1. INTRODUCTION

1.1 Background

Ear shape can be sufficiently unique to identify individuals and has been used in forensic science over the past 40 years [1]. Whereas masks and sunglasses often purposely obscure facial features, ear shape can be all that is required to identify subjects [2, 3]. However, the angle of a shot from a surveillance camera is usually not the same as in a flat 2D image or mug shot. Hence, accounting for differences in shooting angle is necessary.

1.2 Related studies

Most robust ear recognition systems [4, 5] are based on 3D data obtained using laser-range-finders. However, usually mug shots of subjects are 2D images. With that in mind, ways to improve the robustness of the process is important particularly involving shooting angle variation as occurs with flat 2D images. Such recognition systems [6, 7] have been developed although cooperation of the person to be identified was available. Robustness to relatively large pose variations at a level required for forensics is not necessarily considered.

A few studies have treated pose variation [8, 9]. However, these are limited to in-plane rotation; a more general treatment is sought. By extending the methods presented in [8, 9], the authors improved the robustness of the method for off-angle rotation in single-view-based ear recognition [10, 11]. Although this method seems promising, this method must be thoroughly examined in various realistic scenarios, to fill gaps between laboratory settings and real applications.

1.3 Aim of this study

To prove the viability of [10, 11], thorough empirical evaluations were performed in scenarios where camera angles and resolutions are more realistic than those used in [8, 9], namely, with surveillance cameras normally installed on ceilings. A change in shooting angle to an upward direction, for example, has been reported to change the classification of ear lobes (some attached type appear to be middle type and some middle type appear to be free type) [2]. Furthermore, the resolution of surveillance images is usually low. In these cases, careful scrutiny is advised in classifying ear shapes. For example, in regard to scaphoid fossa, an interrupted type (interrupted by the contact between the helix and antihelix) appears to be a continuous type (continues to the ear lobe). The appearance of the intertragic groove (grooves on the lobe starting at intertragic notch) heavily depends on the quality of the image [3]. Hence, an upward change in shooting angle, and low resolution are both considered to be challenging problems even within the forensic community.

In the context of biometrics, these frequent and challenging real scenarios have not yet been fully reviewed in the literature. Therefore, thorough empirical evaluations were performed to examine the robustness in such scenarios using the present authors’ previously-introduced methods [10, 11], demonstrating the viability of the single view-based method as proposed therein.

2. PROPOSED METHOD

2.1 Gabor features of ear minutiae

Let \( \mathbf{x} = (x, y) \) be a point in a plane. A 2D plane wave defined by wavevector \( \mathbf{k} = (k_x, k_y) \) and modified by a
Gaussian function is called a Gabor function (Eq. (1)):

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

(1)

Here $\sigma$ denotes the width of this function determined by the Gaussian function. The factor $\exp(-\sigma^2/2)$ is a compensation term that eliminates averages; this condition is required from wavelet theory, but if $\sigma$ is large enough, this term can be neglected.

Gabor functions are characterized as localized wavy shapes in various directions determined by the plane waves. Gabor filters, i.e., convolutions with these Gabor functions, extract 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 (Fig. 1). Wavy shapes near these feature points are measured and coded using Gabor filters.

Five wavelengths, $4.4\sqrt{2}, 8.8\sqrt{2}, 16$, are adopted for the 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 $\pi/4$ turns are employed. We implemented this Gabor filter by using a mask of $101 \times 101$ pixels for the convolution window, and this convolution is performed using fast Fourier transform.

Using this bank of Gabor filters, Gabor jets are sampled at the feature points indicated in Fig. 1.

2.2 Localization of feature points based on jet space similarity

The localization of feature points consists of an initial coarse step to find ear shapes and then a fine step that locates each feature point of the ear. In preparation, a standard model ear graph, consisting of seven vertices, is created by averaging coordinates from training data after subtracting the center of gravity. Furthermore, principal component analysis (PCA) is applied to both these coordinates and Gabor jets.

In considering computational cost, in the initial coarse step, we used Gabor jets sampled at vertices of sub-grids consisting of three of the seven feature points; the superior antihelix crus, inferior antihelix crus, and the body of the antihelix. This is because successful detection indicates that hair is not heavily occluding the ear. In the fine step, Gabor jets sampled at all seven vertices are used.

In the initial coarse step, the above created standard model ear graph is treated as a “rigid body”. Over the entire input facial image, we scan the point under consideration with spacing of four pixels. The model ear graph consisting of all seven vertexes is arranged so that its center of gravity coincides with the point under consideration.

At each scanning location, Gabor jets at each of the three vertices of the sub-grids of the scanning rigid model ear graph mentioned above are sampled and projected onto the PCA subspace (we examined an 80% contribution ratio) of the Gabor jets for the training data. The correlation between an input Gabor jet and its projection onto the PCA subspace is called a jet space similarity [11].

Candidates for the centers of gravity of the ear graph can be marked as detection points when the jet space similarities at the three feature points are higher than a given threshold. The coordinates of these detection points and the similarity scores are recorded on arrays of 1/4th of the arrays for images. To avoid multiple detections, areas of connected components are computed using labeling and discard the small areas. The centers of gravity of the connected components for the detected regions are computed. Upon computing these centers of gravity, the similarity scores are used as weighting factors. It should also be noted that labeling is carried out on the array recording the detection points, and not the original face images. Thus, only a few candidates for the centers of gravity of the ear graph are determined. For these candidate centers of gravity, the initial points of all the seven vertices are determined by arranging the model ear graph so that its center of gravity coincides with these candidate centers of gravity.

In the fine step, from the initial points of the seven vertices determined above, the shape of the graph is deformed to determine more accurate coordinates for each feature point. This deformation is not performed independently for each point. All seven points are simultaneously deformed in salient directions given by the principal components of the coordinates of the seven vertices in the training data. The expansion coefficients for each ear are estimated by parameter searches using a “for loop”.

The first principal component multiplied by the value of the loop variable corresponding to an expansion coefficient, is added to the averaged shape, creating a deformed model ear graph. For each loop variable, Gabor jets at each vertex are sampled and projected onto the principal component subspace of the training data; jet space similarities are then computed. The first expansion coefficient
is finally estimated as the value of the loop variable where
the jet space similarity is highest. Once the first expansion
coefficient is estimated, the second principal component
multiplied by the loop variable value is added to the
summation of the averaged shape and the first principal
component multiplied by the expansion coefficient esti-
mated above. Similar to the first expansion coefficient, the
second expansion coefficient is estimated using jet space
similarity. Repeating the above for the remaining principal
components (we examined an 80% contribution ratio), a
fine-tuned deformed graph is obtained for each candidate
for the centers of gravity of the ear graph determined in
the initial coarse step. Comparing the jet space similarities
between these fine-tuned deformed graphs, the coordinates
of the seven vertices are finally determined.

At the location of the vertices, the Gabor jets are sampled
and used in the classification stage.

In the standard elastic-bunch graph-matching method
[8], a similarity between sampled Gabor jets and the aver-
gaged or representative Gabor jets of registered images are
used. Using not only averaged but also principal compo-
nents, we expect misdetections because of large individual
variations in ear shape to decrease.

2.3 Estimation of Gabor features after off-angle
rotation

Pose variations within a camera plane can be estimated
by rotating the image (Fig. 2). Such reproductions are not
necessarily accurate, because changes in shading induced
by pose variation are not considered. However, if shading
is relatively light, as a result of using sufficient lighting,
this is a relatively minor issue compared with the diffi-
culty in reproducing images rotated in depth (Fig. 3).

This difficulty stems from the depth of a subject in an
image. Clearly, plane objects without depth are easier to
process (Fig. 4). Reasonable image reproductions rotated
in depth can be obtained using affine transformations.

Figure 2: A subject rotated in a camera plane

Figure 3: A subject rotated in depth

Locally, near each feature point, the surface is approxi-
mated by a tangent plane. The tangent plane does not have
a depth. Hence, the image of this plane rotated in depth can
be estimated (Fig. 5). 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 the above, we explore the
benefits of Gabor jets of subjects rotated in depth. The
following outlines the reproduction method, estimating
Gabor jets of subjects with different poses [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 observ-
ing the transformations of unit vectors, a point on the subject
plane initially at \( u = (x, y) \) is transformed to \( x \) given by

\[
A = \begin{pmatrix}
\cos \phi & \sin \phi & 0 \\
\sin \theta & \cos \theta & 0
\end{pmatrix}
\]

If this plane is initially placed at \((\phi_1, \theta_1)\) and not parallel to
the camera plane, the above transformation is

\[
x = A(\phi_2, \theta_2)A(\phi_1, \theta_1)^{-1}u.
\]

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

\[
j'_k(x) = \int I(A^{-1}x')\Psi_k(x - x')dx'
= \int I(u - u')\Psi_k(Au')|A|du'.
\]

Assuming the following approximation

\[
\Psi_k(Au')|A| \approx \sum_k c_{kk}(A)\Psi_k(u') \tag{2}
\]

Figure 4: A planar object rotated in depth

Figure 5: A tangent plane approximation of a subject rotated
in depth
the Gabor jet transformation is simply written as

\[ j'_k(x) = \sum_{k} \sum_{\lambda} c_{kk}(\lambda) j_k(\lambda) \, . \]

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

\[ j'(x) = C^{(A)} j(x) \, . \] (3)

Matrix \( C \) is obtained by multiplying both sides of Eq. (2) by \( W_k(u) \) and integrating both sides (see Appendix).

This estimation algorithm assumes that the normal vectors of the tangent plane are established in advance so that the angles \( (\phi, \theta) \) between these normal vectors and the vector normal to the camera plane can be determined. In a real scenario, these angles are hard to establish in advance for each individual ear image. Hence, typical normal vectors of each feature point are determined in advance based on an exhaustive search of \( \theta \) and \( \phi \), with which better recognition rates are obtained for each feature point.

### 2.4 Training using estimated Gabor features

In addition to real registration data, the estimated Gabor jets for other poses obtained following the above outline are used as training data, and combined into class information for each individual. Using this class information, we try to improve the robustness of our method against pose variations.

For the training algorithm, multiple discriminant analysis [12] is employed. This algorithm provides coordinate transformations to coordinates where class separations are easier. The matrix \( W \) performing this coordinate change is obtained by maximizing the following function defined by

\[ J(W) = \frac{W^T S_b W}{W^T S_w W} \, . \] (4)

where \( S_b \) is the between-class scatter and \( S_w \) is the within-class scatter. The column vectors \( \omega_i \) of the matrix \( W \) are obtained by solving the following generalized eigenvalue problem

\[ S_b \omega_i = \lambda_i S_w \omega_i \, . \] (5)

where the number of samples is not large enough compared with the dimensions of the data (in our setting the dimension of the Gabor jets), the within-class scatter \( S_w \) is degenerate, and not all vectors \( \omega_i \) are necessarily obtained accurately. To solve this problem, the within-class scatter \( S_w \) was replaced by the total scatter \( S_T \) in [12] and dimensional reduction with PCA is used before the discriminant analysis in [9]. Combining the above two in our experiments, dimensional reduction using PCA is first applied and the total scatter \( S_T \) is used instead of the within-class scatter \( S_w \).

### 2.5 Outline, novelty and expectation of the proposed method

The proposed method is based on the three algorithms mentioned above. To begin, the transformation that enhances robustness against pose variations is obtained. The transformation matrix is calculated based on multiple discriminant analysis, where not only the Gabor jets from the registration image, but also the estimated Gabor jets of different poses obtained through linear jet transformation are used for training data. With this transformation, the Gabor jets of the registered images are transformed to the coordinates where class (individual) separations are easier. With these coordinates, robustness is improved as pose variations are taken into consideration within each class. Given an input image, feature points are then detected using jet space similarity and the Gabor jets are sampled from the feature points. These Gabor jets are also transformed using the transformation matrix enhancing the robustness mentioned above. Finally, the correlation between the Gabor jet of the registered images and the input image are computed to determine similarity.

The novelty of this method is the use of the estimated local features from a single image as training-data for discriminant analysis. Discriminant analysis does not function with one image per person. Hence, the lack of data must be compensated for, and this is resolved by using the estimated training data from this single image. Estimated local features for other poses are subjected to training together with these single registration data.

Similar to the discriminant analysis trained using real data, the discriminant analysis trained from these estimated data is expected to improve the robustness against pose variation.

### 3. EXPERIMENTS

#### 3.1 Images for the experiments

To examine robustness against pose variations, experiments were performed using the human and object interaction processing (HOIP) database [13], which is a database of 300 subjects photographed from 504 (72 yaw angles every 5° and 7 roll angles every 15°) directions (Fig. 6).

In these facial images, the size of the ear approximately fits within a 70×90 pixel window. The feature points of the

---

Figure 6: Ear orientation defined by head pose
ear are located using jet space similarities as described in section 2.2. This process is illustrated in Fig. 7(a) with examples taken from the experiment. The left image illustrates coarse detection using sub-grids, the middle image illustrates super-imposing complementary sub-grids (gray dots) and right image illustrates fine tuning (white dots). Where detection has not been successful, as in Fig. 7(b), errors are manually corrected.

Ear images of a single subject taken from various angles are given in Fig. 8.

In Fig. 8, yaw angles 0°, 90°, and 180° correspond to frontal view, true left profile, and back view. Roll angle 0° corresponds to true left profile. Positive roll angles correspond to images from above.

3.2 Experimental method

We examined two types of image as registration images. The first is an almost true profile (yaw angle = 85°) typical of criminal mug shots. The second is with yaw angle 45°, following the suggestion in [2] (at yaw angles 30° and 45° all of the important tissue in ears are visible in all of the experimental data).

Input images are photographed at yaw angles ranging from 30° to 120°, every 10°. To mimic the shooting angle of a ceiling-attached surveillance camera, images of roll angle 15°, 30° and 45° are also used in addition to 0°. To imitate the low resolution of surveillance camera, images are resized from 10% to 90% and resized back to 100% using linear interpolation (Fig. 9).

To assess our proposed method in circumstances using estimated data for different poses as training data for discriminant analysis, robustness in the following three situations is examined and compared:
1. No training data for other poses are used;
2. Estimated data from other poses are used for training;
3. Real data for other poses are used for training.

Using training data (cases 2,3), other than registration data photographed from (Yaw, Roll) = (85°, 0°) or (45°, 0°), estimated data or real data photographed from (85°±5°,0°), (85°±15°,0°), (85°±5°,15°), (85°±15°,15°) or (45°±5°,0°), (45°±15°,0°), (45°±5°,15°), (45°±15°,15°) were used for training data for discriminant analysis.

Subjects with more than four (out of seven) visible feature points at all of the angles for input, registration and training data were selected. Hence, the number of subjects depends on the yaw angle and roll angle of the input and registration data (Table 1).

In Table 1, N/A signifies experiments not performed as there were less than 20 subjects. This table demonstrates the visibility of the feature points of the ear for the specified camera angles. The larger the number of subjects, the easier it is to find the feature points of the ear.

Accuracy is evaluated using the Rank-1 recognition rate for 1:N recognition and by using the Equal Error Rate (EER) for 1:1 verification as defined in [14].

To be precise, the Rank-1 recognition rate and EER depend on the number of registered subjects. Hence, a comparison between results from different angles may not be very accurate. Nevertheless, comparisons between the three methods are reasonable.

![Figure 7: Examples of ear detection](image1)

(a) Coarse detection, superimposing, and fine tuning

(b) False positives, negatives, and inaccurate tuning

Figure 7: Examples of ear detection

![Figure 8: Ear images taken from various angles](image2)

Figure 8: Ear images taken from various angles

![Figure 9: Example of ear images at various resolutions.](image3)

![Table 1 Number of subjects for each experiment.](image4)
4. EXPERIMENTAL RESULTS

4.1 Experiments at various roll angles

The EER and Rank-1 recognition rate of the three methods are displayed with various yaw-angle input images (Figs. 10 and 11) at each roll angle. Fig. 10 is a result where the pose of registration is \((\text{Yaw}, \text{Roll}) = (45^\circ, 0^\circ)\) following the suggestion in [2] and Fig. 11 corresponds to results for an almost true profile \((\text{Yaw}, \text{Roll}) = (85^\circ, 0^\circ)\).

From Figs. 10 and 11, we find that although the proposed algorithm using estimated data does not surpass the accuracy of the algorithm using real data, trends in

![Figure 10: EER and Rank-1 recognition rate at various angles when the registered image corresponds to \((\text{Yaw}, \text{Roll}) = (45^\circ, 0^\circ)\).](image)

![Figure 11: EER and Rank-1 recognition rate at various angles when the registered image corresponds to \((\text{Yaw}, \text{Roll}) = (85^\circ, 0^\circ)\).](image)
the improvement in robustness are similar.

4.2 Experiments at various resolutions.

The EER and Rank-1 recognition rate of the three methods are displayed with various yaw angled input images (Figs. 12 and 13) at each resolutions.

As for Figs. 10 and 11, we find similar conclusions from Figs. 12 and 13. More specifically, when resolutions are not so high, the improvement is as effective as real data.

![Graphs showing EER and Rank-1 recognition rate at various resolutions](image)

Figures 12(a)-(h) inclusive, the symbols that appear have the following significance:

- ■: no data for other poses
- ●: estimated data for other poses
- ■: real data for other poses

Figure 12: EER and Rank-1 recognition rate at various angles and at various resolutions when the registered image corresponds to (Yaw, Roll) = (45°, 0°).
5. DISCUSSIONS

5.1 Changes in resolution

Regarding resolution, accuracy trends were found to be similar, i.e. up to 20% although these worsen at 10% (Fig. 14). In these images, ears are 7×10 pixels. From Fig. 8, we find classifying scaphoid fossa very hard. Antihelices (angular type, parallel type or arched type [3]) are also not clear. Similar to the suggestion in [3], one criterion that might be applicable to the present method is the visibility of the antihelix or scaphoid fossa.

For Figures 13(a)-(h) inclusive, the symbols that appear have the following significance:

- - no data for other poses
- - estimated data for other poses
--- real data for other poses

Figure 13: EER and Rank-1 recognition rate at various angles and at various resolutions when the registered image corresponds to (Yaw, Roll) = (85°,0°).
5.2 Changes in shooting angles

In [2], a criterion to be able to compare two ear images taken at different camera angles was proposed; both images need to meet two conditions: the scapha must be clearly observed and the external acoustic meatus (Fig. 15) must not be visible.

It is reported that at 90°, 8.9%, at 105°, 17.9% and at 120°, 53.6% of subjects do not show a well-defined scapha because a part of (or all of) the antihelix is occluded by the helix. At 0° and 15°, the scapha of most subjects is not clearly observed because the helix is occluded by the antihelix. Scaphae are only clearly visible at yaw angles between 30° and 75°. At 75°, 42.9% and at 90°, all of the subjects show an external acoustic meatus.

Consistent with the findings in [2], Figs. 10 and 11 indicate that when the yaw angle for registration data are 45°, recognition accuracy in the direction +25° (70°) is better than that in the direction -25° (20°), where many scaphae are not visible. When the yaw angle for registration data are 85°, recognition accuracy in the direction -25° (60°) is better than that in the direction 25° (110°), as many scaphae are not visible and external acoustic meatus are visible.

Regarding changes in roll angles, accuracies are observed to be similar up to 15° for similar yaw angles, but worsen if greater than 30°.

Some of the results demonstrated in Fig. 16 are explained by the (in)visibility of scaphae and external acoustic meatus. However, we believe that accuracies around 30° are difficult to explain just by this argument. We think this can be explained by differences in the number of training angles; we used three times more data on yaw angle change than data on roll angle change owing to limitation in the HOIP database.

Strictly speaking, comparison between different angles may not be appropriate as the EER and rank-1 recognition rate depend on the number of subjects. However, given that better sides in accuracy have the larger number of subjects (Table 1), and that increasing the number of subjects does not increase the performance (selecting correct answers from a larger number of choices is more difficult than from a smaller number of choices), the authors leave the above discussion open until a more rigorous demonstration can be given in the future.

6. CONCLUSIONS

To improve the viability of our single-view-based ear recognition system [10, 11], thorough empirical evaluations were performed in scenarios where camera angles and resolutions are more realistic than those of [10, 11]. Experimental results indicated that the proposed method
has functioned effectively in surveillance-like scenarios where subjects are photographed from above and image resolutions are low.

ACKNOWLEDGEMENTS

Some ear images were taken from HOIP facial database provided by Softopia Japan Foundation with permission. It is strictly prohibited to copy, re-use, or distribute the facial data without permission. This work was supported by KAKENHI 22700219 and 80337609.

REFERENCES


Appendix

Multiplying both sides of Eq. (2) by $\Psi_k^{-1}(u)$ and integrating both sides results in

$$\sum_k c_{kk} \langle A \rangle \int \Psi_k^{-1}(u) \Psi_k(u) du = \int \Psi_k(x) \Psi_k^{-1}(A^{-1}x) dx.$$

Using the following notation,

$$\langle k'|k \rangle = \int \Psi_k(x) \Psi_k^{-1}(A^{-1}x) dx$$

we have

$$\sum_k \langle k'|k \rangle c_{kk} = \langle k|A^{-1}|k \rangle.$$

Rearranging,

$$\begin{bmatrix} \langle k_1|k_1 \rangle & \cdots & \langle k_1|k_N \rangle \\ \vdots & \ddots & \vdots \\ \langle k_N|k_1 \rangle & \cdots & \langle k_N|k_N \rangle \end{bmatrix} \begin{bmatrix} c_{kk} \end{bmatrix} = \begin{bmatrix} \langle k_1|A^{-1}|k_1 \rangle \\ \vdots \\ \langle k_N|A^{-1}|k_N \rangle \end{bmatrix}$$

Denoting the matrix on the left hand side by $T$ and re-arranging,

$$T \begin{bmatrix} c_{kk} \\ c_{kk} \end{bmatrix} = \begin{bmatrix} \langle k_1|A^{-1}|k_1 \rangle \\ \vdots \\ \langle k_N|A^{-1}|k_N \rangle \end{bmatrix}$$

Writing the above as $T = C^T S$, the transformation matrix $C$ for Eq. (3) is given by

$$C = S^T T^{-1}.$$

256
Since $\mathbf{T}, \mathbf{S}$ are determined from Eq. (5), integration of Eq. (5) is possible. It is sufficient to integrate $\langle \mathbf{k}(\mathbf{M}) \mathbf{k}^\dagger \rangle$, for an arbitrary $2 \times 2$ matrix $\mathbf{M} = (a_{ij})$. The function to be integrated is expanded as follows:

$$
\frac{k^2}{\sigma^2} \exp \left( \frac{k^2}{2\sigma^2} \left[ (x^2 + y^2)^2 \right] \right) \cdot \exp \left( i \left[ k_1^* x + k_2^* y \right] \right) - \exp \left( -\sigma^2 / 2 \right)
$$

The function to be integrated is expanded as follows:

$$
\frac{k^2}{\sigma^2} \exp \left( \frac{k^2}{2\sigma^2} \left[ \left( a_{11} x + a_{12} y \right)^2 + \left( a_{21} x + a_{22} y \right)^2 \right] \right)
\times \exp \left( -i \left( k_1 (a_{11} x + a_{12} y) + k_2 (a_{21} x + a_{22} y) \right) \right) \cdot \exp \left( -\sigma^2 / 2 \right)
$$

Integration of Eq. (5) is analytically obtained using the following equality:

$$
\int \exp \left( -ax^2 - by^2 - cxy - dxy \right) dx dy = \frac{2\pi}{\sqrt{4ab-c^2}} \exp \left( \frac{ab^2-cgh+bg^2}{4ab-c^2} \right)
$$

Daishi WATABE
Daishi Watabe is an associate professor of Graduate School of Engineering at Saitama Institute of Technology. He received his M.S. and D.S. degrees from Tohoku University, Japan, in 1998 and 2000, respectively. His research interests include biometrics, statistical machine learning, and numerical analysis. He is a member of the Institute of Electronic Engineers of Japan, the Institute of Image and Television Engineers, and IEEE.

Yang WANG
Yang Wang is a graduate student at the Saitama Institute of Technology, Japan. She received her B.E. degree from University of Science and Technology Liaoning, China, in 2010. Her research interests include biometrics and image processing. She is engaged in graph theory, the application of systems simulation, and soft computing research.

Takanari MINAMIDANI
Takanari Minamidani is a graduate student at the Saitama Institute of Technology, Japan. He received his B.E. degree from Saitama Institute of Technology, Japan, in 2010. His research interests include biometrics and image processing.

Hideyasu SAI
Hideyasu Sai is a researcher supporter of advanced science research laboratory at Saitama Institute of Technology. He received his B.E., M.E., and D.E degrees from Saitama Institute of Technology, Japan, in 2002, 2007, and 2010, respectively. He is currently a graduate student at the institute. His research interests include biometrics and image processing. He is a member of the Institute of Electronic Engineers of Japan and the Institute of Electronics, Information and Communication Engineers.

Takanari MINAMIDANI
Takanari Minamidani is a graduate student at the Saitama Institute of Technology, Japan. He received his B.E. degree from Saitama Institute of Technology, Japan, in 2010. His research interests include biometrics and image processing.

Katsuhiro SAKAI
Katsuhiro Sakai is a professor emeritus at Saitama Institute of Technology, Japan. He received his B.E. and M.S. degrees from Tohoku University, Japan, in 1967 and 1970, respectively. He obtained his D.E. degree from Osaka University, Japan. His research interest includes computational fluid dynamics, especially stable numerical schemes with higher-order accuracy and environment engineering, specifically wind turbines. He is a member of The Japan Society of Fluid Mechanics.

Osamu NAKAMURA
Osamu Nakamura is a professor at the Department of Information and Communication Engineering and the graduate school of Kogakuin University, Japan. He received his B.E. degree in Electronics in 1969, M.E. degree in Electrical Engineering in 1971, and D.E. degree in 1983, all from Kogakuin University. Since 1971, he has been working on learning theory, image processing, and picture coding. From 1989 to 1990, he was an invited professor at INRS-telecommunications, Montreal, Canada. He is a member of the Institute of Electronic Engineers of Japan and IEEE.