2.5D部分顔画像からの顔表情認識

チャロンポン ティーカパン・東海彰吾・長谷博行

福井大学大学院工学研究科 〒910-8507 福井市文京 3-9-1

E-mail theekapun@gmail.com, {hase,tokai}@fuis.fuis.fukui-u.ac.jp

あらまし 本論文では、-45 度から+45 度の視点から撮った 1 枚の 2.5D 部分顔画像から表情を認識する方法を提案する。まず 2.5 部分顔画像から顔対称面を検出して全体顔画像を復元する。復元された3D 顔画像から、顔からの相対位置が安定な顔平面を計算する。顔表面から出る法線と計算された顔平面上の交点を利用して表情認識を二通りの方法で試みた。一つは交点密度を特徴とする方法であり、他は平常顔からの変位を特徴とする方法である。本論文では、顔平面を有効な範囲をいくつかのの小領域に分割し、それぞれの小領域の交点密度と代表変位ベクトルを特徴とした。表情は平常顔、笑顔、怒り顔、驚き顔の 4 種類を対象とした。認識器として SVM を利用した結果、二つの手法とも 4 表情のうち笑顔の認識率が最もよく、また角度が大きくなるに従って認識率は多少下降し 45 度からの部分笑顔では約 90%となった。他の表情は笑顔より数%から 10%程度認識率が低くなった。両特徴の比較では、変位ベクトルによる方法の方が 10%以上良い認識率を示した。

キーワード 顔表情認識、部分顔画像、顔平面、交点、変位ベクトル、SVM

A Novel Method of Facial Expression Recognition from A 2.5D Partial Face Image

Theekapun CHAROENPONG, Shogo TOKAI, Hiroyuki HASE

Dept. of Information Science, University of Fukui, 3-9-1 Bunkyo, Fukui 910-8507, Japan

E-MAIL: theekapun@gmail.com, {tokai,hase}@fuis.fuis.fukui-u.ac.jp

Abstract Due to the existing problem for recognizing facial expression from side view face, this paper propose a new method for recognizing facial expression by using a 2.5D partial face image. The 2.5D partial face image is captured from a viewpoint between -/+45°. This method consists of two parts: crossing points and displacement vector. First, the 3D virtual expression face is reconstructed from a 2.5D partial face image. The facial expression is analyzed in term of the crossing point distribution on the face plane. Second, a displacement vectors are computed by using the movement direction pointing from the crossing point of normal face to the crossing point of expression face. Finally, the crossing point distribution and displacement vectors are used for facial expression recognition by means of support vector machines. The experiments were done for four facial expressions (neutral, anger, surprise and smiling) from 22 persons, based on the assumption that the person has been known. The results show the feasibility of the proposed method.

Keywords Facial Expression; Partial Face Image; Face Plane; Crossing Point; Displacement Vector; Support Vector Machines

1. Introduction

As facial expressions provide significant information such as affective stage, cognitive activity, temperament and personality, truthfulness, and psychopathology, the several researches study on facial perception over the last three decades [1], [2], but these researches perform effectively for recognizing the facial expression from only full frontal facial image. In many situations of robot-vision and human-machine interfaces, only partial face data can be captured. Therefore this paper

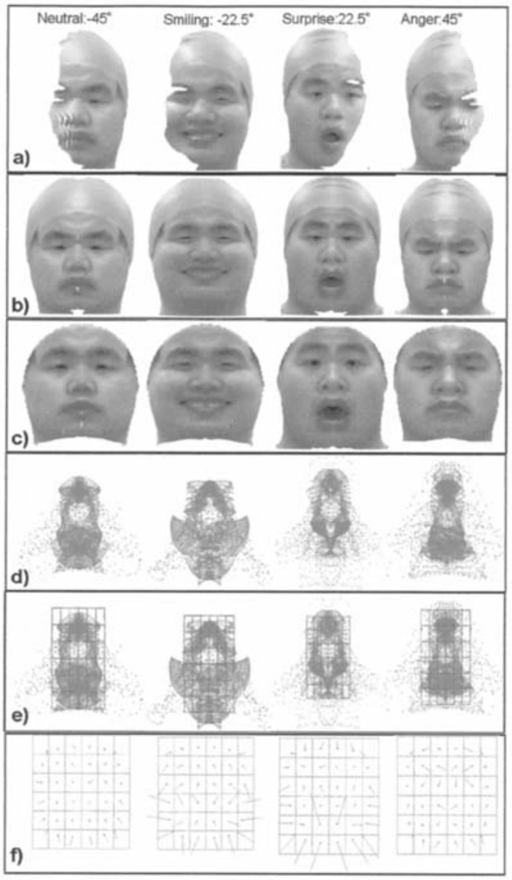

Figure 1. a) Input 2.5D partial face images are captured in different viewpoint starting from -45° (Neutral), -22.5° (Smiling), 22.5° (Surprise) and 45° (Anger). b) The 3D virtual expression face is reconstructed. c) Only face surface is extracted to be used in face plane. d) Crossing point on face plane. e) The crossing point distributions on the face plane and the face profile. f) The displacement vector of each expression face. The 36 vector are used to analyze the facial expression.

propose a facial expression recognition method from a 2.5D partial face data. A 2.5D data is a simplified 3D (x, y, z) surface representation that contains at most one depth value (z-direction) for every point in the (x, y) plane, associated with a registered texture image. Each scan can only provide a single viewpoint of object, instead of the full 3D view [3]. However, the usage of 2.5D image has the advantage over using 2D image for face pose variation. There are many researches proposed as follows.

In the existing researches, the several facial expression recognition methods have been proposed. Most of existing researches can work well with only frontal facial image. M. Pantic and L.J.M Rothkantz [4] proposed a facial expression recognition method by defining their face model as a point-based model composed of the two 2D facial views of the frontal and side view face images. Multiple feature detectors were applied redundantly in order to localize contours of prominent facial features prior to their model. This method needs more than one viewpoint of face image including frontal face image. H., Michael et al. [5] use a generic 3D head model and frontal view image for the synthesis of facial expression, and then the facial expression was classified by local binary pattern techniques. I.A. Essa [6] adopts the optical flow processing for perception and measurement of facial motion. The 3D mesh model of face is fitted by 2D frontal face image. Muscle activation is extracted using a physics-based model of facial deformation, and then the deformation is estimated from the optical flow. C. Li and A. Barreto [7] proposed the framework of 3D face recognition involving an initial expression assessment of unknown face. The smiling and neutral expressions were recognized. The most distinctive features associated with the smile are the bulging of the cheek muscle and the uplift of the corner of the mouth, and then the histograms of the range value for smiling and neutral is used to classify the expression by means of Linear Discriminant Analysis (LDA) and Support Vector Machine (SVM). J. Wang et al. [8] used the rotated 3D frontal face model under viewpoint between -/+40° for each rotation on pitch and yaw rotation to recognize facial expression. Based on the principal curvature information estimated on the 3D triangle mesh model, a surface labeling approach is applied to classify the 3D primitive surface features into twelve basic categories. The statistic histograms of the surface labels of all these regions are combined to construct the specific facial expression feature. Linear Discriminant Analysis (LDA) is used for classification. In this method, the rotated frontal face is used in the experiment but, actually, an only partial face will be obtained in case of capturing the face from side view. These methods are not appropriate for facial expression recognition from a partial face data because the frontal face is necessary.

In the almost existing researches, the only frontal face viewpoint can be used and the feature extraction is always needed for higher-level representation. But actually, the orientation of face is also the one factor to make the recognition result fail and it is difficult to select appropriate features in the 3D facial data. Therefore, in this paper, we propose a novel method for 3D face expression recognition from a 2.5D partial face data. An input 2.5D partial face data is captured from any viewpoint between -45° and +45° around Y axis, and then the 3D virtual expression face is reconstructed according to the 3D face reconstruction algorithm. Only the change of facial skin is considered for facial expression recognition.

By using face plane algorithm, the facial expression is represented in terms of the crossing point distribution and the displacement vector (see Fig. 1). The recognition accuracy by the crossing point distribution and the displacement vector is shown in the experiment. Our method has been developing for recognizing the four facial expressions (neutral, anger, surprise and smiling) by mean of support vector machines.

In the following, our facial expression recognition method is described in section 2. In section 3, the experimental results are shown.

2. Facial Expression Recognition

The process of the facial repression recognition is illustrated in Fig. 1. The 2.5D image acquisition and 3D face reconstruction are prepared in pre-processing process, and then the derivation of the face plane, crossing point distribution and the displacement vector will be described.

2.1 Pre-processing Process

A. 2.5D Image Acquisition

For our 3D facial expression reconstruction algorithm, the 2.5D head image is used. The input data is captured by using VIVID700 and represented by the point set data. The 2.5D image consists of range data and color data. The range image size is 200x200 pixels, the color image has 400x400 pixels in the same scanning region. Unfortunately, VIVID700 is not sensitive for the black color like hair because it uses a red laser. This problem was solved by covering the hair with the cap. However, only facial skin area is used in the recognition system. The input image was captured from a viewpoint around Y axis from -45° to +45° with step 22.5° (see Fig. 2(b)) with four facial expressions (neutral, anger, surprise and smiling).

B. 3D Face Reconstruction

As we aim to recognize facial expression using only a single 2.5D partial face data, the 3D virtual face is reconstructed by the method in ref. [9]. In this subsection, this method will be explained briefly.

The 2.5D partial face data is sprit into N cross section along Y axis. The data on each cross section was fitted by ellipse fitting technique, and then nose tip and nose ridge are detected by correcting the face vectors according to nose ridge detection algorithm. The mirror of real face is computed from the symmetry plane passing the nose tip, the nose ridge and the point on the center line of the face. Finally, 3D virtual face is reconstructed using the real face and its mirror face. The example of reconstructing the 3D virtual expression face is shown in Fig. 1(b).

The nose tip point will be referred in the face plane for reconstructed 3D face algorithm. The face plane which will be explained in the next subsection applied to analyze the facial expression.

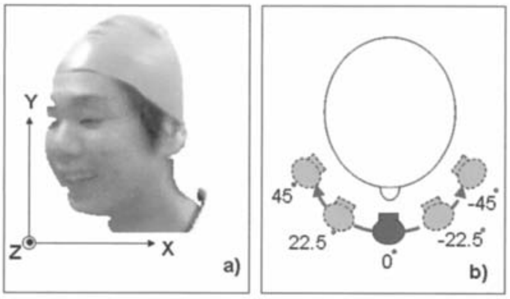

Figure 2. a) The black color surface as hair unmeasured by laser light is covered by cap. b) The viewpoint for capturing the 2.5D image from the top view of human head. 5 viewpoints captured are -45°, -22.5°, 0°, 22.5° and 45°.

2.2. Face Plane Computation

After the virtual face is reconstructed in section 2.1., the facial surface extracted according to color segmentation as shown in figure. 1(c) is considered for recognition. The face plane algorithm is applied for analyzing the facial expression in term of crossing point [10].

A. Derivation of the Face Plane

The face plane is a virtual plane across a head. In face plane algorithm, the normal vectors are used for deriving a face plane. These normal vectors point from the facial surface toward the center of the head. A normal vector is calculated using four neighboring points. The derivation of face plane is described as following.

The three dimensional data on the facial surface is described as $p_i = (a_i, b_i, c_i)^t, (i = 1,\ldots,N)$ (see Fig.3). Let the normalized vector on the normal line passes through a point p_i on the face be $f_i = (j_i, k_i, l_i)^t$, Now suppose a point P inside the head and orthogonal projection from the point P to the normal line. The distance d_i of orthogonal projection from the point to the normal line f_i is as follow:

$$
\begin{aligned}
d_i^2 &= \|P - p_i\|^2 - (P - p_i, f_i)^2 \\
&= (P - p_i)^t (E - f_i \cdot f_i^t)(P - p_i) \qquad (1)
\end{aligned}
$$

where E is the unit matrix and $(\bullet, \bullet)$ is the inner product. The optimization can be carried out by differentiating $Q = \sum_i d_i^2$ with respect to P.

$$
\frac{dQ}{dP} = 2\sum_i (E - f_i \cdot f_i^t)(P - p_i) = 0 \qquad (2)
$$

Then the optimal point P can be obtained.

$$
P = \left[\sum_i (E - f_i \cdot f_i^t)\right]^{-1} \cdot \sum_j (E - f_j \cdot f_j^t) p_j \qquad (3)
$$

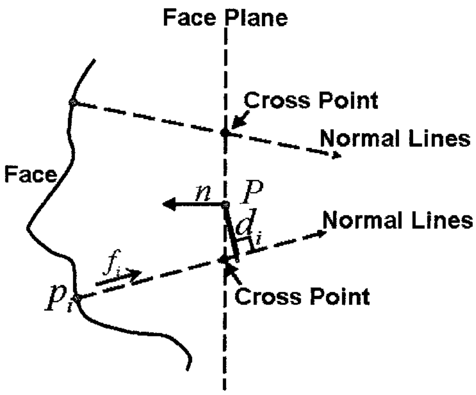

Figure 3. Derivation of face plane

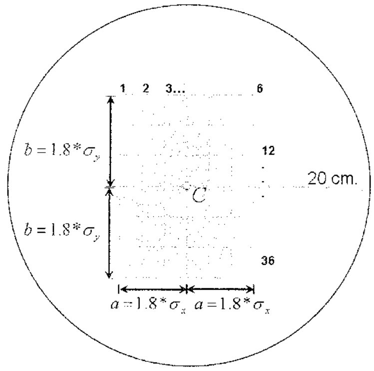

Figure 4. The crossing point distributions in the 36(6x6) subareas

The face plane passes through this point. But the orientation of the plane has not been determined. Therefore, the direction n which makes the next equation maximum is solved. This makes an eigen problem under the condition $\|n\| = 1$.

$$S = \sum_i (f_i, n)^2 = n^t \left(\sum_i f_i \cdot f_i^f \right) n \qquad (4)$$

By solving this problem, the three eigenvalues and three corresponding eigenvectors can be obtained as $\lambda_1 \geq \lambda_2 \geq \lambda_3$ and ϕ_1, ϕ_2, ϕ_3 respectively. ϕ_1 will be chosen as the direction n of the face plane based on the least distance (see Fig. 3). The face plane is fixed by P and n.

B. Crossing Point Computation

As the normalized vector $f_i = (j_i, k_i, l_i)^t$ points from the point $p_i = (a_i, b_i, c_i)^t$, the equation of the normal line is described as follows;

$$h = p_i + t_i f_i \qquad (5)$$

where $h = (x, y, z)^t$ is a point on the normal line, t_i is a real number. Describing Eq. (5) in element-wise, we can obtain the next equation (6).

$$\begin{aligned} x &= a_i + t_i j_i \\ y &= b_i + t_i k_i \\ z &= c_i + t_i l_i \end{aligned} \qquad (6)$$

Then, suppose a virtual plane in a head;

$$n \cdot (X - P) = 0 \qquad (7)$$

where $n = (A, B, C)$ is the normal vector of the face plane, $P = (x_p, y_p, z_p)$ is a point on the face plane. Substituting Eq. (6) in Eq. (7), we can obtain t_i

$$t_i = \frac{(Ax_p + By_p + Cz_p) - (Aj_i + Bk_i + Cl_i)}{(Aa_i + Bb_i + Cc_i)} \qquad (8)$$

By substituting t_i in Eq. (6), we can obtain the crossing point (x_i^c, y_i^c, z_i^c). The crossing point distributions on the face plane will be different according to facial expression change (see Fig. 1.(d)). The advantage of this analysis using the face plane is that the position of crossing points is invariant for the rotation and inclination of the head.

2.3. Crossing Point Distribution Method

In this subsection, the crossing point distribution is divided into 36(6x6) subareas (see Figure. 4). The set of crossing points in each subarea are used as pattern feature. To define the 36 subareas, the center C is defined as the average of crossing points (see Fig. 4). The standard deviation on x-axis (σ_x) and y-axis (σ_y) of crossing points within the area of 20 cm. from the center C are calculated. The width and height of the region of interest (ROI) are defined as $2 * 1.8 * \sigma_x$ and $2 * 1.8 * \sigma_y$ respectively, where the factor 1.8 is determined by a pilot test. The ROI is divided into 36(6x6) subareas. The crossing points in each subarea are counted and normalized as Eq. (9).

$$\begin{aligned} \mathbf{X} &= (x_1, x_2, x_3, \ldots, x_n) \\ \bar{x} &= \frac{1}{n} \sum_{a=1}^{n} x_a \\ V_c &= (x_1 - \bar{x}, x_2 - \bar{x}, x_3 - \bar{x}, \ldots, x_n - \bar{x}) \\ \mathbf{V} &= \left(\frac{x_1 - \bar{x}}{\|V_c\|}, \frac{x_2 - \bar{x}}{\|V_c\|}, \frac{x_3 - \bar{x}}{\|V_c\|}, \ldots, \frac{x_n - \bar{x}}{\|V_c\|} \right) \end{aligned} \qquad (9)$$

where x_a is the number of crossing points in each subarea. a is the subarea number starting from 1 to 36. $\mathbf{V}$ is the normalized vector of feature vector V_c and has 36 components.

For recognition, the facial expression change from neutral face is considered. Therefore, each facial expression is compared with the referential neutral face of the same person (see Eq. (10)). A data used for recognition is:

$$\mathbf{Y}_{\alpha}^{(k)} = \mathbf{V}_{\alpha}^{(k)} - \mathbf{Vn}_{\alpha} \tag{10}$$

where $\mathbf{Y}_{\alpha}^{(k)}$ is the difference vector. $\mathbf{V}_{\alpha}^{(k)}$ is the normalized feature vector of facial expression. $\mathbf{Vn}$ is the feature vector of referent neutral face. α is the data number. k is the category number of facial expression.

2.4. Displacement Vector Method

Based on the assumption that the person is known, each facial expression image is compared with the neutral face of the same person. In order to compute a displacement vector, the size of the expression face is normalized to be the same size as the neutral face. A pair of points on both face surfaces which have the same 2D-coordinate of the neutral face and the expression face are considered. The pair crossing points corresponding to the pair of points on the two face surfaces will be used for computing a displacement vector. The displacement vector is the movement vector of the crossing point pointing from the crossing point of neutral face to the one of expression face (see Fig. 5.).

To compute the displacement vector, the size of neutral face and expression face has to be the same. The two ear positions of both faces are selected by manually for normalization, and then the expression face is normalized to the neutral face. The 3D faces are converted to 2D faces from the front view face. The pair points on the surface of neutral face and expression face which have the same 2D-coordinate is described as p_i^n and p_i^e, (i = 1,...,N), respectively. The two crossing points c_i^n and c_i^e are the positions on the face plane which the normal lines from p_i^n and p_i^e point to the face plane. The displacement vector d_i^e is a vector pointing from c_i^n to c_i^e.

The surface of the expression face is divided into 36 (6x6) areas based on the size of neutral face (see Fig. 1.(f)). The average displacement vector v_m^e, (m = 1,...,36), is the average of vector d_i^e according to the all points p_i^e on the surface of the same area. The set of displacement vector V_e is consisted of 36 average displacement vectors.

$$d_i^e = c_i^e - c_i^n \tag{11}$$

$$v_m^e = \sum_{i=1}^{s} d_i^e \tag{12}$$

$$V^e = \left(v_1^e, \ldots, v_m^e\right) \tag{13}$$

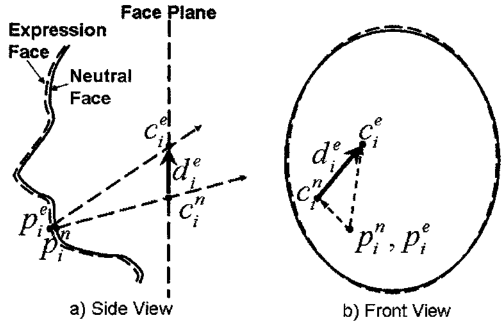

Figure 5. The concept of displacement vector

where c_i is the crossing point, d_i is the displacement vector, v_m is the average of displacement vector according to the points on the surface of same area, V^e is the set of displacement vector, s is the number of crossing points in the area m, e is the category of expression face and n represents neutral face.

3. Experimental Result and Discussion

The experiments were done for four facial expressions (neutral, anger, surprise and smiling) from 22 persons, based on the assumption that the person is known. A 2.5D partial face is captured from a viewpoint from -45° to +45° with step 22.5°. The support vector machines in [11] with the radial basis function is used for recognition. The experimental result is shown in Fig.6. The accuracy of smiling face shows the highest because the motion of crossing point is distinctive from the other. The accuracy of anger face shows the lowest because only partial surface around eyebrows changes from the neutral face.

The performance of displacement vector method and crossing point distribution method in [6] is compared in Figs. 7 and 8. Fig. 7 shows the accuracy for each viewpoint. Fig. 8 shows the accuracy for each expression face. By using displacement vector method, the expression recognition accuracy is improved.

4. Conclusion

In this paper, we have proposed a novel method for facial expression recognition from a 2.5D partial face image. An input 2.5D partial face image captured from any viewpoint from -45° to +45° around Y axis is used. The experiments were done for four facial expressions (neutral, anger, surprise and smiling) from 22 persons, based on the assumption that the person is known. The results showed the feasibility of the proposed method. The advantages of this method are the following: 1) the 2.5D partial face image captured from any viewpoint of face rotation between -45° and + 45° can be used for facial expression recognition, 2) by comparing between our proposed method and the crossing point distribution method, the expression recognition accuracy is improved. For the future work, the additional method such as hierarchical classification algorithm will be used to improve the accuracy of the system.

5. References

[1] M. Pantic, and L.J.M. Rothkrantz, "Automatic Analysis of Facial Expressions: the State of the Art", IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 22, No. 12, pp. 1424-1445, Dec. 2000.

[2] B. Fasel, and J. Luettin, "Automatic Facial Expression Analysis: A Survey", Pattern Recognition, 36(1):259-275, 2003.

[3] X. Lu, A.K. Jain and D. Colbry, "Matching 2.5D Face Scans to 3D Models", *IEEE Trans. on Pattern Anal. Mach. Intell.*; 28(1), pp. 31-43 (2006).

[4] M. Pantic, and L.J.M. Rothkrantz, "Expert System for Automatic Analysis of Facial Expression", Image and Vision Computing Journal, 18(11):881-995, 2000.

[5] H. Micheal, A. Wiratanaya and K.F. Kraiss, "Facial Expression Modeling from Still Images using a Single Generic 3D Head Model", Proceeding of the 28th DAGM Symposium, Berling, pp. 324-333, September, 2006.

[6] I.A. Essa, A.P. Pentland, "Coding, Analysis, Interpretation, and Recognition of Facial Expressions", IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 19, No. 7. pp. 757-763, July, 1997.

[7] Li and A. Barreto, "An Integrated 3D Face-Expression Recognition Approach", Proceeding of ICASSP2006, Toulouse, 2006.

[8] J. Wang, L. Yin, X. Wei and Y. Sun, "3D Facial Expression Recognition Based on Primitive Surface Feature Distribution", Proceeding of CVPR'06, New York, pp. 1399-1406, October, 2006.

[9] Th. Charoenpong, H. Hase and T. Shogo, "A Novel Approach of 3D Face Reconstruction Using Ellipse Fitting", Proceeding of MVA2007, Tokyo, May, 2007.

[10] H. Hase, M. Yoneda, T. Kasamatsu and J. Kato, "A Proposal for a Face Plane", Proceeding of the 7th International Conference on Image Processing, pp. 333-336, Greece, 2001.

[11] C.-C. Chang and C.-J. Lin, "LIBSVM: a library for support vector machines". http://www.csie.ntu.edu.tw/~cjlin/libsvm.

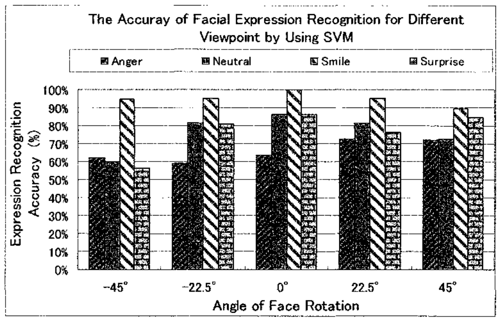

Figure 6. The accuracy of facial expression recognition for different angle of face rotation by using SVM; -45°, -22.5°, 0°, +22.5° and +45°.

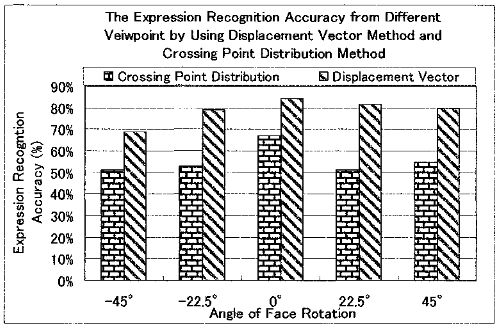

Figure 7. The expression recognition accuracy from different angle of face rotation by using displacement vector method and crossing point distribution method

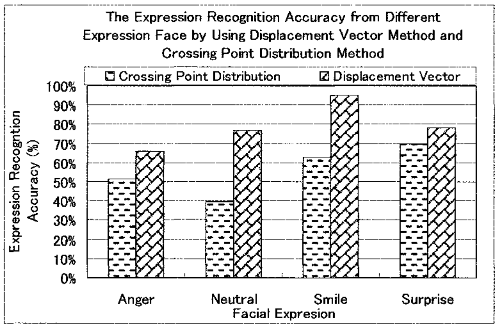

Figure 8. The expression recognition accuracy from different expression face by using displacement vector method and crossing point distribution method