Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 58 (2015) 486 – 491

Procedia Computer Science

Second International Symposium on Computer Vision and the Internet (VisionNet'15)

Facial expression recognition: A survey

Jyoti Kumari[a], R.Rajesh[a], KM.Pooja[a]

[a]Dept. of Computer Science, Central University of South Bihar, India

Abstract

Automatic facial expression recognition system has many applications including, but not limited to, human behavior understanding, detection of mental disorders, and synthetic human expressions. Two popular methods utilized mostly in the literature for the automatic FER systems are based on geometry and appearance. Even though there is lots of research using static images, the research is still going on for the development of new methods which would be quiet easy in computation and would have less memory usage as compared to previous methods. This paper presents a quick survey of facial expression recognition. A comparative study is also carried out using various feature extraction techniques on JAFFE dataset.

Peer-review under responsibility of organizing committee of the Second International Symposium on Computer Vision and the Internet (VisionNet'15)

Keywords: Facial Expression Recognition(FER), LBP, LDP, LGC, HOG;

1. Introduction

One of the non-verbal communication method by which one understands the mood/mental state of a person is the expression of face (for e.g. happy, sad, fear, disgust, surprise and anger)[1,2,3]. Automatic facial expression recognition (FER)[4] has become an interesting and challenging area for the computer vision field and its application areas are not limited to mental state identification[5], security[6], automatic counseling systems, face expression synthesis, lie detection, music for mood[7], automated tutoring systems[8], operator fatigue detection[9] etc.

FER consists of five steps as shown in Fig.1. The noise-removal/enhancement is done in the preprocessing step by taking image or image sequence (a time series of images from neutral to an expression) as an input and gives

* Corresponding author. Tel.: +0-000-000-0000 ; fax: +0-000-000-0000 .
E-mail address: j2kumari@gmail.com, kollamrajeshr@ieee.org (R.Rajesh), poojam102@gmail.com(KM. Pooja)

Peer-review under responsibility of organizing committee of the Second International Symposium on Computer Vision and the Internet (VisionNet'15)
doi:10.1016/j.procs.2015.08.011

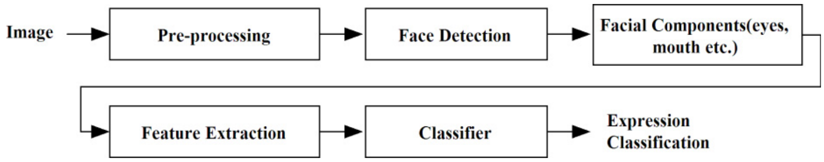

Fig.1. Facial expression classification block diagram

the face for further processing. The facial component detection detects the ROI for eyes, nose, cheeks, mouth, eye brow, ear, fore-head, etc. The feature extraction step deals with the extraction of features from the ROIs. The most popular feature extraction techniques are, but not limited to, Gabor filters[10], Local Binary Patterns (LBP)[11], Principal Component Analysis (PCA)[12], Independent Component Analysis (ICA)[13], Linear Discriminant Analysis (LDA)[14], Local Gradient Code (LGC)[15], Local Directional Pattern (LDP)[16]. The classifier step classifies the features into the respective facial expression classes based on classification methods and some of the most popular classification methods are SVM (Support Vector Machine)[17] and NN (Nearest Neighbor)[18].

The paper proceeds as section 2 showing some feature extraction strategies, section 3 deal with some popular feature extraction techniques, section 4 presents some experimental results. Section 5 concludes the paper.

2. Categorizing facial expression feature

The changes in the facial expression can be either based on minor deformations in wrinkles/bulges[19] or based on major deformations (in eyes, eye-brow, mouth, nose, etc.)[4]. Some of the feature extraction techniques and facial expression categorization includes, Geometric based and Appearance based[20], Action Unit (AU) of individual/group of muscles and Non-AU based[21], Local versus Holistic[4]. In geometric based methods, the position and deformation/displacement information of the facial components are considered[30], whereas appearance methods simply apply a filter.

Facial emotion are categorized into six[21], namely, anger (“combination of brow lowerer, upper lid raiser, lid tightener, lip tightener”), disgust (“combination of nose wrinkler, lip corner depressor, lower lip depressor”), fear (“combination of inner/outer brow raiser, brow lowerer, upper lid raiser, lid tightener, lip stretcher, jaw drop”), happy (“combination of cheek raiser, lip corner puller”), sad (“combination of inner brow raiser, brow lowerer, lip corner depressor”), surprise (“combination of inner/outer brow raiser, upper lid raiser, jaw drop”).

There is still research scope in this area depending on the regional behavior (for e.g., in Indian scenario even eyes alone can tell about anger, surprise and other emotions).

3. Comparison of popular feature extraction method

This section gives a brief description of some of the methods widely used for feature extraction. In this, features are extracted by considering the whole image as a single unit and applying some sort of filters. According to literature, Gabor filters show excellent facial analysis performance with greater computational cost in terms of time and memory usage.

3.1. Local binary pattern

Local Binary Pattern (LBP) was proposed for texture analysis[11] and later extended and applied to other applications. The LBP assigns a label to each pixel in the P-neighborhood (P equally spaced pixel value within a radius (R), denoted by g_p) by thresholding its values with the central value (g_c) and converting these thresholded values into decimal number given by eqn. (1).

$$LBP_{p,R}(X_C, Y_C) = \sum_{p=0}^{p-1} s(g_p - g_c) \quad where,\ s(x) = \begin{cases} 1, & x \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

g_1	g_2	g_3
g_4	g_c	g_5
g_6	g_7	g_8

Fig.2. A 3x3 template for LGC operator

Other related approaches are (i) Local Binary Pattern from Three Orthogonal Planes (LBP-TOP)[22] by Yanjun Guo et al. to detect the micro-expressions from sequential images, and (ii) combination of dynamic edge and texture information named as LOE-LBP-TOP (Local Oriented Edges) by Gizatdinova et al.[23].

3.2. Local gradient code

LGC[15] is based on the relationship of neighboring pixels (while LBP only compares the central pixel value with the neighboring pixel value) given by eqn.(2) [see the position of pixels in Fig. 2]. The optimized LGC-HD (based on principle of horizontal diagonal) and LGC-VD (based on the principle of vertical diagonal) are given by eqn.(3) and (4).

$$LGC = s(g_1 - g_3)2^7 + s(g_4 - g_5)2^6 + s(g_6 - g_8)2^5 + s(g_1 - g_6)2^4 + s(g_2 - g_7)2^3 + s(g_3 - g_8)2^2 + s(g_1 - g_8)2^1 + s(g_3 - g_6)2^0 \tag{2}$$

$$LGC - HD_p^R = s(g_1 - g_3)2^4 + s(g_4 - g_5)2^3 + s(g_6 - g_8)2^2 + s(g_1 - g_8)2^1 + s(g_3 - g_6)2^0 \tag{3}$$

$$LGC - VD_p^R = s(g_1 - g_6)2^4 + s(g_2 - g_7)2^3 + s(g_3 - g_8)2^2 + s(g_1 - g_8)2^1 + s(g_3 - g_6)2^0 \tag{4}$$

3.3. Local directional pattern

In order to get better performance in the presence of variation in illumination and noise, Local Directional Pattern[16] has been developed by Jabid et.al in 2010. In this method, eight Kirsch masks[16] of size 3x3 are convolved with image regions of size 3x3 to get a set of 8 mask values. These mask values are then ranked and the top three will be assigned with one in the 8 bit binary code and the others with zero. The decimal value corresponding to this binary code will be the LDP value for the centre pixel of the selected 3x3 image region. This LDP generated image is divided into blocks and histogram of blocks is concatenated to avail the LDP feature for the image (See fig. 3).

Other variant of LDP are, (i) Local Directional Pattern Variance (LDPv)[24] which combines the texture and contrast, and (ii) LDN[25] which encodes the directional information along with sign.

Fig. 4 shows the original image and the images after applying different methods. Fig. 5 shows the original pixel value and the central pixel value corresponding to the approaches LBP, LDP, LGC.

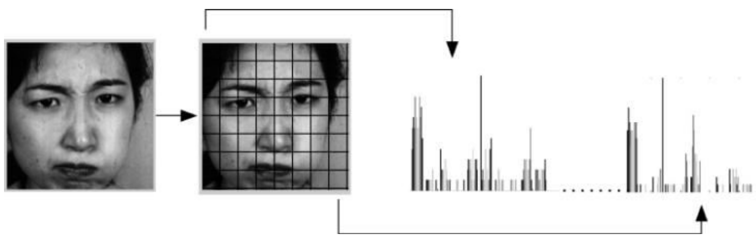

Fig.3. Facial image representation using spatially combined LDP

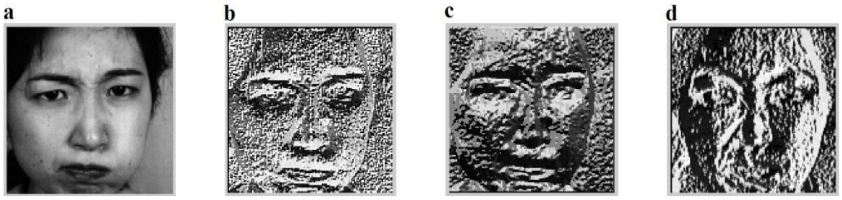

Fig.4. (a)Original Image; (b)LBP Image; (c)LDP Image; (d)LGC Image

146	148	126
145	138	137
145	130	129

1	1	0
1	195	0
1	0	0

0	0	0
0	28	1
0	1	1

1	1	1
0	237	1
1	0	1

Fig.5. Central pixel values corresponding to original, LBP, LDP, LGC image

3.4. Histogram of gradient orientations

Histogram of Oriented Gradients (HOG) 26 is illumination invariant and is found by using magnitude/pixel orientation. Firstly, X and Y gradients of the image is calculated using gradient filter ($Gx = [-1, 0, 1]$, $Gy = [-1, 0, 1]T$). Then using these gradients, corresponding magnitude and angle orientations [ranges $0^\circ - 180^\circ$(unsigned) and $0^\circ - 360^\circ$(signed)] are calculated. The angular orientations are divided into fragments/parts and called as bins. Secondly, the image is divided into cells. The magnitude is binned into corresponding bins depending upon angular section to which it fell. This is done repeatedly by overlapping the cells. Thirdly, the obtained bin values are normalized to cope with contrast problem.

The extensions of HOG can be found in Co-occurrence histograms of oriented gradients (CoHOG)[27] and Coherence Vector of Oriented Gradients (CVOG)[28].

4. Experimental results

The publicly available benchmarking dataset, namely, JAFFE is used to study the performance of various features. There are 213 images of 10 subjects with 7 expressions viz. “anger”, “disgust”, “fear”, “happy”, “neutral”, “sad” and “surprise”. The number of images for one expression of a subject varies from 3 to 4. For our experiment, 196 images are taken and the images are resized into 128×128/256×256 after face detection using Viola-Jones Face Detection algorithm[29].

Table 1. Results of FER on JAFFE dataset for 256 x 256 image sizes with 16 x 16 blocks.

Methods	Feature Length	Recognition Rate
LBP	65536	89.4231
LGC	65536	90.3846
LGC-HD	65536	87.50
LGC-VD	65536	92.3077
HOG	20736	85.7143
LDP	14337	85.2041

Leave-one-out method has been applied where one image is taken as test sample and the rest as training sample. The classification has been done using KNN (K-Nearest Neighbor) classifier. The parameter has been set to default with k value equal to 2.

The recognition rate can be seen as Table 1 on 256×256 image size with 16×16 blocks, Table 2 on 128×128 image size with 16×16 blocks and Table 3 on 128×128 image size with 8×8 blocks.

It can be seen from the results that LGC-VD features are best in emotion recognition. The LBP, LGC and HOG perform equally well, while the worst is LDP.

Table 2. Results for FER on JAFFE dataset for 128 × 128 image sizes, 16 × 16 blocks.

Methods	Feature Length	Recognition Rate
LBP	65536	86.7347
LGC	65536	87.2559
LGC-HD	65536	86.2245
LGC-VD	65536	86.7347
HOG	20736	84.1837
LDP	14337	78.0612

Table 3. Results for FER on JAFFE dataset for 128 × 128 sizes, 8 × 8 blocks.

Methods	Feature Length	Recognition Rate
LBP	16384	88.2653
LGC	16384	88.7755
LGC-HD	16384	84.1837
LGC-VD	16384	85.7143
HOG	5184	86.7347
LDP	3584	64.7959

5. Conclusion

Facial Expression recognition has increasing application areas and requires more accurate and reliable FER system. This paper has presented a survey on facial expression recognition. Recent feature extraction techniques are covered along with comparison. The research is still going on (i) to increase the accuracy rate of predicting the expressions, (ii) to have applications based on dynamic images/sequence of images/videos, (iii) to handle the occlusion.

References

1. Darwin C. The expression of emotions in animals and man. *London: Murray* 1872.
2. Mehrabian A. Communication without words. *Psychological today* 1968;**2**:53–5.
3. Ekman P, Friesen WV. Constants across cultures in the face and emotion. *Journal of personality and social psychology* 1971;**17**:124.
4. Fasel B, Luettin J. Automatic facial expression analysis: a survey. *Pattern recognition* 2003;**36**:259–75.
5. Mandal MK, Pandey R, Prasad AB. Facial expressions of emotions and schizophrenia: A review. *Schizophrenia bulletin* 1998;**24**:399.
6. Butalia MA, Ingle M, Kulkarni P. Facial expression recognition for security. *International Journal of Modern Engineering Research (IJMER),* 2012;**2**:1449–1453.
7. Dureha A. An accurate algorithm for generating a music playlist based on facial expressions. *International Journal of Computer Applications* 2014;**100**:33–9.
8. Wu YW, Liu W, Wang JB. Application of emotional recognition in intelligent tutoring system. In: *Knowledge Discovery and Data Mining, WKDD 2008. First International Workshop on*. IEEE; 2008, p. 449–52.
9. Zhang Z, Zhang J. A new real-time eye tracking for driver fatigue detection. In: *ITS Telecommunications Proceedings, 2006 6th International Conference on*. IEEE; 2006, p. 8–11.
10. Lyons MJ, Budynek J, Akamatsu S. Automatic classification of single facial images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 1999;**21**:1357–62.
11. Ojala T, Pietik¨ainen M, Harwood D. A comparative study of texture measures with classification based on featured distributions. *Pattern recognition* 1996;**29**:51–9.

12. Turk MA, Pentland AP. Face recognition using eigenfaces. In: *Computer Vision and Pattern Recognition. Proceedings, CVPR'91., IEEE Computer Society Conference on*. IEEE; 1991, p. 586–91.
13. Bartlett MS, Movellan JR, Sejnowski TJ. Face recognition by independent component analysis. *Neural Networks, IEEE Transaction on* 2002;**13**:1450–64.
14. Belhumeur PN, Hespanha JP, Kriegman D. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 1997;**19**:711–20.
15. Tong Y, Chen R, Cheng Y. Facial expression recognition algorithm using lgc based on horizontal and diagonal prior principle. *Optik-International Journal for Light and Electron Optics* 2014;**125**:4186–89.
16. Jabid T, Kabir MH, Chae O. Facial expression recognition using local directional pattern (ldp). In: *Image Processing (ICIP),17th IEEE International Conference on*. IEEE; 2010, p. 1605–08.
17. Hsu CW, Chang CC, Lin CJ, et al. A practical guide to support vector classification. 2003.
18. Altman NS. An introduction to kernel and nearest-neighbor nonparametric regression. *The American Statistician* 1992;**46**:175–85.
19. Tian Yl, Kanade T, Cohn JF. Recognizing action units for facial expression analysis. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 2001;**23**:97–115.
20. Tian YL, Kanade T, Cohn JF. Handbook of face recognition. *Ch Facial Expression Analysis, Springer, London* 2005:487–519.
21. Ekman P, Friesen W. Facial action coding system: a technique for the measurement of facial movement. *Consulting Psychologists, San Francisco* 1978.
22. Guo Y, Tian Y, Gao X, Zhang X. Micro-expression recognition based on local binary patterns from three orthogonal planes and nearest neighbor method. In: *Neural Networks (IJCNN), International Joint Conference on*. IEEE; 2014, p. 3473–79.
23. Gizatdinova Y, Surakka V, Zhao G, M¨akinen E, Raisamo R. Facial expression classification based on local spatiotemporal edge and texture descriptors. In: *Proceedings of the 7th International Conference on Methods and Techniques in Behavioral Research*. ACM; 2010, p. 21.
24. KabirMH, Jabid T, Chae O. Local directional pattern variance (ldpv): a robust feature descriptor for facial expression recognition. *Int. Arab J Inf Technol* 2012;**9**:382–391.
25. Ramirez Rivera A, Castillo R, Chae O. Local directional number pattern for face analysis: Face and expression recognition. *Image Processing, IEEE Transactions on* 2013;**22**:1740–52.
26. Dalal N, Triggs B. Histograms of oriented gradients for human detection. In: *Computer Vision and Pattern Recognition (CVPR). IEEE Computer Society Conference on*. IEEE; 2005;**1**: 886–93.
27. Rajesh R, Rajeev K, Gopakumar V, Suchithra K, Lekhesh V. On experimenting with pedestrian classification using neural network. In: *Proc. of 3rd international conference on electronics computer technology (ICECT)*. 2011, p. 107–11.
28. Rajesh R, Rajeev K, Suchithra K, Prabhu LV, Gopakumar V, Ragesh N. Coherence vector of oriented gradients for traffic sign recognition using neural networks. In: *IJCNN*. 2011, p. 907–10.
29. Viola P, Jones M. Rapid object detection using a boosted cascade of simple features. In: *Computer Vision and Pattern Recognition (CVPR). Proceedings of the 2001 IEEE Computer Society Conference on* 2001;**1**: I–511.
30. Chen J, Chen D, Gong Y, Yu M, Zhang K, Wang L. Facial expression recognition using geometric and appearance features. In: *Proceedings of the 4th International Conference on Internet Multimedia Computing and Service. ACM*; 2012, p. 29–33.