TELKOMNIKA, Vol. 11, No. 12, December 2013, pp. 7131~7137
e-ISSN: 2087-278X

Ear Recognition based on Forstner and SIFT

Ma Chi*[1,2],Zhu Yongyong[3],Tian Ying[1]
[1]College of Software, University of Science and Technology LiaoNing, Anshan, China
[2]Dongling School of Economics and Management, University of Science and Technology Beijing, Beijing, China
[3]Department of Economics and Business Administration, Chongqing University of Education, Chongqing, China
*Corresponding author, e-mail: asmachi@126.com

Abstract

Extraction and expression of features are critical to improving the recognition rate of ear image recognition. This paper proposes a new ear recognition method based on SIFT (Scale-invariant feature transform) and Forstner corner detection technology. Firstly, Forstner corner points and SIFT keypoints are detected respectively. Then taking Forstner corner into the SIFT algorithm to calculate their descriptor as the image feature vectors. Finally ear recognition based on these feature is carried out with Euclidean distance as similarity measurement. Abi-directional matching algorithmis utilized for improving recognition rate. Experiments on USTB database show that the recognition rate reaches more 94%. The Experimental results prove the effectiveness of the proposed method in term of recognition accuracy in comparison with previous methods. It is robust to rigid changes of ear image and provides a new approach to the research for ear recognition.

Keywords*: ear recognition, SIFT, forstner*

1. Introduction

Ear recognition has attracted much more public concern in recent years. It is a technology with many advantages: easy collection, easy to accept, low-cost equipment and invulnerable to the environmental impact [1]. Meanwhile, image size of the human ear is directly proportional to the amount of data processing which decides to the color distribution. Image matching of the human ear is an important part of the human ear recognition [2]. Common sense, image matching [3] is a process based on the comparing of the features between known patterns and unknown images.

Feature extraction is one of the most important research areas in image matching.In recent research, according to the extracted features, ear recognition method can be summarized in two categories [2], the method based on the algebraic features [3-5] and the method based on structural features [6-7].

1) Based on the algebraic features: Victor B et al. proposed base on Principal Component Analysis (PCA) for the human ear recognition [8]. Thereafter Dasari et al. proposed Kernel principal component analysis (KPCA) method to extract algebraic features of the human ear, and using the Support Vector Machine for the identification, achieved a recognition rate higher than the PCA method [9]. The Kocaman B et al. used respectively, Principal Component Analysis (PCA) and discriminative common vector analysis (DCVA), Fisher Linear Discriminative Analysis (LDA) and Locality Preserving Projections (LPP) method were used for ear recognition [10].

2) Based on the structural features: These methods by finding the key points of the contour and the internal structure of the human ear, to build the structural features. For example Choras Mproposed geometry-based feature points extraction method [11]. Hurley et al. proposed to extract the structural characteristics of the force field transformation theory of the human ear [12]. The principle is that the pixel in the image of the human ear as a Gaussian attractor, which the human ear image is converted to a force field. Bustard, et al. proposed the ear recognition method based on SIFT features [13].

At present all kinds of methods all need accurate positioning, easy to affected by the change of illumination Angle and so on. So they have a high recognition rate for the standard

Received March 29, 2013; Revised July 2, 2013; Accepted August 27, 2013

ear image. Buttheir recognition rate is not idealfor a change in the Angle of acquisition of ear image. Based on this, this paper puts forward the Forstner-SIFT algorithm which is applied to the ear recognition to improve the recognition rate for the change in the Angle of acquisition of ear image.

SIFT is a local feature descriptors proposed by David Lowe at the university of British Columbia in 1999, and has more in-depth development and perfection in 2004 [14-15]. SIFT feature is that the local characteristics of image, the scaling and brightness to image scale and rotation changes keep invariance, the perspective changes and affine transformation and noise also maintain a degree of stability [16].

Ear recognition based on SIFT features proposed in the literature [13], but for the depth rotation -30°~+30°of the human ear image, recognition rate is only 40% and 34%. The reason is: 1) The shape and structure of ear Image is very simple. It has small changes in gradient and a relatively small amount of information per unit area. SIFT can only find a small number of SIFT keypoint. 2) SIFT feature points is a stable point after the Gaussian convolution, but after a depth rotation, SIFT keypointsof the ear image will be overwritten and the value of the neighborhood will change, so the SIFT feature vector will vary greatly.

The paper is organized as follows: in Section 2, Feature point detection based on SIFT algorithm is described for ear images, briefly discusses the ear recognition based on Forstner-SIFT descriptor. In Section 3, experimental results are demonstrated and finally, the conclusion is made in Section 4.

2. The Ear Recognition Method based on Fusion Forstner and SIFT

In order to overcome the problem of low recognition rate for depth rotate the image for ear recognition based on SIFT.Corner detection method based on image structure is fused into the SIFT algorithm. After theoretical research and experimental, finalize the Forstner corner detection method to be used [17]. The reasons are: 1) Forstner algorithm is accurate, high precision. 2) Forstner are greatly influenced by gray level and contrast. Detection to the ear image using Forstner corner operator is shown in Figure 1.

Combining Forstner algorithm and SIFT for image feature vector generation process is as follows:

1) First load the images of the human ear;
2) Then detection images of SIFT key points and Forstner corner points;
3) Next put the feature points which have been obtained into the point pool;
4) Afterwards, computing the dominant directions of the point in the pool;
5) Finally, generate the points' SIFT feature vector.

The detailed matching process of fusion Forstner corner and SIFT algorithm is described as following.

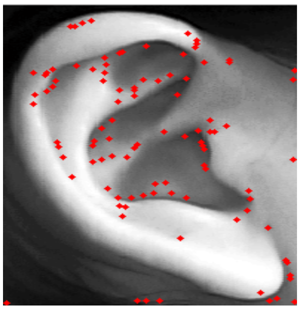

Figure 1. Forstner Operator to the Ear Image Corner Detection

2.1. Detection Image SIFT Keypoints

Under a variety of reasonable assumptions the only possible scale-space kernel is the Gaussian function. Therefore, the scale space of an image is defined as a function, $L(x, y, \sigma)$, that is produced from the convolution of a variable-scale Gaussian, $G(x, y, \sigma)$, with an input image, *I*(*x*, *y*):

$$L(x, y, \sigma) = G(x, y\sigma) * I(x, y) \tag{1}$$

Where * is the convolution operation, $G(x, y, \sigma)$ is a variable-scale Gaussian,

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)} / 2\sigma^2 \tag{2}$$

To efficiently detect stable keypoint locations in scale space, it is required to using scale-spaceextrema in the difference-of-Gaussian function convolved with the image, $D(x, y, \sigma)$:

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) = L(x, y, k\sigma) - L(x, y, \sigma) \tag{3}$$

2.2. Detection Image Forstner Corner Points

Forstner algorithm is to first select a pixel, and then calculates its gradient, and next calculates the gray scale covariance matrix of the window which the point is as the center of it. Ultimately, the least and most close to the graphics points will be the feature points. The calculation steps of Forstner are as follows:

1) Calculate the Robert's gradient of each pixel

$$f_u = \frac{\partial f}{\partial u} = f(i+1, j+1)$$
$$f_v = \frac{\partial f}{\partial v} = f(i, j+1) - f(i+1, j) \tag{4}$$

2) Calculate the gray scale covariance matrix of the 1*1 window

$$Q = N^{-1} = \begin{bmatrix} \sum f_u^2 \sum f_u f_v \\ \sum f_v f_u \sum f_v^2 \end{bmatrix} \tag{5}$$

3) Calculate the interest value

$$w = \frac{1}{trQ} = \frac{DetN}{trN}$$
$$q = \frac{4DetN}{(trN)^2} \tag{6}$$

4) Set the threshold named *Tq* and *Tw*. *Tq* usually between 0.5 and 0.75. Tw is obtained by the formula as follow:

$$Tw = \overline{kw}(k = 0.5 \sim 1.5) \tag{7}$$

5) The least and most close to the graphics points will be the fature points.

2.3. Computing the Dominant Directions

By assigning a consistent orientation to each keypoint based on local image properties, the keypoint descriptor can be represented relative to this orientation, and, therefore, achieve invariance to image rotation. The scale of the keypoint is used to select the Gaussian smoothed image, *L*, with the closest scale, so that all computations are performed in a scale-invariant manner. For each image sample, *L(x, y)*, at this scale, the gradient magnitude, *m*(*x,y*), and orientation, *θ*(*x*, *y*), is precomputed using pixel differences:

$$m(x,y)=\sqrt{(L(x+1,y)-L(X-1,Y))^2+(L(X,Y+1)-l(X,Y-1))^2} \quad (8)$$

$$\theta(x,y)=\alpha\ \tan\ 2((L(x,y+1)-L(x,y-1)/(L(x+1,y)-L(x-1,y))) \quad (9)$$

2.4. The Local Image Descriptor

Generation of keypoint descriptors: The coordinate axis of rotation for the feature points in the direction, in order to ensure its rotation invariance, then make the feature point as the center and take the 8×8 window. Computing the gradient direction histogram is computed in 8 directionson each 4×4 small pieces, and then, marking the values of each gradient direction, thus a descriptor array is formed. In actual calculation process, in order to improve the robustness of matching feature points, 4×4 descriptor array is used to describe the images, a 128-dimensional SIFT feature vector will be got.

The last step of the proposed algorithm generates a pixel neighborhood region feature descriptor based on the SIFT and Forstnerkey points. For the critical point and its neighbor, this descriptor is compressed, with high discrimination described as different (different area) and the shooting angle of illumination conditions, and photo change having the robustness (different images in the same the critical point has a similar representation).

2.5. Matching

After the descriptor of the image keypointsare obtained, it begins to matching. An ear image is matched by individually comparing each feature from the ear image to this previous database and finding candidate matching features based on Euclidean distance of their feature vectors. In order to reduce the mismatch of the keypoints, we control comparative experiment using two-way matching [19]. SUM (col1∩col2), where the col1 is the set of points matching procedure from training image to target image, and the col2 is the set of points matching procedure from target image to training image.

3. Experiments Results and Analysis

In order to detect the performance of the ear recognition algorithm proposed, we do experiments in database of different lighting conditions and rotation angle. The entire experimental process is completely automatic, without any manual intervention.

At present, the ear images used in most works mainly come from three specific databases: UND database collected by University of Norte Dame, UCR database collected by University of California at Riverside and USTB database collected by University of Science and Technology Beijing.

The experiments with the proposed method are accomplished on the ear database of USTB II (University of Science and Technology Beijing). The database consists of 308 ear images of 77 individuals, the four images per person, with a resolution of 300×400 pixels.The first one is ear positive image, second and third one is respectively for the ear, +30 degrees and -30 degrees of depth rotated image, the fourth one is ear positive image of change the lighting conditions.Because the main purpose of this method is to verify the robustness of the image light, rotation angle and scale change, there is no need to advance the image normalization.

First of all, experiment using SIFT-based ear recognition on the image library, where the first image as the training image, other image as the target image. The experiment results are shown in Table 1:

Table 1. The Results of the Experiment of SIFT for the Ear Image Database

distRatio	2nd image	3rd image	4th image	Average
0.4	32.47%	31.17%	85.71%	49.78%
0.6	35.06%	25.97%	85.71%	48.92%
0.8	40.26%	33.77%	80.52%	51.52%

The experimental resultsshow that it can obtain the highest average recognition rate when distRatio=0.8. The results of Two-way matching are shown in Table 2, the vertical axis denotes recognition rate.

Table 2. Results of Two-way Matching

distRatio	2nd image	3rd image	4th image	Average
0.4	15.58%	18.18%	68.83%	34.20%
0.6	41.56%	38.96%	87.01%	55.84%
0.8	44.16%	44.16%	83.12%	57.14%

Table 3. The Result of the Experiment of Forstner-SIFT for the Image Database

distRatio	2nd image	3rd image	4th image	Average
0.6	84.42%	87.01%	96.10%	89.18%
0.8	92.21%	94.81%	98.70%	95.24%
0.9	80.52%	84.42%	94.81%	86.58%

From the experimental results it can be shown that based on SIFT method has the low recognition rate about human ear image. The main reason is that too few keypoints are detected, so causing some mismatch. It reflects the following questions:

1) The main reason for low recognition rate of 2nd and 3rd image is that the SIFT algorithm for image scale, rotation, brightness, and the affine transformation is stable, but if the image is rotated depth, the keypoint may be blocked and the value of the neighborhood will change, in the end, the feature vector vary greatly. Figure 5 shows that SIFT feature vector is generated for the same keypoint in different images. These changes affect recognition stability.

2) The 4th image recognition rate is relatively high, the SIFT algorithm is stable under changing light conditions.

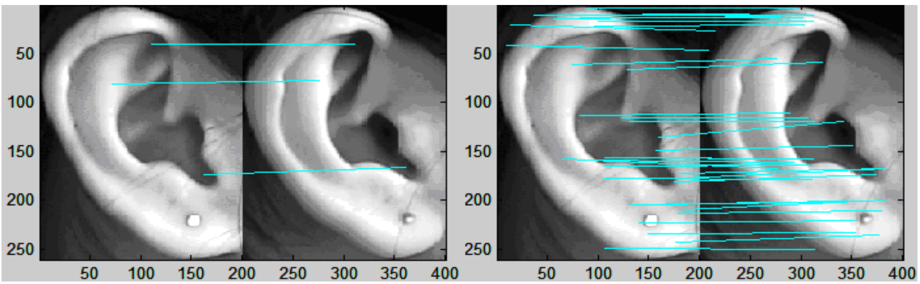

Figure 2. SIFT and Forstner-SIFT Matching Results Contrast

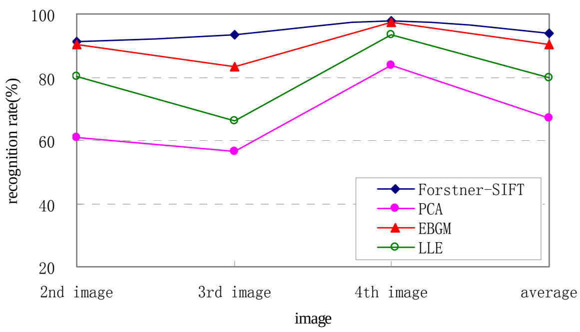

Figure 3. Compare with other Algorithm

3) When the threshold changes that the recognition rate of different images can produce different changes. Matching is more accurate when the threshold is lowered. After depth rotation the image feature vector will change, so the 2nd and 3rd image recognition rate will decline. The fourth image is the highest recognition rate when distRatio = 0.6, which is the same as the threshold recommended by David G. Lowe in his paper. The following experiment is accomplished by using fusion Forstner and SIFT ear recognition method, where the first image as the training image other image as the target image. The experiment results are shown in Table 3.

According to the experimental data we can find that ear recognition based on fusion Forstner and SIFT method has been improved significantly than the SIFT algorithm, the main reason is that the Forstner corner is very stable under the conditions of light and depth change, along with increase of feature points number, the match will be more stable. Figure 2 is comparison the SIFT algorithm with fusion Forstner and SIFT algorithm in matching results, the left picture is about SIFT, the right one shows the proposed fusion Forstner and SIFT.

Ear recognition based on fusion Forstner and SIFT still has some problems: The 3rd image recognition rate significantly higher than the 2nd image, the main reason is that the Forstner corner after a +30 degree rotation of images will be blocked, while the 3rd image in the -30 degree rotation Forstner corner was obscured, as Figure 7. We can see many of the key points in the first image but they were blocked in the second image. This structure-based feature recognition method can not be overcome.

Figure 3 is comparison our method in this paper with other methods using the same human ear experiment database. The vertical axis denotes recognition rate. It is obvious thatproposed method has higher recognition rate than other algorithm.

4. Conclusion

This paper presents an efficient and robust ear recognition method, which uses Forstner corner and SIFT descriptor for feature extraction. SIFT descriptor contains a wealth of local image information, with rotation, zoom scale invariance, affine transformation and so on, and to noise has a certain stability, but has low recognition rate of the image after depth rotation. In this paper, the reasons that Forstner and SIFT can be combined is analyzed, to some extent, this method can overcome the shortcoming of SIFT, by using two-way matching can obtain higher recognition rate. Experiment result shows robust performance of the method, the recognition performs with an accuracy of more than 94% by using two-way matching. This proves that the method can be developed for actual application.

Acknowledgement

Thanks for the image library of Ear Recognition Laboratory at USTB. This paper is supported by (1) The National Natural Science Foundationof China No. 61202315, (2) Foundation of Liaoning Educational committeeunder the Grant No.L2010194.

References

[1] M Burge, W Burger. *Ear Biometrics for Machine Vision*. Proceedings of 21st Workshop of the Austrian Association for Pattern Recognition. 1997: 275-282.
[2] WangXiaoyun. Study on Human Ear Recognition by Algebraic Methods. PhD Thesis, ShenYang University of Techinology. 2011.
[3] Wang Xiaoyun, GuoJinyu, Human ear recognition based on statistic features of local information. *Computer Engineering and Applications*. 2012; 9: 141-144.
[4] Wang Yu, Xue Hong. Application of strengthened authentication vectors to multimodal recognition of ear and face. *Computer Engineering and Applications*. 2012; 48(12): 29-33.
[5] LiYibo, Cao Jingliang, ZHANG Haijun. Improved method fear recognition based on tensor PCA. *Computer Engineering and Applications*. 2011; 47(25): 171-174.
[6] Feng Jun, Liang Xiaoxia, Mu Zhichun. Study On Ear Recognition By Using Histogram Of Oriented Gradient Features And Subregion Fuzzy Fusion. *Computer Applications and Software*. 2012; 48(4): 79-82.
[7] Lu Xiu-Li, ShenHong-Feng, Zhao Li-Hua. Ear Recognition Based on 2DLDA and FSVM. *Science Technology and Engineering*. 2012; 20(12): 2852-2855.

[8] Victor B, Bowyer K, Sarkar S. *An evaluation of face and ear biometrics*. International Conference on Pattern Recognition. 2002: 429-432.
[9] Shailaja D, Gupta P. *A simple Geometric Approach for Ear Recognition.* International Conference on Information Technology. 2006: 164-167.
[10] Kocaman B, Kirci M Gunes EO, et al. *On ear biometrics*. IEEE International Conference on EUROCON '09. 2009: 327-332.
[11] Choras M, Choras RS. *Geometrical algorithms of ear contour shape representation and feature extraction*. Proceedings of Six Conference on International System Design and Applications. 2006: 451-456.
[12] DJ Hurley, MS Nixon, JN Carter. Force Field Feature Extraction for Ear Biometrics. *Computer Vision and Image Understanding*. 2005; 98(3): 491-512.
[13] Bustard JD, Nixon MS. *Robust 2D Ear Registration and Recognition Based on SIFT Point Matching*, 2th IEEE International Conference on Biometrics: Theory. Applications and Systems. 2008: 1-6.
[14] Lowe DG. *Object recognition from local scale-invariant features*. International Conference on Computer Vision. 1999: 1150-1157.
[15] Lowe DG. Distinctive image features from scale-invafiant key points. *International Journal of Computer Vision*. 2004; 60(2): 91-110.
[16] Zeng Luan, Wang Yuan-Qin, Tan Jiu-Bin. Improved algorithm for SIFT feature extraction and matching. *Editorial Office of Optics and Precision Engineering*. 2011; 19(6): 1391-1397.
[17] Zhang Yan, Zhang Zhicheng. Hybrid Forstner algorithm and SIFT gray image feature extraction. *Bulletion of science and technology*. 2012; 28(10).
[18] Yao Y. Image Registration and Application Based on Feature-point Detection. *National University of Defense Technology*. 2010.
[19] Anam Tariq, M Usman Akram. Personal Identification Using Ear Recognition. *TELKOMNIKA*. 2012; 10(2): 321~326.
[20] Sun Nong-Liang, Li Huan-Huan, Yang Ning, et al. Identification of Correspongding Points Based on SIFT Twice Matching Method. *Computer Engineering.* 2012; 4: 155-157.
[21] Wiskott L, Fellous JM, Kruger N, et al. Face recognition by elastic bunch graph matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1997; 19(7): 775-779.
[22] HendraKusuma, Wirawan, Adi Soeprijanto. Gabor-based Face Recognition with IlluminationVariation using Subspace-Linear Discriminant Analysis. *TELKOMNIKA*. 2012; 10(1): 119-128.
[23] Xie Zhao-xia, Mu Zhi-chun, XieJian-jun. Multi-pose ear recognition based on locally linear embedding. *Transactions on Intelligent Systems*. 2008; 3(4): 321-327.