Facial Expression Recognition Based on Geometric Features and Geodesic Distance

Kezheng Lin, Weiyue Cheng and Jingtian Li

Department of Computer Science and Technology, Harbin University of Science and Technology Harbin, 150080China
E-mail: link@hrbust.edu.cn

Abstract

The paper mainly studies static 2D face images through reconstructing 3D model by a specific algorithm. First, the paper need collect geometric features, and obtain the three-dimensional space of false geodesic distance. Those are to emotional changes. Second, remove the relative feature extraction. Finally, compares the test sample and the training samples about the Mahalanobis distance. The experimental results show the recognition rate reaching to ninety percent. Results illustrate the validity of the algorithm.

Keywords: geometric features; false geodesic distance; relative feature extraction; Mahalanobis distance

1. Introduction

Emotional information is a important way of information transmission in interpersonal communication. It can convey a lot of information which words cannot convey. [1] 54% of emotional information is communicated through facial expressions [2]. The Information contained in human face accounts for a substantial proportion in image understanding. Facial expression recognition is the process of extraction and classification about facial expression information by computer. The computer can infer a person's mental state according to the facial expression. And it completes the interaction of intelligent human-computer [3]. Facial expression recognition system consists of four basic parts: expression image acquisition, image preprocessing, face expression feature extraction and classification. So, in recent years, the facial expression analysis has attracted attentions from many computer vision researchers. The facial emotions recognition has been one of the dynamic research interests in the field of pattern recognition. People extract the stretch eyebrows, frowned, giggle, half smile, frown, smiling and common human facial expressions in much rich expression. People solve a series of problems of facial expression by environmental interference. The passions will be reflected in facial expression. Expression solves many practical problems especially in face expression. In a broad sense, the facial emotion recognition based on static images includes: pretreatment, face detection, extraction of emotional characteristics and emotions classification [4, 5]. But the traditional expression recognition of static images based on geometrical characteristics is generally classified by obtaining facial characteristics through the Euclidean distance on the 2D plane. This paper represents an expression recognition method by manifesting the facial characteristics through false geodesic distance and remodeling the facial expressions in the 3D space and verifies this method being able to identify different expressions from static facial images in a more effective and remarkable manner. The experiments demonstrate that

this method achieves better results of expression recognition than traditional methods and shows stronger robustness to changes of illumination.

2. Computation Based on False Geodesic Distance

First, select 6 facial images from the Yale faces collection, each containing 5 emotions: normal expression (NOR), happiness (HA), surprise (SU), frown (WI) and sadness (SA). See Figure 1.
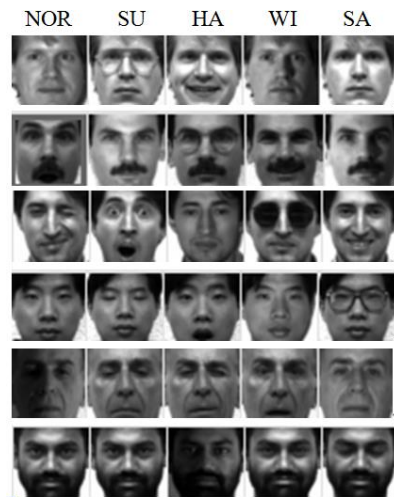

Figure 1. Samples of Faces Collection

Knowledge guidance is an important theory in computer vision. Through analysis of the characteristics of coordinate distributions on human faces in long-term facial metrology experiments, the facial organs distribution rule can be summarized as a "three-part five-eye" distribution structure as shown in Figure 2, *i.e.*, vertically, approximately identical distances from chignon to glabellum, from glabellum to tip of nose and from tip of nose to lower jaw which are called "three parts"; laterally, the distance between the cheekbones can be generally divided from the middle by the width between the two eyes [6-8].
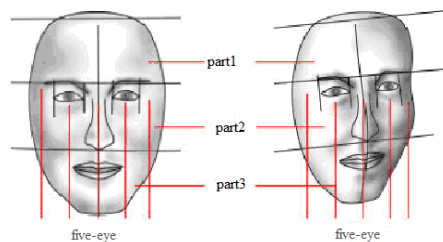

Figure 2. Three-part Five-eye Diagram

Suppose the entire facial plane is referred to as E, nose, cheekbone and forehead are respectively referred to as A^0, B^0 and C^0, and function F is the function for heights of

arbitrary points x and y within the plane, develop 3D facial model per the following 4 restrictions.

1. $\left\{ F(x_A, y_A) < F(x_0, y_0) \mid \forall (x_A, y_B) \in A^O, (x_0, y_0) \notin A \right\} I \left\{ F(x_A, y_A) < \alpha \mid \exists \alpha > 0, \forall (x_A, y_B) \in A^O \right\}$

2. $\left\{ F(x_B, y_B) < F(x_A, y_A) \mid \forall (x_A, y_A) \in A^o, (x_B, y_B) \in B^o \right\} I$
   $\left\{ \left| F(x_B, y_B) - F(x_C, y_C) \right| < \beta \mid \exists \beta > 0, \forall (x_B, y_B) \in B^o, (x_C, y_C) \in C^o \right\} I$
   $\left\{ F(x_B, y_B) > F(x, y) \mid \forall (x_B, y_B) \in B^o, (x, y) \in Cu_E (A \cup B \cup C) \right\}$

3. Establish $\left(x^*, y^* \right)$ as the centre point of section E, where $X = \left(x^*, y^* \right), Y = (x_1, y_1), Z = (x_2, y_2)$

   then $\left\{ F(x_1, y_1) > F(x_2, y_2) \mid \|X - Y\| > \|X - Z\|, X, Y, Z \in E \right\}$ and
   $\left\{ F'(x_A, y_A) > F'(x_B, y_B) > F'(x_C, y_C) \mid (x_A, y_A) \in A^o, (x_B, y_B) \in B^o, (x_C, y_C) \in C^o \right\}.$

4. Establish $\left\{ (x, y) \mid \min F(x, y) = \gamma, (x, y) \in Cu_E (A \cup B \cup C) \right\}.$ As the set of points representing the

minimum value of the function, so it comes to $\left\{ \left| F(x, y) - \gamma \right| < \varepsilon \mid \exists \varepsilon > 0, \forall (x, y) \in Cu_E (A \cup B \cup C) \right\}.$

Obtain the CA part through 2D wavelet decomposition and carry on an image enhancement as shown in Figure 3 to provide convenience in seeking for A, B and C zones [9, 10].
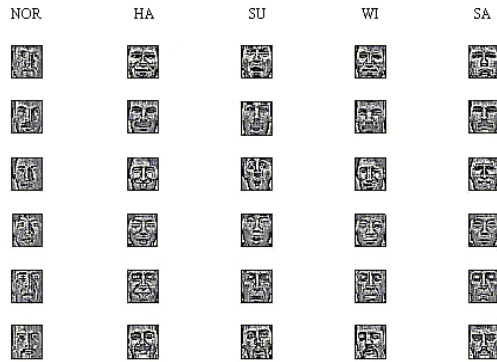
Figure 3. Decomposition Results

Identify A, B and C zones based on the decomposition results. Reduce the impact of illuminations and establish 3D model per restrictions [11, 12]. See Figure 4.
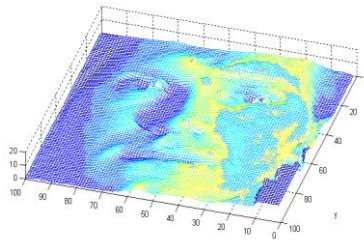
Figure 4. Addition of 3D Model Developed per Restrictions

Extract from the developed 3D image the following attribute value [8] through false geodesic distance: length of face (L_F), width of face (W_F), brow width (W_B), brow height (H_B), vertical distance between centre of brow and that of eyes (D_{BE}), eye width (W_E), distance between upper and lower eyelids (D_E), mouth width (W_M) and the distance between upper and lower lips [13, 14]. Table 1 shows the attribute values obtained by locating different characteristics in the first set of images.

Table 1. First Set of Facial Attribute Values

	L_F	W_F	W_B	H_B	D_{BE}	W_E	D_E	W_M	D_M
NOR	49	46	14	3	4	9	3	16	0
HA	49	48	13	2	5	11	4	23	4
SU	49	45	15	3	7	9	4	19	7
WI	49	46	12	2	2	11	0	20	0
SA	49	47	13	3	4	9	3	25	0

3. Emotion Recognition Based on PCA and False Geodesic Distance

Suppose there are n samples of facial models, each with p attributes observed, the following initial data information matrix is achieved. Then take p vectors of the data matrix as aggregate indicator vectors: $F_1=a_{11}X_1+a_{21}X_2+…+a_{p1}X_p$, $F_2=a_{12}X_1+a_{22}X_2+…+a_{p2}X_p$, ……

$F_p=a_{1p}X_1+a_{2p}X_2+…+a_{pp}X_p$. Recorded as: $F_i=a_{1i}X_1+a_{2i}X_2+…+a_{pi}X_p$.

The above equations set require: $\alpha_{1i}^2 + \alpha_{2i}^2 + L + \alpha_{pi}^2 = 1$ where $i = 1, 2, L , p$.

Coefficient a_{ij} is determined by the following principle:

1. F_i and F_j are uncorrelated $i \neq j; i, j = 1, 2, ..., p$.

2. F_1 holds the biggest variance among all linear combination. Similarly, F_p is the pth biggest variance among all linear combinations of $X_1, X_2, ..., X_p$ which are entirely uncorrelated with $F_1, F_2, ..., F_{p-1}$.

Set X as the Characteristic value of covariance matrix: $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_p$, hence $\text{var} F_1 \geq \text{var} F_2 \geq ... \geq \text{var} F_p > 0$

Supposing the contribution rate of the kth principle component is $\lambda_k / (\lambda_1 + ... \lambda_j + ... \lambda_p)$ and the accumulated contribution rate of the first k principle components. See Figure 5. If the accumulated contribution rate exceeds 85%, this proves that the information obtained from the first k principle components has basically covered the information contained in all measurement indicators that reduce the number of variables and eliminate the dependencies between variables. Table 2 represents all principle components' parameters.
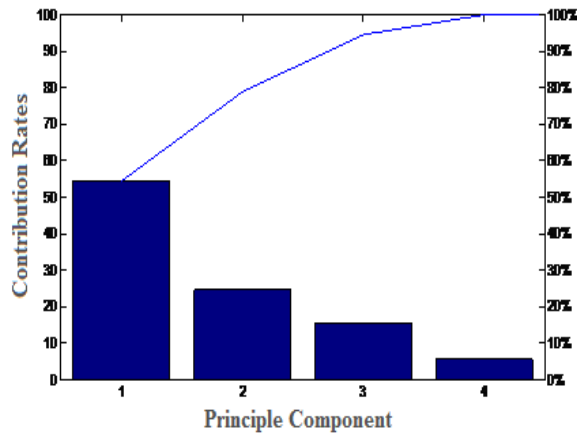
Figure 5. Histogram of Principle Component Contribution Rate

Table 2. Table of Parameters of all Principle Components

Principle Components	Characteristic Value	Variance Contribution Rate	Accumulated Contribution Rate
1	4.31	53.88	53.88
2	2.01	25.12	79.00
3	1.24	15.50	94.50
4	0.44	5.50	100
5	0	0	100
6	0	0	100
7	0	0	100
8	0	0	100

Table 3. Characteristic Vector; Corresponding to Principle Components

	No.2 Characteristic Vector	No.3 Characteristic Vector
0.25	0.49	0.34
-0.45	-0.01	-0.10
-0.38	-0.25	0.45
-0.42	0.32	-0.13
0.38	0.25	-0.45
-0.35	0.42	0.18
0.17	0.47	0.44
-0.31	0.37	-0.47

G_1 and G_2, representing two totals, are the random variables obtained from R^n. Their mathematic expectation and variance are: $EG_1 = \mu^{(1)}, EG_2 = \mu^{(2)}$, $\text{var}\,G_1 = S^{(1)}, \text{var}\,G_2 = S^{(2)}$

Suppose there is sample $x \in R^p$, the principle to judge if X is involved in G_1 or G_2: If X is closer to G_1, then X is involved in G_1; if X is closer to G_2, X is involved in G_2.

If Mahalanobis Distance $d(X, G_i)$ is employed, then: $d^2(X, G_i) @ (X - \mu^{(i)})^T S^{-1} (X - \mu^{(i)})$ i=1,2

The square difference of distances between X and G_1 and X and G_2 is:

$$d^2(X, G_1) - d^2(X, G_2) @ (X - \mu^{(1)})^T S^{-1}(X - \mu^{(1)}) - (X - \mu^{(2)})^T S^{-1}(X - \mu^{(2)}) = -2(X - \mu)^T S^{-1}(\mu^{(1)} - \mu^{(2)})$$

In which $\mu = (\mu^{(1)} + \mu^{(2)})/2$. Defined function: $U(X) = -2(X - \mu)^T S^{-1}(\mu^{(1)} - \mu^{(2)})$.

If $d(X, G_1) \le d(X, G_2)$, then $X \in G_1$;

If $d(X, G_2) \le d(X, G_1)$ $X \in G_2$, i.e. $U(X) \ge 0$, then $X \in G_1$; if $U(X) \le 0$, then $X \in G_2$.

4. Experiment Result and Analysis

Select from Yale faces collections facial images as shown in Figure 6 as test samples, manifest the characteristics of the samples based on the 2D plane distance and the false geodesic distance, carry out principle component computation, compare against training samples and figure out the Mahalanobis Distance between the two types of samples and finally identify what categories the two to-be-tested samples belong to.

Figure 6. Facial Images to Be Identified

The results are shown in Table 4 and Table 5.

Table 4. Results of Emotion Recognition by PCA Method based on False Geodesic Distance

	F1	F2	F3	Actual Category	Identification Category
Sample1	4.02	37.89	14.81	3	3
Sample2	7.10	34.68	17.16	1	1
Sample3	8.61	37.56	22.93	5	5
Sample4	7.56	33.63	17.64	4	4
Sample5	8.55	36.30	17.94	2	2

Table 5. Results of Emotion Recognition by PCA method based on 2D Plane Distance

	F_1	F_2	F_3	Actual Category	Identification Category
Sample1	3.14	40.73	18.92	3	3
Sample2	7.24	32.68	17.11	1	4
Sample3	9.16	40.15	19.58	5	5
Sample4	8.65	34.29	17.16	4	2
Sample5	7.56	33.29	17.44	2	4

Map the three principle components of the training samples and "to-be-tested sample 1" from the two methods into 3D coordinates so as to read the results more intuitively. Figure 7 and Figure 8 shows the projection of the 3D space coordinates of all samples on the ZOY plane with X as the sample to be tested.
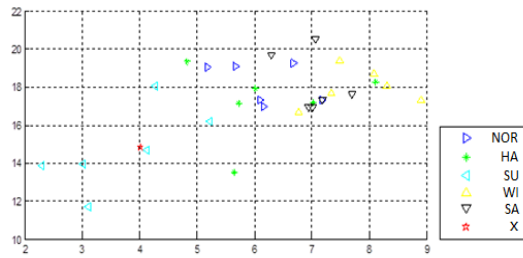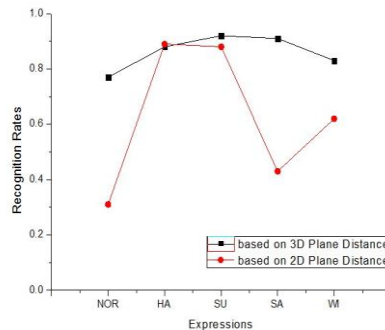
Figure 7. Results based on 2D Plane Distance
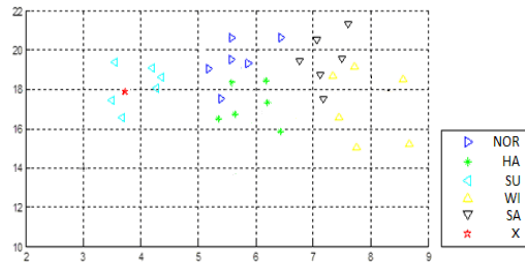
Figure 8. Results based on 3D Distance

Figure 9 shows the rate of recognition of different expressions from 30 static facial images based on the two methods.

Figure 9. Recognition Rate in Terms of Different Expressions

Figure 9 shows that the traditional methods produce a relatively low accuracy in identifying NOR and SA expressions. As we can directly see from NOR and SA expressions in the static facial images that they vary little in terms of geometric configurations, it is difficult for the traditional methods to discriminate this minor difference. However, the 3D-based method as described in this article may properly discriminate such difference and classify the expressions, displaying a more advantageous performance.

5. Conclusion

Nowadays, facial expression recognition is an active direction in the field of human-computer interaction. Research on facial expression recognition in general is equivalent to the data dimension reduction. The paper observes static face image about subtle

changes in geometry. It need extract 30 samples from the YALE face database to test. The experimental results show that the method can accurately identify the static face image, and is more effective than the traditional method.

Acknowledgements

References

[1] X. Hu, "Facial Affective Recognition Based on Gabor Wavelet Transformation and Fractal Dimension", Changsha University of Science & Technology, **(2008)**.

[2] W. Tian, "Feature Extraction Based on Depth Image of 3D Face", Beijing Jiaotong University, **(2009)**.

[3] K. Wang and J. Zhu, "Face Feature Extraction Based On Wavelet Analysis and Discrete Cosine Transform", Pattern Recognition and Simulation, vol. 28, no. 4, **(2009)**, pp. 65-67.

[4] P. Fei and W. Wang, "Edge Detection Based on Image Enhancement of Wavelet Transform and Modified Multi-scale Morphological Gradient", Topography and Information Science of Wuhan University, vol. 02, **(2007)**, pp. 120-123.

[5] G. Wang and X. Gong, "Perception based 3D face modeling and its applications", Journal of Chongqing University of Posts and Telecommunications (Natural Science Edition), vol. 04, **(2009)**, pp. 558-559.

[6] X. Chai and S. Shan, "Pose and Illumination Invariant Face Recognition Based on 3D Face Reconstruction", Journal of Software, vol. 3, **(2006)**, pp. 527-531.

[7] F. Lavagetto and R. Pockaj, "An efficient use of MPEG-4 FAP interpolates for facial animation at 70 bits/frame", Circuits and Systems for Video Technology, vol. 11, no. 10, **(2001)**, pp. 1085-1095.

[8] S. Sebastian and A. Fink Gernot, "A New Method for Combined Face Detection and Recognition Using Interest Point Descriptors", In International Conference on Automatic Face & Gesture Recognition and Workshops, **(2011)**, pp. 219-523.

[9] P. Lu and Y. Chen, "Fast USAN-AAM method for locating facial features", Application Research of Computers, vol. 27, no. 8, **(2010)**, pp. 3188-3190.

[10] M. Pantic, "Facial expression recognition", Encyclopedia of Biometrics, **(2009)**.

[11] D. V. Jadhav and R. S. Holambe, "Feature extraction using radon and wavelet transform with application to face recognition", Neuro-computing, vol. 72, no. 7-9, **(2009)**, pp. 1951-1959.

[12] R. Wei, L. Jiang and L. Tao, "Facial expression recognition system based on multiple feature integration", Chinese Journal of image and graphics, vol. 14, no. 5, **(2009)**, pp. 792-880.

[13] Z. Qing, D. Rui, Z. Xue-ying and W. Sui, "Facial Expression Geometrical Feature Extraction Based on Chain Code", Computer Engineering, vol. 38, **(2012)**, pp. 156-159.

[14] X. Rui, Z. Qijun and Z. David, "Facial Expression Recognition on Multiple Manifolds", Pattern Recognition, vol. 44, no. 1, **(2011)**, pp. 107-116.