with linear kernel to classify data into 7 categories
- Used LibSVM library for Matlab
- Used multiclass SVM (1 vs. rest) approach with linear kernel
- Final decision based on margin of classification and not just voting

MAP decision with parameter estimation using MLE – Baseline classifier
**DATASET**

- Extended Cohn-Kanade CK+ Dataset
  - 593 posed sequences from 123 subjects.
  - Each sequence starts with a neutral expression and terminates with the peak expression.
  - 327 of the 593 sequences are emotion labeled
  - 7 expressions present in the database: Angry, Disgust, Fear, Happy, Sadness, Surprise, Neutral
## RESULTS

<table>
<thead>
<tr>
<th>Method (Feature Selection + Classifier)</th>
<th>Accuracy (10 fold cross validation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCA + SVM (1 vs. 1)</td>
<td>71.08%</td>
</tr>
<tr>
<td>PCA + SVM (1 vs. rest)</td>
<td>72.19%</td>
</tr>
<tr>
<td>PCA + Baseline</td>
<td>80.45%</td>
</tr>
<tr>
<td>None + SVM (1 vs. 1)</td>
<td>75.39%</td>
</tr>
<tr>
<td>None + SVM (1 vs. rest)</td>
<td>88.87%</td>
</tr>
<tr>
<td>Adaboost + SVM (1 vs. 1)</td>
<td>80.43%</td>
</tr>
<tr>
<td>Adaboost + Baseline</td>
<td>86.64%</td>
</tr>
<tr>
<td>Adaboost + SVM (1 vs. rest)</td>
<td><strong>94.72%</strong></td>
</tr>
</tbody>
</table>
## Per Emotion Accuracies

<table>
<thead>
<tr>
<th>Emotion</th>
<th>No feature selection</th>
<th>Adaboost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neutral</td>
<td>97.5%</td>
<td>98.05%</td>
</tr>
<tr>
<td>Angry</td>
<td>91.65%</td>
<td>95.26%</td>
</tr>
<tr>
<td>Disgust</td>
<td>98.04%</td>
<td>99.72%</td>
</tr>
<tr>
<td>Fear</td>
<td>96.1%</td>
<td>98.04%</td>
</tr>
<tr>
<td>Happy</td>
<td>98.6%</td>
<td>98.89%</td>
</tr>
<tr>
<td>Sadness</td>
<td>94.16%</td>
<td>94.99%</td>
</tr>
<tr>
<td>Surprise</td>
<td>97.78%</td>
<td>99.17%</td>
</tr>
</tbody>
</table>
**Comparison**

**Accuracy on CK+**

- **PC + LDA**
- **Gabor Filter + SVM**
- **Boosted LBP + SVM**
- **Gabor Filter + AdaSVM**
- **This method**

Accuracy:
RESPONSES ON VIDEOS

- Obtained English responses on 40 videos from 4 different emotion categories – Angry, Happy, Sad, Surprise
- Participants correctly identified the emotion almost all the time.
- 6 subjects – 10 responses each
- Responses transcribed into English
- Keywords observed – Distressed, Unhappy, Sad, Amazed, Extreme happiness, Frowned
- Problems
  - Posed expression dataset. Expressions don’t seem natural.
TO DO

- Try to automatically identify the keywords in the responses and figure out the correct expression
- Obtain a rough classification on the basis of responses only
- If sufficient descriptive adjectives are obtained, I will try to assign different intensities to various images and try to find a correlation between high intensity images (or low intensity) in the same expression.
REFERENCES

- Recognizing facial expression: Machine learning and application to spontaneous behavior – Bartlett et al. – CVPR 2005
- The extended Cohn-Kanade dataset (CK+): A complete dataset for action unit and emotion-specified expression – Lucey et al. – CVPRW 2010