Original Article

pISSN 2383-5389 / eISSN 2383-8116

Journal of International Society for Simulation Surgery 2015;2(1):7-9 http://dx.doi.org/10.18204/JISSiS.2015.2.1.007

Facial Feature Extraction with Its Applications

Minkyu Lee, Sangyoun Lee

Department of Electrical and Electronic Engineering, Yonsei University, Seoul, Korea

Purpose In the many face-related application such as head pose estimation, 3D face modeling, facial appearance manipulation, the robust and fast facial feature extraction is necessary. We present the facial feature extraction method based on shape regression and feature selection for real-time facial feature extraction.

Materials and Methods The facial features are initialized by statistical shape model and then the shape of facial features are deformed iteratively according to the texture pattern which is selected on the feature pool.

Results We obtain fast and robust facial feature extraction result with error less than 4% and processing time less than 12 ms. The alignment error is measured by average of ratio of pixel difference to inter-ocular distance.

Conclusion The accuracy and processing time of the method is enough to apply facial feature based application and can be used on the face beautification or 3D face modeling.

Key Words Facial Feature Extraction · Shape Regression · Feature Selection · Statistical Face Model.

Received: May 10, 2015 / **Revised:** May 12, 2015 / **Accepted:** May 25, 2015
Address for correspondence: Sangyoun Lee
Department of Electrical and Electronic Engineering, 50 Yonsei-ro, Seodaemun-gu, Seoul 120-752, Korea
Tel: 82-2-2123-5768, **Fax:** 82-2-362-5563, **E-mail:** syleee@yonsei.ac.kr

Introduction

With the development of biometrical techniques, various face-related applications are emerged for the purpose of authentication, modeling, and entertainment. The greater part of these applications is based on the facial features. For example, 3D face reconstruction from 2D face image (1, 2), 3D head pose estimation (3) and face beautification methods (4, 5) use the labeled facial feature.

In the recent facial feature extraction methods, they are classified with two categories which one is part-based method and the other is shape regression based method.

Part-based methods (6, 7, 8) learn patterns of each facial part e.g. nose, mouth, and eyes, and geometrical relations between each parts. The estimated locations of each facial feature are determined by combination of texture patterns of each part and their geometrical relations with maximum posterior likelihood.

Shape regression methods (9, 10) learn motion vectors from the texture features over the training data. Thus the initial points are moved to desired location using motion vector.

In this paper, we present shape regression based facial feature extraction method and evaluate the performance. The workflow is as in the following. First, the texture patterns are selected from the feature pool. Second, motion vectors whose displacement from previous shape to ground truth is learned according to the selected texture patterns. Finally, with the learned motion vector, we can detect the facial feature from the initial shape model.

Materials and Methods

The main objective function is to minimize the difference between initial shape and ground truth (1).

$$f(X) = X^* - g(X), \tag{1}$$

where X is previous shape, X^* is ground truth shape, and g is updated shape from previous shape.

In the facial feature initialization process, the initial facial feature points are generated by statistical model of training data (Fig. 1).

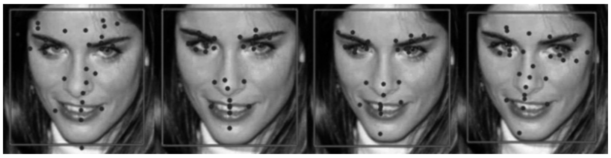

Fig. 1. Facial Feature Initialization.

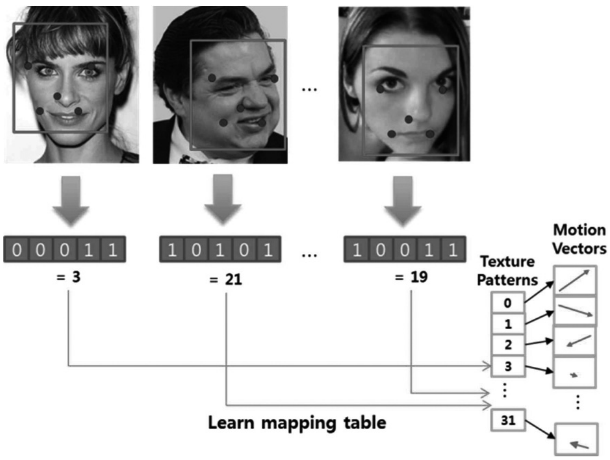

Fig. 2. Learning mapping table from texture patterns to motion vectors.

After the initialization, the mapping table is learned from texture patterns to motion vectors (Fig. 2). The texture patterns are generated by combination of several binary features, ϕ, computed from two points on the relative position to previous shape of facial features with some random displacement (2).

$$\phi(p_i, p_j) = \begin{cases} 1, \text{ } if, p_i - p_j > \theta \\ 0, \text{ } O.W. \end{cases} , \quad (2)$$

where p_i is the pixel value of i-th location and θ is threshold. From these binary features, some features highly correlated to motion vectors are selected and combined to a single texture pattern.

The motion vectors v_k of k-th facial feature are computed with minimum alignment error over the training data (3).

$$v_k = \underset{v}{\operatorname{argmin}} \sum_{n=1}^{N})|x^*_{kn} - (x_{kn} + v)|, \quad (3)$$

where N is the number of training data, x_k is the previous location of k-th facial feature and x^*_k is the ground truth location. In the statistical view of equation (3), the motion vector with minimum error is the average of motion vector over the training samples. Finally the previous shapes are updated using learned motion vectors. The process of the feature selection, learning motion vector, and update shape is repeated T times.

Table 1. Average error and processing time with respect to the number of initialization on LFPW.

# of initialization	Average error	Average error for success samples (<10%)	# of failed sample (> 10%)	Processing time (ms)
5	3.72	3.44	7	6.05
10	3.48	3.31	4	11.63
20	3.38	3.18	5	22.07
30	3.31	3.14	4	33.30
40	3.31	3.10	5	43.38

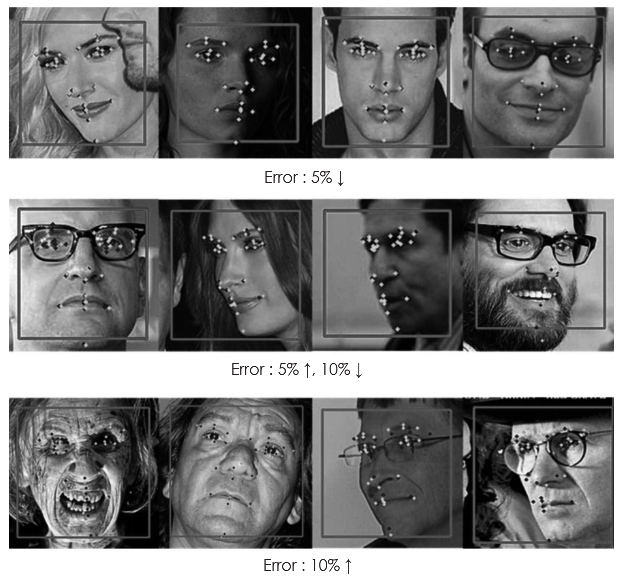

Fig. 3. Facial feature extraction results.

Results

We evaluate the performance of facial feature extraction method with LFPW (Labeled Face Part in the Wild) DB. LFPW DB contains 35 labeled facial features with large pose, occlusion and illumination variation and total 997 samples which are composed of 785 training samples and 212 test samples. Since the annotation is not well located on the original LFPW, we re-annotated all feature points and use only 29 facial features without ears.

For evaluation, alignment error is computed by an average ratio of alignment error to inter-ocular distance (4).

$$\text{error} = \frac{1}{ZL}\sum_{k=1}^{L}(x^*_k - x_k), \quad (4)$$

where Z is inter-ocular distance and L is the number of facial features. We consider any error above 10% to be fail case and also compute the error without fail case.

When we evaluate test sample, multiple initialization is used to avoid bad initializations and achieve more stable and robust alignment result. The evaluation results are described on the table with respect to the number of initialization (Table 1).

Fig. 3 shows facial feature extraction results on the test images using 10 initializations. The white circles are ground truth points and the black circles are estimated points.

Discussion

Based on the experiments over the wild dataset, only 186 of 212 test images get the error below 5%. We realize that the rest test samples showing large pose variation, occlusion and uncommon texture patterns like hooked nose. Since the number of training data set does not enough to span all challenging cases the performance is highly biased to the weak variation cases. However, for the face-related application using near frontal faces, the errors would be acceptable for the application such as 3D face modeling, 3D head pose estimation, and facial appearance manipulation.

Conclusion

In this paper, we present shape regression based facial feature detection method. By the experiment, we show that the shape regression is robust and fast and would be applied to face-related applications.

Acknowledgements

This research was supported by the MKE (The Ministry of Knowledge Economy), Korea and Microsoft Research, under IS/SW Creative research program supervised by the NIPA (National IT Industry Promotion Agency) (NIPA-2014-H0510-14-1034).

REFERENCES

1. Blanz V, Vetter T. A morphable model for the synthesis of 3D faces. In Proceedings of SIGGRAPH, 1999:187-194
2. Han H, Jain AK. 3D face texture modeling from uncalibrated frontal and profile images. IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS), 2012
3. Leyvand T, Cohen-Or D, Dror D, Lischinski D. Data-driven enhancement of facial attractiveness. ACM Transactions on Graphics (TOG) 2008:38
4. Wang Q, Wang Y, Wang Z. Online Smart Face Morphing Engine with Prior Constraints and Local Geometry Preservation. Multimodal Pattern Recognition of Social Signals in Human-Computer-Interaction. Springer International Publishing 2015:130-140
5. Kim WW, Hwang JK, Park SH, Lee SY. Automatic head pose estimation from a single camera using projective geometry. IEEE conference on Information, Communications and Signal Processing (ICICS), 2011
6. Belhumeur PN, Jacobs DW, Krieman DJ, and Kumar N. Localizing parts of faces using a consensus of exemplars. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2011
7. Uřičář M, Franc V, and Hlaváč V. Detector of facial landmarks learned by the structured output SVM. VISAPP 2012:547-556
8. Zhu X, Ramanan D. Face detection, pose estimation, and landmark localization in the wild. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2012
9. Xiong X, De la Torre F. Supervised descent method and its applications to face alignment. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2013
10. Cao X, Wei Y, Wen F, Sun J. Face alignment by explicit shape regression. International Journal of Computer Vision 2014:177-190