Estimating a Shooting Angle in Ear Recognition

Daishi Watabe, Takanari Minamidani, Hideyasu Sai, Taiyo Maeda, Takaharu Yamazaki, Jianting Cao

Estimating a shooting angle in ear recognition

Daishi Watabe, Takanari Minamidani, Hideyasu Sai, Taiyo Maeda,Takaharu Yamazaki and Jianting Cao

Saitama Institute of Technology, 1690 Fusaiji, Fukaya, Saitama 369-0293, Japan
email: dw@sit.ac.jp

Abstract. To improve on our earlier work on single-view-based ear biometrics, an estimation method is presented for the shooting angle of an ear image based on the summation of similarity scores over a threshold within a database of known shooting angles. Experimental results indicate that the estimation method can improve the robustness of ear recognition in varied poses.

Keywords: ear recognition, normalizing shooting angle, Gabor feature

1 Introduction

1.1 Background

Ear shape is unique to every individual and has been used in forensic science over the past 40 years [1]. In particular, ear prints left on walls have been used in identification of criminals, most notably in the Netherlands [2]. A detailed survey on using ear shapes for forensic purposes is available in [2], where historical studies and present issues are detailed. Furthermore, detailed surveys of automatic ear recognition systems are available in [3][4][5], where databases, algorithms, experimental conditions, and accuracies are presented. Whereas masks and sunglasses often purposely obscure facial features, ear shapes can be all that is required to identify subjects. However, the shooting angle of an ear from a surveillance camera is usually not the same as that for a facial image. Hence, accounting for such differences is necessary[6][7].

1.2 Related studies

Moriyoshi [6][7] thoroughly investigated the effect of differences in the shooting angle in the context of forensic science. As far as image processing and computer vision are concerned, a few studies have variations in the angle of the shot [8][9]. However, these studies were limited to in-plane rotations. In single-view-based ear recognition studies, the authors improved the robustness of the method for off-angle rotation [10][11], in which feature vectors of various poses are estimated from a single-view image and recognition processes are performed without using prior knowledge of shooting angles of the input images. This is done by taking a correlation against the averaged estimated feature vectors over various poses. Although this averaging process may contaminate final accuracy, we did not use information on the

shooting angle of input ear images because the estimation method of the shooting angle of an ear image is not established. One may point out the usability of the shooting angle of face images, for which various estimation methods are well-established. However, because individual variations in the ear overhang angle are considerable, the use of the estimated face angles as an ear angle is not feasible. We do not know, however, how promising it is to pursue the direction of seeking a method of estimating the shooting angle of an ear. If such an estimation method is established, it will be possible to improve the accuracy of the single-view-based ear recognition system, by taking the correlation against estimated feature vectors of a specified angle, but not against contaminated averaged feature vectors of unspecified shooting angles.

1.3 Aim of this study

An initial attempt to estimate the shooting angle of an ear image is presented. Using these estimated shooting angles, a few estimation methods for the feature vector of other shooting angles are compared experimentally. We examine the possibility of improving the robustness of ear recognition by estimating the shooting angle of an ear image.

2 Proposed method

2.1 Outline

In Subsections 2.2 and 2.3, our method for estimating the shooting angle of an input image is explained. For completeness, the method we used for ear recognition is summarized briefly in Subsections 2.4, 2.5, and 2.6.

2.2 Gabor features of ear minutiae

To fix a baseline, we used the Gabor features for the various methods described below: Let $\mathbf{x}=(x,y)$ be a point in a plane. A 2D plane wave defined by wave vector $\mathbf{k}=(k_x,k_y)$ and modified by a Gaussian function is called a Gabor function (Eq. (1)):

$$\psi(\mathbf{x})=\frac{|\mathbf{k}|^2}{\sigma^2}\exp\left(-\frac{|\mathbf{k}|^2|\mathbf{x}|^2}{2\sigma^2}\right)\left[\exp\left(i\mathbf{k}\cdot\mathbf{x}\right)-\exp\left(-\frac{\sigma^2}{2}\right)\right]. \tag{1}$$

Here σ denotes the width of this function determined by the Gaussian function. The factor $\exp\left(-\sigma^2/2\right)$ is a compensation term that eliminates averages. This condition is required from wavelet theory, but if σ is large enough, the factor can be ignored. Gabor functions are characterized as localized wavy shapes in various directions determined by plane waves. Gabor filters, i.e., convolutions with these Gabor functions, extract the direction and wavelength of these localized wavy shapes of an image near the point under consideration. Wavy shapes in various directions also characterize the outer ear. Thus, endpoints, junctions, and protuberances of the ridges of the outer ear are selected as feature points. Wavy shapes near these feature points are measured and

coded using Gabor filters.

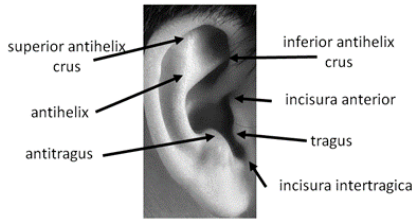

Fig. 1. Feature points of an ear.

2.3 Gabor configurations

Five wavelengths, $4,4\sqrt{2},8,8\sqrt{2},16$, were adopted as Gabor filters to cover the various widths of ridges along the ear that appear in the experimental data. Furthermore, to cover all directions evenly, eight directions corresponding to $\pi/8$ rotations are employed. To realize these settings, we set $\sigma = 2\pi$ and

$$\mathbf{k} = \left(k_x, k_y\right) = \frac{2\pi}{\lambda}\left(\cos\frac{\pi}{8}\mu\ , \sin\frac{\pi}{8}\mu\ \right), \tag{2}$$

for $\lambda = 4, 4\sqrt{2}, 8, 8\sqrt{2}, 16$ and $\mu = 0, \ldots, 7$ in Eq. (1).

We implemented these Gabor filters using a mask of 101×101 pixels for the convolution window, and this convolution was performed using the fast Fourier transform. Using this bank of Gabor filters, Gabor feature vectors were sampled at the feature points as indicated in Fig. 1 and then stacked into one vector with maximum dimensions of 560 (=80×7). Furthermore, when phases were ignored, that is only taking absolute values, these vectors became 280 (=40×7) dimensional.

2.4 Estimating shooting angles of an input image

Given an input ear image of unknown shooting angle, we can compute the similarity scores between this input ear image and the ear images of a known shooting angle from a database. Where this angle is close to the unknown shooting angle of an input ear image, it is anticipated that the number of ear images with higher scores for their similarities may be large. Based on this concept, we examined the following algorithm:

1. First, similarity scores between an input ear image and images of known shooting angles in a database are computed.
2. The summation scores for these similarity scores above a given threshold within a shooting angle are obtained. If there is no sample with a similarity score higher than the threshold, this algorithm returns a failure for the shooting angle estimation.
3. The above process 1-2 is repeated through various shooting angles in the database.
4. Finally, the shooting angle with the maximum summation is returned as an estimated shooting angle of an input image.

Similarity scores are given by normalized-cross-correlations to the phase-ignored Gabor feature vectors. A threshold is employed in order not to contaminate the estimation accuracy of shooting angle through using the lower scores of the similarities of non-similar ears. This threshold is obtained by maximizing the estimation accuracies of shooting angles through a survey of threshold-values using leave-one-out cross validation strategy.

2.5 Estimation of Gabor features after off-angle rotation for a single registration image using a linear jet transformation

For completeness, the method used in [11] is outlined. Locally, near the feature points, the subject is approximated by a tangent plane. The tangent plane does not have depth. Hence, the image of this plane rotated in depth can be estimated. This estimated image reflects local features under pose variations near the feature points. Similar to the tangent plane, Gabor jets only represent local features. Motivated by this, we explore the benefits of Gabor jets for subjects rotated in depth. The following outlines the reproduction method using Gabor jet estimates of subjects with different poses [11][10]. Let the *x*-*y* coordinates be set on the camera plane and the *z*-axis set perpendicular to this plane. Suppose that a subject plane, initially placed parallel to the camera plane, is rotated by ϕ around its *x*-axis and then θ around its *y*-axis. By observing the transformations of unit vectors, a point on the subject plane initially at $\mathbf{u}=(x,y)$ is transformed to $\mathbf{x}$ given by:

$$\mathbf{x}=\mathbf{A}\mathbf{u}, \qquad \mathbf{A}=\begin{pmatrix}\cos\phi & \sin\theta\sin\phi\\ 0 & \cos\theta\end{pmatrix}. \tag{3}$$

If this plane is initially placed at (ϕ_1,θ_1) and not parallel to the camera plane, the above transformation is:

$$\mathbf{x}=\mathbf{A}(\phi_2,\theta_2)\mathbf{A}(\phi_1,\theta_1)^{-1}\mathbf{u}. \tag{4}$$

Under this transformation, the transformation of the Gabor jets corresponding to the pose change can then be estimated. In what follows, $\mathbf{A}(\phi_2,\theta_2)\mathbf{A}(\phi_1,\theta_1)^{-1}$ is denoted as $\mathbf{A}$ for simplicity. Components of the transformed Gabor jets are obtained by convoluting the Gabor function with the transformed image $I(\mathbf{A}^{-1}\mathbf{x})$. Using $\mathbf{x}=\mathbf{A}\mathbf{u},\ \mathbf{x}'=\mathbf{A}\mathbf{u}'$, this is

$$\begin{aligned} j'_{\mathbf{k}}(\mathbf{x}) &= \int I(\mathbf{A}^{-1}\mathbf{x}')\Psi_{\mathbf{k}}(\mathbf{x}-\mathbf{x}')d\mathbf{x}' \\ &= \int I(\mathbf{u}-\mathbf{u}')\Psi_{\mathbf{k}}(\mathbf{A}\mathbf{u}')\,|\,\mathbf{A}\,|\,d\mathbf{u} \end{aligned}. \tag{5}$$

Assuming the following approximation:

$$\Psi_{\mathbf{k}}(\mathbf{A}\mathbf{u}')\,|\,\mathbf{A}\,|\approx\sum_{\mathbf{k}'}c_{\mathbf{k}\mathbf{k}'}(\mathbf{A})\Psi_{\mathbf{k}'}(\mathbf{u}'), \tag{6}$$

the Gabor jet transformation is simply written as:

$$j'_{\mathbf{k}}(\mathbf{x})\approx\sum_{\mathbf{k}'}\mathbf{c}_{\mathbf{k}\mathbf{k}'}(\mathbf{A})j_{\mathbf{k}'}(\mathbf{u}). \tag{7}$$

Once $\mathbf{C}^{(\mathbf{A})} = (c_{\mathbf{kk}'}(\mathbf{A}))$ is obtained, the transformation of the Gabor jets can be estimated using:

$$\mathbf{j}'(\mathbf{x}) \approx \mathbf{C}^{(\mathbf{A})}\mathbf{j}(\mathbf{u}) \quad . \tag{8}$$

Matrix $\mathbf{C}$ is obtained by multiplying both sides of Eq. (6) by $\overline{\Psi_{\mathbf{k}''}(\mathbf{u}')}$ and integrating both sides. Two of the variables are difficult to determine. One is (ϕ_2, θ_2), which depends on the poses of input images. This is an unknown in real scenarios. In [11], we solved this issue by producing the Gabor feature of many other poses in advance. The other unknown variable is (ϕ_1, θ_1), which represents the normal vector of the tangent plane at each feature point. Because this variable is difficult to determine from a single-view-image, some type of statistical modeling is necessary. In [11], this model was produced using an exhaustive search of smaller equal error rates in the variable ϕ and θ using a five-fold cross validation strategy.

2.6 Estimation of Gabor features after off-angle rotation for a single registration image using principal component analysis (PCA)

To estimate the Gabor feature vectors for other poses, the feature vectors taken at the registration and input angles are stacked into one feature vector for the same person in a training set. Because phases are ignored and absolute values taken, a 560 (=40×7×2) dimensional vector is obtained. Such stacked feature vectors are created for all training datasets and subjected to PCA. For testing the sample, the Gabor feature vector at the registration angle and the null data are stacked into one vector. Using the principal component subspace, the Gabor feature vector at the input angles are estimated as a sub-vector of the back-projected stacked feature vector. Similar to Subsection 2.5, a five-fold cross validation is used to create training and test sets. The principal component subspaces serves 3D statistical modeling for estimating feature vectors of other poses.

2.7 Estimation of Gabor features after off-angle rotation for a single registration image using multiple regression analysis (MRA)

To estimate the Gabor feature vectors taken from different shooting angles, the normal equation is solved to obtain the regression coefficients that describe each component of the Gabor feature as a linear combination of the components of Gabor features of the registration angle. With phases ignored, a set of 280(=40×7) normal equations is solved for training sets and used to estimate the Gabor features of the input angles for test sets. Similar to Subsection 2.5, a five-fold cross validation is used to create training and test sets. The regression coefficients obtained serve 3D statistical modeling that can be used to estimate feature vectors for other poses.

2.8 Creating the linear discriminant analysis matrix using the estimated features for input images with unknown poses

To fix a baseline for comparison, similar to the method described in [12], all the estimated feature vectors, as illustrated in Subsections 2.5, 2.6 and 2.7, are subject to multiple linear discriminant analysis (LDA), thereby creating the LDA matrix of discriminant vectors. After applying this LDA matrix to both the input and registration feature vectors, a normalized correlation is computed to obtain the similarity between the input and registration feature vectors taken from different shooting angles.

3 Experiment

3.1 Database of feature vectors for the experiment

Experiments were performed using the database of Gabor feature vectors at ear feature points of images from the human and object interaction processing (HOIP) database [12] obtained in our previous studies [11]. The HOIP database consisted of facial images of 300 subjects photographed from 504 (72 yaw angles every 5° and 7 roll angles every 15°) directions, where the size of the ear fitted approximately within a 70×90 pixel window. By mirror reflecting images, left profile images of 600 people were subjected to Gabor feature computations. Thus a database of Gabor feature vectors for 600 people was obtained.

3.2 Shooting angles and number of visible feature points in the experiments.

To examine robustness against yaw-angle pose variations, verification experiments were performed using an ear image of the true left profile of a registration image taken from 85°. Input data were taken from yaw angles varying from 40° to 80°, every 10°.

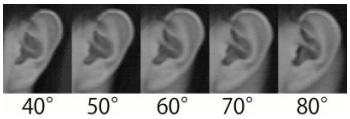

Fig. 2. Example of input images.

(Yaw angles 0°, 90°, and 180° corresponding to frontal face, true left profile, and back)

One hundred and sixty-two subjects corresponding to images with seven visible feature points at all angles were selected for input, registration, and training data. In these datasets, there was a single biometric sample for each identity in each angle.

3.3 Experiment for estimating the shooting angle.

For the algorithm presented in Subsection 2.4, a threshold was determined by maximizing the estimation accuracy of the shooting angle using a leave-one-out cross validation strategy as follows:

1. First, the set of images at a 40° angle were selected as input images. An image from this image set was selected and treated as an input image of an unknown shooting angle. All other images with the same identity as this input image were removed from the image sets of yaw angles varying from 40° to 80°, every 10°.
2. Second, the shooting angle for this input image was estimated using the algorithm presented in Section 2.4.
3. Repeating this process over all the images in the image sets at 40° angles provided an estimation accuracy at a 40° angle.
4. Performing the above steps 1-3 similarly on sets of images at 50° to 80° angles provided estimation accuracies for each angle. The accuracies from 40° to 80° angles were averaged.
5. A survey search of threshold-values that maximized the averaged accuracy was performed based on the coarse-to-fine approach.

3.4 Experiment for examining robustness

The effect of our proposed method using the estimated shooting angle, robustness was examined as follows:

1. From the feature vectors of registration data taken from 85°, feature vectors for the yaw angles 40°, 50°, 60°, 70° and 80° were estimated using algorithms LJT, PCA and MRA as demonstrated in 2.5-2.7.
2. As in Subsection 2.8, the LDA matrix was created using these estimated and registration datasets. Using this matrix, the registration, input and estimated datasets were all transformed into a coordinated dataset where discrimination was easier.
3. Similarity scores were obtained using normalized cross correlations taken against the 85° angle registration feature vector, and the estimated feature vector of the estimated shooting angle. Equal error rates were obtained from ROC using the algorithm in Section 3.5.

In summary, the following six cases were compared;

- LJT, 85° angle registration of the feature vector
- LJT, estimated feature vector of the estimated shooting angle.
- PCA, 85° angle registration of the feature vector
- PCA, estimated feature vector of the estimated shooting angle
- MRA, 85° angle registration of the feature vector
- MRA, estimated feature vector of the estimated shooting angle

Similarity scores obtained using the 85° registration feature vector corresponded to our previous method in [11]. Where the shooting angle estimation failed, the registration angle for 85° was used instead for correlation computation. A small number of such cases depended on the shooting angle (~2%).

3.5 The validity metrics for the experiments

As validity metrics for the verification experiments, the ROC, and the equal error rate EER are commonly used (10.6.3 of [13]). For computing these metrics, we used the algorithm recommended in Annex. F.1–2 of [13].

3.6 Results of the experiment for estimating shooting angle.

The accuracy at each threshold is demonstrated in Fig. 3. When the threshold value is 0.88, the maximum averaged accuracy is 46.6%. This is a somewhat encouraging result as an initial attempt, because this accuracy far exceeds that of a random answer (20%) to the question of five selective answers (40°, 50°, 60°, 70°, 80°). However, there seems to be considerable room to improve the estimation accuracy of the shooting angle.

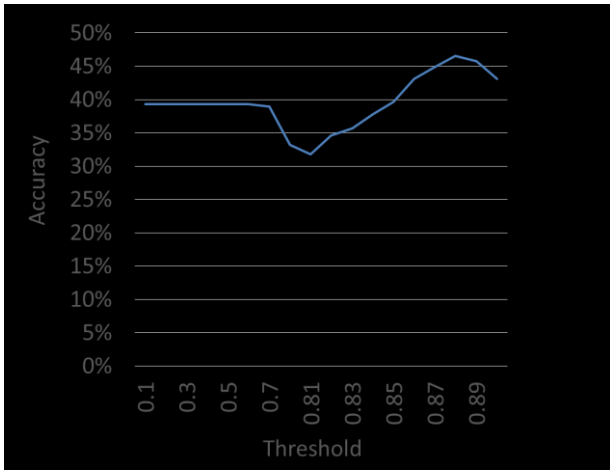

Fig. 3. Accuracy at various thresholds.

3.7 Result of the experiment to examine robustness

Using the estimated shooting angle as determined in Subsection 3.6, equal error rates at various input yaw angles were obtained, as in Fig. 4.

.

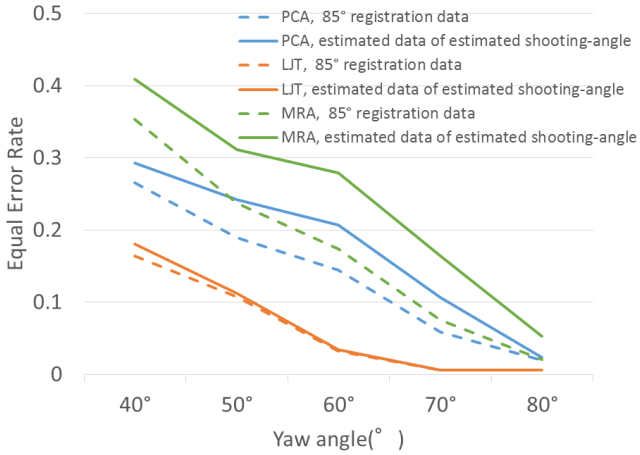

Fig. 4. Equal error rate at various yaw angles.

The results using PCA and MRA were not particularly good. Equal error rates for the estimated data of the estimated shooting angles were worse than the equal error rates using registration data without shooting angle normalization. Similar to our previous report [14], the number of subjects may not be sufficient for accurately determining the principal component subspace and regression matrix. However, using LJT, the estimated data for the estimated shooting angle perform as accurately as our previous method using averaged estimated feature vectors of various poses.

4 Discussion and Conclusions

An initial attempt to estimate the shooting angle of an ear image is presented. Although the estimation accuracy was 48.8% and far exceeds the accuracy of a random answer (20%) to a question of five selective answers, there seems to be considerable room to improve it because the presented estimation algorithm for the shooting angle is not sophisticated. Using this estimated shooting angle, the estimation method of feature vector of other poses—PCA MRA and LJT—are examined. Although none perform beyond the accuracy of our previous method using averaged estimated feature vectors of various poses, LJT performs as accurately as our previous result. Hence, using LJT and refining the accuracy of the estimated shooting angle by improving the algorithm, there may be a chance to improve the robustness of single-view-based ear recognition.

5 Acknowledgments

The facial database was used with the permission of Softopia Japan Foundation. This work was supported by KAKENHI 22700219 and 80337609.

6 References

1. Iannarelli, A. V.: Iannarelli system of ear identification. Foundation Press, Brooklyn, NY (1989).

2. Meijerman, L., Thean, A., Maat, G.: Earprints in forensic investigations. Forensic Sci. Med. Pathol. 1, 247–256 (2005).

3. Abaza, A., Ross, A., Hebert, C., Harrison, M.A.F., Nixon, M.: A Survey on Ear Biometrics. ACM Trans. Embed. Comput. Syst. 9, 39:1–39:31 (2010).

4. Pflug, A., Busch, C.: Ear biometrics: a survey of detection, feature extraction and recognition methods. IET Biometrics. 1, 114 (2012).

5. Yuan, L., Mu, Z.C., Yang, F.: A Review of Recent Advances in Ear Recognition. Proc. 6th Chinese Conf. Biometric Recognit. CCBR. 252–259 (2011).

6. Moriyoshi, H., Miyoshi, M., Hino, D., Nakayama, H., Morikawa, T., Nakaki, S., Itohara, K.: Morphological Classification of the Ear Components and the Correlation between Morphological Characteristics in Each Component. Japanese J. Forensic Sci. Technol. 14, 61–70 (2009).

7. Moriyoshi, H., Miyoshi, M., Hino, D., Nakayama, H., Morikawa, T., Nakaki, S., Itohara, K.: Personal Identification from Ear Images in the Different Orientations. Japanese J. Forensic Sci. Technol. 12, 27–34 (2007).

8. Watabe, D., Sai, H., Ueda, T., Sakai, K., Nakamura, O.: ICA , LDA , and Gabor Jets for Robust Ear Recognition , and Jet Space Similarity for Ear Detection. Int. J. Intell. Comput. Med. Sci. Image Process. 3, 9–29 (2009).

9. Watabe, D., Sai, H., Sakai, K., Nakamura, O.: Ear biometrics using jet space similarity. In: Proc. of the IEEE Canadian Conference on Electrical and Computer Engineering, 2008. CCECE 2008. pp. 1259–1264. , Niagara Falls. Canada (2008).

10. Watabe, D., Wang, Y., Minamidani, T., Sai, H., Sakai, K., Nakamura, O.: Empirical Evaluations of a Single-view-based Ear Recognition when Rotated in Depth. Kansei Eng. Int. J. 11, 247–257 (2012).

11. Watabe, D., Minamidani, T., Sai, H.: Improving the Robustness of Single-View-Based Ear Recognition When Rotated in Depth. In: International Conference on Neural Information Processing, Lecture Notes in Computer Science. pp. 177–187. Springer-Verlag (2012).

12. Yamamoto, K., Niwa, Y.: Human and Object Interaction Processing (HOIP) project. In: Joho Shori Gakkai Shinpojiumu Ronbunshu (in Japanese). p. 379–.384 (2002).

13. ISO/IEC: Information technology — Biometric performance testing and reporting — Part 1: Principles and framework. In: ISO/IEC 19795-1 (2006).

14. Watabe, D., Minamidani, T., Sai, H., Cao, J.: Comparison of Ear Recognition Robustness of Single-View-Based Images Rotated in Depth. In: 2014 Fifth International Conference on Emerging Security Technologies (EST),. pp. 19 – 23. IEEE, Alcala de Henares (2014).