The Eighteenth International Symposium on Artificial Life and Robotics 2013 (AROB 18th '13),
Daejeon Convention Center, Daejeon, Korea, January 30-February 1, 2013

Face identification insensitive to facial expression in the crowded people scene

Qingyu Shu[1], Tetsuo Hattori[1], Saki Masunari[1], Tetsuya Izumi[2],
Hiromichi Kawano[3], Bahram Javidi[4]
[1]Kagawa University, Japan
[2]Micro Technica Co., Ltd., Japan
[3]NTT Advanced Technology, Japan
[4]University of Connecticut, USA
Corespondence:hattori@eng.kagawa-u.ac.jp

Abstract: Face identification in the crowded people with various facial expressions is one of the most difficult themes in the pattern recognition field. In addition, the recognition of facial expression after the identification is very important as an advanced pattern recognition problem in Kansei Engineering field. Currently, there have been developed some methods at the level of practical use such as personal authentication. However, neither the theory nor the method has been established yet for the partial recognition problem that means the identification of a specific person's face from the crowd people scene. In this paper, we propose a novel face identification method insensitive to facial expressions using the notion of weighted vector field and weighted similarity. Also, we show the effectiveness by showing the experimental results.

Keywords: face identification, KL Expansion, region weighted vector field, spatial correlation, weighted similarity

1 INTRODUCTION

Although research on identification of person by face has being widely done ([1]-[6]), the recognition for arbitrary plural people image under genaral background is still difficult problem.

For such a problem, we have proposed a recognition method using a vector field ([7]-[12],[15]), that judges whether the face of acquaintance exists or not in the crowded people image, and also detects and extracts the face region from the image. In the method, there is an assumption that the pattern to be recognized is perturbed by some variation which is roughly equals to be an Affine transform (rotated, enlarged/reduced, and translated) from a registered original face pattern.

However, since the recognition method using the vector field method is sensitive to the facial expression, when the expression is big, there occurs a case where the method does not work well.

Then in this paper, we propose a region weighted vector field (RWVF) method insensitive to the facial expression. The RWVF itself is obtained by emphasizing some characteristic parts in the vector field that is constructed by making a normalized gradient vector field (NGVF) of grey level image ([13],[14]). This method is also aimed for the recognition of the kansei (feeling) information from the facial expression. That is, the method is proposed not only for the identification but also for the recognition of feeling after the identification processing.

This paper also describes the effectiveness of this method, by showing the experimental results for people images that are naturally acquired indoor and outdoor.

2 RECOGNITION METHOD

This method is composed of 2 stages, one is Learning and Analysis stage as a preprocessing one and another is Recognition stage. The input images are monochrome grayscale (or gray level) images with a size of 256x256 pixels. Color images are converted into grayscale images before input to recognition processing.

Here, we use the following terminology.

(a)Original registered pattern: the original face image to be recognized, also called original pattern.

(b)Registered pattern: the variation of original pattern obtained by Affine transform.

(c)Reference pattern: the eigenvector obtained from a set of several registered patterns by Karhunen-Loeve (KL) expansion. It will be used for similarity computation with input image.

This method automatically generates reference patterns at the learning and analysis stage, and after that, it executes the recognition process as recognition stage.

2.1. Learning and analysis stage as preprocessing

The flow of this processing is as follows.

STEP1 : Make plural registered patterns from an original face pattern by Affine transform (enlarged or reduced, rotated).

STEP2 : Divide registered patterns into some groups, and construct the Normalized Gradient Vector Field (NGVF) expression. Then, apply the KL (Karhunen- Loeve) expansion to every group, and make the reference pattern that has common component.

2.2. Generation of registered patterns

Applying Affine transform to given original registered pattern, our method automatically makes some registered patterns by enlargement/reduction and rotation (Fig.1).

As an experiment, this method totally generates 240 registered patterns, changing original registered pattern by every 3% enlargement/reduction from 50% to 150%, and by every 3 degrees rotation from -10 degrees to +10 degrees. Then it can be divided them into 20 groups, where each group contains 12 registered patterns.

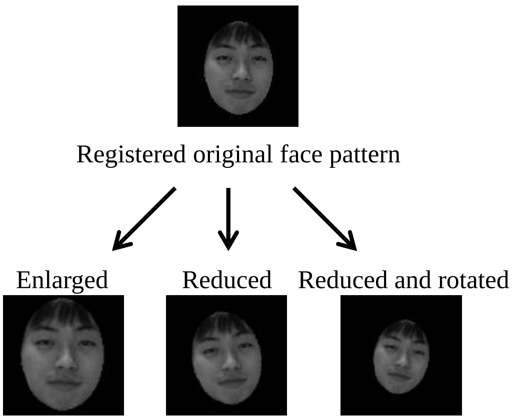


Fig.1. Example of registered patterns ([12],[15]).

2.3. Normalized Vector Field（NGVF）

Gradient vector is obtained at each point $(x,\ y)$ by partial differentiation of the gray level image as Eq.(1). To obtain the NGVF, normalize the size (norm) of the gradient vector to 1.0 if the norm is greater than a threshold, else to 0.

$$\mathbf{grad}\ f(x,y)=\left(\frac{\partial f}{\partial x},\frac{\partial f}{\partial y}\right)=\|\mathbf{grad}\ f(x,y)\|\cdot(\cos\theta,\sin\theta) \quad (1)$$

where θ is the phase of the vector.

2.4. Making reference patterns

By using Karhunen-Loeve(KL) expansion for each group of registered patterns represented as NGVF, we can obtain the reference patterns (eigenvectors in KL expansion) which has the common feature among patterns in each group. These reference patterns are used in the subsequent spatial correlation and similarity computation.

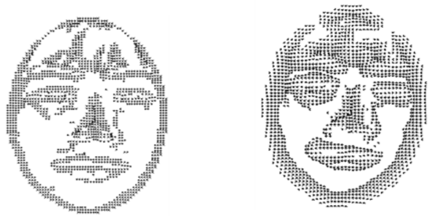


Fig.2. (a) Example of NGVF. (b) Example of reference pattern ([12],[15]).

2.5. Recognition stage

The flow of this recognition processing is as follows.

STEP1: Calculate the similarity which are based on spatial correlation between input image and the reference pattern (pattern matching), and find similarity distribution with the reference pattern in each point in an input image.

STEP2: Based on similarity distribution, detect a candidate locations where the face pattern of interest may exist in the input image, and roughly estimate the scaling and rotation parameters.

STEP3: Repeatedly calculating the similarity, perform the adjustment (fine estimation) of parameter values by means of correction formulas based on Taylor expansion of the affine transforms.

STEP4: Recognize the biggest similarity using the corrected parameters, and extract the face patterns exhibiting the maximum similarity.

2.6. Similarity computation

In order to detect the existence of a target face pattern, our method computes the similarity between input image and reference pattern both of which are represented as NGVF, based on the spatial correlation as shown in Eq.(2). The input image g ($g(x,y)$) and the reference pattern ev ($ev(x,y)$) are defined as complex valued functions.

The similarity between reference pattern ev and input g is given by Eq.(2).

$$Similarity=\frac{|\mathbf{Re}\{\langle g|ev\rangle\}|}{\|ev\|\cdot\|g_s\|} \quad (2)$$

where $\langle g|ev\rangle$ is the inner product between input image and reference pattern, and the inner product is defined as $\langle g|ev\rangle=\iint g(x,y)\cdot\overline{ev(x,y)}dxdy$. The notations **Re** and | | stand for real number of complex number and absolute value, respectively. || || denotes norm of pattern, and it is defined as $\|ev\|=\sqrt{\langle ev\,|\,ev\rangle}$. $\|g_s\|$ represents the norm of the subpattern g_s in input image g, which is defined on the overlapped region between *g* and *ev*. The spatial correlation can be considered as inner product $\langle g|ev\rangle$ at every pixel when the reference pattern *ev* is shifted.

(a) Original face. (b) Input image.(c) Recognition and extraction result.

Fig.3. Recognition processing using NGVF ([12],[15]).

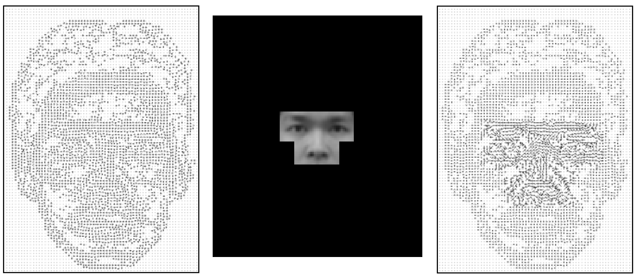
2.7. Region Weighted Vector Field (RWVF)

In RWVF, the absolute of the vector is normalized using some weights (e.g., 1.5, 2.0), depending on face regions. At the pixel other than the weighted region, the norm becomes 1.0 (i.e. the same as the usual NGVF). The purpose of using this **RWVF** is to emphasize the characteristic part of the face (i.e. eyes, nose) and to reduce the lighting influence.

In the similarity calculation, we use a spatial weight function $w(x, y)$ (real variable function). And so, the inner product between input image and reference pattern is shown as Eq.(3). Then, the similarity is defined as Eq.(4).

$$\iint w(x,\ y)\cdot g(x,\ y)\cdot \overline{ev(x,\ y)}dxdy = \iint g(x,\ y)\cdot \overline{w(x,\ y)ev(x,\ y)}dxdy = \langle g|w\cdot ev\rangle \qquad (3)$$

$$Similarity = \frac{\left|\mathbf{Re}\{\langle g|w\cdot ev\rangle\}\right|}{\|w\cdot ev\|\cdot\|g_s\|} \qquad (4)$$



(a) (b) (c)

Fig.4. (a)NGVF, (b)Weighted region, (c)RWVF.

3 EXPERIMENTATION

3.1. Environment

Using the RWVF method, we perform some experiments. The experimental environment is in the following.

OS: Microsoft Windows XP Professional. CPU: Intel(R) Core(TM)2. Drive frequency: 2.40GHz. Memory storage: 1.85GB. Programming Language: Borland C++ 5.02J.

3.2. Experimentation of comparison between the NGVF method and RWVF one

(i) Outline of the experiment

We compare the recognition rate and recognition time by the two methods, NGVF method and RWVF one by experimenting for 100 pieces of input images that are taken indoors and outdoors. There are 2-5 people in the images. The weight in RWVF is 2.0, that is, the vector norm on the weighted region is 2.0, otherwise the same as NGVF.

(ii) Experimental results

In this time experimentation, recognition rate has been almost same level. But the failed images are different (Fig. 5 and 6). In Fig.5 and Fig.6, the persons in the red circle and blue circle are the target to be recognized and misrecognized person, respectively. From this result, the recognition failure by the usual NGVF method can be improved by RWVF method. The results are shown as in the following.

[NGVF]

Accurate recognition rate: 99/100 (99%).

Average recognition time: 3.6 sec.

(Rough estimation: 2.6 sec. Fine estimation: 1 sec.)

[RWVF (weight:2.0)]

Accurate recognition rate: 98/100 (98%).

Average recognition time: 3.8 sec.

(Rough estimation: 2.5 sec. Fine estimation: 1.3 sec.)

Fig.5. Recognition failed image by NGVF.

Fig.6. Recognition failed images by RWVF method in the case where the weight is 2.0.

3.3. Relation between weight and recognition rate by RWVF method

Based on the result of the comparative experiment, we change the weight of the RWVF method and investigate the relation between the weight and the recognition rate for the same 100 pieces of images. We have changed the weight from 1.0 to 2.0. Fig.7 shows that the recognition rate becomes 100% when the weight lies in the range from 1.09 to 1.46. Then we consider that RWVF with suitable weight is effective to the facial expression variations.

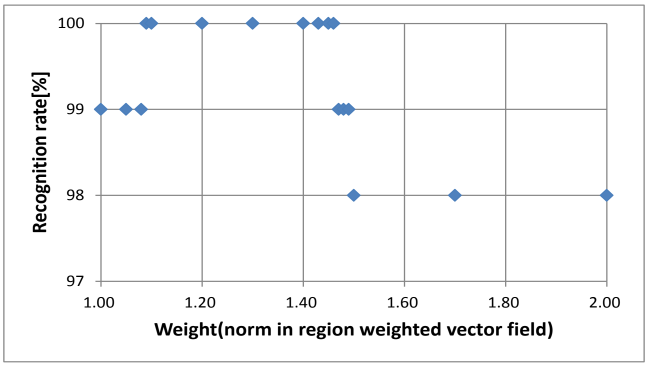


Fig.7. Relationship between weight and recognition rate.

4 CONCLUSION

The aim of our study is not only the identification of the person by face but also to read the Kansei (or feeling) information from the facial expressions. However, because our previous method is sensitive to the facial expression, it does not work so well if the expression is big. This paper has proposed RWVF method insensitive to the facial expression. The experimental results for 100 images have shown that, if the weight value is in the range from 1.09 to 1.46, the face identification is very successful (100% recognition rate).

REFERENCES

[1] M.A.Turk and A.P.Pentland, "Face recognition using eigenfaces," Proc. of IEEE conference on Computer Vision and Pattern Recognition, pp.586-591, June, 1991.
[2] Murase and S.K.Nayar, "Learning object models from appearance," Proc. of AAAI-93, American Association for Artificial Intelligence, pp.836-843, July, 1993.
[3] A.Pentland, B.Moghaddam, T.Starner, "View based and modular eigenspace for face recognition," Proc. of IEEE conference on Computer Vision and Pattern Recognition, pp.84-91, 1994.
[4] Hiroshi Murase and Shree K.Nayar, "Illumination Planning for object recognition using parametric eigenspace," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol.16, No.12, pp.1219-1227, 1994.
[5] H.Wechsler, P. Jonathon Phillips, V. Bruce, F.F. Soulie, T.S. Huang, ed., Face Recognition --From Theory to Applications--, Springer-Verlag, 1998.
[6] M. Weber, W. Einhauser, M. Welling, P. Perona, "Viewpoint-Invariant Learning and Detection of Human Heads," Proc. of the International Conference on Automatic Face and Gesture Recognition 4TH, pp.20-27, 2000.
[7] T. Hattori, W. Fukuda, K. Takeda, and T. Yamasaki, "Face Recognition among Others as an Affine Transformed Subpattern," Intelligent Autonomous Systems, Y. Kakazu et al.(Eds.), IOS Press, pp.791-796, ISBN:9051993986, June, 1998.
[8] Tetsuo Hattori, Seiu Shu, Waka Fukuda, Toshinori Yamasaki, "New Face Recognition Method among Others in Complex background," Proc. of 2000 NOLTA2000, IEICE, pp.201-204, ISBN: 3933592844, Sept., 2000.
[9] Tetsuo Hattori, Seiu Shu, Toshinori Yamasaki,Waka Fukuda,"Face Pattern Recognition under Complex Background", CD ROM Proceedings of IEEE International Conference on SMC, pp.3890-3895, ISBN:0-7803-6586-0, Oct. 2000.
[10] Tetsuo Hattori, Hiroyuki Kitajima, Toshinori Yamasaki, "Face Recognition in the Crowded People under General Background," Proc. of IEEE IECON2002, vol.2/4, pp.1428-1433, ISBN: 0780374754, Nov., 2002.
[11]Tetsuya Izumi, Tetsuo Hattori, Shu Qing Yu, Hiroyuki Kitajima, Toshinori Yamasaki, "Face Identification for People Image with General Background," Proc. of ISHS'06, pp.140-145, Busan, Korea, Oct., 2006.
[12]Tetsuya Izumi, Tetsuo Hattori, Shu Qing Yu, Hiroyuki Kitajima, Toshinori Yamasaki, "Face Recognition for Image of People under General Background Using Approximate Parameter Estimation of Affine Transform", IEEJ Trans. EIS, Vol.127, No.4, pp.623-629, April 2007 (in Japanese).
[13] Shu Qing Yu, Tetsuo Hattori, Tetsuya Izumi, Hiroyuki Kitajima, and Toshinori Yamasaki: "Face Recognition Method for People Image Using Vector Phase Field and Region-weighted Similarity", Proc. of The 3rd International Symposium on Humanized Systems (ISHS'07), Japan Society of Kansei Engineering, Chinese Association for Artificial Intelligence, The Institute of Control, Automation, and System Engineers (Korea), pp.115-118, Muroran Japan, Sept.13-15, 2007.
[14] Qingyu Shu, Tetsuo Hattori, Tetsuya Izumi, Hiroyuki Kitajima, Toshinori Yamasaki: "Face Identification for People Image with General Background Using Vector Phase Field," CD Proc. of IEEE ICIT2008, WP1-D5, pp.1-5, IEEE Catalog Number: CFP08CIT-CDR, ISBN:978-1-4244-1706-3, Chengdu, China, 2008.
[15]Tetsuya Izumi, Tetsuo Hattori, Shu Qing yu, Hiroyuki Kitajima, Toshinori Yamasaki: "Face recognition for images of persons against a general background using approximate parameter estimation of affine transform", Journal of Electronics and Communications in Japan, John Wiley & Sons, ISSN 1942-9541, Vol.92, No.11, pp.1-8, Nov. 2009.