Face Atlas Construction and Spatial Representation for Dynamic Expression Recognition

Jefrin Infant Sindhya R[1] Miss. Muthu Ramya S. C[2]
[1]PG Scholar [2]Professor
[1,2]Department of Electronics and Communication Engineering
[1,2]SCAD College of Engineering and Technology, Cheranmahadevi-627414

Abstract— Dynamic facial expression recognition is subject-specific. It varies from person to person and cannot be estimated with a thresholding technique. Facial movements of different expressions are described by a diffeomorphic growth model. Salient longitudinal facial expression atlas is built for each expression by a sparse group wise image registration method, which can describe the overall facial feature changes among the whole population and can suppress the bias due to large inter-subject facial variations. The image appearance information in spatial domain and topological evolution information in temporal domain are used to guide recognition by a sparse representation method. One limitation of the proposed method is that it is still not robust enough to overcome challenges of strong illumination changes. The main reason is that the LDDMM registration algorithm do not support strong illumination changes.

Key words: Dynamic Facial Expression, Diffeo-Morphic Growth Model, Groupwise Registration, Sparse Representation

I. Introduction

Automatic facial expression recognition (AFER) has essential real world applications. Its applications include, but are not limited to, human computer interaction (HCI), psychology and telecommunications. It remains a challenging problem and active research topic in computer vision, and many novel methods have been proposed to tackle the automatic facial expression recognition problem Facial expression analysis and recognition has been one of the fast analysis area due to its wide range of application areas such as emoticon analysis, bio metrics and image retrieval. Facial recognition has been used in various real life applications such as security systems, computer graphics, psychology and computer vision. Facial expression recognition has been used in various other fields such as in projects for example, to identify driver’s expression to decide whether he is tired or not during a long drive therefore an alert should be displayed to fulfill the requirements of a safe drive.

In this paper, a new dynamic facial expression recognition method is presented. It is motivated by the fact that facial expression can be described by diffeomorphic motions of muscles beneath the face. Intuitively, ‘diffeomorphic’ means the motion is topologically preserved and reversible. Different from previous works by using pairwise registration to capture the temporal motion, this method considers both the subject-specific and population information by a groupwise diffeomorphic registration scheme. Moreover, both the spatial and temporal information are captured with a unified sparse representation framework. Our method consists of two stages: atlas construction stage and recognition stage. Atlases, which are unbiased images, are estimated from all the training images belonging to the same expression type with groupwise registration.

In the atlas construction stage, a diffeomorphic growth model is estimated for each image sequence to capture subjectspecific facial expression characteristics. To reflect the overall evolution process of each expression among the population, longitudinal atlases are then constructed for each expression with groupwise registration and sparse representation. In the recognition stage, we first register the query image sequence to atlas of each expression. Then, the comparison is conducted from two aspects: image appearance information and temporal evolution information. The preliminary work has been reported in.

For the proposed method, a more advanced atlas construction scheme is used. In previous method, the atlases are constructed using the conventional groupwise registration method, thus lots of subtle and important anatomical details are lost due to the naive mean operation.

To overcome this shortage, a sparse representation based atlas construction method is proposed in this paper. It is capable of capturing subtle and salient image appearance details to guide recognition, and preserving common expression characteristics. In the recognition stage, the previous method in compared image differences between the warped query sequence and atlas sequence, which is based on image appearance information only. In this paper,the temporal evolution information is also taken into account to drive the recognition process. It has shown to provide complementary information to image appearance information and can significantly improve the recognition performance.

II. Literature Survey

Intensive studies have been carried out on AFER problem in static images during the last decade [1], [2]: Given a query facial image, estimate the correct facial expression type, such as anger, disgust, happiness, sadness, fear or surprise. It mainly consists of two steps: feature extraction and classifier design. For feature extraction, Gabor wavelet [3], local binary pattern (LBP) [4], and geometric features such as active appearance model (AAM) [5] are in common use. For classifier, support vector machine is frequently used. Joint alignment of facial images under unconstrained condition has also become an active research topic in AFER [6].

In this paper, a new dynamic facial expression recognition method is presented. It is motivated by the fact that facial expression can be described by diffeomorphic motions of muscles beneath the face. Intuitively, ‘diffeomorphic’ means the motion is topologically preserved and reversible. The formal definition of ‘diffeomorphic’ transformation is given in Section II. Different from previous works [10], by using pairwise registration to capture the temporal motion, this method

D. Rueckert et al(1999)., proposed 'Nonrigid registration using free-form deformations: Application to breast MR images", Here Free From Deformation(FFD) technique is used to capture motions between images that means recording the movement of objects. This method aims to model temporal evaluation process of facial expressions.

Y. Zhang et al(2005)., proposed a 'Active and dynamic information fusion for facial expression understanding from image sequences', In this dynamic facial expression deals with the identification of facial expression of the given image.By systematically representing DBN structure, we can model the relationship between different action unit.By using DBN we can get facial feature points and action units in the output and we can able to recognize expression by model relationship between facial feature point and action unit.

Y. Chang et al(2006).,proposed 'Manifold based analysis of facial expression,'Facial action coding system(FACS) is a system used to categorize human facial movements by their appearance on the face. FACS can detects faces in videos and extracts geometrical features of faces.Movements of individual facial muscles are encoded by Facial action coding system from slight changes in facial appearance.

Z.Zeng et al(2009)., proposed 'A Survey Of Affect Recognition Methods: Audio, Visual, and Spontaneous Expressions'.Examine available approaches for solving the problem of machine understanding of human affective behaviour . Outline some of the scientific and engineering challenges of advancing human affect sensing technology.Large collection of acted affective displays are shared by the researchers in the field

S. Yousefi et al(2010) .,proposed' Facial Expression Recognition based on Diffeomorphic matching', Here diffeomorphic mapping is used in shape analysis, shape modeling an medical image registration. Here diffomorphic mapping is used for facial expression recognition. F irstLanmarks are selected from faces based n manual and automatic method. All the landmark from different images are registered to a reference landmark using rigid registration algorithm.The distance between all sets of landmarks are then computed using diffeomorphic matching. Finally k-Nearest neighbor classifier is used to classify a query images. The classifier is trained based on the images in the data then tested on new query images.

Y. Guo et al (2012)., proposed a 'Dynamic facial expression recognition using longitudinal facial expression atlases', Here facial atlases are constructed using greedy iterative algorithm method. It is easy to understand but here lots of important anatomical details are lost due to its simple operation and here the recognition is based on image appearance information only.

Z. Wang et al(2013)., proposed 'Capturing complex spatio-temporal relations among facial muscles for facial expression recognition. Interval temporal beysian network is proposed here to moel facial expressions. Facial expressions are the outcome of set of muscle motions over a time interval. This movement can create different expressions. Interval temporal Bayesian network is time interval based technique, which allows us to model the relation among both sequence and overlapping temporal events.

III. PROPOSED METHOD

We propose a new dynamic facial expression recognition method. There are mainly two stages: atlas construction stage and recognition stage. In the atlas construction stage, atlas sequence is built where salient and common features for each expression among the population are extracted. Meanwhile, the variations due to inter-subject facial shapes can be suppressed. In the recognition stage, expression type is determined by comparing the corresponding query sequence with each atlas sequence.

In the atlas construction stage, longitudinal atlas of different facial expressions are constructed based on sparse representation groupwise registration. The constructed atlas can capture overall facial appearance movements for a certain expression among the population. In the recognition stage, both the image appearance and temporal information are considered and integrated by diffeomorphic registration and sparse representation. Our method has been extensively evaluated on five dynamic facial expression recognition databases. The experimental results show that this method consistently achieves higher recognition rates than other compared methods.

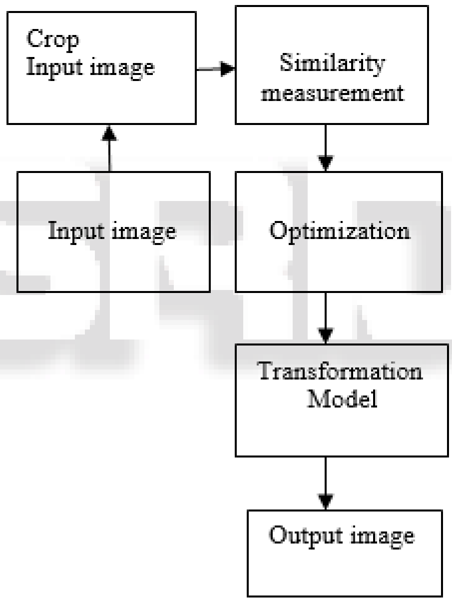

Fig. 3.1: Block Diagram of Proposed Method

As facial expression process is topologically preserved and reversible, it can be considered as a diffeomorphic transformation of facial muscles. Therefore, the diffeomorphic transformation during the evolution process of facial expression can be used to reconstruct facial feature movements and further guide the recognition task. Given P facial expression images $I_1, \ldots\ldots I_p$, a straightforward solution to transform them to a common space is to select one image as the template, then register the remaining P - 1 images to the template by applying P-1 pairwise registration. However, the registration quality is sensitive to the selection of template. Therefore, the idea of groupwise registration was provided, where the template is estimated to be the Frechet mean on the Riemannian manifold whose geodesic distances are measured based on diffeomorphisms.

The diffeomorphic groupwise registration problem can be formulated as the optimization problem. The neutral face gradually evolves to the apex state, and then facial muscles get it back to another neutral state. Therefore, this can be considered as a diffeomorphic transformation process (i.e., topologically preserved and reversible).

In diffeomorphic groupwise registration the estimated template, which is also named atlas, represents overall facial feature changes of a specific expression among the population. The atlas is unbiased to any individual subject and reflects the general expression information. Our dynamic facial expression recognition framework is based on diffeomorphic groupwise registration.

IV. RESULT AND DISCUSSION

The proposed procedure was implemented and tested with set of images.

A. Test Image

In Fig 4.1 Test image is shown, it is given as input. For the given input image, it is necessary to determine the facial expression type.

Fig. 4.1: Test image

B. Crop Input Image

Cropped input image is shown above. From the input image given in 4.2, only face part of the human is cropped to determine the expression effectively.

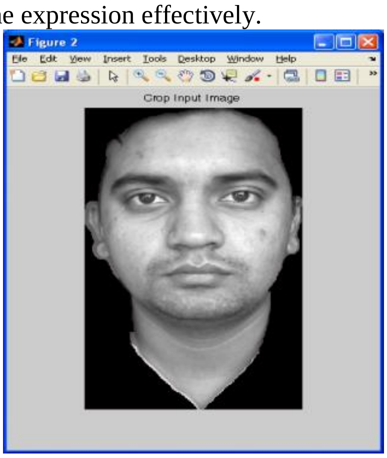

Fig. 4.2: Crop Input Image

C. Similarity Measurement

In similarity measurement test image input is compared with each atlases, for example Happiness atlas, Sad atlas, Surprise atlas, Fear atlas, Anger atlas.

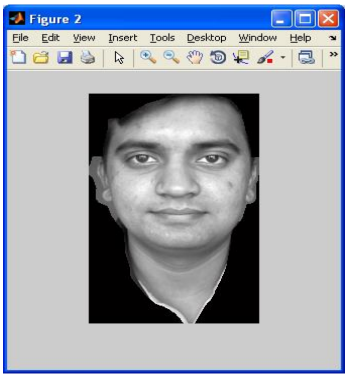

Fig. 4.3: Similarity Measurement

D. Ooptimization

It optimizes the atlas image and only provides the image which are similar to input image. In this project, Input image is compared with 50 atlas images. After optimization, it provides 31 similar images.

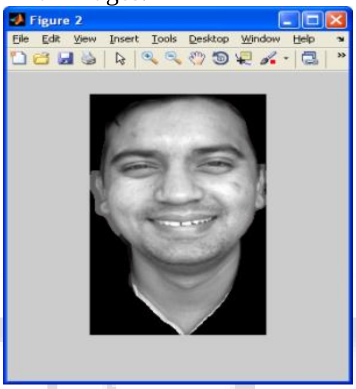

Fig. 4.4: Optimization

E. Transformation Model

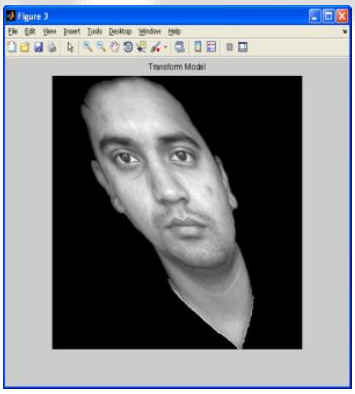

Fig. 4.5: Transformation Model

Based on certain parameters transformation model is generated as shown above.

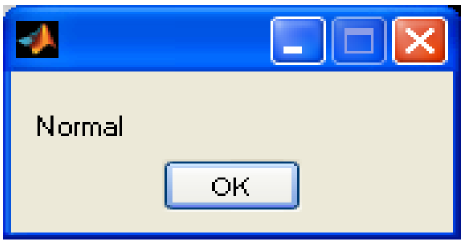

Fig. 4.6: Output Image

Finally, the expression type of given input image is determined and listed as normal facial expression. Similarly, for different facial expressions, we can identify the expression types by using this project.

V. Conclusion

A new way to tackle the dynamic facial expression recognition problem is proposed. It is formulated as a longitudinal atlas construction and diffeomorphic image registration problem. Our method mainly consists of two stages, namely atlas construction stage and recognition stage. In the atlas construction stage, longitudinal atlas of different facial expressions are constructed based on sparse representation groupwise registration. The constructed atlas can capture overall facial appearance movements for a certain expression among the population. In the recognition stage, both the image appearance and temporal information are considered and integrated by diffeomorphic registration and sparse representation. Our method has been extensively evaluated on five dynamic facial expression recognition databases. The experimental results show that this method consistently achieves higher recognition rates than other compared methods. One limitation of the proposed method is that it is still not robust enough to overcome challenges of strong illumination changes. The main reason is that the LDDMM registration algorithm used in this paper may not compensate strong illumination changes. One possible solution is to use complex image matching metrics in the LDDMM framework, such as localized correlation coefficient and localized mutual information which have some degrees of robustness against illumination changes. This is one possible future direction for this study.

References

[1] M. Bartlett, G. Littlewort, M. Frank, C. Lainscsek, I. Fasel, and J. Movel- lan, “Recognizing facial expression: machine learning and application to spontaneous behavior,” in IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005, pp. 568-573.
[2] Y. Chang, C. Hu, R. Feris, and M. Turk, “Manifold based analysis of facial expression,” Image and Vision Computing, vol. 24, no. 6, pp. 605¬614, 2006.
[3] G.David, “Distinctive image features from scale-invariant keypoints,”International Journal of Computer Vision, vol. 60, no. 2, pp. 91–110, 2004.
[4] Y. Guo, G. Zhao, and M. Pietikainen, “Dynamic facial expression recognition using longitudinal facial expression atlases,” in European Conference on Computer Vision, 2012, pp. 631-644.
[5] I. Kemelmacher-Shlizerman, S. Suwajanakorn, and S. Seitz, “Illumination-aware age progression,” in IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2014, pp. 3334–3341.
[6] P. Lucey, J. Cohn, K. Prkachin, P. Solomon, S. Chew, and I. Matthews, “Painful monitoring: Automatic pain monitoring using the unbcmcmastershoulder pain expression archive database,” Image and Vision Computing, vol. 30, pp. 197–205, 2012.
[7] D. Rueckert, L. Sonoda, C. Hayes, D. Hill, M. Leach, and D. Hawkes, “Nonrigid registration using free-form deformations: Application to breast MR images,” IEEE Transactions on Medical Imaging, vol. 18, pp. 712-721, 1999.[16].
[8] J. Saragih, S. Lucey, and J. Cohn, “Deformable model fitting by regularized landmark mean-shift,” International Journal of Computer Vision, vol. 91, no. 2, pp. 200–215, 2011.
[9] S. Seitz, “Collection flow,” in IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2012, pp. 1792–1799.
[10] M. Valstar, M. Mehu, B. Jiang, M. Pantic, and S. K., “Meta-analysisof the first facial expression recognition challenge,” IEEE Transactions on Systems, Man, and Cybernetics Part B, vol. 42, no. 4, pp. 966–979, 2012.
[11] Z. Wang, S. Wang, and Q. Ji, “Capturing complex spatio-temporal relations among facial muscles for facial expression recognition,” in IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2013, pp. 3422-3429. [26].
[12] S. Yousefi, P. Minh, N. Kehtarnavaz, and C. Yan, “Facial expression recognition based on diffeomorphic matching,” in International Conference on Image Processing, 2010, pp. 4549-4552.