Research Paper
Impact Factor : 3.567
Peer Reviewed Journal

IJMDRR
E- ISSN –2395-1885
ISSN -2395-1877

FACIAL EMOTION RECOGNITION PERFORMANCE RATE ANALYSIS ON JAFFE AND CK DATABASES

K. Kalasha* **V. Sujitha****
**Head, PG and Research Department of Computer Science, Government Arts College, Thiruvannamalai.*
***Research Scholar, Government Arts College, Thiruvannamalai.*

Abstract
Facial expression is non verbal message which means appearance on face, arm actions, or tone of voice which illustrate the feel about something without using words. Facial Expression used to many applications. Facial expressions recognition gives larger attention to various fields. Such as chemical and pharmaceutical science, computer science, biotechnology and psychology. HCI research used this facial expression for getting improved results. Facial expression recognition provides perfect emotional features extraction. FER approaches in static images have not fully considered and utilized the features of facial element and muscle movements, which represent static and dynamic, as well as geometric and appearance characteristics of facial expression. By solving this limitation using salient distance features, which are obtained by extracting patch-based 3D Gabor features, selecting the salient patches, and performing patch matching operations. The experimental results produce the output in the form of CRR (Correct Recognition Rate), significant performance improvements due to the consideration of facial element and muscle movements, promising results under face registration errors, and fast processing time. The difference of the state-of-the-art performance confirms that the proposed approach give the highest CRR on the JAFFE database and is among the top performers on the Cohn-Kanade database.

Key Words - HCI - Human Computer Interaction, FER - Facial Expression Recognition System, CRR - Correct Recognition Rate.

I. Introduction

Facial expressions obtained from facial muscle movements. The facial expression recognition system consists of four steps. First process is face detection phase that detect the face from an image or video. Second process is normalization phase that remove the noise and normalize the face against brightness and pixel position. Third process phase features are extracted and irrelevant features are removed. Final step basic expressions are divided into six basic emotions like anger, fear, disgust, sadness, happiness and surprise. Facial expression recognition has been dramatically developed in recent years, thanks to the advancements in related fields, especially machine learning, image processing and human cognition. Accordingly, the impact and potential usage of automatic FER have been growing in a wide range of applications, including human-computer interaction, robot control and driver state surveillance. However, to date, robust recognition of facial expressions from images and videos is still a challenging task due to the difficulty in accurately extracting the useful emotional features. These features are often represented in different forms, such as static, dynamic, point-based geometric or region-based appearance. Facial movement features, which include feature position and shape changes, are generally caused by the movements of facial elements and muscles during the course of emotional expression. The facial elements, especially key elements, will constantly change their positions when subjects are expressing emotions.

II. Emotion Recognition

In 1884, William James gives the important physiological theory of emotion that is in a person emotions are rooted in the bodily experience. First we perceive the object then response occurs and then emotions appear. For example, when we see a tiger or we are in the dangerous position we begin to run and then we fear. Each emotion has its own characteristics and appearance figures. Six basic emotions i.e. fear, surprise, sadness, happiness, anger and disgust are universally accepted. Basic emotions can be distinguished as negative and positive emotions.

III. Techniques

In this section provides an overview and comparison of various techniques that can be used for facial expression recognition. Principal Component Analysis (PCA) is a technique that reduces the dimensionality of image and provides the effective face indexing and retrieval. It is also known as the Eigen face approach [1]. Linear projection is used in PCA, which maximize the projected sample scattering [2]. Imaging conditions like lighting and viewpoint should not be varied for better performance. Fisher's Linear Discriminate is another approach that reduces the projected sample scattering and has better performance than PCA [2].

An ideal emotion detection system should recognize expressions regardless of gender, age, and any ethnicity. Such a system should also be invariant to different distraction like glasses, different hair styles, mustache, facial hairs and different

IJMDRR
E- ISSN –2395-1885
ISSN -2395-1877

lightening conditions. It should also be able to construct a whole face if there are some missing parts of the face due to these distractions. It should also perform good facial expression analysis regardless of large changes in viewing condition and rigid movement [3]. Achieving optimal feature extraction and classification is a key challenge in this field because we have a huge variability in the input data [4]. For better recognition rates most current facial expressions recognition methods require some work to control imaging conditions like position and orientation of the face with respect to the camera as it can result in wide variability of image views. More research work is needed for transformation-invariant expression recognition.

The vast majority of the past work on FER does not take the dynamics of facial expressions into account.

- Some efforts have been made on capturing and utilizing facial movement features, and almost all of them are video-based.
- These efforts try to adopt either geometric features of the tracked facial points (e.g. shape vectors, facial animation parameters, distance and angular, and trajectories) or appearance difference between holistic facial regions in consequent frames (e.g. optical flow, and differential-AAM) or texture and motion changes in local facial regions (e.g. surface deformation, motion units, spatiotemporal descriptors, animation units, and pixel difference).
- Although achieved promising results, these approaches often require accurate location and tracking of facial points, which remains problematic.

IV.Recognition Performance

JAFFE Database

CRR Correct recognition rate of all sets in 10 leave-one set- out cross-validations. Results obtained using four SVMs and four distances.

Fig 1: Expressions Recognition Images

We can see that the proposed approach performs the best with a CRR of 92.93% using DL2 and linear SVM. Regarding the performance of distances, DL2 achieves higher CRRs than the other three distances for all SVMs. When L1 is used, sparse distances outperform dense distances for linear, RBF and sigmoid SVMs. On the contrary, when L2 is used, dense distances outperform spare distances for all SVMs (note that the CRR of DL2 and sigmoid SVM is not shown). For both sparse and dense distances, L2 performs better than L1 for all SVMs. Among four SVMs, linear and RBF outperform polynomial and sigmoid for all distances. More exactly, the best performance is obtained by linear, which is followed by RBF, whereas sigmoid ranks the least.

Research Paper
Impact Factor : 3.567
Peer Reviewed Journal

IJMDRR
E- ISSN –2395-1885
ISSN -2395-1877

TABLE-1

OVERLAPPING PATCHES ON JAFFE ND CK DATABASES

	JAFFE	CK
Patch Size	5(4*4): 1(6*6): 2(8*8)	2(2*2):1(8*8)
Patch Scale	1(3rd): 2(6th): 2(7th): 3(8th)	3(4th)
Emotion Pair	3(AN-AN): 2(DI-DI): 1(FE-FE): 1(FE-SA). 1(SA-SA)	2(AN-AN):1(AN-DI)
Total Number	8	3

TABLE-2

CRRS OF SIX EMOTIONS ON JAFFE DATABASE

	DL1	DL2	SL1	SL2
Linear	81.52%	92.93%	87.50%	88.59%
Polynomial	54.89%	64.13%	45.65%	60.33%
RBF	76.63%	89.67%	82.07%	87.50%
Sigmoid	25.00%	-	26.09%	29.35%

The CRR of DL2 and sigmoid SVM is not shown.

TABLE-3

CONFUSION MATRIX OF SIX EMOTIONS ON JAFFE DATABASE

	AN	DI	FE	HA	SA	SU	Overall
AN	29	1	0	0	0	0	96.67%
DI	2	27	0	0	0	1	90.00%
FE	2	0	30	0	0	0	93.75%
HA	1	0	1	29	2	0	93.55%
SA	0	0	1	1	29	0	93.55%
SU	0	1	1	1	0	27	90.00%

TABLE-4

CRRS OF SIX EMOTIONS ON CK DATABASE

	DL1	DL2	SL1	SL2
Linear	90.20%	93.36%	83.67%	86.71%
Polynomial	87.73%	91.22%	65.43%	80.97%
RBF	92.34%	**94.48%**	80.07%	86.26%
Sigmoid	26.46%	-	37.39%	66.67%

IJMDRR
E- ISSN –2395-1885
ISSN -2395-1877

TABLE-5

CONFUSION MATRIX OF SIX EMOTIONS ON CK DATABASE

	AN	DI	FE	HA	SA	SU	Overall
AN	81	1	2	0	9	0	**87.10%**
DI	4	92	2	2	1	1	90.20%
FE	0	4	138	7	0	1	92.00%
HA	0	2	2	203	0	0	98.07%
SA	6	0	2	0	118	3	91.47%
SU	0	0	0	0	0	207	**100%**

Table-3 demonstrates the confusion matrix of six emotions using DL2 and linear SVM. Observed from this table, disgust and surprise belong to the most difficult facial expressions to be correctly recognized with the same CRR of 90.00%, whereas anger is the easiest one with a CRR of 96.67%. Regarding the misrecognition rate, anger contributes the most; as a result, it has a major negative impact on the overall performance. The emotion that follows in misrecognition rate is fear.

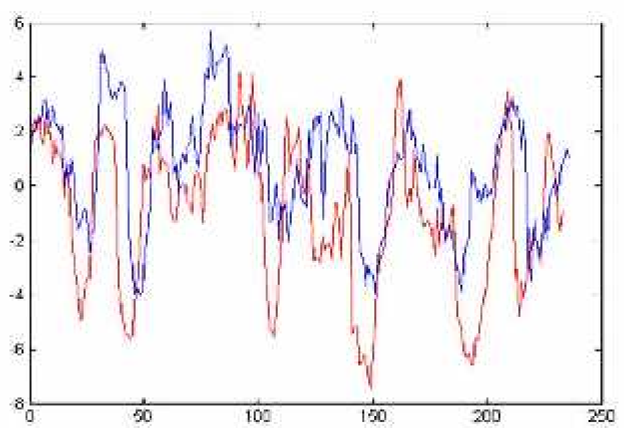

Fig 2: MFCC Trace for Happiness and Anger Utterances

The fig 2 shows the following content. In addition 7 statistical features were computed for every utterance, obtaining a 27 length feature vector. Given a digital image, or a region within an image, the feature extraction task implies the taking out of a quantity of information capable to characterize the original image. Working with two-dimensional signals, the number of samples is much bigger than in the case of a one dimensional signal, thus the necessity of information quantity reduction is obvious.

On our images of 256x256 pixels, jpg format, the wavelet transform was computed by the iterative method of Mallat, yielding wavelet coefficients at each level in a resolution pyramid, where at each successive level the image resolution is decreased by factor 2. Five iteration steps were applied, and the approximation coefficients together with the seven moments of Hu, subjected to classification analyses. For every image the computed vector dimension is 8x8+7.

CK Database

The CRRs using four SVMs and four distance metrics are shown in which the proposed approach obtains the highest CRR of 94.48% using DL2 and RBF SVM. Regarding the performance of distances, DL2 keeps the highest CRRs for all SVMs (note that the CRR of DL2 and sigmoid SVM is not shown). Moreover, dense distances have a higher overall performance than sparse distances. This reflects that emotional information in the CK images is distributed over all orientations rather than the dominant orientation of Gabor features. As for SVMs, RBF performs the best for dense distances, while linear performs the best for sparse distances. This confirms with the results in that RBF and linear perform better than polynomial on the CK database.

The confusion matrix of six emotions using DL2 and RBF SVM. As can be seen, surprise performs the best with a CRR of 100%, the following one is happy with a CRR of 98.07%. On the other hand, anger is the most difficult facial expression to be correctly recognized with a CRR of only 87.10%. The performance of surprise and anger on CK contrasts with that on JAFFE, in which surprise and anger are the most difficult and easiest emotions respectively. The reason probably is that surprise images on CK are often characterized as an exaggerated “open mouth”, while those on JAFFE are normally with a “close or slightly open mouth”. This can be seen from that the selected patches for CK focus on the mouth region, but those for JAFFE are mainly distributed around the eyes regions. Similarly, anger images on JAFFE are better expressed by the selected patches in mouth region than the selected patches are all over the face region those on CK. Among six emotions, anger and sad contribute most to the misrecognition rate.

Conclusion

In this paper, explores the issue of facial expression recognition using facial movement features. The effectiveness of the proposed approach is testified by the recognition performance, computational time, and comparison with the state-of-the-art performance. The experimental results also demonstrate significant performance improvements due to the consideration of facial movement features, and promising performance under face registration errors. The results indicate that patch-based Gabor features show a better performance over point-based Gabor features in terms of extracting regional features, keeping the position information, achieving a better recognition performance, and requiring a less number. Different emotions have different ‘salient’ areas; however, the majority of these areas are distributed around mouth and eyes.

In addition, these ‘salient’ areas for each emotion seem to be not influenced by the choice of using point-based or using patch-based features. The ‘salient’ patches are distributed across all scales with an emphasis on the higher scales. For both the JAFFE and CK databases, DL2 performs the best among four distances. As for emotion, anger contributes most to the misrecognition. The JAFFE database requires larger sizes of patches than the CK database to keep useful information. The proposed approach can be potentially applied into many applications, such as patient state detection, driver fatigue monitoring, and intelligent tutoring system. In our future work, we will extend our approach to a video based FER system by combining patch-based Gabor features with motion information in multi-frames. Recent progress on action recognition and face recognition has laid a foundation for using both appearance and motion features.

References

1. L. Sirovich and m. Kirby, “low dimensional procedure for characterization of human faces”, j. Optical soc. Am., vol. 4, pp. 519-524, 1987.
2. P. N. Belhumeur, j. P. Hespanha, and d. J. Kriegman, “eigenfaces vs. Fisherfaces: recognition using class specific linear projection”, ieee transactions on pattern analysis and machine intelligence, vol. 19, no. 7, pp. 711–720, jul. 1997.
3. D. Arumugam, s. Purushothaman, “emotion classification using facial expression”, international journal of advanced computer science and applications, vol. 2, no. 7, pp. 92-98, 2011.
4. A.k. Jain, r.p.w. Duin, j mao, "statistical pattern recognition: a review", ieee trans. Pattern analysis and machine intelligence, vol. 22, no. 1, pp. 4-37, 2000.
5. K.schindler and l. Van gool, “action snippets: how many frames does human action recognition require?”, in computer vision and pattern recognition, 2008. Cvpr2008. Ieee conference, 2008, pp.18.
6. A.hadid and m.pietikainen, “combining appearance and motion for face and gender recognition from videos”, pattern recognition, vol.42, pp.2818-2827, 2009.
7. M. Armbrust, a. Fox, r. Griffith, a. D. Joseph, r. H. Katz, a. Konwinski, g. Lee, d. A. Patterson, a. Rabkin, i. Stoica, and m. Zaharia, “above the clouds: a berkeley view of cloud computing”, university of california, berkeley, tech. Rep. Usb-eecs-2009-28, feb 2009.
8. “the apache cassandra project”, http://cassandra.apache.org/.
9. l. Lamport, “the part-time parliament”, acm transactions on computer systems, vol. 16, pp. 133–169, 1998.
10. n. Bonvin, t. G. Papaioannou, and k. Aberer, “cost-efficient and differentiated data availability guarantees in data clouds”, in proc. Of the icde, long beach, ca, usa, 2010.
11. O. Regev and n. Nisan, “the popcorn market. Online markets for computational resources”, decision support systems, vol. 28, no. 1-2, pp. 177 – 189, 2000.
12. A. Helsinger and t. Wright, “cougaar: a robust configurable multi agent platform”, in proc. Of the ieee aerospace conference, 2005.
13. D. Arumugam, s. Purushothaman, “emotion classification using facial expression”, international journal of advanced computer science and applications, vol. 2, no. 7, pp. 92-98, 2011.