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

Methodology

Face detection (Viola Jones)

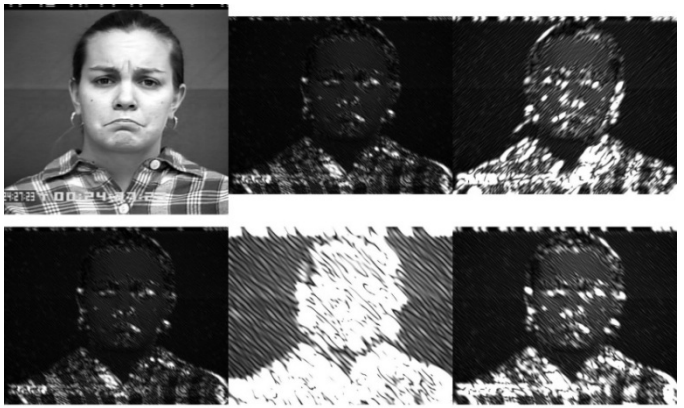
Feature Extraction using Gabor Filters

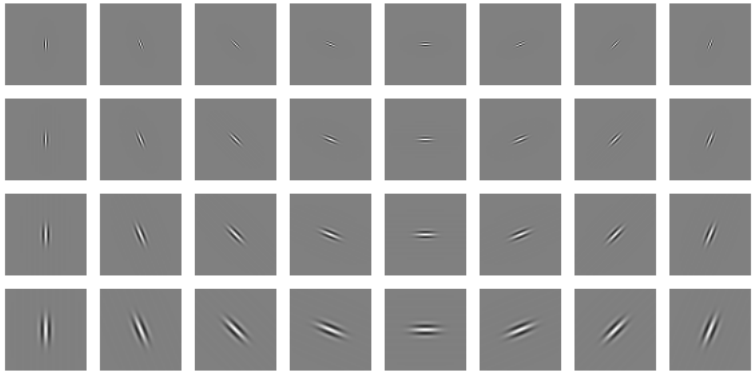
Dimensionality Reduction/Feature Selection

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

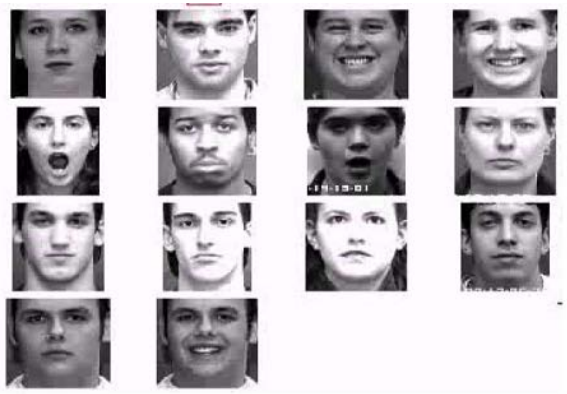


Figure 3: Images from the Cohn-Kanade Dataset

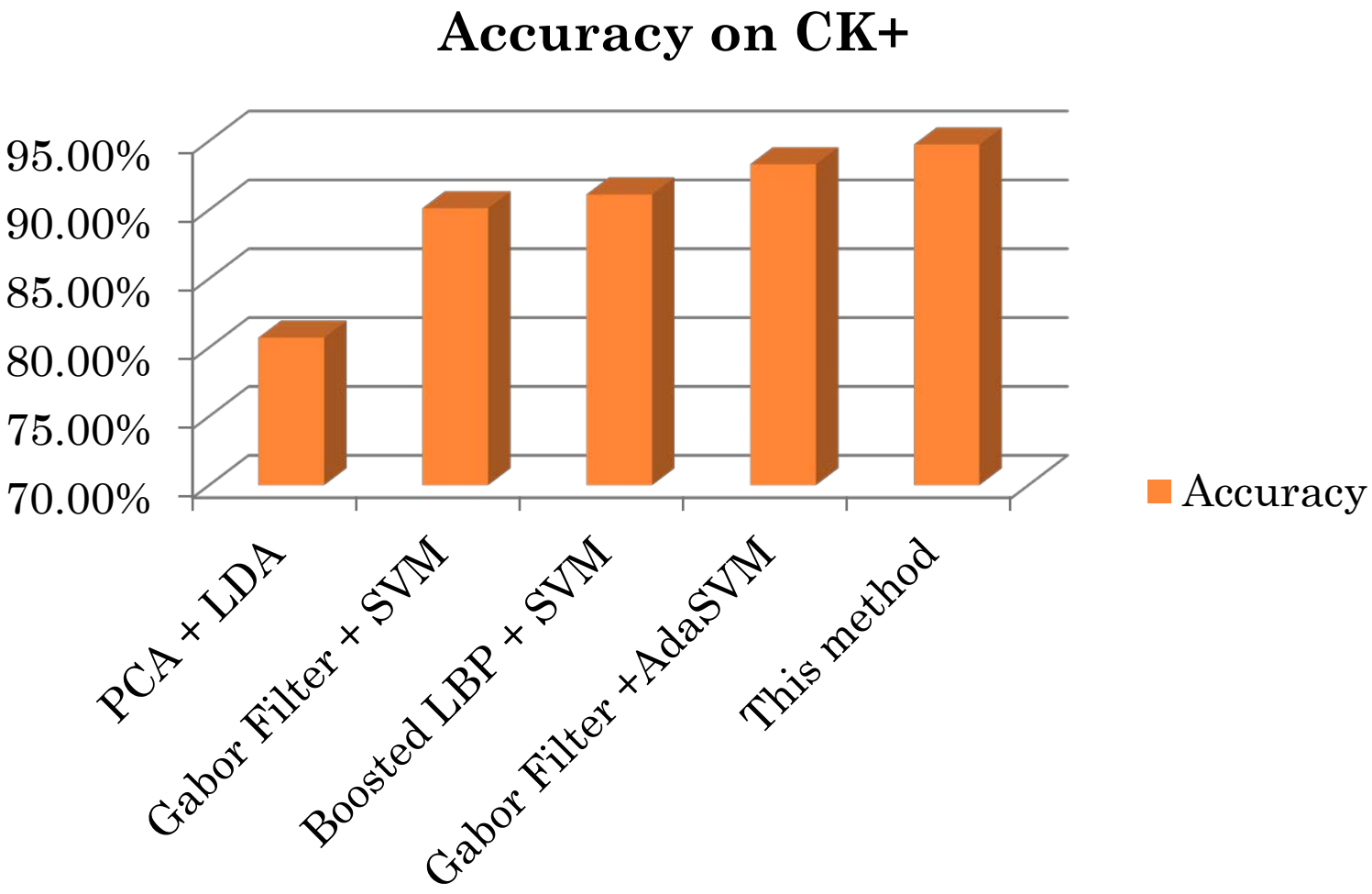
Results

Method (Feature Selection + Classifier)	Accuracy (10 fold cross validation)
PCA + SVM (1 vs. 1)	71.08%
PCA + SVM (1 vs. rest)	72.19%
PCA + Baseline	80.45%
None + SVM (1 vs. 1)	75.39%
None + SVM (1 vs. rest)	88.87%
Adaboost + SVM (1 vs. 1)	80.43%
Adaboost + Baseline	86.64%
Adaboost + SVM (1 vs. rest)	**94.72%**

Per Emotion Accuracies

Emotion	No feature selection	Adaboost
Neutral	97.5%	98.05%
Angry	91.65%	95.26%
Disgust	98.04%	99.72%
Fear	96.1%	98.04%
Happy	98.6%	98.89%
Sadness	94.16%	94.99%
Surprise	97.78%	99.17%

Comparision



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