Electronic Preprint for Journal of Information Processing Vol.26

Regular Paper

Personal Identification by Integrating a Number of Features from Iris and Periocular Region Using AdaBoost

Shintaro Oishi[1,a)] Yoshihiro Shirakawa[1] Masatsugu Ichino[1] Hiroshi Yoshiura[1]

Received: October 16, 2017, Accepted: March 6, 2018

Abstract: A personal identification method has been developed for searching for a specific person among other people that fuses iris and periocular features using AdaBoost. It effectively integrates scores for many features using an AdaBoost configuration in which feature selection corresponds to weak classifier selection. We found three interesting facts of evaluation. First, evaluation using up to eight features showed that identification accuracy increased with the number of features used. The lowest equal error rate (EER) was 1.3% when eight features were used, and the highest identification rate was 94.1% when eight features were used. Second, the advantage of the proposed method over a weighted sum method increased with the number of features used. The difference in EER was 1.1% when eight features were used due to the generation of a nonlinear decision boundary, and the difference in the identification rate was 1.8% when eight features were used, again due to the generation of a nonlinear decision boundary. Finally, using an effective combination of information from both eyes further improved the accuracy (the difference in EER between the four-feature case and the eight-feature case was 0.7%, and the difference in identification rate between the four-feature case and the eight-feature case was 4.6%).

Keywords: Iris, Periocular region, Personal Identification

1. Introduction

Biometrics based on physical or behavioral characteristics has the advantage of not requiring people to carry anything around or remember passwords and can be used to identify people in a wide range of fields. In particular, iris biometrics has a very high level of accuracy. However, it requires a high-quality iris image, which means that it has to be captured at a very short distance, which is highly intrusive on the person being identified and increases the burden and limits the action of users. A less intrusive approach would be to capture iris images with identification equipment situated further away, but this is liable to result in poor-quality images and lower identification accuracy.

A method that combines iris biometrics with periocular biometrics (identification based on images of the area surrounding the eye) has been proposed. Images of the iris and the region surrounding the eye are obtained simultaneously from a greater distance such as about 2 to 3 meters, and periocular biometrics is used to compensate for the loss of iris biometrics accuracy that results from acquiring images from a greater distance. This increased distance reduces the burden on the person to be identified.

Previous research on the fusion of iris and periocular features falls short in three ways. First, many features can be used for recognition. In iris biometrics, Daugman's iris code is mainly used for recognition. However, periocular biometrics can use many features, and the features used determine the effectiveness of recognition. Each periocular feature has particular advantages and disadvantages. For example, the local binary pattern (LBP) feature is robust against lighting variation but fragile against geometric transformation. Previous research did not provide a systematic method to find the effective combination of many features. Second, to the best of our knowledge, scores from both sides of the face (those from left iris and left periocular region and those from right iris and right periocular region) have not been integrated. It should be possible to improve identification performance by using the best combination of features from both sides. Third, a weighted sum is often used to integrate the iris and periocular feature scores. However, the greater the number of features used, the more complicated the decision boundary on the score map. Weighted sum simply creates a linear decision boundary. We have developed a personal identification method that overcomes these shortfalls and is thus well suited for searching for a specific person among other people. It is based on the fusion of iris and periocular features using AdaBoost [1].

The rest of this paper is organized as follows. In Section 2, we summarize related work, and in Section 3, we describe our proposed feature fusion method. We describe our evaluation in Section 4 and present the results in Section 5. In Section 6, we discuss the results of our evaluation, and in Section 7, we summarize the key points and mention future work.

2. Related Work

2.1 Iris Biometrics

Daugman's iris code is commonly used in iris biometrics. However, it requires a high resolution image of an iris. The lower the image resolution, the lower the recognition accuracy [2], [3].

[1] The University of Electro-Communications, Chofu, Tokyo 182–8585, Japan
[a)] s-oishi@uec.ac.jp

2.2 Periocular Biometrics

In periocular biometrics, a variety of features are used for recognition. **Table 1** summarizes the features used for periocular biometrics. A variety of features have been used for a variety of data sets. However, the features that are effective for recognition have not been determined. Moreover, each feature has particular advantages and disadvantages. For example, LBP is robust against lighting variation but fragile against geometric transformation while scale invariant feature transformation (SIFT) is invariant to scale and rotation. The features used are experimentally selected by the researchers.

2.3 Fusion of Iris and Periocular Features

As shown in Table 1, a weighted sum is often used for fusing the iris and periocular features [4], [5], [6], [7], [8], [9], [10], [11]. The iris score and periocular score are represented as S_{iris} and $S_{periocular}$. They are min-max normalized using

$$x_{new} = \frac{x_{old} - x_{\min}}{x_{\max} - x_{\min}}, \tag{1}$$

where $x_{\max}$ is the maximum value in score set x, $x_{\min}$ is the minimum value in score set x, x_{old} is the score before normalization, and x_{new} is the score after normalization. The basic integration score S is calculated using weights ω_1 and ω_2 for the iris score and the periocular score:

$$S = \omega_1 S_{iris} + \omega_2 S_{periocular}, \qquad \omega_1 + \omega_2 = 1. \tag{2}$$

The performance is optimized by adjusting ω_1 and ω_2, for example, in the range 0.1 to 0.9 [6]. The performance is better with fusion than with either iris or periocular modality alone [4], [5], [6]. Tan et al. reported an identification rate for the iris features of 55% [4], for the periocular region features of 60%, and for fusion of the iris and periocular features of 80%.

2.4 Problems with Previous Fusion Methods

As mentioned above, a variety of features have been used in periocular biometrics. The recognition algorithm should automatically select features that are effective in the combination with the features from iris. It should also automatically select features that are effective for the target data set, which includes images affected by various factors, for example lighting conditions and geometric distortion. If a weighted sum is used in the recognition algorithm, the features must be selected in advance.

If both sides of the face are included in the image, more features must be dealt with, which increases the number of dimensions of the feature vectors. The greater the number of dimensions, the more complicated the distribution of data for the genuine and imposter. This results in a nonlinear boundary on the score map to distinguish the genuine and imposter. A weighted sum, on the other hand, simply creates a linear decision boundary, so the distinction between the genuine and imposter may be more difficult.

Table 1 Summary of related work.

Reference	Dataset	Periocular features	Fusion method of iris and periocular
[12] [13]	FRGC2.0	LBP HOG SIFT	NA
[14]	UBIRIS.v2	GIST CLBP	NA
[15]	FRGC FERET	LBP+GEFE	NA
[16]	FRGC	LBP+LBP Walsh Masks+LBP Law's Masks+LBP DCT+LBP DWT+LBP Force Field Transform+LBP GaborFilter+LBP Log Filter+LBP SIFT, SURF	NA
[17]	FOCS	HOG m-SIFT PDM	NA
[18]	CASIA v4 -distance	LBP m-SIFT Local phase feature	NA
[19]	One's own data	LBP HOG SIFT	NA
[20]	One's own data	LBP HOG SIFT	NA
[21]	Transgender dataset	TPLBP LBP HOG SIFT	NA
[22]	One's own data	LBP HOG SIFT	NA
[4] [5]	UBIRIS v2 FRGC CASIA v4 -distance	LBP HOG DSIFT GIST LMF	Weighted Sum
[6]	MBGC	LBP	Weighted Sum
[7]	FOCS	Optimal Trade-off Synthetic Discriminant Function (OTSDF) Correlation Filter	Weighted Sum
[8]	FOCS	LBP SIFT Gabor Wavelets	Weighted Sum
[9]	One's own data	LBP+SRC	Weighted Sum
[10]	One's own data	BSIF+SRC	Weighted Sum
[11]	One's own data	PerioCode	Weighted Sum

NA in fusion method column means recognition using only periocular features.

3. Proposed Method

3.1 Requirements

Three requirements were identified for our method. First, it must select effective features automatically because the periocular regions and irises from both sides of a face have many features, each of which has particular advantages and disadvantages for a target data set. Second, it must distinguish data for the genuine from data for imposter even if the data has a complex distribution in a high-dimensional vector space. Third, the tradeoff between recognition accuracy and speed should be controllable to make the method suitable for practical use.

3.2 Using AdaBoost

To meet these requirements, we use the AdaBoost algorithm [1], which is the most frequently used method for boosting classification performance. It does this by using combinations of weak classifiers. AdaBoost is effective for our purpose because it has following merits.

- AdaBoost can create non-linear or piecewise linear decision boundaries, which meets the second requirement.
- AdaBoost can reduce the number of errors exponentially by increasing the number of weak classifiers used, thus controlling the trade-off between recognition accuracy and recognition speed by setting the number of weak classifiers, which meets the third requirement.

In a typical AdaBoost configuration, each weak classifier uses a feature vector that consists of all features. Each weak classifier must therefore use all features and cannot select effective features. We use a different configuration-each classifier uses a feature vector consisting of only one feature. This configuration enables automatic selection of effective features by selecting the corresponding classifiers, which meets the first requirement.

3.3 Methods of Training and Testing

[Training]

B is the number of boosting rounds. J is the number of features used.

(1) Initialization of weights

The number of imposter data samples is much larger than that of genuine data samples. Sample weights $D_1(t)$ are used to treat the genuine and imposter data equally:

$$D_1(t) = \begin{cases} 1/\text{number of genuine data samples} \\ \text{(for genuine data sample weight)} \\ 1/\text{number of imposter data samples} \\ \text{(for imposter data sample weight)} \end{cases} \quad (3)$$

The sum of the weights is normalized to 1 ($\sum_{t=1}^{M} D_1(t) = 1$). M represents the number of training data samples.

(2) Training of weak classifier

i. Steps **ii**. to **vi.** are processed for each round $b = 1, 2, \cdots, B$.

ii. The following process is done for the features $j = 1$ to $j = J$.

For training data samples $t = 1$ to $t = M$, identification is performed using the t-th data sample x_t. The j-th weak classifier wc_{bj} is calculated using x_t. The error rate ε_{bj} of the training samples is calculated using

$$\varepsilon_{bj} = \sum_{t:y_t \neq wc_{bj}(x_t)} D_b(t), \quad (4)$$

where y_t is the correct class.

The weak classifier is designed using

$$wc_{bj}(x_t) = \begin{cases} +1 & \text{if } p \cdot FV_j(x_t) > p \cdot \theta \\ -1 & \text{otherwise,} \end{cases} \quad (5)$$

where $FV_j(x_t)$ is the value of the j-th feature for the t-th sample, and p is a variable that determines the inequality direction by comparing the feature value and the threshold θ (p takes a value of $+1$ or -1). The p can be used to change the inequality orientation by using the position relationship of the genuine samples and the imposter samples in the feature space. If $p = 1$, $wc_{bj}(x_t) = +1$ when $FV_j(x_t) > \theta$. And $wc_{bj}(x_t) = -1$ when $FV_j(x_t) \leq \theta$. If $p = -1$, $wc_{bj}(x_t) = +1$ when $FV_j(x_t) \leq \theta$. And $wc_{bj}(x_t) = -1$ when $FV_j(x_t) > \theta$. The θ and p are calculated such that the error rate is minimum.

iii. Weak classifer $h_b = wc_{bq}$ and corresponding feature q are selected such that

$$h_b = \underset{wc_{bj}}{\operatorname{argmin}}\, \varepsilon_{bj}, \qquad (1 \leq j \leq J). \quad (6)$$

That is, $h_b(x_t)$ is selected such that the error rate (ε_{bj} given by Eq. (4)) is minimum.

q is the feature selected such that the error rate ε_{bj} is minimum in J features. q is the feature selected in each round of AdaBoost training.

iv. When feature q is selected, the error rate is put in ε_{bq}. Next, α_b is calculated using

$$\alpha_b = \frac{1}{2} \log\left(\frac{1 - \varepsilon_{bq}}{\varepsilon_{bq}}\right). \quad (7)$$

α_b is weight (degree of confidence) of each feature (weak classifier).

v. The sample weight is renewed using

$$D_{b+1}(t) = D_b(t) \exp\left(-\alpha_t y_i h_b(x_t)\right). \quad (8)$$

For a correctly identified sample, $D_{b+1}(t) = D_b(t) \exp(-\alpha_t)$. For an incorrectly identified sample, $D_{b+1}(t) = D_b(t) \exp(\alpha_t)$.

vi. The weights of the samples are normalized so that the sum of the weights is 1.

(3) Construction of strong classifier

The strong classifier $H(x_t)$ is given by

$$H(x_t) = \sum_{b=1}^{B} \alpha_b h_b(x_t). \quad (9)$$

[Testing]

Testing is done using the strong classifier, which uses $h_b(x_t)$ and α_b calculated in the training phase.

4. Evaluation

4.1 Overview

An overview of the experiment used for evaluation is shown in **Fig. 1**. We first extracted iris and periocular region features from images of left and right eyes. We then used iris code calculated using Daugman's algorithm for the iris biometrics. We used the LBP, histogram of oriented gradients (HOG), and SIFT features for the periocular biometrics because they are widely used for periocular biometrics. We used them to calculate the similarity scores between two images. Next, we performed AdaBoost training and created a strong classifier. We then evaluated the identification accuracy by using the test data. Finally, we compared the proposed method with a weighted sum method (performed using min-max normalization).

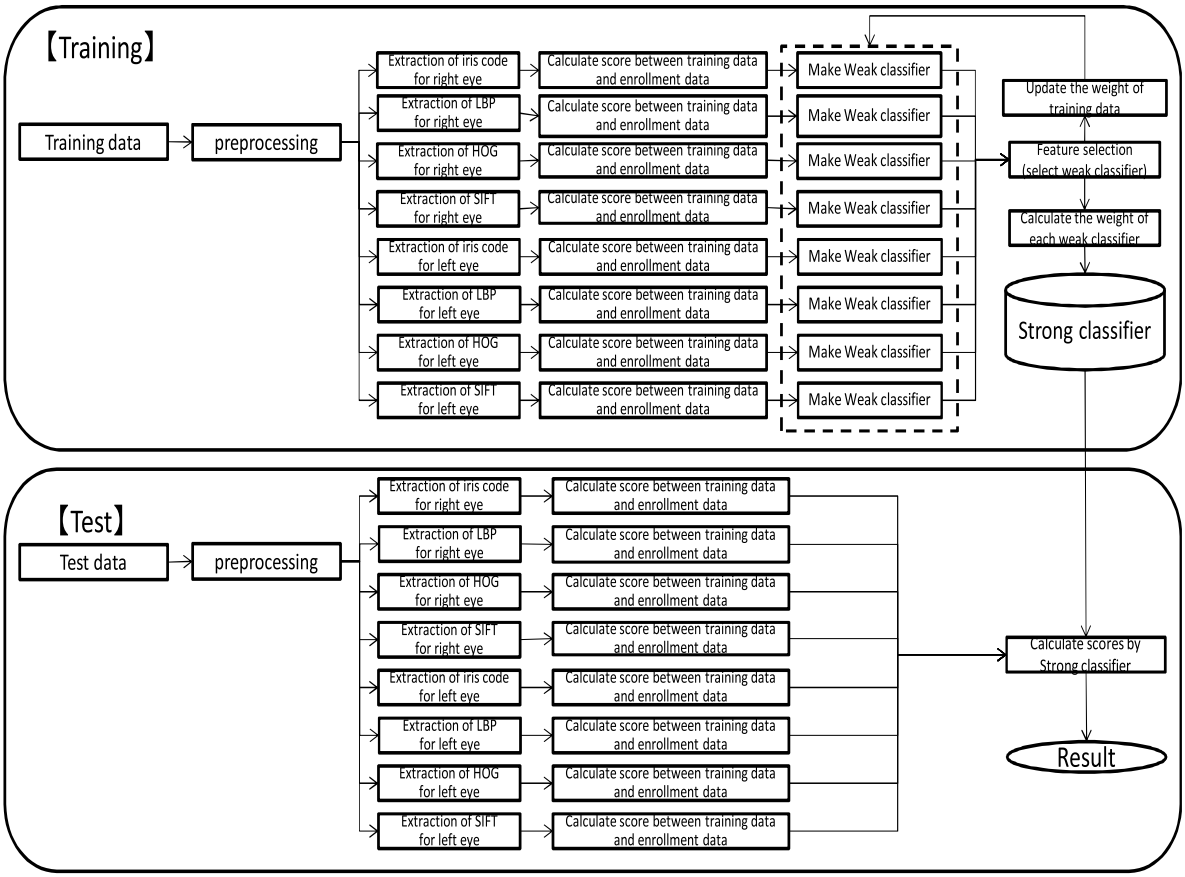


Fig. 1 Outline of experiment.

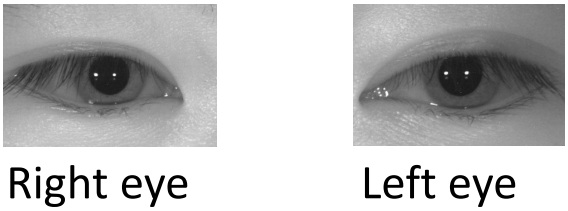


Fig. 2 Example preprocessed images.

4.2 Preprocessing

We detected the left and right eye areas by using the OpenCV. The pupil of each eye was detected using the Hough transform. The center of the pupil was automatically relocated to the center of the image, and the image was clipped to a size of 600×400 pixels. Example preprocessed images are shown in **Fig. 2**.

4.3 Iris Biometrics

For Daugman's iris recognition algorithm, we used OSIRIS V.4.1 [23]. The matching score was obtained by calculating the Hamming distance.

4.4 Periocular Biometrics

We normalized the images by referring to Tan's method [4] while maintaining the iris radius.

For periocular identification, we normalized each preprocessed image. First, scale factor $S_f = r_{norm}/r_{iris}$ was obtained; r_{iris} is the iris radius, and r_{norm} is the radius after normalization. Each image was normalized using the bicubic method to achieve an image size of S_f (600×400 pixels). As a result of this normalization, the iris radius was equal to r_{norm}. Then, the center of the iris was shifted to the center of the image, and the image size was cropped again to $6r_{norm} \times 4r_{iris}$ [4]. An r_{norm} value of 64 was used in accordance with the size of the polar coordinate display images in OSIRIS. Therefore, the size of a normalized image was 384×256.

Next we extracted the LBP, HOG and SIFT features as the features to be used. The advantages of these features are summarized as follows.

- LBP
  LBP is robust against lighting variation. We obtained the LBP for the entire image after applying a Gaussian filter with a standard deviation $\sigma = 4.0$ to the image. The LBP pixels were then partitioned into 96 blocks of $0.5r_{norm} \times 0.5r_{norm}$ [12], [13]. Next, a brightness histogram was calculated for each block and quantized into eight values. These 8 values were then concatenated into a 768-dimensional feature vector. The matching score was obtained by calculating the Manhattan distances.
- HOG
  HOG is robust against lighting variation and geometric transformation. We obtained the HOG for the entire image after applying a Gaussian filter with a standard deviation $\sigma = 4.0$. The HOG pixels were then partitioned into 96 cells of $0.5r_{norm} \times 0.5r_{norm}$. We normalized the image size with a block of 3×3 cells by shifting one cell. The number of normalizations based on this block size was 10×6

Table 2 Combinations of features.

No	Eye	Features	Total no. of features
1	either	iris (iris code) and periocular (LBP)	2
2	either	iris (iris code) and periocular (HOG)	2
3	either	iris (iris code) and periocular (SIFT)	2
4	either	iris (iris code) and periocular (LBP, HOG, SIFT)	4
5	both	iris (iris code) and periocular (LBP, HOG, SIFT)	8

either:either right eye or left eye; both: both right eye and left eye

= 60. Next, the normalized cells were concatenated into a 4860-dimensional feature vector. The matching score was obtained by calculating the Manhattan distances.

- SIFT

  SIFT is invariant to scale and rotation. The preprocessed images were subjected to histogram equalization to emphasize the contrast, and the SIFT key points of each image were detected. The SIFT feature extraction and description of OpenCV were used to calculate the scores. However, some key points were corresponded incorrectly. We compared the Euclidian distance $dist1$ between the closest key points to the Euclidian distance $dist2$ between the second-closest key points. The key points that did not meet the following condition were removed.

  $dist1 < dist2 \times 0.75$

  We corresponded the only nearest key point to meet above requirement. The matching score S was obtained using $S = N \times \frac{1}{\frac{\sum_{j=1}^{N} d_j}{N}}$, where N represents the number of corresponding points between the images and d_j represents the distance between the j-th corresponding points between the images.

4.5 Combinations of Features

We evaluated five combinations of features, as shown in **Table 2**.

4.6 Experimental Dataset

We obtained eye images from the CASIA-Iris-Distance database [24]. The images had been taken at a distance of 3 meters using a near-infrared camera. We used images for 122 people. All images were of the naked eye (without eyeglasses). For each person, we used eight images (four for training and four for testing). The training and testing datasets thus each contained 488 images (122 × 4). The number of matching pairs in each dataset was ${}_{488}C_2$ = 118,828 (genuine - genuine: ${}_4C_2 \times 122$ people = 732; genuine - imposter: 118,096). We calculated the iris and periocular matching scores of these pairs. Then we applied the proposed method and the weighted sum method. We performed this experiment 70 (${}_8C_4$) times with different training and testing images and averaged the results.

5. Results

Table 3 shows the equal error rate (EER) and identification rate for each feature for the proposed method. The EER is the value when the False Accept Rate (FAR) and False Reject Rate (FRR) are equal. The EER and receiver operating characteristic (ROC) curve show the accuracy of 1:1 matching (matching between input and a particular template), and the identification rate shows

Table 3 EER and identification rate for each feature.

Feature		EER (%)		Identification rate (%)	
		left	right	left	right
iris		7.6	7.5	60.6	62.1
periocular	LBP	8.0	9.3	70.1	65.3
	HOG	9.1	8.4	60.7	60.3
	SIFT	6.6	6.0	65.7	70.0

that of 1:N matching (matching between input and all templates). The identification rate is the percentage of instances in which a person is correctly classified as belonging to the input person's template.

$$\text{Identification rate} = \frac{\text{number of person data classified correctly}}{\text{total number of person data}} \times 100[\%] \quad (10)$$

Since our objective here is to present a recognition method integrating many features from both eyes by using AdaBoost and to demonstrate its effectiveness quantitatively, we evaluated its effectiveness by comparing its EER with those when two features (lines 1, 2, and 3 in **Table 4**) and four features (line 4 in Table 4) were used. We also compared its EER with the EERs when the weighted sum was used as the algorithm for integrating the score.

Table 4 shows the EER and identification rate for each feature combination for both the proposed method using AdaBoost and the weighted sum method. The EERs when two features (lines 1, 2, and 3 in Table 4) were used were better than when one feature (Table 3) was used. The EER when four features (line 4 in Table 4) were used was better than when two features (lines 1, 2, and 3 in Table 4) were used. The EER when eight features (line 5 in Table 4) were used was better than when four features (line 4 in Table 4) were used and when two features (lines 1, 2 and 3 in Table 4) were used. The EERs when AdaBoost and two features were used were slightly better than when the weighted sum and two features were used. This advantage became larger when four features were used. It became still larger when eight features were used. The EER when AdaBoost and eight features were used was the lowest in Table 4. An EER of 1.3% when AdaBoost and eight features were used is roughly equivalent to the EERs of other approaches [29] and reasonable value here.

Since our objective here is to present a recognition method integrating many features from both eyes by using AdaBoost and to demonstrate its effectiveness quantitatively, we evaluated its effectiveness by comparing its identification rate with those when two features (lines 1, 2, and 3 in Table 4) and four features (line 4 in Table 4) were used. We also compared its identification rate with the identification rates when the weighted sum was used as the algorithm for integrating the score.

The identification rates when two features (lines 1, 2, and 3 in Table 4) were used were better than when one feature (Table 3)

Table 4 EER and identification rate for each feature combination for proposed and weighted sum methods.

Features combination	EER (%)				Identification rate (%)			
	AdaBoost		Weighted sum		AdaBoost		Weighted sum	
	left	right	left	right	left	right	left	right
1. iris+periocular region (LBP)	3.9	3.8	4.2	4.2	80.1	79.5	82.4	80.9
2. iris+periocular region (HOG)	4.7	4.1	4.8	4.2	76.4	77.9	78.0	79.8
3. iris+periocular region (SIFT)	2.8	2.6	2.7	2.4	82.5	84.5	86.1	88.0
4. iris+periocular region (LBP,HOG,SIFT)	2.2	2.0	2.7	2.4	89.5	89.2	91.0	89.9
5. iris+periocular region (LBP, HOG, SIFT) [left+right]	1.3		2.4		94.1		92.3	

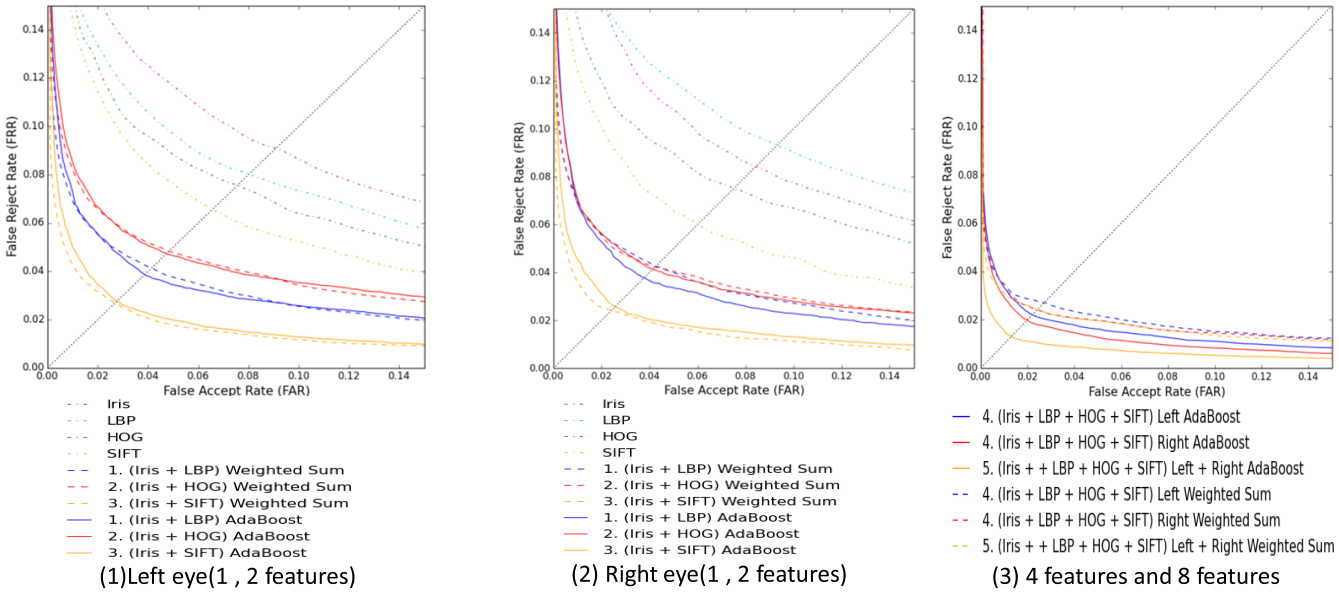


Fig. 3 ROC curves.

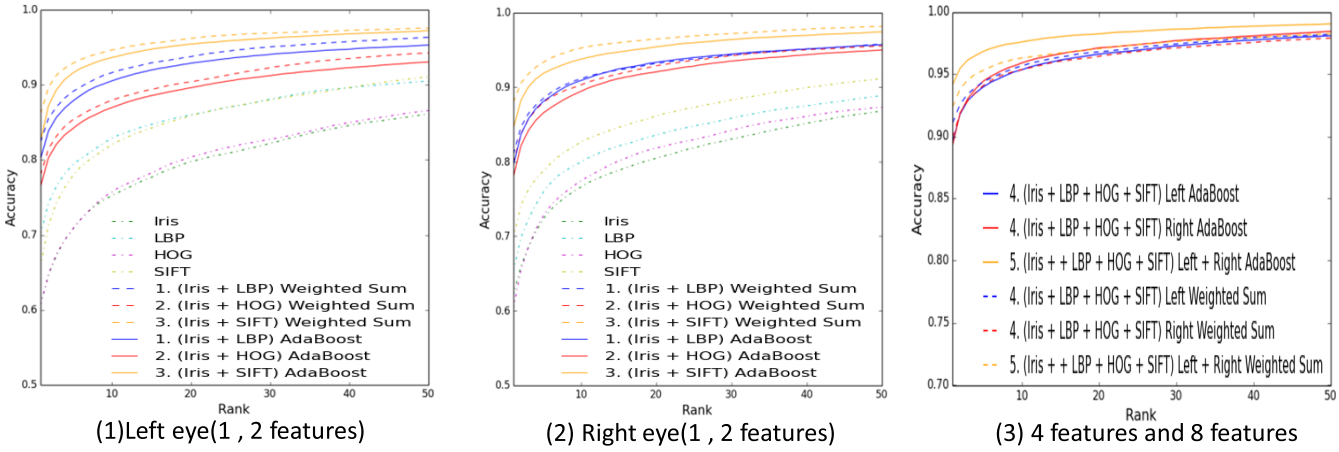


Fig. 4 CMR curves.

was used. The identification rate when four features (line 4 in Table 4) were used was better than when two features (lines 1, 2, and 3 in Table 4) were used. The identification rate when eight features (line 5 in Table 4) were used was better than when four features (line 4 in Table 4) were used and when two features (lines 1, 2, and 3 in Table 4) were used. The identification rate when AdaBoost and eight features were used was better than when the weighted sum and eight features were used. The identification rate when AdaBoost and eight features were used was the highest in Table 4.

Figure 3 shows the ROC curves. The vertical axis represents the FRR, and the horizontal axis represents the FAR. The performance of the proposed method when eight features (line 5 in Table 4) were used was better than when two or four features were used. It was also better than that of the weighted sum method for all combinations.

Figure 4 shows the cumulative match characteristic curves. The vertical axis represents the cumulative match rate (CMR), and the horizontal axis represents the number of candidates (rank). Given that the person is genuine for a range of candidate rankings up to rank N (an integer equal to or greater than 1), a cumulative match characteristic curve represents recognitions deemed successful. For example, N = 1 means an identification rate as shown in Tables 3 and 4 while N = 2 means that not only

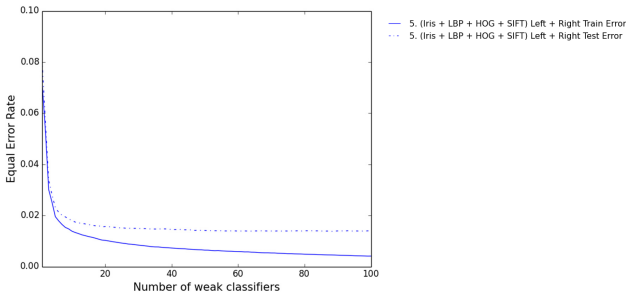


Fig. 5 Training and testing errors.

the first-ranked candidate but also the second-ranked candidate are considered to be the genuine. The performance of the proposed method when eight features (line 5 in Table 4) were used was better than when two or four features were used. It was also better than that of the weighted sum method for all combinations.

The training and testing errors when using eight features are plotted in **Fig. 5**. The vertical axis represents the EER, and the horizontal axis represents the number of weak classifiers. A solid line represents the training error and a dotted line represents the testing error. The figure shows that recognition time and recognition accuracy can be controlled flexibly in accordance with the application because recognition accuracy can be controlled by adjusting the number of weak classifiers.

6. Discussion

6.1 Features Selected by AdaBoost

The feature corresponding to the weak classifier with the lowest error rate is selected in each round of AdaBoost training. The features selected by AdaBoost in each round when eight features were used are listed in **Table 5**. The features for which the single-feature recognition accuracy (see Table 3) was high were preferentially selected. The weights of the features selected in the early rounds were larger than those selected in the later rounds. This indicates that automatic feature selection using AdaBoost is an effective approach. For fusion of the left eye and right eye features, the features extracted from the left and right images were selected interchangeably. This means that it is effective to fuse the left and right eye features because the features extracted from the left and right images were used to improve recognition accuracy.

6.2 Decision Boundary

We investigated the decision boundaries generated by the AdaBoost and weighted sum methods. The score map and decision boundary for the iris and periocular region (LBP feature) are shown in **Fig. 6**. The vertical axis represents the scores output by the classifier using the LBP feature of the periocular region. The horizontal axis represents the scores output by the classifier using iris feature. The blue points represent the scores for the genuine, and the red points represent those for imposter. The decision boundaries are shown for the EER. The distributions of the scores overlap, and the decision boundary between genuine and imposter is nonlinear. However, the decision boundary for the weighted sum method is linear while that for AdaBoost is nonlinear. This means that the decision boundary for AdaBoost more clearly identifies the scores for the genuine. The proposed method is thus better than a weighted sum method for identifying the genuine.

Table 5 Features selected by AdaBoost.

Round	Feature selected	Feature weight (α)
1	right SIFT	1.39
2	left iris	1.05
3	right iris	0.97
4	left SIFT	0.80
5	right HOG	0.62
6	left LBP	0.46
7	right SIFT	0.47
8	left SIFT	0.44
9	right iris	0.41
10	left SIFT	0.40

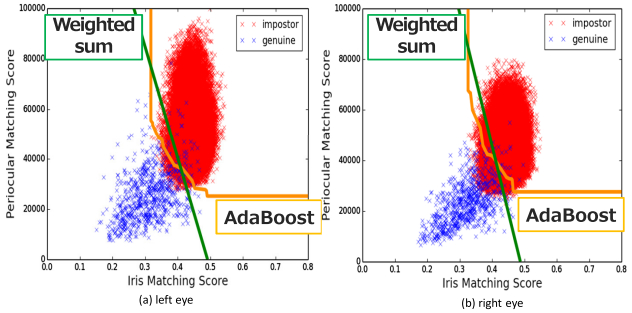


Fig. 6 Score map and decision boundary for iris and periocular region (LBP feature).

6.3 Factors Affecting Iris and Periocular Recognition

In this subsection, we describe the factors to affect iris and periocular recognition.

Iris recognition can be affected by contact lenses. Recently, the use of contact lens becomes more prevalent. While contact lenses are generally used to correct eyesight as a replacement for glasses, they are increasingly being used for cosmetic reasons. For example, the texture and color of the iris region can be changed by wearing a thin textured lens. A study on the effect of color contact lenses on the performance of iris recognition [25] showed that the use of a color cosmetic lens changes the appearance and texture of the eye in both the visible and near-infrared spectrum. This suggests that both transparent (prescription) and color cosmetic lenses (textured) significantly affect recognition accuracy.

A method proposed for reducing the effect of color contact lens [25] uses a lens detection algorithm to first reject cases in which there are obfuscated iris patterns. Only cases in which there is no lens or there is only a transparent lens are accepted. This method improves iris recognition performance. Lens detection algorithms were proposed [26]: one uses iris edge sharpness and the other uses textural features based on a co-occurrence matrix.

Periocular recognition can be affected by make-up, which has many variations. These variations include differences in color, texture, and application method. Cosmetics are commonly used to cover facial flaws and to make the wearer look more attractive. Unlike digital enhancement, make-up actually changes the wearer's physical appearance. A previous study [27] showed that make-up was listed as misleading for other features. When a sub-

ject changed the make-up, recognition was harder.

Mitigating the effect of make-up is essential. Make-up can be used to impersonate someone else or change one's appearance. Face recognition must thus take into account the effect of make-up. Hu et al. proposed using canonical correlation analysis (CCA) to mitigate the effect of make-up by learning the meta subspace [28], which can maximize the correlation of feature vectors belonging to an individual. This approach should be helpful in mitigating the effect of make-up in periocular recognition.

6.4 Additional Experiment

6.4.1 Evaluation and Results

To provide more evidence for the effectiveness of the proposed method, we conducted the same experiments using the CASIA-Iris-Lamp dataset.

Using the experimental processing described in Section 4, we obtained human eye images from the CASIA-Iris-Lamp database [24]. Multiple images of each person had been taken using a near-infrared camera with a lamp close to the person either turned on or turned off. We used images of the naked eye (without eyeglasses) for 408 people. For each person, we used eight images (four for training and four for testing). The training and testing datasets thus each contained 1,632 images (408 × 4). The number of matching pairs in each dataset was ${}_{1632}C_2$ = 1,330,896 (genuine - genuine: ${}_4C_2 \times 408$ people = 2,448; genuine - imposter: 1,328,448). We calculated the iris and periocular matching scores for these pairs and then applied the proposed method and the weighted sum method. We performed this experiment 20 times with different training and testing images and averaged the results.

Table 6 shows the EER and identification rate for each feature for the proposed method. **Table 7** shows the EER and identification rate for each feature combination for both the proposed method using AdaBoost and the weighted sum method. The EERs when AdaBoost and two features (lines 1, 2, and 3 in Table 7) were used were better than when one feature (Table 6) was used. The EER when AdaBoost and four features (line 4 in Table 7) were used was better than when AdaBoost and two features (lines 1, 2, and 3 in Table 7) were used. The EER when AdaBoost and eight features (line 5 in Table 7) were used was better than when AdaBoost and four features (line 4 in Table 7) were used and when AdaBoost and two features (lines 1, 2, and 3 in Table 7) were used. The EERs when AdaBoost and two features were used were slightly better than when the weighted sum and two features were used. This advantage became larger when four features were used. It became still larger when eight features were used. The EER when AdaBoost and eight features were used was the lowest in Table 7. We think that LBP and HOG provided a poor performance in the additional experiment because images from the CASIA-Iris-Lamp database had been taken using a near-infrared camera.

The identification rate when AdaBoost and two features (lines 1, 2, and 3 in Table 7) were used were better than when one feature (Table 6) was used. The identification rate when AdaBoost and four features (line 4 in Table 7) were used was better than when AdaBoost and two features (lines 1, 2, and 3 in Table 7) were used. The identification rate when AdaBoost and eight features (line 5 in Table 7) were used was better than when AdaBoost and four features (line 4 in Table 7) were used and when AdaBoost and two features (lines 1, 2, and 3 in Table 7) were used. The identification rate when AdaBoost and eight features were used was better than when the weighted sum and eight features were used. The identification rate when AdaBoost and eight features were used was the highest in Table 7.

Figure 7 shows the ROC curves. The vertical axis represents the FRR, and the horizontal axis represents the FAR. The performance of the proposed method when eight features (line 5 in Table 7) were used was better than when two or four features were used. It was also better than that of the weighted sum method for all combinations.

Figure 8 shows the cumulative match characteristic curves. The vertical axis represents the CMR (accuracy), and the horizontal axis represents the number of candidates (rank). The performance of the proposed method when eight features (line 5 in Table 7) were used was better than when two or four features were used. It was also better than that of the weighted sum method for all combinations.

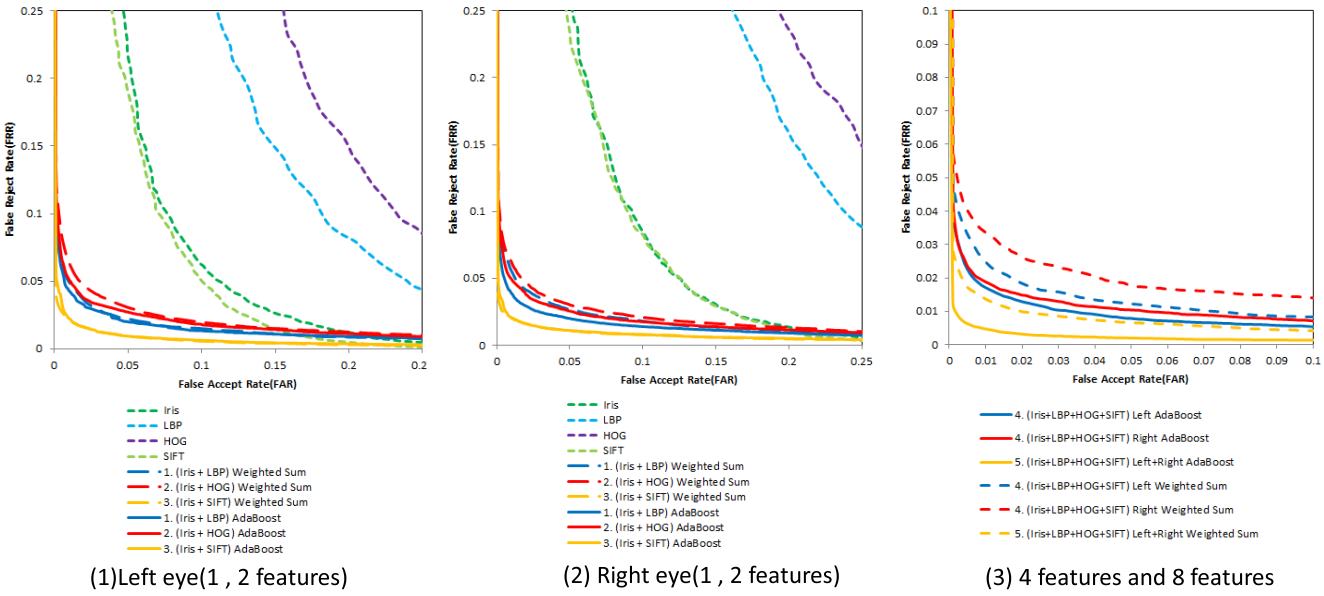
The training and testing errors when using eight features are plotted in **Fig. 9**. The vertical axis represents the EER, and the horizontal axis represents the number of weak classifiers. A solid line represents the training error and a dotted line represents the testing error. The figure shows that recognition time and recognition accuracy can be controlled flexibly in accordance with the application because the recognition accuracy can be controlled by adjusting the number of weak classifiers.

Table 6 EER and identification rate for each feature (CASIA-Iris-Lamp).

Feature		EER (%)		Identification rate (%)	
		left	right	left	right
iris		8.5	9.4	61.6	58.9
periocular	LBP	15.0	18.8	28.8	24.0
	HOG	17.8	21.2	23.5	18.5
	SIFT	8.0	9.2	56.4	49.4

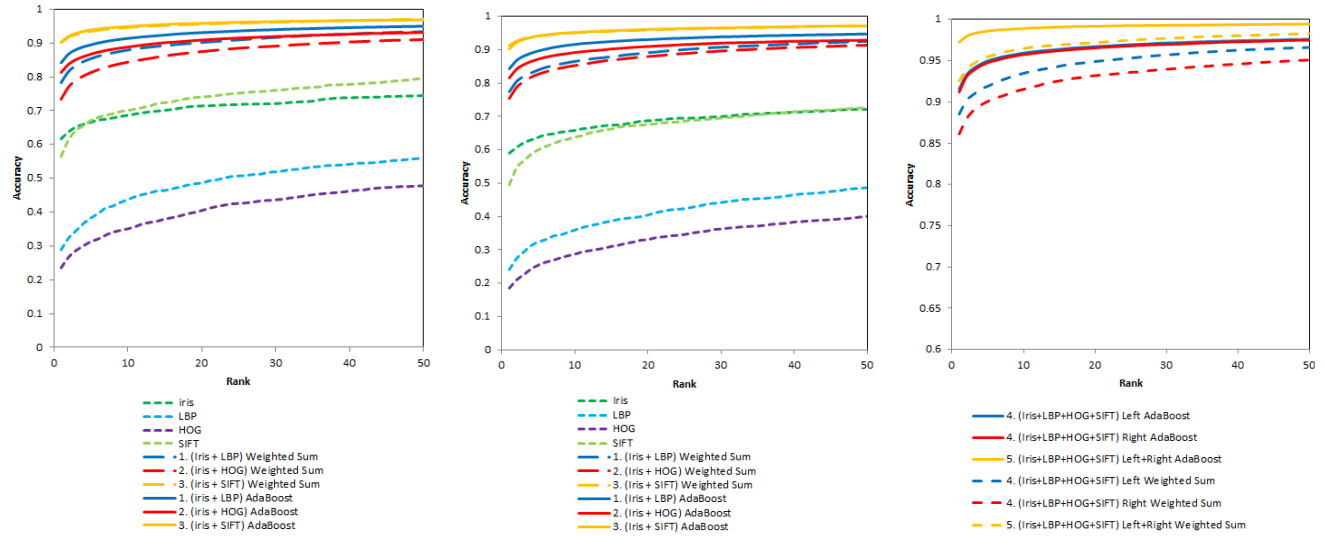
Table 7 EER and identification rate for each feature combination for proposed and weighted sum methods (CASIA-Iris-Lamp).

Features combination	EER (%)				Identification rate (%)			
	AdaBoost		Weighted sum		AdaBoost		Weighted sum	
	left	right	left	right	left	right	left	right
1. iris+periocular region (LBP)	2.8	2.6	2.9	3.3	84.1	84.2	78.2	77.3
2. iris+periocular region (HOG)	3.2	3.1	3.6	3.5	81.3	81.5	72.2	75.3
3. iris+periocular region (SIFT)	1.7	1.7	1.7	1.7	90.2	90.2	90.1	91.2
4. iris+periocular region (LBP, HOG, SIFT)	1.6	1.6	1.9	2.4	91.5	91.2	88.5	86.1
5. iris+periocular region (LBP, HOG, SIFT) [left+right]	0.6		1.2		97.2		92.5	



(1)Left eye(1 , 2 features) (2) Right eye(1 , 2 features) (3) 4 features and 8 features

Fig. 7 ROC curves (CASIA-Iris-Lamp).



(1)Left eye(1 , 2 features) (2) Right eye(1 , 2 features) (3) 4 features and 8 features

Fig. 8 CMR curves (CASIA-Iris-Lamp).

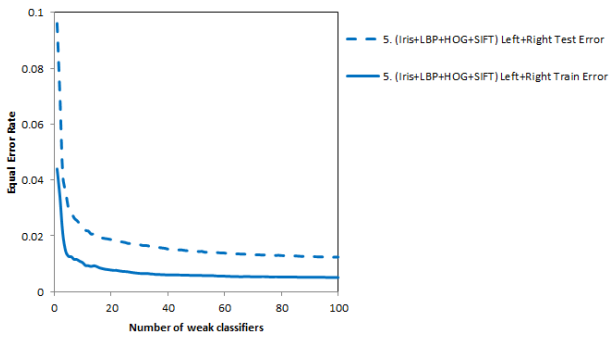


Fig. 9 Training and testing errors (CASIA-Iris-Lamp).

Table 8 Features selected by AdaBoost (CASIA-Iris-Lamp).

Round	Feature selected	Feature weight (α)
1	left SIFT	1.26
2	right iris	1.06
3	left iris	0.77
4	right SIFT	0.64
5	left LBP	0.50
6	right SIFT	0.32
7	left iris	0.31
8	right LBP	0.29
9	right HOG	0.29
10	left SIFT	0.28

6.4.2 Effectiveness of Applying AdaBoost

The feature corresponding to the weak classifier with the lowest error rate is selected in each round of AdaBoost training. The features selected by AdaBoost when eight features were used are listed in **Table 8**. The features for which the single-feature recognition accuracy (see Table 6) was high were preferentially selected. The weights of the features selected in the early rounds were larger than those selected in the later rounds. This indicates that automatic feature selection using AdaBoost is an effective approach. For fusion of the left eye and right eye features, the features extracted from the left and right images were selected

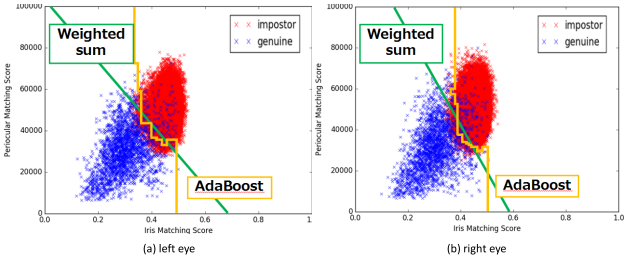


Fig. 10 Score map and decision boundary (CASIA-Iris-Lamp).

interchangeably. This improved EER and indentification rate, meaning that fusing the left and right eye features is an effective approach.

We investigated the decision boundaries generated by the AdaBoost and weighted sum methods. The score map and decision boundary for the iris and periocular region (LBP feature) are shown in **Fig. 10**. The vertical axis represents the scores output by the classifier using the LBP feature of the periocular region. The horizontal axis represents the matching scores output by the classifier using the iris feature. The blue points represent the scores for the genuine images, and the red points represent those for the imposter images. The decision boundaries are shown for the EER. The distributions of the scores overlap, and the decision boundary between genuine and imposter is nonlinear. However, the decision boundary for the weighted sum method is linear while that for AdaBoost is nonlinear. This means that the decision boundary for AdaBoost more clearly identifies the scores for the genuine images. The proposed method is thus better than a weighted sum method for identifying genuine images.

7. Conclusion

Our proposed identification method for searching for a specific person among other people is based on fusion of the iris and periocular features using AdaBoost. Effective features are automatically selected by defining a feature vector for each weak classifier that consists of only one feature. This enables effective features to be selected by selecting effective classifiers. Evaluation using eight features (iris_left, iris_right, periocular region (LBP_left, LBP_right, HOG_left, HOG_right, SIFT_left, SIFT_right)) in comparison with a method that uses a weighted sum to integrate the scores revealed three interesting results. First, the more features that were used, the more precise the identification due to selecting effective features. The lowest EER was 1.3% for eight features, and the highest identification rate was 94.1% for eight features. Second, the more features that were used, the better the relative performance of the proposed method (difference in EER of 1.1% for eight features; difference in identification rate of 1.8% for eight features) due to the generation of a nonlinear boundary. Third, using an effective combination of information from both eyes improved the accuracy (difference in EER between four-feature case and eight-feature case was 0.7%; difference in identification rate between four-feature case and eight-feature case was 4.6%). Experiments conducted using the CASIA-Iris-Lamp dataset provided more evidence for the effectiveness of the proposed method.

Future work includes evaluating the performance when the distance between the person and camera is variable.

Acknowledgments This work was supported by JSPS KAKENHI Grant Number 16K18103.

References

[1] Freund, Y. and Schapire, R.E.: A decision theoretic generalization of on-line learning and an application to boosting, *Journal of Computer and System Science*, Vol.55, No.1, pp.119–139 (1997).
[2] Umeoka, Y., Kaneko, N., Saito, T. and Sumi, K.: A relation between Resolution and Recognition Rate in Iris Recognition, IEICE Technical Report, Vol.BioX2014-5, pp.25–30 (2014). (in Japanese)
[3] Matey, J.R., Naroditsky, O., Hanna, K., Kolczynski, R., LoIacono, D., Mangru, S., Tinker, M., Zappia, T. and Zhao, W.Y.: Iris on the Move: Acquisition of Images for Iris Recognition in Less Constrained Environments, *Proc. IEEE*, Vol.94, pp.1936–1947 (2006).
[4] Tan, C.W. and Kumar, A.: Human identification from at-a-distance images by simultaneously exploiting iris and periocular features, *International Conference on Pattern Recognition 2012*, pp.553–556 (2012).
[5] Tan, C.W. and Kumar, A.: Towards online iris and periocular recognition under relaxed imaging constrains, *IEEE Trans. Image Processing*, Vol.22, pp.3751–3765 (2013).
[6] Woodard, D., Pundlik, S., Miller, P., Jillela, R. and Ross, A.: On the fusion of periocular and iris biometrics in non-ideal imagery, *International Conference on Pattern Recognition 2010*, pp.201–204 (2010).
[7] Boddeti, V.N., Smereka, J.M. and Kumar, B.V.K.V.: A comparatice evaluation of iris and ocular recognition methods on challenging ocular images, *International Joint Conference on Biometrics 2011*, pp.1–8 (2011)
[8] Klontz, J. and Burge, M.J.: Periocular recognition from low-quality iris images, *Springer Handbook of Iris Recognition*, pp.309–319 (2013).
[9] Raghavendra, R., Raja, K.B., Yang, B. and Busch, C.: Combining iris and periocular recognition using light field camera, *2013 2nd IAPR Asian Conference on Pattern Recognition*, pp.155–159 (2013).
[10] Raja, K.B., Raghavendra, R. and Busch, C.: Binarized statistical features for improved iris and periocular recognition in visible spectrum, *2014 International Workshop on Biometrics and Forensics* (*IWBF*), pp.1–6 (2014).
[11] Xiao, L., Sun, Z. and Tan, T.: Fusion of iris and periocular biometrics for cross-sensor identification, *Biometric Recognition Lecture Notes in Computer Science*, Vol.7701, pp.202–209 (2012).
[12] Park, U., Jillela, R., Ross, A. and Jain, A.: Periocular Biometrics in the visible spectrum: A feasibility study, *Biometrics: Theory, Applications and Systems 2009*, pp.1–6 (2009).
[13] Park, U., Jillela, R., Ross, A. and Jain, A.: Periocular biometrics in the visible spectrum, *IEEE Trans. Information Forensics and Security*, Vol.6, pp.96–106 (2011).
[14] Bharadwaj, S., Bhatt, H., Vatsa, M. and Singh, R.: Periocular biometrics : When iris recognition fails, *Biometrics: Theory, Applications and Systems 2010*, pp.1–6 (2010).
[15] Simpson, L., Dozier, G., Adams, J., Woodard, D., Miller, P., Bryant K. and Glenn, G.: Genetic amp; evolutioonary type ii feature extraction for Periocular -based biometric recognition, *IEEE Congress on Evolutionary Computation* (*CEC*), pp.1–4 (2010).
[16] Juefei-Xu, F., Cha, M., Heyman, J., Venugopalan, S., Abiantun, R. and Savvides, M.: Robust local binary pattern feature sets for Periocular biometric identification, *Biometrics: Theory, Applications and Systems 2010*, pp.1–8 (2010).
[17] Ross, A., Jillela, R., Smereka, J.M., Boddeti, V.N. and Kumar, B.V.K.V., Barnard, R., Xiaofei, H., Pauca, P. and Plemmons, R.: Matching highly non-ideal ocular images: An information fusion approach, *2012 5th IAPR International Conference on Biometrics*, pp.446–453 (2012).
[18] Aoyama, S., Kusanaga, D., Ito, K. and Aoki, T.: A Study on Matching Ocular Images Using Local Phase Features, *The 31st Symposium on Cryptography and Information Security*, pp.1–7 (2014). (in Japanese)
[19] Hollingsworth, K.P., Darnell S.S., Miller, P.E., Woodard, D.L., Bowyer, K.W. and Flynn, P.J.: Human and machine performance on periocular biometrics under near-infrared light and visible light, *IEEE Trans. Information Forensics and Security*, Vol.7, pp.588–601 (2012).
[20] Padole, C.N. and Proenca, H.: Periocular recognition: Analysis of performance degradation factors, *International Conference on Biometrics 2012*, pp.439–445 (2012).
[21] Mahalingam, C. and Ricanek, Jr., K.: Is the eye region more reliable than the face? A preliminary study of face-based recognition on a transgender dataset, *Biometrics: Theory, Applications and Systems 2013*, pp.1–7 (2013).
[22] Mary, P.F.G., Paul, P.S.K. and Dheeba, J.: Human identification using

periocular biometrics, *International Journal of Science, Engineering and Technology Research 2013*, Vol.2, pp.1047–1053 (2013).
[23] OSIRIS: OSIRIS version 4.1 (2013), available from ⟨http://svnext.it-sudparis.eu/svnview2-eph/ref_syst//Iris_Osiris_v4.1/⟩.
[24] CASIA: CASIA Iris Image Database V4.0, available from ⟨http://www.idealtest.org/dbDetailForUser.do?id=4⟩.
[25] Kohli, N., Yadav, D., Vatsa, M. and Singh, R.: Revisiting Iris Recognition with Color Cosmetic Contact Lenses, *International Conference on Biometrics 2013*, pp.1–7 (2013).
[26] Wei, Z., Qiu, X., Sun, Z. and Tan, T.: Counterfeit iris detection based on texture analysis, *International Conference on Pattern Recognition 2008*, pp.1–4 (2008).
[27] Hollingsworth, K., Bowyer, K.W. and Flynn, P.J.: Identifying Useful Features for Recognition in Near-Infrared Periocular Images, *International Conference on Biometrics: Theory Applications and Systems 2010*, pp.1–8 (2010).
[28] Hu, J., Ge, Y., Lu, J. and Feng, X.: Makeup-robust face verification, *International Conference on Acoustics, Speech and Signal Processing 2013*, pp.2342–2346 (2013).
[29] Nguyen, K., Fookes, C., Jillela, R., Sridharan, S. and Ross, A.: Long range iris recognition: A survey, *Pattern Recogn.*, Vol.72, pp.123–143 (2017).

Shintaro Oishi received his B.E. degree in engineering from the University of Electro-Communications, Tokyo, Japan in 2013 and M.E. in the Graduate School of Electro-Communications, the University of Electro-Communications, Tokyo, Japan in 2015.

Yoshihiro Shirakawa received his B.E. degree in engineering from the University of Electro-Communications, Tokyo, Japan in 2016. He is a master's degree student at the Graduate School of Informatics and Engineering, the University of Electro-Communications. His research interests include biometrics.

Masatsugu Ichino received his B.E. degree in electronics, information and communication engineering from Waseda University, Tokyo, Japan in 2003, and M.E. and Ph.D. degrees in computer science and engineering from Waseda University, Tokyo, Japan, in 2005 and 2008, respectively. He is currently an Associate Professor at the Graduate School of Informatics and Engineering, the University of Electro-Communications, Tokyo, Japan. His research interests include biometrics, network security and pattern recognition. He is a member of IEICE and IPSJ.

Hiroshi Yoshiura received his B.S. and D.Sc. from University of Tokyo, Japan in 1981 and 1997. He is currently a Professor in the Graduate School of Informatics, the University of Electro-Communications. Before joining UEC, he had been at Systems Development Laboratory, Hitachi, Ltd. He has been engaged in research on information security and privacy and a member of Japan Society of Security Management, Information Processing Society of Japan, The Institute of Electronics, Information and Communication Engineers, The Japanese Society for Artificial Intelligence, The Institute of Systems, Control and Information Engineers, and The Institute of Electrical and Electronics Engineers. He is an IPSJ fellow.