International Journal of Research in Engineering, Science and Management
Volume-2, Issue-10, October-2019
www.ijresm.com | ISSN (Online): 2581-5792

Tracking and Recognition of Facial activities from Images or Videos

Nikhil S. Kale[1], S. A. Vyawhare[2]
[1]*M.E. Student, Department of CSIT, Sanmati Engineering College, Washim, India*
[2]*Professor, Department of CSIT, Sanmati Engineering College, Washim, India*

***Abstract*: The tracking and recognition of facial activities from images or videos have attracted great attention in computer vision field. The facial feature tracking and expression recognition represent the facial activities in three levels from local to global, and they are interdependent problems. However, most current methods only track or recognize the facial activities in one or two levels, and track them separately, either ignoring their interactions or limiting the interaction to one way. Generally, methods such as Multi feature Discriminant Analysis, dynamic Bayesian network; gradient orientation pyramid, etc. are used for feature extraction. The modified PCA algorithm called the Self-PCA is used for creating the Eigen Face for feature extraction for the recognition process. The Self-PCA can be used in order to consider distinctiveness of the effects of facial expressions of a person for face recognition. A facial expression is one or more motions or positions of the muscles beneath the skin of the face. These movements convey the emotional state of an individual to observers. Facial expressions are a form of nonverbal communication. Humans can adopt a facial expression voluntarily or involuntarily, and the neural mechanisms responsible for controlling the expression differ in each case. The universally accepted categories of emotion are: Sadness, Anger, Happiness, Fear, Disgust and Surprise. In previous days much work has been carried out for facial feature extraction and expression recognition using various techniques for improving the accuracy rate however the aim of proposed work is to achieve high level accuracy reducing the computational time. Instead of the work done yet there is needed to achieve huge level accuracy by reducing the computational time. The community of computer vision has attracted the attention of facial emotion recognition over the last decade A vast amount of work has been done and is in progress to make life easy for the disabled (e.g. blind, dumb) and aged people by the help of improving all aspects of interaction between computers and human beings. The recognition of face expression is an important way in interaction between human and computer. Mathematical model of facial expression has been done by extracting the features from face. Many researchers have been proposed to analyze facial expression.**

The facial feature tracking and expression recognition represent the facial activities in three levels from local to global, and they are interdependent problems. For example, facial feature tracking can be used in the feature extraction stage in expression recognition, and expression recognition results can provide a prior distribution for facial feature points. current methods only track or recognize the facial activities in one or two levels, and track them separately, either ignoring their interactions or limiting the interaction to one way. In addition, the estimates obtained by image-based.

***Keywords*: Preprocessing, Facial Feature Extraction and Facial Expression Recognition**

1. Introduction

The facial feature tracking, AU recognition and expression recognition represent the facial activities in three levels from local to global, and they are interdependent problems. For example, facial feature tracking can be used in the feature extraction stage in expression/AUs recognition, and expression/AUs recognition results can provide a prior distribution for facial feature points. However, most current methods only track or recognize the facial activities in one or two levels, and track them separately, either ignoring their interactions or limiting the interaction to one way. In addition, the estimates obtained by image-based methods in each level are always uncertain and ambiguous because of noise, occlusion and the imperfect nature of the vision algorithm. In this paper, in contrast to the mainstream approaches, we build a probabilistic model based on the Dynamic Bayesian Network (DBN) to capture the facial interactions at different levels. Hence, in the proposed model, the flow of information is two-way, not only bottom-up, but also top-down. In particular, not only the facial feature tracking can contribute to the expression/AUs recognition, but also the expression/AU recognition helps to further improve the facial feature tracking performance. Given the proposed model, all three levels of facial activities are recovered simultaneously through a probabilistic inference by systematically combining the measurements from multiple sources at different levels of abstraction.

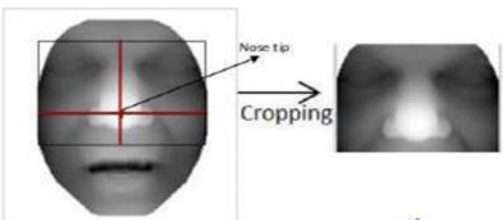
2. Facial feature extraction techniques

A. *Dynamic Bayesian Network (DBN)*

In contrast to the mainstream approaches, a probabilistic model based on the Dynamic Bayesian Network (DBN) has been built to capture the facial interactions at different levels. Hence, the flow of information is two-way, not only bottom-up, but also top-down. In general, not only the facial feature tracking can contribute recognition, but the expression/AU recognition also to the expression/AUs helps to improve the facial feature tracking performance. All three levels of facial activities are recovered simultaneously through a probabilistic inference by systematically combining the measurements from multiple sources at different levels of abstraction. The facial feature point measurements has been tracked through an active shape model (ASM) based approach [1], which first searches each point locally and then constrains the feature points based on the ASM model, so that the feature points can only deform in specific ways found in the training data. All the 26 facial feature point positions are manually labeled in each training image

B. Crop faces and Interpolation

The face is cropped to isolate such regions. The cropping process was done with respect to the nose region and cropping points were chosen along the nose tip vertical and horizontal axes [2]. This cropping process could include regions of the background, as shown in Figure 2(a), with the top left and top right corners representing intersection regions between face and background.

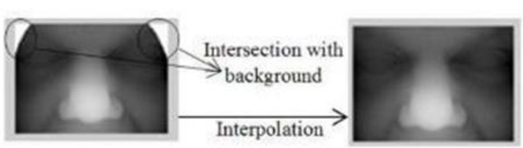


(a) (b)

Fig. 1. a) Face before cropping. b) Face after cropping originally shown in [2]

The Curvelet transform considers the intersection region as significant regions due to the high variance at edges (curves). To reduce the Curvelet sensitivity to these regions were interpolated face regions over extracted background regions.

After the completion of cropping process, each cropped face was decomposed into four levels of scales using the Curvelet transform. Scale 1 captures the lowest frequency components with high variance while Scale 4 captures the highest frequency components with low variance. However, they do not strongly capture geometric features, due to the absence of angle decomposition. Scale 2 and scale 3 have 16 and 32 sub bands of Curvelet coefficients, respectively. Thus, the Curvelet transform was applied to cropped faces before extracting the features (coefficients) of each wanted region separately.



(a) (b)

Fig. 2. a) Cropped face before interpolation. b) Cropped face after interpolation originally or actually shown in [2]

3. Proposed work and objectives

In contrast to the mainstream approaches, which usually only focus on one or two levels of facial activities, and track (or recognize) them separately, this paper introduces a unified probabilistic framework based on the dynamic Bayesian network to simultaneously and coherently represent the facial evolvement in different levels, their interactions and their observations. Advanced machine learning methods are introduced to learn the model based on both training data and subjective prior knowledge. Given the model and the measurements of facial motions, all three levels of facial activities are simultaneously recognized through a probabilistic inference. Extensive experiments are performed to illustrate the feasibility and effectiveness of the proposed model on all three level facial activities.

4. Classification techniques

A. Principle Component Analysis (PCA)

PCA is one of the most widely used methods in image recognition and compression. PCA aims to obtain a set of mutually orthogonal bases that describe the global information of the data points in terms of variance. PCA has been successfully applied to discover the subspace of the face space, which is termed as eigen faces.

Scatter maximization is the main drawback in PCA which is not only due to the between-class scatter that is useful for classification, but also to the within-class scatter. Maximization of within-class scatter includes unwanted information to the classification process [2]. Hence, PCA might be optimal in terms of dimensionality reduction; however, it may not be optimal in terms of discriminating images of one class from images of other classes.

B. Grassmannian Manifold

A Grassmannian manifold is a space of all d-dimensional linear subspaces. The existing clustering techniques on the Grassmannian manifold need to compute the distance as well as the mean of the points on the manifold [4], for every iteration of the clustering algorithm (e.g. K-means). Methods of computations of the mean and distance on the Grassmannian manifold can be broadly categorized as intrinsic and extrinsic

C. Local Fisher Discriminant Analysis

The technique of facial expression recognition in the encrypted domain based on LFDA. Basically, LFDA divides image samples in each class into multiple local classes in the higher dimensional image space [5]. It then projects images belonging to a local class closer to each other while keeping projected images of other local classes apart. However, FLDA and LFDA project samples of each class separately in the feature space as both of them are supervised feature extractors. Both PCA and FLDA smear the samples of both classes together while LFDA preserves the local structure by projecting them into separable region(s) in the feature space.

5. Conclusion

After the investigation of different techniques for facial feature extraction for facial expression recognition, it has been concluded that though different techniques like crop faces and interpolation, Region of Interest, Pseudo Zernike Moments, FACE, discriminative features, video patches etc. used for facial feature extraction for expression recognition provide different level of accuracy but still there is a need to achieve high level of accuracy rate by reducing the computation time.

Acknowledgement

It gives The making of this work needed co-operation and guidance of a number of people. I therefore consider it my prime duty to thank all those who had helped me through their venture. It is my immense pleasure to express my sincere thanks to the Head of Department and guide Prof. S. A. Vywahare who provided me constructive and positive feedback. I express my sincere thanks to the head of department and all other staff members of CSE department for their kind co-operation.

I am thankful to my friends and library staff members whose encouragement and suggestion helped me to complete my proposed work.

References

[1] [1] G. Donato, M. S. Bartlett, J. C. Hager, P. Ekman, and T. J. Sejnowski, “Classifying facial actions,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 21, no. 10, pp. 974–989, Oct. 1999.
[2] Y. Tian, T. Kanade, and J. F. Cohn, “Recognizing action units for facial expression analysis,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 23, no. 2, pp. 97–115, Feb. 2001.
[3] G. Zhao and M. Pietikainen, “Boosted multi-resolution spatiotemporal descriptors for facial expression recognition,” Pattern Recognit. Lett., vol. 30, no. 12, pp. 1117–1127, 2009.
[4] M. Valstar and M. Pantic, “Combined support vector machines and hidden Markov models for modeling facial action temporal dynamics,” in Proc. IEEE Int. Conf. Human-Comput. Interact., vol. 4796. Oct. 2007, pp. 118–127.
[5] M. Valstar and M. Pantic, “Fully automatic facial action unit detection and temporal analysis,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshop, Jun. 2006, p. 149.
[6] A. Kapoor, Y. Qi, and R. W. Picard, “Fully automatic upper facial action recognition,” in Proc. IEEE Int. Workshop Anal. Model. Faces Gestures, Oct. 2003, pp. 195–202.
[7] M. Pantic and I. Patras, “Dynamics of facial expressions: Recogniton of facial actions and their temporal segments from face profile image sequences,” IEEE Trans. Syst., Man, Cybern. B, Cybern., vol. 36, no. 2, pp. 433–449, Apr. 2006.
[8] Z. Zeng, M. Pantic, G. I. Roisman, and T. S. Huang, “A survey of affect recognition methods: Audio, visual, and spontaneous expressions,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 31, no. 1, pp. 39–58, Jan. 2009.
[9] J. J. Lien, T. Kanade, J. F. Cohn, and C. Li, “Detection, tracking, and classification of action units in facial expression,” J. Robot. Auto. Syst., vol. 31, no. 3, pp. 131–146, 2000.
[10] Y. Tong, J. Chen, and Q. Ji, “A unified probabilistic framework for spontaneous facial activity modeling and understanding,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 32, no. 2, pp. 258–273, Feb. 2010.
[11] Y. Tong, W. Liao, and Q. Ji, “Facial action unit recognition by exploiting their dynamic and semantic relationships,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 29, no. 10, pp. 1683–1699, Oct. 2007.
[12] F. Dornaika and F. Davoine, “Simultaneous facial action tracking and expression recognition in the presence of head motion,” Int. J. Comput. Vis., vol. 76, no. 3, pp. 257–281, 2008.
[13] K. Schwerdt and J. L. Crowley, “Robust face tracking using color,” in Proc. IEEE Int. Conf. Autom. Face Gesture Recognit., Mar. 2000, pp. 90–95.
[14] J. L. Minoi and D. F. Jupit, “Facial Expression Reconstruction of 3D Faces based on real Human Data,” IEEE Conf. on Cybern. Com., pp. 185-189, 2012.
[15] K. T. Song and S. C. Chein, “Facial Expression Recognition Based on Mixture of Basic Expression & Intensities,” IEEE Conf. on Sys., Man & Cybern., pp. 3123-3128, Oct 14-17, 2012.
[16] Marc Mehu, Bihan Jiang and Maja Pantic, “Meta- Analysis of the First Facial Expression Recognition Challenge,” IEEE Trans. On Sys., Man & Cybern. Part B: Cybern, Vol 42, No. 4, pp. 966-979, August 2012.
[17] S. Yang and B. Bhanu, “Understanding Discrete Facial Expression in video using in Emotion Avatar Image,” IEEE Trans. On Sys., Man & Cybern., Vol 42, No. 4, pp. 980-992, August 2012.
[18] R. A. Khan, Alexandre Meyer, Hubert Konik and Saida Bouakaz, “Human Vision Inspired Framework for Facial Expressions Recognition,” IEEE Conf., pp. 2593-2596, 2012.
[19] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews, “The extended Cohn-Kande dataset (CK+): A complete facial expression dataset for action unit and emotion-specified expression,” in Proc. 3rd IEEE Int. Conf. Comput. Vis. Pattern Recognit., pp. 94–101, Jun. 2010.
[20] T. Kanade, J. Cohn, and Y. L. Tian, “Comprehensive database for facial expression analysis,” in Proc. 4th IEEE Int. Conf. Autom. Face Gesture Recognit., pp. 46–53, Mar. 2000.
[21] X. Zhao and S. Zhang, “Facial Expression Recognition Based on Local Binary Patterns and Kernel Discriminant Isomap,” Sensors, vol. 11, no. 10, pp. 9573-9588, 2011.