Facial Mood Expression Recognitions: A Review

Md. Mosharraf Hossain
ID: 16155031
Dept. of Information Technology
University of Information Technology & Sciences

Abstract—With the unambiguous statement on the importance of facial configuration for the judgments of human emotion, facial expression recognition has been the utmost research in affective computing in the recent years. The motive is to empower computer so that it could be adaptive to extracting and classifying various user emotions into the “universal facial expressions” or its subsets. This project reviews a number of methodological approaches relative to real-time facial expression recognitions systems and proposes further research areas that require more attention towards the successful implementation of a more efficient channel for machine - emotion interaction.

I. INTRODUCTION

A challenging research issue and one that has been of growing importance to those working in human-computer interaction is to endow a machine with an emotional intelligence [1]. In other to increase the said human-computer interaction, computers must be given some abilities to sense and respond to what it senses. This research area that addresses issues relating to emotion in computing is refers to as affective computing [2].

Human-machine emotion recognition involves series of stages, hence He et al. (2008) [3] conceptualized facial expression recognition into three folds: face detection, discriminative information extraction and expression classification. Research focus for the last decade has been on recognizing the universal facial expressions which are happiness, surprise, anger, sadness, fear and disgust or a subset of these prototypic facial expressions from images [3, 4, 5, 6].

This review firstly crystallizes some of the reported approaches on real-time face detection and tracking. Since the accuracy of the face expression recognition systems depends solemnly on the effectiveness of the adopted feature extractions techniques, various publications on facial feature extraction will be reviewed. Although some methods can possibly belong to more than one section, some documented methods employed in facial expression classification will equally be analyzed. The last section of the paper summarizes the generally identified loop holes in the various reviewed real-time facial expression recognition systems with suggestions on possible future research areas to explore.

II. FACE DETECTION AND TRACKING

The first and the fundamental step in any facial expression is face detection and tracking. The goal is to detect and track unknown number of faces either in a still or video image. With the existence of human variability in size, shape, color, and texture, different approaches are used in the detection and tracking stage by many researchers. Froba & Kublbeck in 2002 [7] used edge detection from the classified feature-based approach to face detection which was later appended with the research of Suzuki & Shibata in 2004 [8]. Hao & Wang [9] introduced texture-based approach in 2002 while the use of color for the face detection were researched by Liu et al. (2005) [10], Tsalakanidou et al. (2005) [11] and others.

Researches have shown that motion-based approach is efficient to detect faces or heads when video sequences are available, hence Faundzzanuy & Ortega (2004) [12], Deng et al. (2005) [13], Lie´vin et al. (2004) [14] and Loutas et al. (2004) [15] implemented motion-based approach to face detection.

Active Appearance Model and Active Shape Model are commonly used as deformable model and recently being used by researchers for face detection, Choi Oh (2007) [16], Mahoor & Mottaleb†(2006) [17].

Neural Networks, Support Vector Machines, Probabilities Model, STAAM & MVUAM, Principal Components Analysis are series of appearance-based model that are also implemented by a number of recent researches, Ulrich et al. (2008) [18], Sung et al.(2006) [19], Kotsia & Pitas (2005) [20], Wang et al. (2004) [21], Viola & Jones (2004) [22] and others.

Table I below provides a summary of the reviewed face detection and tracking approaches.

TABLE I
Face Detection & Tracking Approaches

Authors	Approaches	Comments
Suzuki & Shibata (2004) Froba& Kublbeck (2002)	Edges	Features can be easily detected within a very short time but not robust for face detection in complex environment.
Hao & Wang (2002)	Texture	For detecting lips, pupils and eyebrow since they always darker than other region around.

Z. Liu, J. Yang, N.S. Peng,(2005) F. Tsalakanidou, S. Malassiotis, M.G. Strintzis, (2005) M. Soriano, B. Martinkauppi, S. Huovinen, M. Laaksonen, (2003) R.L. Hsu, M. Abdel-Mottaleb, A.K. Jain, (2003)	Colour	Using color model like RGB, YIQ.
Espinosa-Duro, Faundzzanuy & Ortega (2004) Deng, Su, Zhau & Fu (2005) M. Lie´vin, F. Luthon, (2004) E. Loutas, I. Pitas, C. Nikou,(2004)	Motion	Appropriate when video sequences are available.
Mohammad H. Mahoor and Mohamed Abdel-Mottaleb†(2006) Hyun-Chul Choi and Se-Young Oh (2007)	Deformable Model	Generally used for face model (AAM, ASM)
R. Xiao, M. Li, H. Zhang,(2004) Y. Li, S. Gong, J. Sherrah, H. Liddell, (2004) P. Viola, M.J. Jones,(2004) L. Huang, A. Shimizu, Y. Hagihara, H. Kobatake,(2003) Yubo Wang, Haizhou AI, Bo Wu, Chang Huang (2004) Jaewon Sung, Sangjae Lee, Daijin Kim (2006) I. Kotsia and I. Pitas (2005) Hoffmann, Ulrich ; Naruniec, Jacek; Yazdani, Ashkan ; Ebrahimi, Touradj (2008) Zhou et al. (2004)	Appearance-Based	Mostly most effective with superior performance. But are difficult to apply to real time.

III. FACIAL FEATURE EXTRACTION

Facial feature extraction is the next step to designing a facial expression recognition system. Meanwhile this area has also drawn a lot of attention of researchers in the recent years. In 2007, Shen et al. [23] achieved this using Gobor Wavlets while Zhao et al. [24] also came up with Linear Laplacian Discrimination which was an extension to Linear Discriminant Analysis (LDA). Fishers Linear Discriminant Analysis (FLDA) were also adopted by Ling Guan (2005) [25], and Principal Component Analysis by Park et al. (2005) [26] while Zheng et al. (2004) [27] introduced Gram-Schmidt Orthogonalization for linear discriminant analysis (GSLDA) for facial feature extraction.

Table II summarizes some of the reviewed researches on approaches employed for facial feature extraction.

Table II
Facial Feature Extraction Approaches

Authors	**Approaches**	**Comments**
LL Shen, L Bai, M Fairhurst (2007)	Gobor Wavelets	Robust against facial expression, illumination, and can be also effect if the person wearing glassed, and so on.
MS Park, JH Na, JY Choi (2005)	PCA	Is an holistic methods, which uses the whole face region as raw input to the recognition system
Yongjin Wang Ling Guan (2005)	FLDA	It involves a long computational time, based on a large-sized covariance matrix
D Zhao, Z Lin, R Xiao, X Tang (2007)	LLD, LDA	LDA is problematic in terms of computational demands and numerical accuracy when the size of the matrices is large while LLD is an extension of LDA.
Wenming Zheng, Cairong Zou, Li Zhao (2004)	GSLDA	Able to avoid the large matrices computation and achieves better recognition performance than LDA.

IV. FACIAL EXPRESSION CLASSIFICATIONS

An effective facial feature extraction approach enhances accuracy of expression classifications. Recent researches are employing Neural Networks (Tivive & Bouzerdoum, 2004) [28], Support Vector Machines, SVM, (Shih & Liu, 2005) [29], Adaptive boosting (Adaboost) (Hayashi & Hasegawa, 2006) [30] and Eigenfaces (Wong et al., 2001) [31] to institute a more functional expression system.

Table III
Reviewed Facial Classification Approaches

Authors	Approaches	Comments
Tivive & Bouzerdoum, (2004)	Neural Networks	Model that Emulates biological neural network.
Shih & Liu, (2005)	SVM	Relies on results from statistical learning theory to guarantee high generalisation performance.
Hayashi & Hasegawa, (2006)	Adaboost	Sensitive to noisy data and outliers.
Wong, Lam, Siu & Tse, (2001)	Eigenfaces	first successful example of facial recognition technology, complemented with eigenfeatures.

V. DISCUSSION

Real-time facial expression recognition has a quite challenging left area in affective computing especially for certain unconstrained environments. Most of the reported real-time facial expression systems with optimum accuracy are restricted to recognizing either only frontal views of expressions at a single scale; relevant when user sit in front of a personal computer, or constrained to a pre-defined distance between user's face and the camera. Infact, almost all such facial expression recognition system require the whole capturing of the user's face before classification can be carried out successfully.

In the same vein, visual images have been identified with some loop holes. According to Shiqian et al. (2007) [32], IR imaging is independent of external illumination. While visual images represent the reflectance information of the facial surface, IR facial images contain more information about faces themselves, such as anatomical information. The thermal characteristics of faces with variations in facial expression and make-up remain nearly invariant, and the tasks of face detection, location, and segmentation are relatively easier and more reliable than those in visual images.

VI. CONCLUSION

Optimistically, the use of an IR facial image will enhance the implementation of a more effective real-time facial expression recognition system applicable in a more human-natural situation.

Hence, future research on such area will cater for the present identified loop holes in the real- time facial expression recognition systems.

REFERENCES

[1] Ginevra Castellano, Loic Kessous, and George Caridakis (2008): Emotion Recognition through Multiple Modalities: Face, Body Gesture, Speech" Affect and Emotion in HCI, LNCS 4868, pp. 92–103.

[2] Keith Anderson and Peter W. McOwan (2006): "A Real-Time Automated System for the Recognition of Human Facial Expressions" IEEE Transactions on Systems, Man, and Cybernetics Part B VOL. 36, NO. 1.

[3] He Li, Jos´e M. Buenaposada and Luis Baumela (2008): "Real-Time Facial Expression Recognition with Illumination-Corrected Image Sequences" IEEE.

[4] Y. Chang, C. Hu, and M. Turk (2004): "Probabilistic expression analysis on manifolds. In Proc. of CVPR, volume 2, pages 520–527.

[5] C. Shan, S. Gong, and P. W. McOwan (2006): "Dynamic facial expression recognition using a bayesian temporal manifold model" In Proc. BMVC, volume 1, pages 297–306.

[6] G. Zhao and M. Piettikainen (2007): "Dynamic texture recognition using local binary patterns with an application to facial expressions" Trans. on PAMI, 29(6):915–928.

[7] Froba, B., & Kublbeck C. (2002): "Robust face detection at video frame rate based on edge orientation features." Proceedings of the International Conference on Automatic Face and Gesture Recognition, 327-332.

[8] Suzuki, Y., & Shibata T. (2004): "Multiple-clue face detection algorithm using edge-based feature vectors." Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, 5, V-737-40.

[9] Hao, W., & Wang K. (2002): "Facial feature extraction and image-based face drawing." Proceedings of the International Conference on Signal Processing, 699-702.

[10] Z. Liu, J. Yang, N.S. Peng (2005): "An efficient face segmentation algorithm based on binary partition tree." Signal Processing: Image Communication Vol 20.

[11] F. Tsalakanidou, S. Malassiotis, M.G. Strintzis (2005): " Face localization and authentication using color and depth images." IEEE Transactions on Image Processing 14 (2) 152–168.

[12] Espinosa-Duro, V. , Faundez-Zanuy M., & Ortega J. A. (2004): "Face detection from a video camera image sequence." Proceedings of International Carnahan Conference on Security Technology, 318-320.

[13] Deng Y-F., Su G-D., Zhou J. & Fu B. (2005): "Fast and robust face detection in video." Proceedings of the International Conference on Machine Learning and Cybernetics, 7, 4577-4582.

[14] M. Lie´vin, F. Luthon (2004): "Nonlinear color space and spatiotemporal MRF for hierarchical segmentation of face features in video" IEEE Transactions on Image Processing 13 (1) 63–71.

[15] E. Loutas, I. Pitas, C. Nikou (2004): " Probabilistic multiple face detection and tracking using entropy measures" IEEE Transactions on Circuits and Systems for Video Technology 14 (1) (2004) 128–135.

[16] Hyun-Chul Choi and Se-Young Oh (2007): "Facial Identity and Expression Recognition by using Active Appearance Model with Efficient Second Order Minimization and Neural Networks" Proceedings of the 2007 IEEE International Symposium on Computational Intelligence in Robotics and Automation Jacksonville, FL, USA, June 20-23.

[17] Mohammad H. Mahoor and Mohamed Abdel-Mottaleb† (2006): "Facial Features Extraction in Color Images Using Enhanced Active Shape Model" Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition (FGR'06)

[18] Hoffmann, Ulrich ; Naruniec, Jacek ; Yazdani, Ashkan ; Ebrahimi, Touradj (2008): "Face Detection Using Discrete Gabor Jets and a Probabilistic Model of Colored Image Patches, In: Communications in Computer and Information Science (CCIS series).

[19] Jaewon Sung, Sangjae Lee, Daijin Kim (2006): "A Real-Time Facial Expression Recognition using the STAAM" Proceedings

of the 18th International Conference on Pattern Recognition (ICPR'06)

[20] I. Kotsia and I. Pitas (2005): “Real Time Facial Expression Recognition from image sequences using support vector machines” IEEE.

[21] Yubo Wang, Haizhou AI, Bo Wu, Chang Hung (2004): “Real Time Facial Expression Recognition with Adaboost” Proceedings of the 17th International Conference on Pattern Recognition (ICPR’04).

[22] P. Viola, M.J. Jones (2004): “ Robust real-time face detection, International Journal of Computer Vision 57 (2) 137–154.

[23] LL Shen, L Bai, M Fairhurst (2007): “Gabor wavelets and General Discriminant Analysis for face identification and verification” Image and Vision Computing.

[24] D Zhao, Z Lin, R Xiao, X Tang (2007): “Linear Laplacian discrimination for feature extraction” International Conf. on Computer Vision and Pattern 2007.

[25] Yongjin Wang Ling Guan (2005): “Recognizing human emotion from audiovisual information Acoustics, Speech, and Signal Processing, 2005. Proceedings. (ICASSP '05). IEEE International Conference Volume: 2, page(s): ii/1125- ii/1128 Vol. 2.

[26] MS Park, JH Na, JY Choi (2005): “PCA-based feature extraction using class information” IEEE International Conference on Systems, Man and Cybernetics.

[27] Wenming Zheng, Cairong Zou, Li Zhao (2004): “Real-Time Face Recognition Using Gram-Schmidt Orthogonalization for LDA” Proceedings of the 17th International Conference on Pattern Recognition (ICPR’04).

[28] Tivive, F. H. C., & Bouzerdoum A. (2004): “A face detection system using shunting inhibitory convolutional neural networks.” Proceedings of IEEE International Joint Conference on Neural Networks, 4, 2571-2575.

[29] Shih, P., & Liu C. (2005): “Face detection using distribution-based distance and support vector machine.” Proceedings of the International Conference on Computational Intelligence and Multimedia Applications, 327-332.

[30] Hayashi, S., & Hasegawa O. (2006): “A Detection Technique for Degraded Face Images.” Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1506-1512.

[31] Wong K. W., Lam K. M., Siu W. C., & Tse K. M. (2001): “Face segmentation and facial feature tracking for videophone applications.” Proceedings of the International Symposium on Intelligent Multimedia, Video and Speech Processing, 518-521.

[32] Shiqian Wu, Zhenghui Gu, Kia Ai Chia, Sim Heng Ong (2007): “Infrared Facial Recognition Using Modified Blood Perfusion” Information, Communications & Signal Processing, 2007 6th International Conference onVolume , Issue , 10-13 Page(s):1 - 5