

Fuzzy-C-Mean Determines the Principle Component Pairs to Estimate the Degree of Emotion from Facial Expressions

M. Ashraful Amin¹, Nitin V. Afzulpurkar², Matthew N. Dailey³, Vatcharaporn Esichaikul¹ and Dentcho N. Batanov¹

¹Department of CS & IM, ² Department of MT & ME,
Asian Institute of Technology, Thailand
amin021us@yahoo.com, nitin@ait.ac.th, vatchara@cs.ait.ac.th,
batanov@cs.ait.ac.th

³Sirindhorn International Institute of Technology, Thammasat University, Thailand
mdailey@siit.tu.ac.th

Abstract. Although many systems exist for automatic classification of faces according to their emotional expression, these systems do not explicitly estimate the strength of given expressions. This paper describes and empirically evaluates an algorithm capable of estimating the degree to which a face expresses a given emotion. The system first aligns and normalizes an input face image, then applies a filter bank of Gabor wavelets and reduces the data's dimensionality via principal components analysis. Finally, an unsupervised Fuzzy-C-Mean clustering algorithm is employed recursively on the same set of data to find the best pair of principle components from the amount of alignment of the cluster centers on a straight line. The cluster memberships are then mapped to degrees of a facial expression (i.e. less Happy, moderately happy, and very happy). In a test on 54 previously unseen happy faces., we find an orderly mapping of faces to clusters as the subject's face moves from a neutral to very happy emotional display. Similar results are observed on 78 previously unseen surprised faces.

1 Introduction

A significant amount of research work on facial expression recognition has been performed by researchers from multiple disciplines [1], [2], [3]. In this research, we build on existing systems by applying a fuzzy clustering technique to not only determine the category of a facial expression, but to estimate its strength or degree. The clustering is also used to choose the best description of faces in a reduced dimension.

Very few researchers have considered the problem of estimating the degree or intensity of facial expressions. Kimura and Yachida [4] used the concept of a potential network on normalized facial images to recognize and estimate facial expression and its degree respectively. Pantic and Rothkrantz [5] used the famous Ekman [1] defined FACS (Facial Action Coding System) to determine facial expression and its intensity.

2 The Facial Expression Degree Estimation System

Our implementation of the facial expression recognition & degree estimation system involves four major steps (Fig-1).

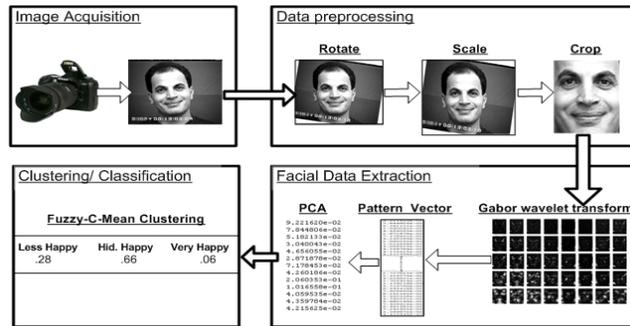


Fig. 1. The facial expression degree estimation system

2.1 Facial Data Acquisition

The training and testing data for our experimental facial expression recognition & degree estimation system is collected from the Cohn-Kanade AU-Coded Facial Expression Database [6].

2.2 Data Preprocessing

Two main issues in image processing will affect the recognition results: the brightness distribution of the facial images and facial geometric correspondence to keep face size constant across subjects. To ensure the above-mentioned criterions in facial expression images, an affine transformation (rotation, scaling and translation) is used to normalize the face geometric position and maintain face magnification invariance and also to ensure that gray values of each face have close geometric correspondence [7].

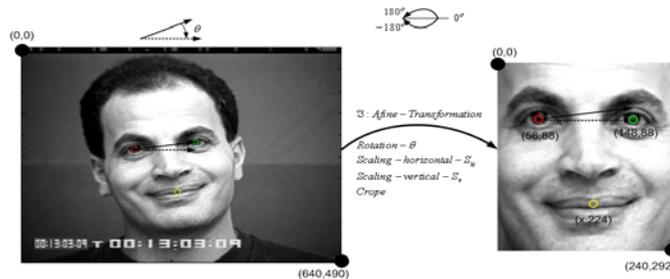


Fig. 2. Normalization of Face image using the affine-transformation \mathfrak{A}

2.3 Facial Data Extraction

The Gabor wavelets, whose kernels are similar to the 2D receptive field profiles of the mammalian cortical simple cells, exhibit desirable characteristics of spatial locality and orientation selectivity and are optimally localized in the space and frequency domains [8]. The Gabor wavelet (kernel, filter) can be defined as follows:

$$\psi_{\mu,\nu}(z) = \frac{\|k_{\mu,\nu}\|^2}{\sigma^2} e^{-\frac{\|k_{\mu,\nu}\|^2 \|z\|^2}{2\sigma^2}} \left[e^{ik_{\mu,\nu}z} - e^{-\frac{\sigma^2}{2}} \right] \quad (1)$$

where μ and ν define the orientation and scale of the Gabor kernel, $z = (x, y)$, $\|\bullet\|$ denotes the Euclidean norm operator, and wave vector $k_{\mu,\nu}$ is defined as follows:

$$k_{\mu,\nu} = k_\nu e^{i\phi_\mu} \quad (2)$$

where $k_\nu = k_{\max} / f^\nu$ (here $\nu = \{0,1,\dots,4\}$) and $\phi_\mu = \pi\mu/8$ (here $\mu = \{0,1,2,\dots,7\}$), here k_{\max} is the maximum frequency, and f is the spacing factor between kernels in frequency domain [9]. We employ a lattice of phase-invariant filters at five scales, ranging between 16 and 96 pixels in width, and eight orientations, 0 to $7\pi/8$.



Fig. 3. The magnitude of Gabor kernels at five different scales

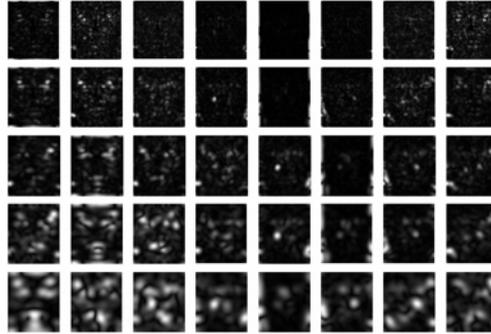


Fig. 4. Gabor magnitude representation of the face from Figure-2

$$\sigma = 2\pi, k_{\max} = \pi/2, \text{ and } f = \sqrt{2}$$

Principle Component Analysis (PCA): Even with sub-sampling, the dimensionality of our Gabor feature vector is much larger to classify. Principal components analysis is a simple statistical method to reduce dimensionality while minimizing mean squared reconstruction error [11].

2.4 Clustering and Classification

Fuzzy C-Means (FCM) [10] is one of the most commonly used fuzzy clustering techniques for different degree estimation problems. Its strength over the famous k-Means algorithm [11] is that, given an input point, it yields the point's membership value in each of the classes. In one dimension, we would expect the technique to yield a membership function as shown in Fig. 5.

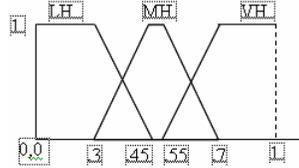


Fig. 5. An ideal membership function for the degree of an emotion

The aim of FCM is to find cluster centers (centroids) that minimize a dissimilarity function. The membership matrix U is randomly initialized as:

$$\sum_{i=1}^c u_{ij} = 1, \text{ for } \forall j = 1, \dots, n \quad (3)$$

The dissimilarity function that is used in FCM is given as:

$$J(U, c_1, c_2, \dots, c_c) = \sum_{i=1}^c J_i = \sum_{i=1}^c \sum_{j=1}^n u_{ij}^m d_{ij}^2 \quad (4)$$

Here, $u_{ij} \in [0, 1]$, c_i is the centroid of i^{th} cluster, d_{ij} is the Euclidian distance between i^{th} centroid and j^{th} data point and $m \in [1, \infty]$ is a weighting exponent.

To reach a minimum of dissimilarity function there are two conditions. These are given in Equation (5) and (6).

$$c_i = \frac{\sum_{j=1}^n u_{ij}^m x_j}{\sum_{j=1}^n u_{ij}^m} \quad (5)$$

$$u_{ij} = \frac{1}{\sum_{k=1}^c \left(\frac{d_{ij}}{d_{kj}} \right)^{2/(m-1)}} \quad (6)$$

Pseudo code for Fuzzy-C-Means follows:

- I.** Randomly initialize the membership matrix (U) that has constraints in Equation (3).
- II.** Calculate centroids (c_i) by using Equation (5).
- III.** Compute dissimilarity between centroids and data points using Equation (4). Stop if its improvement over previous iteration is below a threshold.
- IV.** Compute a new U using Equation (6) Go to Step 2.

3. Results & Observations

3.1 Neutral-Happy Faces

We applied the above-mentioned steps on 945 facial images which are classified as neutral-happy sequences portrayed by 50 actors. Initially we divided the data into two random groups to use in the PCA stage: 200 faces are used to compute the covariance matrix, and then the remaining 745 faces are projected onto this covariance matrix's principal components.

Out of these 745 faces we held out 4 randomly-selected subjects (54 faces) to test our clustering approach. We used FCM to cluster the remaining 691 faces into 3 fuzzy clusters. From this clustering we obtained the clusters. In Fig. 6., we present the scatter plot of the first three clustering, where the cluster centers are the black dots.

In the clustering process we used different combinations of principal components. We considered the principle components in groups of two in sequence: (1st, 2nd), (2nd, 3rd), (3rd, 4th), (4th, 5th) and so on in different combinations until we get a satisfactory result. Later in this chapter we describe what satisfactory result means.

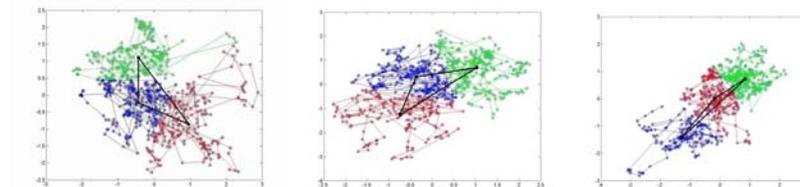


Fig. 6. Plot for principle component pairs; (1st,2nd),(2nd,3rd),(3rd,4th) of Neutral-Happy faces.

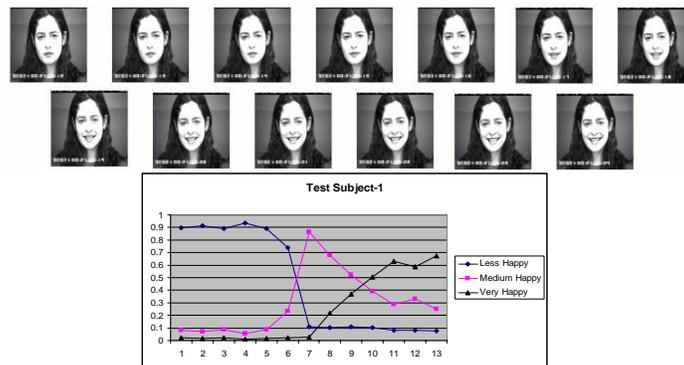


Fig.7. Neutral-Happy photo sequence of Test Example-1; Plot of membership values.

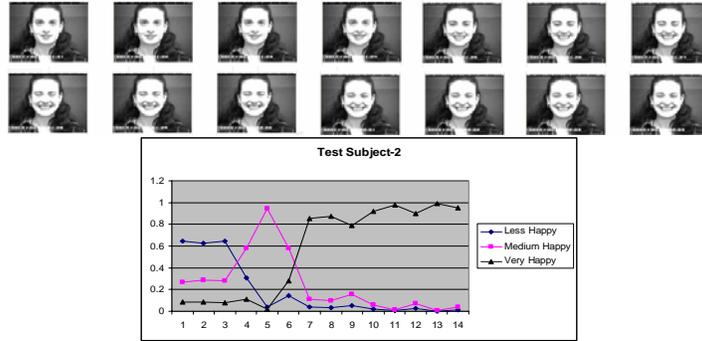


Fig.8. Neutral-Happy photo sequence of Test Example-2; Plot of membership values

Two individual subjects Neutral-Happy facial sequences are presented in Fig.7. and Fig.8., along with the membership values for each facial image that are projected using the (3rd, 4th) principle component pairs. Few interesting points from these two figures need to be noticed, one is that the membership function is similar to the ideal trapezoidal fuzzy membership function given in Fig. 5.. The other observation is that when we use the winner-take-all criteria to assign the absolute membership; the first person goes slowly to the maximum intensity on the other hand the second individual remains longer in the maximum intensity (Table 1). Similar results are presented on Fig.9. (Due to space constrains the image sequence is skipped).

Winner-take-all strategy is applied on all the fuzzy membership values for all 4 subjects for (3rd, 4th) principle component and the result is provided in Table 1. If any face is classified as less where it is appearing after the sequence medium then it is consider as an uneven assignment of class. It absolutely follows the desired sequencing for a fuzzy clustering in the fuzzy membership representation and also reflected in the absolute class assignment.

Test subjects are viewed in trajectory plotted using the 3rd and 4th principle component cluster center. Here notice that, the individuals faces are moving from near of one cluster center to the other cluster centers in Fig.10.. This again proves that the system is able to correctly capture the fuzzy characteristic that is embedded in the degree estimation process of facial expression.

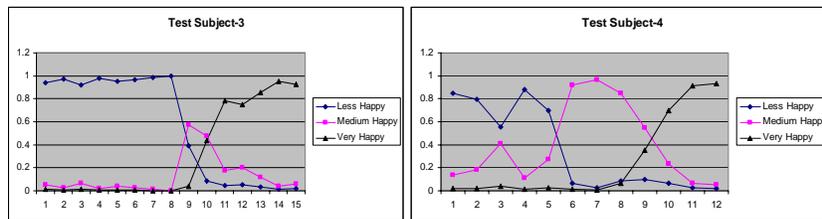


Fig.9. Plot of Neutral-Happy sequence of Test Example-3 (left) & Example-4 (right) membership values

Table 1. Absolute membership assigned to each image for Neutral-Happy sequence

Subject	Color in Fig. 11.	Less Happy (LH)	Medium Happy (MH)	Very Happy (VH)
Test Subject-1	Green	1-6	7-9	10-13
Test Subject-2	Blue	1-3	4-6	7-14
Test Subject-3	Red	1-8	9-10	11-15
Test Subject-4	Black	1-5	6-9	10-12

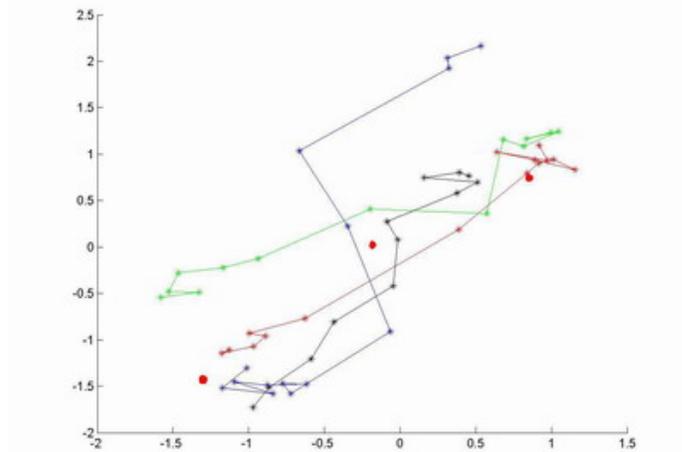


Fig.10. Neutral-Happy (3rd, 4th principle component) trajectory of the 4 test individuals.

3.2 Neutral-Surprise Faces

We applied the similar steps on 1173 facial images which are classified as Neutral-Surprise sequences as portrayed by 63 actors. This time 209 faces are kept for PCA. The remaining 964 faces are projected onto this covariance matrix's principal components. Out of these 964 faces we held out 4 randomly-selected subjects (72 faces) to test our clustering approach. Cluster center is calculated using other 892 faces (Fig. 11.). The best clustering is achieved for 2nd, 3rd principle components.

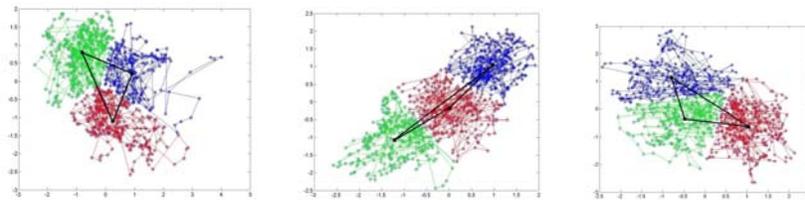


Fig. 11. Plot for (1st, 2nd), (2nd, 3rd), (3rd, 4th) principle components of Neutral-Surprise faces.

In Fig.12. and Fig.13. notice that the membership curve is rather an x-curve then a trapezoidal one. But also notice that it still follows the concept of fuzzy class assignment as only two consecutive classes are present for one individual.

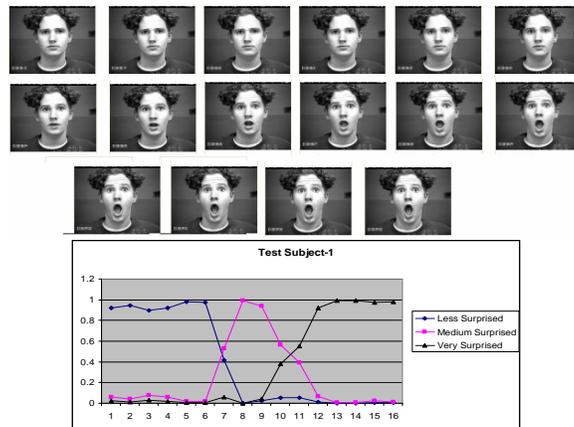


Fig.12. Neutral-Surprise photo sequence of Test Example-1; Plot of membership values

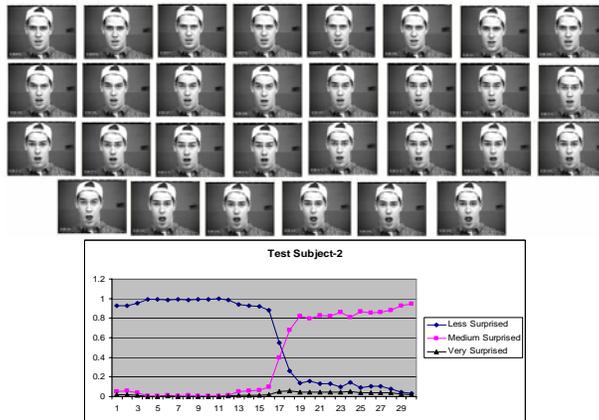


Fig.13. Neutral-Surprised photo sequence of Test Example-2; Plot of membership values.

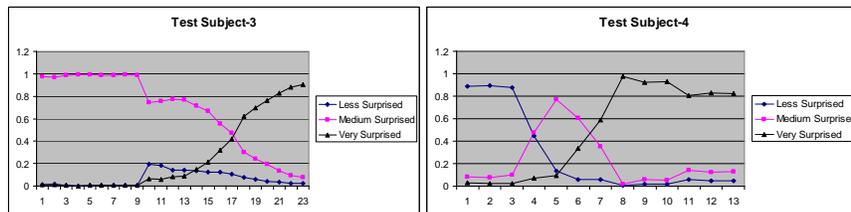


Fig.14. Plot of Neutral-Surprise sequence of Test Example-3 (left) & Example-4 (right) membership values.

Table 2. Absolute membership assigned to each image for Neutral-Surprised sequence

Subject	Color in Fig. 17.	Less Surprise (LS)	Medium Surprise (MS)	Very Surprise (VS)
Test Subject-1	Green	1-6	7-10	11-16
Test Subject-2	Blue	1-17	18-30	-----
Test Subject-3	Red	-----	1-17	18-23
Test Subject-4	Black	1-3	4-6	7-13

Notice that in Table 2. for subject-2 the Very Surprised category is empty. And for the 3rd subject the Less Surprised is empty, which also is trivial from the membership curves. More over the result could be more suitably projected from Fig.15., here notice that two subject ends their trajectory substantially earlier.

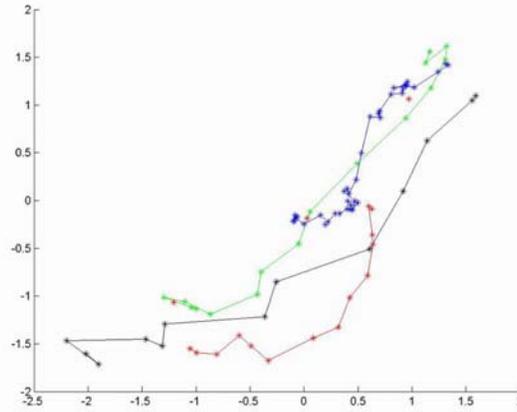


Fig.15. Neutral-Surprise (2nd, 3rd principle component) trajectory of the 4 test individuals.

3.3 Best Cluster Criterion

From the above example and evidences it is clear that our proposed system works satisfactorily. But interesting point to be noted is that how to find out the principle components that best captures the fuzzyness of data. As we could see the principle component pair for different emotion is different (Happy 3rd, 4th, Surprised 2nd, 3rd). We had to check as many of the combination possible.

The best pair of principle components is chosen depending on **minimum distance criterion (MDC)**. The criterion is, out of all combinations of principle component pairs, we cluster them using the similar initial condition and record the distance of the middle cluster center from the middle of the other two centers connecting line. This is provided in the following equation:

$$D = \sqrt{(x_2 - (x_1 + x_3/2))^2 + (y_2 - (y_1 + y_3/2))^2} \quad (7)$$

Where (x_1, y_1) , (x_2, y_2) and (x_3, y_3) are cluster centers and D is the distance of the middle cluster center from the midpoint of the connecting line of other two.

4. Conclusion & Future Work

Here we have shown that fuzzy clustering is a promising approach to estimating the degree of intensity of a facial expression, when a face is characterized with Gabor kernels and projected into a low-dimensional space using PCA. The best result is achieved when (3rd,4th) and (2nd, 3rd) principal components are used to describe the Neutral-Happy and Neutral-Surprise faces consecutively. The best suitable principle component is selected depending on the MDC (Minimum Distance Criteria). Satisfactory results are observed in the experimentation process.

Presently we are doing experiment on other prototypic emotions using the same approach and in future we will expand this to more sophisticated facial expressions that are already known as hard problem to characterize with computers.

Reference

1. Ekman, P. and Friesen, W.: The Facial Action Coding System. Consulting Psychologists Press, San Francisco, USA, 1978.
2. Pantic, M. and Rothkrantz, L. J. M.: Automatic Analysis of Facial Expression: the State of Art, IEEE Trans. Pattern analysis and machine intelligence, Vol. 24, NO. 1, 2000, 1424-1445.
3. Cowic, R., Douglas, E., Tsapatsoulis, N., Votsis, G., Kollias, S., Fellenz, W. and Taylor, J. G.: EMOTION RECOGNITION in Human-Computer Interaction IEEE Signal Processing Magazine, no. 1, 2001, 32-80.
4. Kimura, S. and Yachida, M.: Facial Expression Recognition and Its Degree Estimation, Proc. Computer Vision and Pattern Recognition, 1997, 295-300.
5. Pantic, M. and Rothkrantz, L. J.M.: An Expert System for Recognition of Facial Actions and their Intensity, Proc. 12th International Conference on Innovative Applications of Artificial Intelligence, 2000, 1026-1033.
6. http://vasc.ri.cmu.edu/idb/html/face/facial_expression/, 15 May 2004.
7. Jain, A. K.: Fundamentals of Digital Image Processing, Prentice-Hall of India Private Limited, New Delhi, India, 2003.
8. Daugman, J. G.: Complete Discrete 2-D Gabor Transform by Neural Networks for Image Analysis and Compression, IEEE Trans. Acoustics, Speech and Signal Processing, vol. 36, no. 7, 1988, 1169-1179.
9. Lee, T. S.: Image Representation Using 2D Gabor Wavelets, IEEE Trans. PAMI, Vol. 18, no. 10, 1996, 959-971.
10. Höppner, F., Klawonn, F., Kruse, R. and Runkler, T.: Fuzzy Cluster Analysis, Wiley, 1999.
11. Dubes, R. C.: Cluster analysis and related issues, Handbook of Pattern Recognition & Computer Vision, Chen, C. H., Pau, L. F., and Wang, P. S. P. (Eds.): World Scientific Publishing Co., Inc., River Edge, NJ, 3-32.