

Fiducial Point Location Algorithm for Automatic Facial Expression Recognition

D. Malathi¹, A. Mathangopi², Dr. D. Rajinigrinath³

¹PG Student, ²Assistant Professor, ³HOD, Department of CSE,
^{1,2,3}Sri Muthukumar Institute of Technology, Chikkarayapuram, Chennai, Tamil Nadu, India

How to cite this paper: D. Malathi | A. Mathangopi | Dr. D. Rajinigrinath "Fiducial Point Location Algorithm for Automatic Facial Expression Recognition" Published in International Journal of Trend in Scientific Research and Development (ijtsrd), ISSN: 2456-6470, Volume-3 | Issue-3, April 2019, pp.290-293, URL: <http://www.ijtsrd.com/papers/ijtsrd21754.pdf>



IJTSRD21754

ABSTRACT

We present an algorithm for the automatic recognition of facial features for color images of either frontal or rotated human faces. The algorithm first identifies the sub-images containing each feature, afterwards, it processes them separately to extract the characteristic fiducial points. Then Calculate the Euclidean distances between the center of gravity coordinate and the annotated fiducial points' coordinates of the face image. A system that performs these operations accurately and in real time would form a big step in achieving a human-like interaction between man and machine. This paper surveys the past work in solving these problems. The features are looked for in down-sampled images, the fiducial points are identified in the high resolution ones. Experiments indicate that our proposed method can obtain good classification accuracy.

Keywords: Pattern Analysis and Recognition, Facial Feature Detection, Face Recognition, Image Processing

Copyright © 2019 by author(s) and International Journal of Trend in Scientific Research and Development Journal. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0) (<http://creativecommons.org/licenses/by/4.0>)



I. INTRODUCTION

The algorithms reported in literature can be classified into color-based and shape-based. The first class of methods characterizes the face and each feature with a certain combination of colors [4]. This is a low-cost approach, but, not very robust. The shape-based approaches look for specific shapes in the image adopting either template matching (with deformable templates [5] or not [6]), graph matching [7], snakes [8], or the Hough transform [9]. Although these methods give good results, they are computationally expensive and they often work only under restricted assumptions (regarding the head position and the illumination conditions).

In this paper we describe a technique which uses both color and shape information to automatically identify a set of feature fiducial points with great reliability. Results on a database of 200 color images, taken at different orientations, illumination conditions and resolution are reported and discussed.

The feature extraction methods are commonly divided into two major categories: texture features (i.e. Gray intensity [1], Discrete Cosine Transform (DCT) [2], Local Binary mode (LBP) [3], Gabor Filter output [4], etc) and geometric features (i.e. Facial feature lines [5], FAUS [6], Active Shape Model (ASM) [7], etc).

The terms "face-to-face" and "interface" indicate that the face plays an essential role in interpersonal communication. The face is the mean to identify other members of the species, to interpret what has been said by the means of lipreading, and to understand someone's emotional state and intentions on the basis of the shown facial expression. Personality, attractiveness, age, and gender can also be seen from someone's face. Considerable research in social psychology has also shown that facial expressions help coordinate conversation [4], [22], and have considerably more effect on whether a listener feels liked or disliked than the speaker's spoken words [15]. Mehrabian indicated that the verbal part (i.e., spoken words) of a message contributes only for 7 percent to the effect of the message as a whole, the vocal part (e.g., voice intonation) contributes for 38 percent, while facial expression of the speaker contributes for 55 percent to the effect of the spoken message [55]. This implies that the facial expressions form the major modality in human communication.

II. FACIAL EXPRESSION ANALYSIS

In the case of static images, the process of extracting the facial expression information is referred to as localizing the face and its features in the scene. In the case of facial image sequences, this process is referred to as tracking the face and

its features in the scene. At this point, a clear distinction should be made between two terms, namely, facial features and face model features. The facial features are the prominent features of the face: eyebrows, eyes, nose, mouth, and chin. The face model features are the features used to represent (model) the face. The face can be represented in various ways, e.g., as a whole unit (holistic representation), as a set of features (analytic representation) or as a combination of these (hybrid approach). The applied face representation and the kind of input images determine the choice of mechanisms for automatic extraction of facial expression information. The final step is to define some set of categories, which we want to use for facial expression classification and/or facial expression interpretation, and to devise the mechanism of categorization.

Our aim is to explore the issues in design and implementation of a system that could perform automated facial expression analysis. In general, three main steps can be distinguished in tackling the problem. First, before a facial expression can be analyzed, the face must be detected in a scene. Next is to devise mechanisms for extracting the facial expression information from the observed facial image or image sequence.

A. Face Detection

The first preparatory step consists to locate the face in the input image. The face detection is performed by the traditional Viola-Jones object detection framework. The Viola-Jones framework consists of two main steps: a) Haar-like features extraction and b) Adaboost classifier [12].

B. Facial Feature Extraction

The next step consists to extract the facial features using the Active Shape Models (ASM) proposed by Cootes et al. [13]. Typically the ASM works as follows: each structured object or target is represented by a set of landmarks manually placed in each image of the training set. Next, the landmarks are automatically aligned to minimize the distance between their corresponding points. The ASM creates a statistical model of the facial shape which iteratively deforms to fit the model in a new image.

C. Facial Expression Classification

As previously shown in the related works, several classifiers have been used to predict facial expressions. In this work, the proposed system is evaluated with three different classifiers: ANN, LDA and KNN. The goal is to determine which of the three classifiers achieves the best results for the seven facial expressions: happiness, anger, sadness, surprise, disgust, fear and neutral. In the next section, the experimental results of the proposed system are shown.



Fig I. Examples of ASM fiducial points location

The fiducial points' location results are shown as Fig.I. Two images of one neutral and one surprise expression face

images. As can be seen, because of the different expressions, there are different deformation of a face shape, especially in the facial components.

D. Model selection and parameter selection

It is suggested that if we don't know which kernel function is the most suitable, we always choose RBF as the first choice. The RBF kernel nonlinearly maps instances to a higher dimensional space, unlike the linear kernel function; it can handle the case when the relation between class labels and feature attributes is nonlinear. LIBSVM also provides a parameter selection tool using the RBF kernel: cross validation via parallel grid search. So, the parameters of our experiments is the two: c and r corresponding the C-SVC SVMs. Note that now under the one-against-one method, the same pair wise parameters (c, r) is used for our experiments' $7*(7-1)/6$ binary C-SVC SVMs.

III. AUTOMATIC FACIAL EXPRESSION ANALYSIS

For its utility in application domains of human behavior interpretation and multimodal/media HCI, automatic facial expression analysis has attracted the interest of many computer vision researchers. Since the mid-1970s, different approaches are proposed for facial expression analysis from either static facial images or image sequences. In 1992, Samal and Iyengar [19] gave an overview of the early works. This paper explores and compares approaches to automatic facial expression analysis that have been developed recently, i.e., in the late 1990s. Before surveying these works in detail, we are giving a short overview of the systems for facial expression analysis proposed in the period of 1991 to 1995.

3.1 Face Detection

Summary of the Methods for Automatic Face Detection

	Reference	View	Method	Comments
Facial images				
Holistic approach	Huang [32]	Frontal view	Canny edge detector PDM model fitting	No rigid head rotations
	Pantic [62]	Dual view	Image histogram analysis Thresholding	Mounted camera on the subject's head
Analytic approach	Hara [42]	Frontal view	Brightness distribution	No rigid head motions Real-time process
	Yoneyama [104]	Frontal view	-	-
	Kimura [41]	Frontal view	Integral projection [99] Potential Net fitting	No rigid head rotation
Arbitrary images				
Holistic approach	Hong [30]	Frontal view	Steffens et al. [81]: Spatio-temporal filtering Stereo algorithm Color detector Convex region detector Linear predictive filter	Complex background Slight head motions
	Essa [24]	Frontal to profile view	Pentland et al. [65]: Spatiotemporal filtering Eigenfaces Eigenfeatures	Complex background Rigid head motions Faces with facial hair Faces with glasses Real-time process

Table 1

Independently of the kind of input images-facial images or arbitrary images-detection of the exact face position in an observed image or image sequence has been approached in two ways. In the holistic approach, the face is determined as a whole unit. In the second, analytic approach, the face is detected by detecting some important facial features first (e.g., the irises and the nostrils). The location of the features in correspondence with each other determines then the overall location of the face. Table 1 provides a classification of facial expression analyzers according to the kind of input images and the applied method.

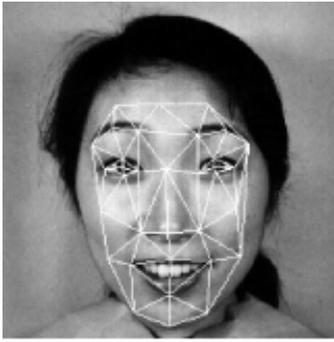


Fig. II. Fiducial grid of facial points

IV. DISCUSSION

We believe that a well-defined and commonly used single database of testing images (image sequences) is the necessary prerequisite for "ranking" the performances of the proposed systems in an objective manner. Since such a single testing data set has not been established yet, we left the reader to decide the ranking of the surveyed systems according to his/her own priorities and based on the overall properties of the surveyed systems.

The experimental results have shown that the LDA classifier has the best hit rate: 99.7% for MUG database and 99.5% for FEEDTUM database. In addition, LDA is less sensitive than ANN classifier. As we saw in experimental results, the ANN shows higher hit rate variations given the number of hidden neurons, an issue that is absent in LDA classifier. Moreover, in Section IV-B we have shown that LDA gets high hit rate starting with 24 landmarks. However, KNN has left a great deal to be desired, getting inferior results than ANN and LDA classifiers.

V. CONCLUSION

Analysis of facial expressions is an intriguing problem which humans solve with quite an apparent ease. We have identified three different but related aspects of the problem: face detection, facial expression information extraction, and facial expression classification. Capability of the human visual system in solving these problems has been discussed. It should serve as a reference point for any automatic vision-based system attempting to achieve the same functionality. Among the problems, facial expression classification has been studied most, due to its utility in application domains of human behavior interpretation and HCI. Most of the surveyed systems, however, are based on frontal view images of faces without facial hair and glasses what is unrealistic to expect in these application domains. Also, all of the proposed approaches to automatic expression analysis perform only facial expression classification into the basic emotion categories defined by Ekman and Friesen [20]. Nevertheless, this is unrealistic since it is not at all certain that all facial expressions able to be displayed on the face can be classified under the six basic emotion categories. Furthermore, some of the surveyed methods have been tested only on the set of images used for training. We hesitate in belief that those systems are person independent what, in turn, should be a basic property of a behavioral science research tool or of an advanced HCI. All the discussed problems are intriguing and none has been solved, in the general case. We expect that they would remain interesting to the researchers of automated vision based facial expression analysis for some time.

REFERENCES

- [1] J.N. Bassili, "Facial Motion in the Perception of Faces and of Emotional Expression," *J. Experimental Psychology* 4, pp. 373-379, 1978.
- [2] M.J. Black and Y. Yacoob, "Tracking and Recognizing Rigid and Non-Rigid Facial Motions Using Local Parametric Models of Image Motions," *Proc. Int'l Conf. Computer Vision*, pp. 374-381, 1995.
- [3] L.S. Chen, T.S. Huang, T. Miyasato, and R. Nakatsu, "Multimodal Human Emotion/Expression Recognition," *Proc. Int'l Conf. Automatic Face and Gesture Recognition*, pp. 366-371, 1998.
- [4] T.F. Cootes, C.J. Taylor, D.H. Cooper, and J. Graham, "Active Shape Models-Training and Application," *Computer Vision Image Understanding*, vol. 61, no. 1, pp. 38-59, 1995.
- [5] DeCarlo, D. Metaxas, and M. Stone, "An Anthropometric Face Model Using Variational Techniques," *Proc. SIGGRAPH*, pp. 67- 74, 1998.
- [6] Donato, M.S. Bartlett, J.C. Hager, P. Ekman, and T.J. Sejnowski, "Classifying Facial Actions," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 21, no. 10, pp. 974-989, Oct. 1999.
- [7] G.J. Edwards, T.F. Cootes, and C.J. Taylor, "Face Recognition Using Active Appearance Models," *Proc. European Conf. Computer Vision*, vol. 2, pp. 581-695, 1998.
- [8] I.Essa and A. Pentland, "Coding, Analysis Interpretation, Recognition of Facial Expressions," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 757-763, July 1997.
- [9] A.J. Fridlund, "Evolution and Facial Action in Reflex, Social Motive, and Paralanguage," *Biological Psychology*, vol. 32, pp. 3- 100, 1991.
- [10] R.J. Holt, T.S. Huang, A.N. Netravali, and R.J. Qian, "Determining Articulated Motion from Perspective Views," *Pattern Recognition*, vol. 30, no. 9, pp. 1435-1,449, 1997.
- [11] C.L. Huang and Y.M. Huang, "Facial Expression Recognition Using Model-Based Feature Extraction and Action Parameters Classification," *J. Visual Comm. and Image Representation*, vol. 8, no. 3, pp. 278-290, 1997.
- [12] C.E. Izard, "Facial Expressions and the Regulation of Emotions," *J. Personality and Social Psychology*, vol. 58, no. 3, pp. 487-498, 1990.
- [13] Kawakami, M. Okura, H. Yamada, H. Harashima, and S. Morishima, "3D Emotion Space for Interactive Communication," *Proc. Computer Science Conf.*, pp. 471-478, 1995.
- [14] A.Pentland, B. Moghaddam, and T. Starner, "View-Based and Modular Eigenspaces for Face Recognition," *Proc. Computer Vision and Pattern Recognition*, pp. 84-91, 1994.
- [15] J.A. Russell, "Is There Universal Recognition of Emotion from Facial Expression?" *Psychological Bulletin*, vol. 115, no. 1, pp. 102- 141, 1994.
- [16] M.J. Lyons, J. Budynek, and S. Akamatsu, "Automatic Classification of Single Facial Images," *IEEE Trans.*

- Pattern Analysis and Machine Intelligence, vol. 21, no. 12, pp. 1,357-1,362, 1999.
- [17] M. Rosenblum, Y. Yacoob, and L. Davis, "Human Emotion Recognition from Motion Using a Radial Basis Function Network Architecture," Proc. IEEE Workshop on Motion of Non-Rigid and Articulated Objects, pp. 43-49, 1994.
- [18] H.A. Rowley, S. Baluja, and T. Kanade, "Neural Network-Based Face Detection," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 20, no. 1, pp. 23-38, Jan. 1998.
- [19] A. Samal and P.A. Iyengar, "Automatic Recognition and Analysis of Human Faces and Facial Expressions: A Survey," Pattern Recognition, vol. 25, no. 1, pp. 65-77, 1992.
- [20] K.K. Sung and T. Poggio, "Example-Based Learning for ViewBased Human Face Detection" IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 20, no. 1, pp. 39-51, Jan. 1998.
- [21] D. Terzopoulos and K. Waters, "Analysis and Synthesis of Facial Image Sequences Using Physical and Anatomical Models," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 15, no. 6, pp. 569-579, June 1993.
- [22] N.M. Thalmann, P. Kalra, and M. Escher, "Face to Virtual Face," Proc. IEEE, vol. 86, no. 5, pp. 870-883, 1998.

