Recognizing Faces Showing Expressions

Yaser Yacoob, Heung-Man Lam and Larry S. Davis
yaser@cs.umd.edu, lsd@umiacs.umd.edu
Computer Vision Laboratory
Center for Automation Research
University of Maryland
College Park, MD 20742-3275

To appear in International Workshop on Automatic Face- and Gesture-Recognition, Zurich, 1995

Abstract

This paper describes work in progress that evaluates the performance of alternative approaches for face recognition when applied to static and dynamic imagery of people expressing emotion. The work is carried out on a database of 130 people that includes over 125,000 frames.

1 Introduction

Face recognition continues to be a challenging problem in computer vision despite the variety of approaches proposed in recent years [BRU93,MAN91,PEN94] (for a comprehensive review of the literature see [CHE94]). Most research has focused on static imagery, often assuming frontal view and a neutral facial expression. Some approaches are biased towards recognizing males over females, or particular race and skin color.

There are many applications of face recognition in dynamic scenes such as at airports, banks, video analysis etc. to identify or verify the identity of individuals. In these cases, the expression and viewpoint of an individual are unpredictable and uncontrollable. Our experiments, with eigenfaces and graph matching, attempt to determine the utility of these two popular approaches to face recognition in dynamic environments.

Our current research on face recognition is focused on the following goals:

- To carry out extensive comparative empirical studies of both template based and feature-based approaches. The *eigenface* approach proposed in [TUR91] and the *feature-graph* based approach [MAN91] have been selected as representatives. These two approaches have been extensively tested although never under similar conditions. These algorithms were not designed to handle faces expressing emotions; therefore, our experiments should shed some light on how they perform if active faces are encountered.
- To evaluate the sensitivity of face recognition with respect to facial expression and minor head motions; these issues have not been addressed by earlier research, yet they might have an adverse impact on the performance of these algorithms. We chose to ignore gross head motions since these may be compensated by a face tracking module such as the one proposed by Black and Yacoob [BLA94].

  To address these goals we created a large database of active face images. Specifically, we intend to record image sequences of 250 people (so far 130 people are included in our database) in a laboratory environment where facial expression, minor head motions, and slight illumination changes are the only factors that vary. For each individual we record about 1300 images that include facial deformations (a total of over 300,000 images is expected).

In this paper we provide some preliminary results that will be expanded upon in the period before the workshop.

2 Overview of the approaches

In this subsection we provide a review of the two approaches employed. For a general review of feature-based approaches versus correlation approaches see [BRU93,CHE94], and for more details on each algorithm see the respective publications.

The eigenface approach measures the degree of correlation between an image and a set of images that constitute the face database. Computing this correlation is carried out by projecting the face image onto the face space that encodes the variation among the images of the database. The face space is defined by the eigenvectors of the set of face images.

The eigenface approach is a global operator. Generally, it is expected to be robust to local changes (a fact that can play against it if the similarity between the faces is high).

The most extensive testing of this approach was reported by Pentland *et. al* [PEN94]. The database included 7562 images of about 3000 people. The reported performance was around 90% correct classification on people in the database.

The feature-graph based approach proceeds in three stages. The first selects feature points using a Gabor wavelet transform that identifies points with high curvature changes. The second stage constructs a graph where features serve as nodes and the directional edges are constructed based on the distances between nodes and limited by the possible number of neighbors for each node. The third stage performs a simple graph matching algorithm. This approach was originally tested on a database of 303 images of 86 people with success rate of 86%.

The feature-graph based approach is local, and thus is sensitive to local changes while tolerant of translations.

3 Methodology

The experiments we performed are based on video-clips of volunteers at the University of Maryland. We collected pictures from all those who volunteered regardless of how able were they to express emotions in front of a camera. Subjects were requested to select any of the six principle emotions and display them as they wish (we gave no information on how expressions should be shown).

Since the two approaches for face recognition are sensitive to scale changes in the face the subjects were requested to maintain a constant distance from the camera. Nevertheless, some minor scale changes occurred; however these did not affect the performance when judged over the entire database of subjects.

4 Results of eigenface-based recognition

A major difficulty with this approach is the requirement to align the face image to the faces in the database, quite a difficult task when people are active in front of a camera. To achieve this our subjects were requested to minimize head motion while we ensured that the first image in the sequence is aligned with the data-base. Unavoidably, some minor motion still occurred when people were expressing emotions. To compensate for this, we developed an automatic alignment algorithm that uses the centroid of the gradient images to register consecutive images.

Since our database does not have all the faces aligned, we perform initially a manual alignment of all faces before starting (an imprecise process).

The computation of the eigenfaces was computed on the database of 130 subjects. One neutral expression frame was chosen from the sequence of each subject.

Figure 1 shows recognition results for several image sequences. The horizontal axes show the temporal frame and the vertical axes show the distance values of the lowest three face scores (the dot, empty circle, and the cross denote the lowest, second lowest and third lowest distances). The lower the distance values, the better the confidence of recognition. Whenever the system did not rank the correct match first the information is not plotted. If the experiments were run on static images of neutral faces, most of the scores would be well below 1000 on the scale shown.

The results in Figure 1 show segmentation of the input sequence into parts where facial expressions occur (see [YAC94] for a psychological review and further information on an automatic approach for face expression recognition). These graphs show a consistent degradation in the distance results in the presence of facial expressions. The degradation is not so large that recognition fails (i.e., that the smallest distance corresponds to the correct individual). However, the choice of a threshold for a minimal acceptable distance (needed to reject faces not in the database) becomes a sensitive parameter due to the wide variation in the values encountered.

It should be noticed that during a no-expression period the distance is not always less than 1000 since a small degree of motion during expression affects the scale and the rotation of the face. This is an inherent difficulty with the eigenface approach. Overall, these results are consistent with the conjecture that the eigenface approach is only slightly sensitive to local variations from the database [PEN94].

5 Results of feature-graph based approach

The feature-graph approach consists of two levels of processing that make the approach more sensitive to facial expressions. The selection of the feature points can be altered by the expressions due to changing the topography of the surface of the intensity image. In addition, the graph matching can be affected by these variations in the matched point sets.

Figure 2 corresponds to the same faces used in Figure 1. The recognition results are less accurate than the eigenface approach when computed over the whole sequence. The distances between the first three ranking faces is closer than the similar distances in the eigenface approach.

One advantage of this approach, however, is that it is insensitive to translation of the face in the image plane.

6 Statistical analysis

The results shown in Figures 1 and 2 emphasize the importance of the threshold value of the distance parameter used for the recognition and rejection of faces. Figures 3-5 show the reject probabilities of known faces for the feature-graph approach (top left) and eigenfaces (bottom left), the accept probabilities for unknown faces for the feature-graph approach (top center)and eigenfaces (bottom center), and the Receiver Operating Characteristic graph for the feature-graph approach (top right) and eigenfaces (bottom right). Notice the scale change on the eigenface ROC. The role the distance parameters play can be reflected in the computation of several measures

- The probability of rejecting a face that is part of the database. Figure 3 (left column) shows the rejection probability as a function of the distance threshold for the two algorithms. The graphs show similar qualitative behavior.

- The probability of accepting an unknown face as a known one. Figure 3 (center column) shows the mis-recognition probability as a function of the distance threshold for the two algorithms. Here too, we observe a similar qualitative behavior for both algorithms.

- The receiver operating characteristic graph (ROC). Figure 3 (right column) shows the tradeoff between mis-recognition and non-recognition for both algorithms. The results indicate that the eigenface approach has better behavior since it minimizes both the probability of accepting an unknown person and rejecting a known person simultaneously.

Figures 4 and 5 show the separation of the statistical results into segments with expressions and segments with neutral expressions, respectively. These figures elaborate on the role of expressions in the recognition. In both algorithms it is observed that the graphs are pushed and stretched rightward which indicates worsening of the performance on expression segments compared to neutral segments.

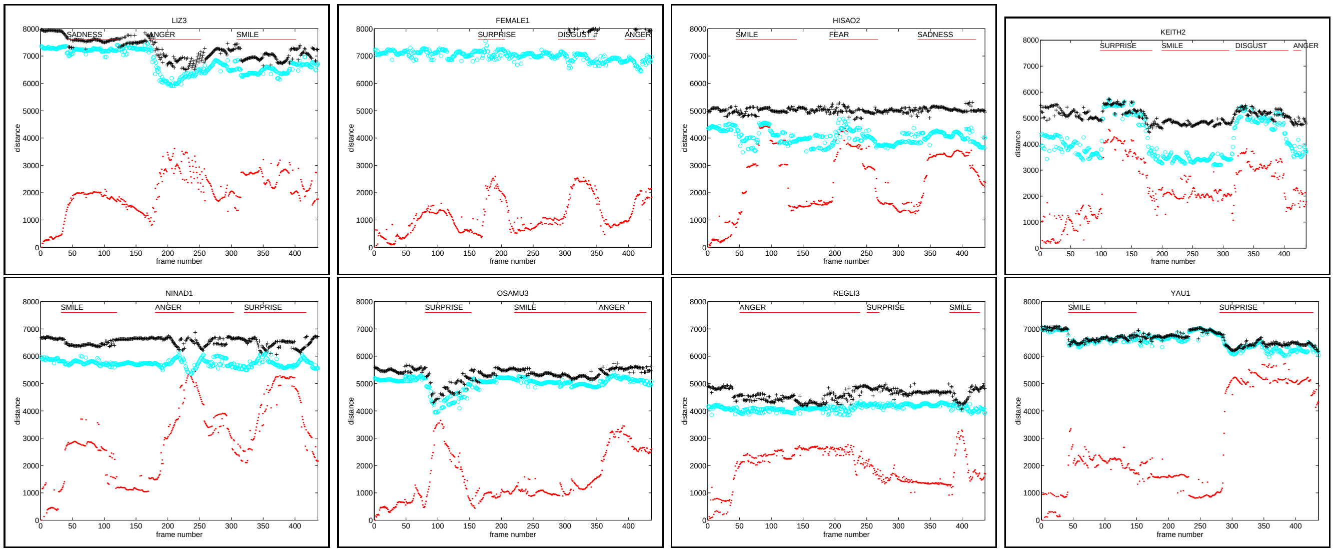

Figure 1: Results of eigenface approach on a sample of eight sequences

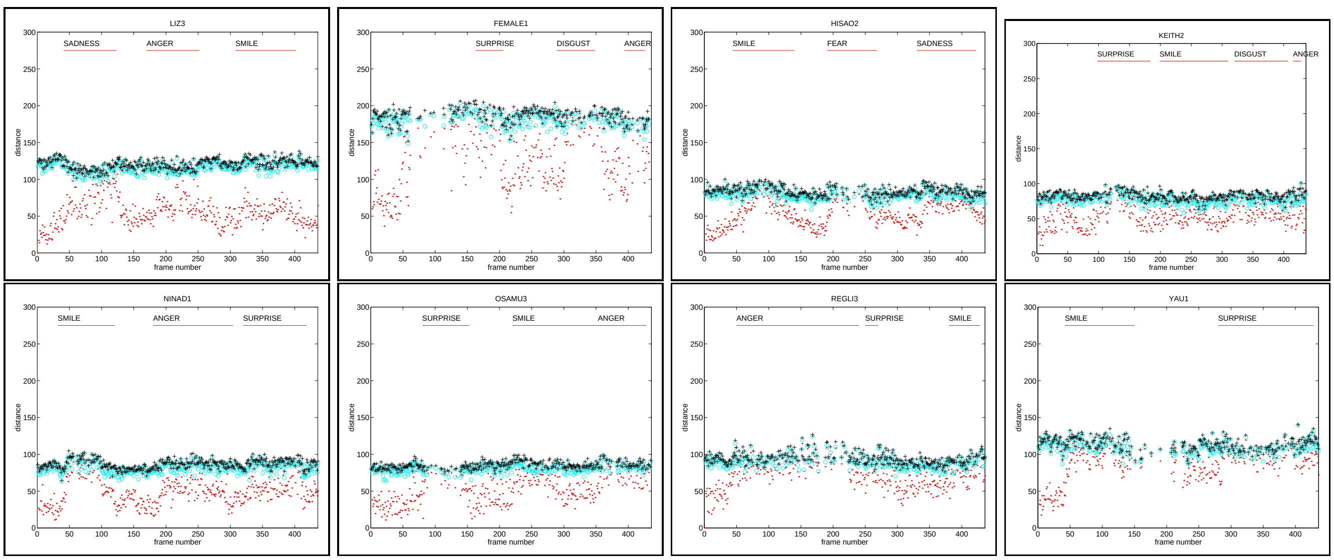

Figure 2: Results of feature-graph approach on a sample of eight sequences

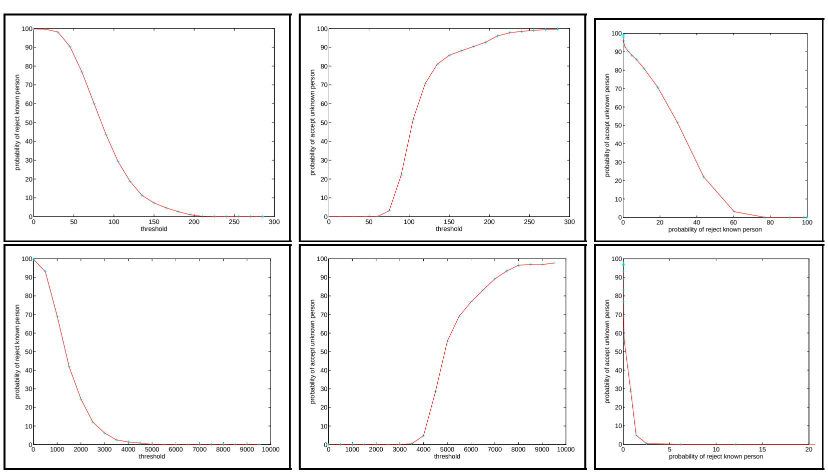

Figure 3: Statistical analysis for the entire sequence

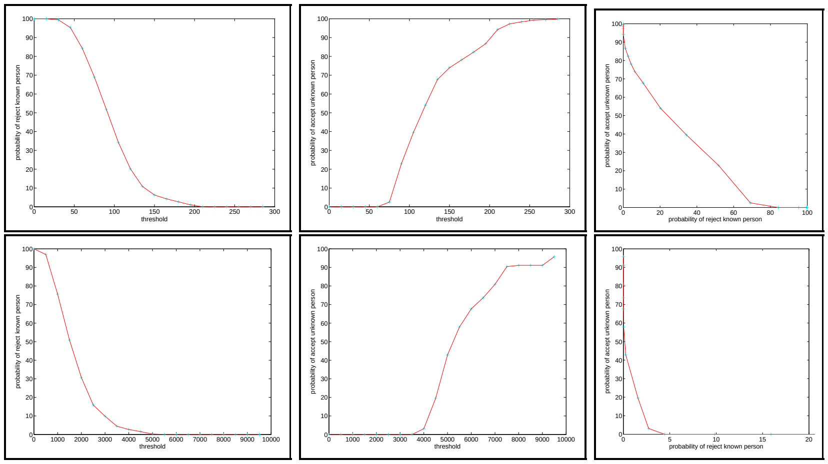

Figure 4: Statistical analysis for the segments of expressions within sequence.

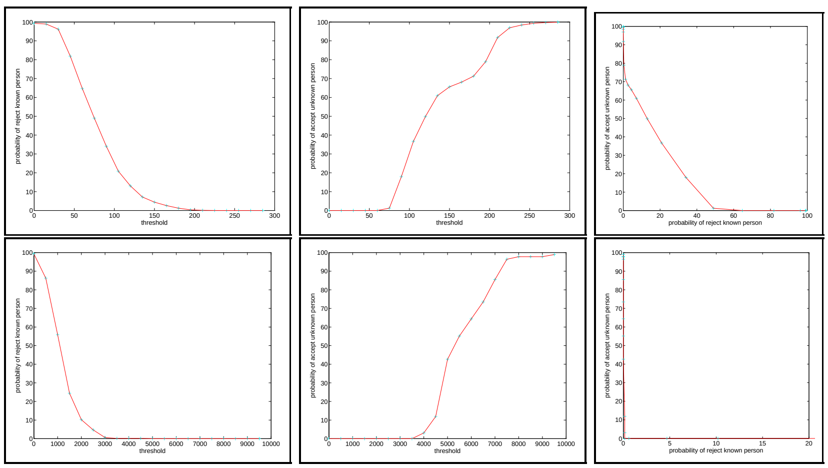

Figure 5: Statistical analysis for the segments of neutral expressions within sequence.

7 Performance by gender

Performance of face recognition approaches across gender has received little attention. The differences we expect from such comparison are due to

- Hair. The length and arrangement of the hair are a critical part of recognition. Changes in the arrangement of the hair can easily fool most face recognition algorithms. For example, for the eigenface approach, long hair would occupy a relatively large area of the correlated image. Thus, it is expected to bias the results accordingly.
- Makeup. Wearing makeup affects the appearance of the face.

Here we provide some preliminary results on the performance of the two algorithms on male and female databases. We focus on the role of the presence or absence of long hair in recognition. In later experiments we will evaluate the role of makeup and hair arrangement.

Figure 6 provides the recognition results for a male-only database of 103 subjects (top row) and a mixed database of 130 subjects (second row) for the eigenface approach. The third and fourth row provide the equivalent graphs for the feature-graph approach. The differences in performance are minor, thus suggesting that for short hair the inclusion of long hair in the database does not affect performance.

Figure 7 provides recognition results for a female-only database of 23 subjects (top row) and a mixed database of 130 subjects (second row) for the eigenface approach. The third and fourth row provide the equivalent graphs for the feature-graph approach. In the eigenface approach we observe better distances for the female-only database, thus suggesting that long hair is used in the recognition. The performance degraded somewhat when a mixed database was used since the hair became a less strong component in the recognition.

8 Conclusions

In this paper we reported preliminary results on the effect of facial expression and gender on the recognition rates of two algorithms. Our evaluation suggests that there is need to incorporate dynamic analysis of facial expressions in future face recognition systems to better recognize faces.

Acknowledgments

The authors are thankful to M. Turk, A. Pentland, M. Manjunath, R. Chellappa and C. Malsburg for providing the code of their face recognition algorithms. Also, we wish to thank S. Sirohey for assisting in porting part of the software.

References

[BLA94] M. Black and Y. Yacoob, *Tracking and*

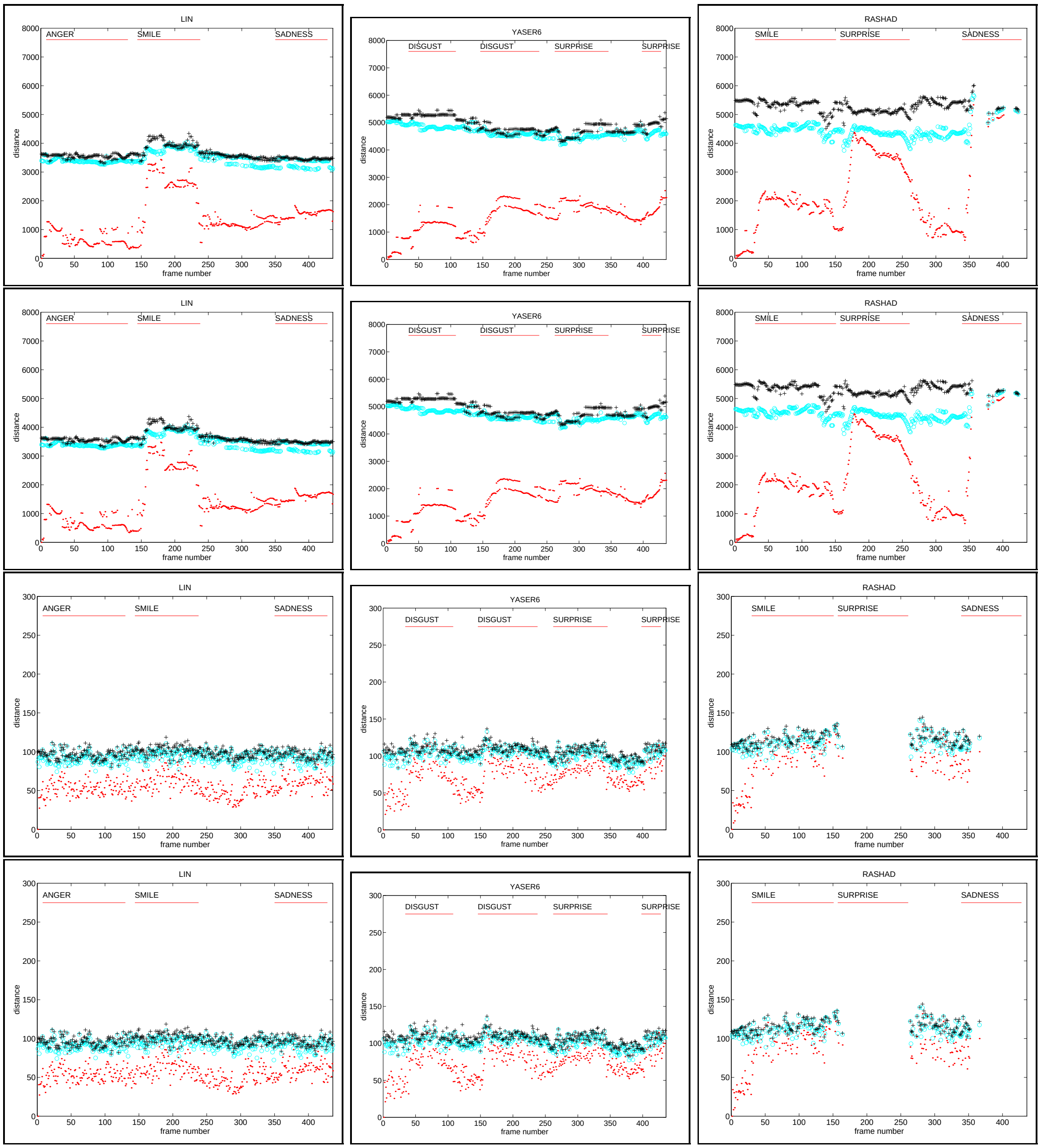

Figure 6: Males statistics

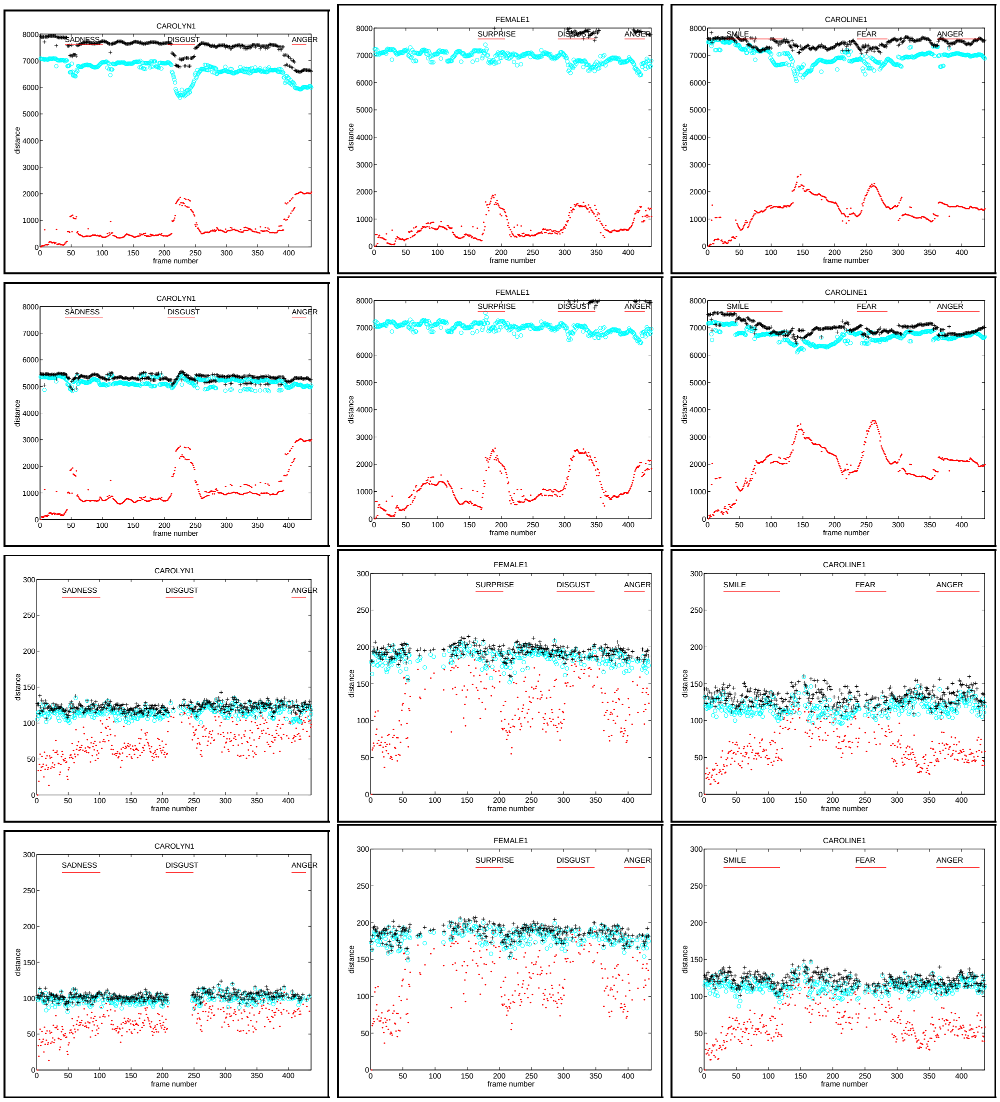

Figure 7: Females statistics

Recognizing Rigid and Non-Rigid Facial Motions using Local Parametric Models of Image Motion, in preparation.

[BRU93] R. Brunelli and T. Poggio, "Face recognition: Features versus templates," *IEEE PAMI*, Vol. 15, No. 10, 1042-1052, 1993.

[CHE94] R. Chellappa, S.A. Sirohey, C.L. Wilson, and C.S. Barnes, *Human and Machine Recognition of Faces: A Survey,* Center for Automation Research, University of Maryland, Technical Report CAR-TR-731, August, 1994.

[MAN91] M.S. Manjunath, R. Chellappa, and C.v.d. Malsburg, "A feature based approach to face recognition," *IEEE CVPR*, 373-378, 1992.

[PEN94] A. Pentland, B. Moghaddam, and T. Starner, "View-based and modular eignspaces for face recognition," *IEEE CVPR,* 84-91, 1994.

[TUR91] M. Turk and A. Pentland, "Eigenfaces for Recognition," *Journal of Cognitive Neuroscience*, Vol. 3, No. 1, 71-86, Winter 1991.

[YAC94] Y. Yacoob, and L.S. Davis, "Computing Spatio-Temporal Representations of Human Faces," *IEEE CVPR*, 70-75, 1994.