ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/343443531

Facial Emotion Recognition: A Brief Review

Conference Paper · August 2020

CITATIONS
4

READS
12,945

2 authors:

Illiana Azizan
Universiti Putra Malaysia
7 PUBLICATIONS 7 CITATIONS

SEE PROFILE

K. Fatimah
Universiti Putra Malaysia
71 PUBLICATIONS 390 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project

Region-based Surveillance Video Indexing and Retrieval Framework Using Video Abstraction View project

All content following this page was uploaded by K. Fatimah on 23 October 2020.

The user has requested enhancement of the downloaded file.

International Conference on Sustainable Engineering, Technology and Management (*ICSETM* -2018), Dec. 20, 2018, Negeri Sembilan, Malaysia

Facial Emotion Recognition: A Brief Review

Illiana Azizan
Faculty of Computer Science and Information Technology,
Universiti Putra Malaysia, Serdang, Selangor, Malaysia
illianaazizan@gmail.com

Fatimah Khalid
Faculty of Computer Science and Information Technology,
Universiti Putra Malaysia, Serdang, Selangor, Malaysia
fatimahk@upm.edu.my

Abstract- **This paper give an overview of current Facial Emotion Recognition (FER) stages, techniques, and datasets. FER has been recognized for a decades and it is a vital topic in the fields of computer vision and machine learning. Automatic FER is useful in most of the applications such as healthcare, teaching, criminal investigation, Human Robot Interface (HRI), etc. This paper is aim to understand the basic principles of FER and make a comparison of current research.**

Keywords-Facial Expression Recognition, Facial Feature Extraction, Expression Classification.

I. Introduction

Facial emotion recognition has become an important issue in many application nowadays. In recent years, the research on facial emotion recognition has become extensive. The aim of facial emotion recognition is to help identify the state of human emotion (eg; neutral, happy, sad, surprise, fear, anger, disgust, contempt) based on particular facial images. The challenge on facial emotion recognition is to automatically recognize facial emotion state with high accuracy. Therefore, it is challenging to find the similarity of the same emotion state between different person since they may express the same emotion state in various ways. As an example, the expression may varies in different situations such as the individual's mood, their skin color, age, and environment surrounds.

The acronym for Facial Emotion Recognition (FER) is different in every paper, such as Facial Emotion Recognition and Facial Expression Recognition. In this paper the acronym FER is refer to Facial Emotion Recognition.

Generally FER is devided into three major stages as shown in Figure 1: (i) Face Detection, (ii) Feature Extraction, and (iii) Emotion Classification. At first stage, which is a pre-processing stage, an image of a face is detected and facial components of the face will be detected from the region. The facial components can be an eyes, brows, nose, and mouth. In the second stage, an informative features will be extracted from different parts of the face. In the last stage, a classifier need to be trained before been used to generate labels for the Emotions using the training data.

Figure 1. Facial Emotion Classification Stages.

Reflects to feature extraction stage, there is another approach for facial expression analysis [19], which is Facial Action Coding System (FACs). This approach has been introduced by Ekman [14] as a popular facial coding system to classify emotions based on movements of certain facial muscles. Facial actions are classified into different Action Units (AUs) and emotions are categorised using collections of AUs [20].

Deep learning is a part of machine learning approaches which can be adapted to emotion recognition and facial expression analysis. However, deep learning depends on data size which may affect on its performance [19].

II. Face Detection

Face detection is a pre-processing phase to recognize facial expressions of human. An image is segmented into two parts which have faces and other non-face regions [27]. There are numerous methods used for face detection. Figure 2 shows an example of methods used for face detection in real-time.

Algorithm	Accuracy	Performance in real-time
Haar classifier	Accuracy is high for face detection due to suitable Haar features	Computational complexity is very less due to a set of features which contribute the maximum, for the face detection problem in a training phase
Adaptive skin colour	Accuracy is good as skin colour in identified easily but fails in different levels of illumination	Approaches such as adaptive gamma corrective method is used to get rid of illumination problem which leads to high computational complexity and is not suitable in real-time environment
Adaboost contour points	High detection accuracy due to strong classifier single face is detected using contour points due to which the accuracy is good	Computational cost is less due to trained model low computational complexity due to less number of features

Figure 2. Face Detection Algorithm in real-time [27].

A. *Haar Classifier*

Haar features can be measured by expanding or reducing the size of the pixel group. It used Haar-like features to detect an image. This method will allow objects to be detected in various sizes. Haar classifier will identify a set of features which are most contributing for the face detection problem in training phase itself. Therefore, it is suitable for face detection in training phase as it may indicates to high detection accuracy since the computation complexity is small [27].

B. Adaptive Skin Colour

Adaptive skin-colour model used face detection method based on skin-color model to detect the face region. This algorithm shows a high accuracy as skin color is used for segmentation. Hence it can be easily to differentiate the face region and non-face region. However, this algorithm does not work with different levels of illumination. To avoid this problem, adaptive gamma corrective method is suitable to be used (but not suitable in real-time environment due to high computational complexity) [27].

C. Adaboost Contour Points

Adaboost is suitable to detect face in a real-time environment due to low computational complexity and high accuracy. In this technique, several classifiers can be cascaded. It trained the faces and built a strong classifier which leads to high detection accuracy. Then the new face will be compared with the model built by the classifier. It also used contour points to detect face. The contour points may give a good accuracy and performance because the features extracted are less which leads to low computational complexity [27].

III. Feature Extraction

Feature extraction converts a pixel data of the face region into a higher-level representation of shape, colour, texture, and spatial configuration of the face or its components. Feature extraction will reduce the dimension of the input space while keeping the important information. Feature extraction is important in formulating a better emotion categorization as the extracted facial feature given inputs to classification module which finally it categorizes different emotions.

Feature extraction can be divided into two categories which are; (i) feature base (ii) appearance base. Feature base can be divided into geometric feature and appearance feature. Geometric feature will recognize face principal feature such as eyes, nose, and lips. Geometric features can be classified into permanent features and transient features. Meanwhile Appearance feature will identify the skin texture, wrinkle, and furrows.

Feature extraction can be performed in many techniques. Figure 2 shows some of the techniques for feature extraction techniques.

A. Local Binary Pattern (LBP)

LBP is a visual descriptor used in classification for computer vision. LBP originally proposed for texture analysis. It is the primary technique in face recognition. Most of the enhancements among the face recognition techniques are created based on LBP. Later, it has been applied in facial images analysis [13]. Paper [13] has made a comparison between several techniques to view the performance of FER using LBP features. Their study determined that in a compact representation, LBP features able to retain discriminative facial information and derived faster in a single scan of raw images plus in low-dimensional feature space [13]. However, the data collected must be in a highly controlled environment with high resolution of frontal faces [28].

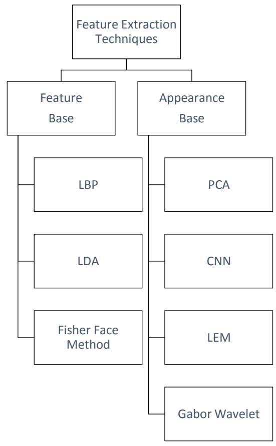

Figure 3. Feature Extraction Techniques.

B. Linear Discriminant Analysis (LDA)

LDA is a supervised subspaces learning technique. Paper [16] has applied a simple LDA-based classification scheme for FER as it can be trained quickly [16]. Meanwhile, paper [13] further adopted LDA to recognize expressions by using LBP features. In paper [13], LDA is used to search for the projection of axes on which the data points of different classes are far from each other while requiring data points of the same class to be close to each other.

C. Fisher Face Method

Fisher Face method required precise normalization and and registration of facial internal features [16]. Peter N. Belhumeur, Jo~ao P. Hespanha, and David J. Kriegman [17] has proposed Fisherface method which shows a result of error rates that are lower than Eigenface which been tested on Harvard and Yale face databases. Fisherface is derived from Fisher's Linear Discriminant (FLD) method. It will maximize the ratio of between-class scatter to that of within-class scatter [17].

D. Principle Component Analysis (PCA)

PCA is also the most widely used techniques. It is a linear appearance based face recognition method. PCA is based on information theory approach that decomposes input face images into a small set of characteristic feature images. The feature image is call an "Eigen Face" which is a principal components of the preliminary training set of face images. PCA also used in handwriting analysis, lip reading, voice recognition, hand gesture and medical imaging analysis [29].

E. Convolutional Neural Network (CNN)

Currently, CNN is one of the most mainstream approach in the deep learning techniques [23]. It use a variation of multilayer perceptrons designed to require minimal pre-processing.

F. Line Edge Map (LEM)

LEM proposed by [25] extracts lines from a face edge map as features. LEM is a combination between template matching and geometrical feature. It has a high recognition performance of template matching. It also has advantages for feature-based approaches (invariance to illumination and low memory requirement) [24].

G. Gabor Wavelet

Gabor Wavelet is one of the image analysis techniques which characterize the image as a selective localized orientation. The face region (such as eyes, nose, mouth, wrinkles, dimples, scars, etc) are enhances as key features to represent the face in high dimensional space. In some degree, the Gabor wavelet is robust to misalignment because it captures the local tecture characterized by spatial frequency, spatial position and orientation.

IV. Expression Classification

This stage is performed by a classifier. There are various classifications methods used to extract expressions (Figure 3). Ekman defined two main types of classes used in facial expression recognition are action units and prototypic facial expressions [14].

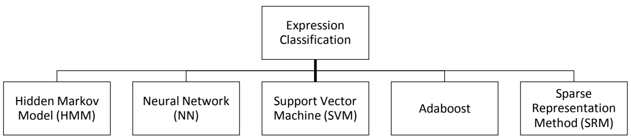

Figure 4. Expression Classification.

A. Hidden Markov Model (HMM)

HMM use a set of statistical model to describe the statistical behaviour of a signal [21]. In paper [21], they investigates five choices of HMM models which are; (i) Left-Right HMM, (ii) Ergodic HMM, (iii) Emotion-Specific HMMs, (iv) Multilevel HMM and (v) Mixture of HMM and neural network. As conclusion, the HMM achieves better classification either 3 or 5 state model used independent expression or combination of expression using multilevel state models.

B. Neural Network (NN)

NN executes a nonlinear reduction of the input dimensionality. It formulates a statistical decision about the category of the expression that has been observed. Every output unit will estimates on the probability of the examined expression belongs to the associated category [22].

C. Support Vector Machine (SVM)

SVM is one of the famous statistical techniques used in machine learning to analyze data used for classification and regression analysis. SVM used different kernel function to map data in input space into high-dimensional feature spaces [21].

D. Bayesian Network (BN)

BN is a graphical model which able to show a clear and intuitive relationships by using prababilistic approach among a set of attribute [21].

E. Sparse Representation Method (SRM)

SRM use a linear combination of all the training samples to "sparsely" represent and classify the face image. "Sparsely" is refering to some coefficients of the linear combination are equal or close to zero [18].

V. Dataset

To perform an experiment of FER, a standard database is required. The data can be perceive as primary or secondary. A primary dataset consume a long period to be completed with dataset collection. For study in FER, a variety of dataset available currently (Cohn Kanada (CK+), JAFFE, MMI etc).

VI. Comparision of Current Research Papers on FER

Table 1 shows a comparison of current research papers on FER. The comparison is made randomly based on the expressions compared, methods and techniques used for face detection, feature extraction and expression classifications. It shows a different measurement on the accuracy of the detection been achieved.

Table 1. Comparison of current research papers on FER

Paper	**Expressions**	**Face Detection**	**Feature Extraction**	**Expression Classification**	**Dataset**	**Accuracy**
[3]	Happy, Anger, Sad, Surprise, Normal	Luxand Face Recognition	Angle And Distance Method	SVM	Real-time	85.6%
[5]	Anger, Contempt, Disgust, Fear, Happy, Sadness, Surprise, Neutral	AAM, Hog, PCA	DSAE	Softmax Classifier	CK+ JAFFE	95.79 %
[7]	Happy, Sadness, Surprise, Anger, Disgust, Fear, Neutral	Define Distance between Two Centers of Both Eyes	Wavelet Entropy	Jaya Algorithm	[15]	96.8%
[8]	Happy, Sadness, Surprise, Anger, Disgust, Fear, Neutral	Viola-Jones's Haar-like feature cascade detector	PCA, Fisher Face + HOG	SVM	CK+	81%
[9]	Neutral, Fatigue	Hog	Facial Landmarks Points (Cascade of Regression Tree)	SVM	CK+ PICS	82.79 %

[11]	Joy, sad, surprise, fear, anger, disgust, neutral	Hog	Face Alignment with Regression Tree (Landmark Detection Algorithm)	Multilayer Perceptron	JAFFE	88.03 %

VII. Conclusion

Emotion expression is important in communication, hence improving the quality of interaction between human. Furthermore, the study of facial emotion recognition may contribute a better feedback to the society and also interaction between Human Robot Interface (HRI) in a near future. The emotion detection mostly involves with the geometric part of the face (eg; eyes, eyebrow, and mouth). The review takes consideration of experiment which been conducted in a controlled environment, real-time, and wild images. The main issues are highlighting the accuracy of the techniques chosen in their research.

References

[1] B. Ko, "A Brief Review of Facial Emotion Recognition Based on Visual Information," *Sensors*, vol. 18, no. 2, p. 401, 2018.

[2] R. C. Chivers, V. V Ramalingam, and A. Pandian, "Facial Emotion Recognition System – A Machine Learning Approach Facial Emotion Recognition System ± A Machine Learning," 2018.

[3] M. H. Siddiqi, M. Alruwaili, J. Bang, and S. Lee, "Real Time Human Facial Expression Recognition System using Smartphone," vol. 17, no. 10, pp. 223–230, 2017.

[4] A. Savva, V. Stylianou, K. Kyriacou, and F. Domenach, "Recognizing Student Facial Expressions: A Web Application," *IEEE Glob. Eng. Educ. Conf. EDUCON*, vol. 2018–April, pp. 1459–1462, 2018.

[5] N. Zeng, H. Zhang, B. Song, W. Liu, Y. Li, and A. M. Dobaie, "Facial expression recognition via learning deep sparse autoencoders," Neurocomputing, vol. 273, pp. 643–649, 2018.

[6] M. U. Ahmed, K. J. Woo, K. Y. Hyeon, M. R. Bashar, and P. K. Rhee, "Wild facial expression recognition based on incremental active learning," Cogn. Syst. Res., vol. 52, pp. 212–222, 2018.

[7] S. H. Wang, P. Phillips, Z. C. Dong, and Y. D. Zhang, "Intelligent facial emotion recognition based on stationary wavelet entropy and Jaya algorithm," Neurocomputing, vol. 272, pp. 668–676, 2018.

[8] J. J. Pao, "Emotion Detection through Facial Feature Recognition," p. 6, 2018.

[9] M. Merlin Steffi and J. John Raybin Jose, "Comparative Analysis of Facial Recognition involving Feature Extraction Techniques," Int. J. Comput. Sci. Eng. Open Access Rev. Pap., no. 6, 2018.

[10] N. Irtija, M. Sami, and A. R. Ahad, "Fatigue Detection Using Facial Landmarks," pp. 1–6, 2018.

[11] Alvarez, V. M., Velazquez, R., Gutierrez, S., & Enriquez-Zarate, J. (2018). A Method for Facial Emotion Recognition Based on Interest Points. 2018 International Conference on Research in Intelligent and Computing in Engineering (RICE), 1–4.

[12] Ekman, P. (2013). *Emotion in the Human Face.* USA: Malor Books.

[13] C. Shan, S. Gong, and P. W. McOwan, "Facial Expression Recognition Based on Local Binary Patterns: A comprehensive study," *Image Vis. Comput.*, vol. 27, no. 6, pp. 803–816, 2009.

[14] Ekman, P., W. Friesen, Measuring Facial Movement, Environmental Psychology and Noverbal Behavior 1(1), Fall, 1976.

[15] Y. Zhang *et al.*, "Facial Emotion Recognition Based on Biorthogonal Wavelet Entropy , Fuzzy Support Vector Machine , and Stratified Cross Validation," pp. 8375–8385, 2016.

[16] M.J. Lyons, J. Budynek, S. Akamatsu, "Automatic Classification Of Single Facial Images", IEEE Transactions on Pattern Analysis and Machine Intelligence 21 (12) (1999) 1357–1362.

[17] P. Hespanha, D. J. Kriegman, and P. N. Belhumeur, "Eigenfaces vs . Fisherfaces : Recognition Using Class Specific Linear Projection," vol. 19, no. 7, pp. 711–720, 1997.

[18] Y. Xu, Q. Zhu, Z. Fan, D. Zhang, J. Mi, and Z. Lai, "Using the idea of the sparse representation to perform coarse-to-fine face recognition," *Inf. Sci. (Ny).*, vol. 238, pp. 138–148, 2013.

[19] D. A. Pitaloka, A. Wulandari, T. Basaruddin, and D. Y. Liliana, "Enhancing CNN with Preprocessing Stage in Automatic Emotion Recognition," *Procedia Comput. Sci.*, vol. 116, pp. 523–529, 2017.

[20] S. L. Happy, A. Dasgupta, P. Patnaik, and A. Routray, "Automated alertness and emotion detection for empathic feedback during E-learning," *Proc. - 2013 IEEE 5th Int. Conf. Technol. Educ. T4E 2013*, pp. 47–50, 2013.

[21] M. H. Siddiqi *et al.*, "A Brief Review of Facial Emotion Recognition Based on Visual Information," *2018 IEEMA Eng. Infin. Conf. eTechNxT 2018*, vol. 5, no. 1, pp. 196–201, 2018.

[22] M. Pantic, M. Pantic, and M. Rothkrantz, "Automatic Analysis of Facial Expressions," *IEEE Trans. Pattern Recognit. Mach. Intell.*, vol. 22, no. 12, pp. 1424 – 1444, 2000.

[23] Z. J. Wang *et al.*, "Article in press," vol. 0, pp. 1–11, 2017.

[24] A. S. Tolba, A. H. El-baz, and A. A. El-Harby, "Face Recognition : A Literature Review," *Int. J. Signal Process.*, vol. 2, no. 2, pp. 88–103, 2006.

[25] Y. Gao and K.H. Leung, "Face recognition using line edge map," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 6, June 2002.

[26] S. Du, Y. Tao, and A. M. Martinez, "Compound Facial Expressions of Emotion," *Proc. Natl. Acad. Sci.*, vol. 111, no. 15, pp. E1454–E1462, 2014.

[27] S. Deshmukh, M. Patwardhan, and A. Mahajan, "Survey on Real-Time Facial Expression Recognition Techniques," *IET Biom.*, pp. 1-9, 2015.

[28] Y. Tian, "Evaluation of Face Resolution for Expression Analysis", *CVPR Workshop on Face Processing in Video*, 2004.

[29] M. Rahman, "A comparative study on face recognition techniques and neural network," *arXiv Prepr. arXiv1210.1916*, vol. 3, no. 5, pp. 155–160, 2012.

View publication stats