Disguised-Face Discriminator for Embedded Systems

Woo-han Yun, DoHyung Kim, Ho-Sub Yoon, and Jaeyeon Lee

In this paper, we introduce an improved adaptive boosting (AdaBoost) classifier and its application, a disguised-face discriminator that discriminates between bare and disguised faces. The proposed classifier is based on an AdaBoost learning algorithm and regression technique. In the process, the lookup table of AdaBoost learning is utilized. The proposed method is verified on the captured images under several real environments. Experimental results and analysis show the proposed method has a higher and faster performance than other well-known methods.

Keywords: Disguised face, discriminator, AdaBoost.

Manuscript received Mar. 15, 2010; accepted June 10, 2010.

This work was supported by the R&D program of the Korea Ministry of Knowledge and Economy (MKE) and the Korea Evaluation Institute of Industrial Technology (KEIT) [Project KI001813, Development of HRI Solutions and Core Chipsets for u-Robot].

Woo-han Yun (phone: +82 42 860 5804, email: yochin@etri.re.kr), DoHyung Kim (email: dhkim008@etri.re.kr), Ho-Sub Yoon (email: yoonhs@etri.re.kr), and Jaeyeon Lee (email: leejy@etri.re.kr) are with the IT Convergence Technology Research Laboratory, ETRI, Daejeon, Rep. of Korea.

doi:10.4218/etrij.10.1510.0139

I. Introduction

In the security field, disguised-face discrimination is helpful for improving the overall safety of a system, such as that in an ATM machine. It is also important for the automatic face identification used in a door lock system. Disguised-face discrimination, which occurs prior to the main face identification, could improve the robustness of the overall identification system because it could filter out disguised faces and pass bare faces into the main process.

Essentially, we are concerned with a simple classification problem between disguised and bare faces. We can apply several well-known feature extraction and classification methods, such as principal component analysis (PCA), Fisher discriminant analysis (FDA), and support vector machine (SVM) [1]-[3]. In previous research, a discrimination technique using several face-component detectors based on a revised modified census transform and AdaBoost algorithm was proposed [4].

Another factor to be considered is the complexity or processing time of the algorithm. A disguised-face discrimination algorithm is sometimes loaded on embedded systems such as wall-pads. Thus, the algorithm has to work with low complexity and high performance.

In this paper, we present an adaptive boosting (AdaBoost) classifier that is improved by regression techniques. For training and testing, an input query image is processed with modified census transform (MCT) preprocessing [5]. The feature vector of the query image is extracted using the lookup table stemmed from the AdaBoost learning. The score of the input query is drawn from the feature vector using a robust linear regression method. The input query image is determined to be a normal or disguised face by comparing the score with a predefined threshold. Figure 1 illustrates the overview of the algorithm. The detailed process is described in section II.

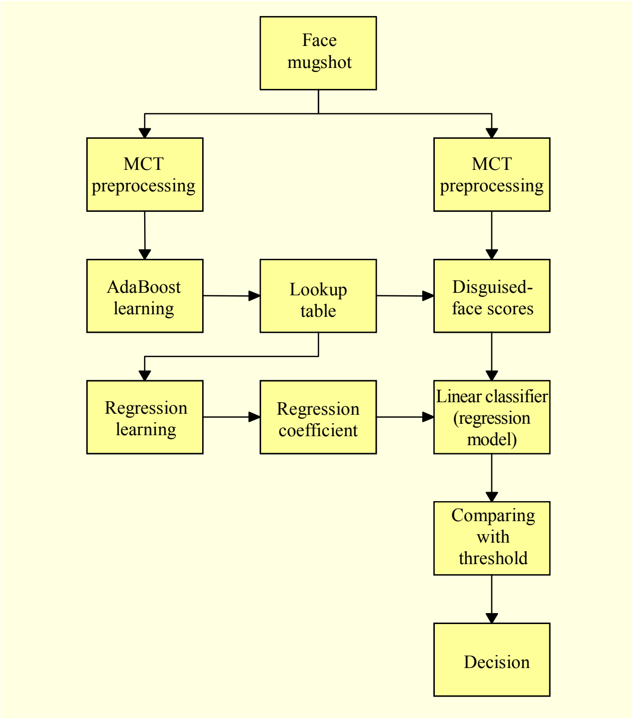

Fig. 1. Overall process of the approach.

II. Approach

1. AdaBoost-Based Classifier

Together, a modified MCT and an AdaBoost algorithm create a lookup table. The modified MCT algorithm is used for preprocessing and the function of the AdaBoost learning algorithm is self explanatory. MCT is a non-parametric local transform used to obtain a local feature from a local patch of an image. This feature shows the local spatial information of oriented edges, junctions, and line segments. This feature has proven its performance for face detection and recognition [5], [6]. The AdaBoost algorithm, formulated by Freund and Schapire, has proven its performance in object detection in studies by Jones and Viola [7], [8]. Froba and Ernst improved a face detector with an MCT in the preprocessing method [5]. Differing from other object detection methods, we adopt only one cascade step. Overall, the earlier process is similar to the face detection we used.

What follows is the approach in detail. An input mugshot is processed with MCT to extract robust features for illumination. After the MCT process, the value of each pixel ranges from 0 to 511 [5]. Then, we train the AdaBoost-based classifier. The training procedure is described in detail in [5]. Let $v(x, y)$ be the value of the (x, y) position in the MCT feature. In t iterations of the AdaBoost learning step, we obtain the best position (x_t, y_t) and lookup table $L(x_t, y_t, v)$, $v \in \{0, \ldots, 511\}$. After all the T iterations of the algorithm, we obtain the best positions $\{(x_1, y_1)\ldots(x_T, y_T)\}$ and lookup table $L(x, y, v)$ with a size of $T \times 512$ and threshold TH. We could consider the lookup table L as a transfer function $L(x, y, v)$ from the MCT space to a score space. The strong classifier H is the combination of transfer function L.

$$H(I) = \begin{cases} 1, & \sum_{t=1}^{T} L(x_t, y_t, v_t) < TH, \\ 0, & \text{otherwise,} \end{cases} \tag{1}$$

where v_t is the value of position (x_t, y_t) of the MCT feature and I is an input query mugshot.

Instead of the classification result of classifier $H(I)$, we use the extracted scores from a lookup table in the next step for constructing a stronger classifier.

2. From Lookup Table to Linear Classifier

From the lookup table, we extract the scores of the input query face mugshot and construct the score vector $V = [L(x_1, y_1, v_1), \ldots, L(x_T, y_T, v_T)]^{\mathrm{T}}$. We transfer the vector into a single scalar score with a robust linear regression model. The coefficients $\{\alpha_0 \ldots \alpha_T\}$ of the regression model were learned in the training phase with a training data set [9]. The final score is obtained with a combination of regression coefficients and the score vector of the lookup table using

$$s(I) = \alpha_0 + \sum_{t=1}^{T} \alpha_t \times L(x_t, y_t, v_t). \tag{2}$$

Figure 2 shows the lookup table scores and weighted lookup table scores of bare and disguised faces of the sample queries. We normalize the scores into values of 0 to 255 to represent the gray scale. In the figure, the bright pixels represent high scores, and dark pixels represent low scores. The bare face has higher scores in the mouth region than the disguised face with a mask. The ambiguous scores with gray pixels in the lookup table (center in the figure) are altered into more definite scores in the weighted lookup table using a regression model (right in the figure). The bare face has higher scores in the weighted lookup table than in the lookup table. The disguised face has lower scores in the weighted lookup table than in the lookup table.

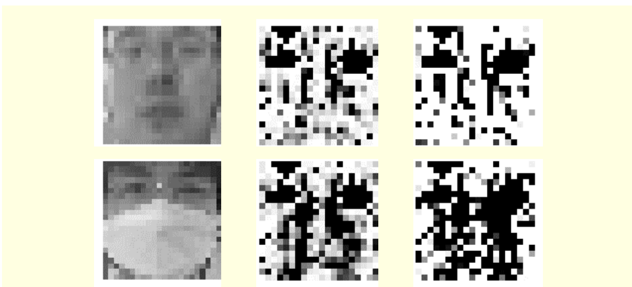

Fig. 2. Lookup table scores (center) and weighted lookup table scores (right) of a bare face (upper left) and disguised face (lower left) of the sample query images.

3. Decision

To discriminate an input query mugshot, the final score is compared with a predefined threshold TH_{final}, and the input query is determined to be a bare or disguised face.

In the experiment, we train the linear regression model with zero for disguised faces and one for bare faces. Then, we obtain the weighted combination result ranging from zero to one. The more similar the input query mugshot is with a disguised face, the more likely the result of linear regression is to be zero. In our case, we set the predefined threshold TH_{final} at 0.5.

III. Experimental Results

1. Database and Experimental Settings

There is no popular disguised-face database, so we acquired a database under various environments to validate the proposed algorithm. We collected negative images which include disguised faces with various masks or caps. The total number of images in the negative set includes 10,496 photos of 5 persons with 7 kinds of masks and 2 kinds of caps. Figure 3 shows the sample images of the negative set.

We also collected a positive set that includes 27,961 images of five persons with bare faces in various environments. The sample images in the positive set are in shown in Fig. 4. To balance the number of positive and negative images, we randomly sampled 13,231 positive images and 10,496 negative images. We chose 30% of them for the training set, and the others were used for the test set. We detected the faces using a revised MCT (RMCT) and AdaBoost-based face detector [10] and resized the cropped-face mugshots to 22 × 22 pixels.

We tested the algorithm on a desktop computer with an Intel 2.67 GHz CPU and 3.0 GB of RAM. The number of trained features in the classifier based on AdaBoost is 350.

Fig. 3. Sample images in a negative set. Several images contain backlighting over the frontal face images. Some masks have face-like patterns.

Fig. 4. Sample images in the positive set.

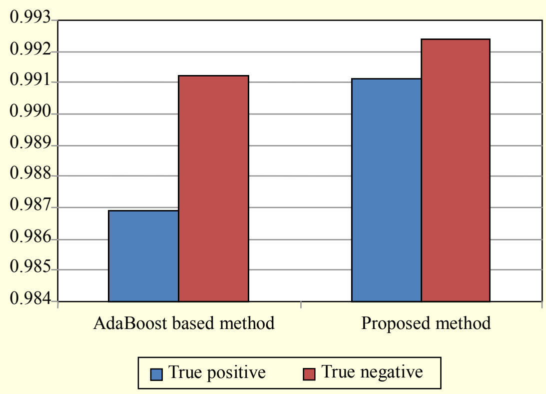

Fig. 5. True positive and true negative rates of AdaBoost-based method and proposed method.

2. Results and Analysis

We measured the classification rate of the AdaBoost-based classifier and the proposed method. The results of the comparison are shown in Fig. 5. The lookup table and regression-based classifier (right) have higher classification rates than the ordinary classifier (left). This result comes from the postprocessor regression method. In Fig. 2, the weighted lookup table scores (right) show more definite results than the ordinary lookup table scores (center). The weights of robust linear regression make the scores into more definite results.

We also compared the proposed method with other well-known classifiers such as PCA/FDA feature extraction and kNN classifier, FDA + SVM with RBF kernel, and decision tree. The results are shown in Fig. 6. In the comparison, we selected a reduced dimension of PCA for preservation of 95% of the data. The proposed method shows 99.11% true positive rates and 99.24% true negative rates. These are better classification rates than other methods we compared. The classifiers using FDA (FDA+kNN, FDA+SVM) reported lower rates than others because we trained FDA with a two-class problem (bare and disguised faces) in spite the negative database having various accessories including caps and masks.

We also tested the effectiveness of the feature scores from the lookup table in the proposed classifier. In Fig. 7, the comparison results of the overall methods are improved over the previous results in Fig. 6. The results show that the lookup

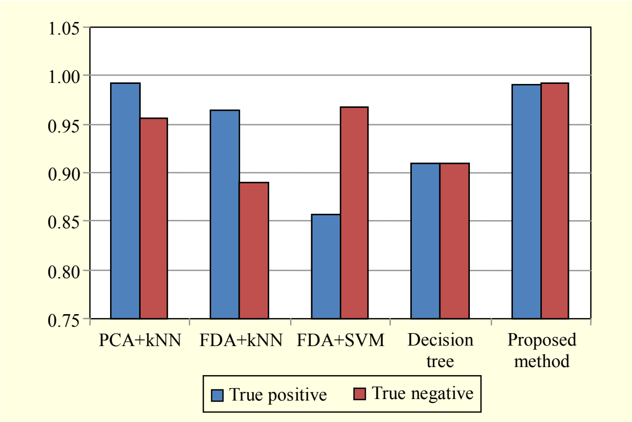

Fig. 6. Comparison results of proposed method and other classifiers.

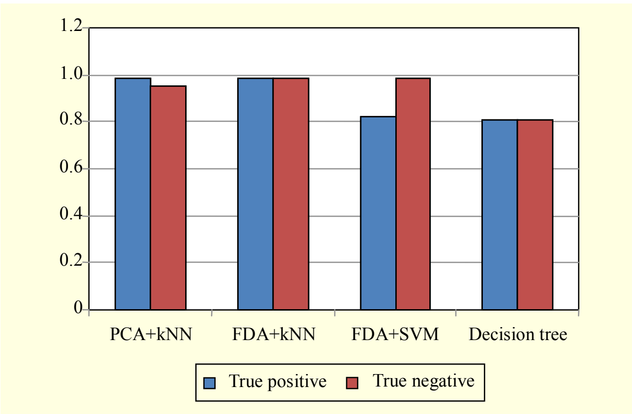

Fig. 7. Results of classifiers using lookup table scores.

table score, trained using AdaBoost, has effectiveness for classification. The rates of FDA+kNN, which show 96.38% as true positives and 89.07% as true negatives in Fig. 6, are improved to 98.19% and 98.37%, respectively, in Fig. 7. The use of lookup table scores is helpful for classification performance.

3. Processing Time and Implementation on Embedded System

In a real application, the speed of the algorithm is a crucial factor for application in embedded systems. We compared the speed of the proposed algorithms with other methods in Fig. 8. A decision tree and the proposed classifier needed processing times under 0.1 ms for one mugshot. The overall results show that the proposed method is better than other classifiers in classification rates and processing times.

We implemented this method and a face detector [10] on an embedded system with an ARM 11 CPU. The embedded system performed the face detection and discrimination at about 4 frames per second for a 320 × 240 sized image. In the processing time, the face detector and disguised-face discriminator occupied 99% and 1% of the resources, respectively.

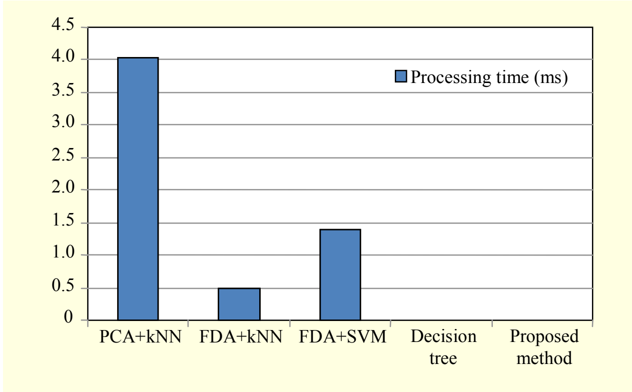

Fig. 8. Processing times.

IV. Conclusion

In this paper, we proposed a disguised-face discrimination algorithm. The classifier consists of MCT preprocessing, AdaBoost-based lookup table, and robust linear regression. In the experiment, we showed that the proposed method is an accurate and fast classifier for embedded systems. This algorithm can be supplemented with a face tracking algorithm for speed improvement. Alternatively, the algorithm could use an eye detection method to improve future performance [11], [12].

References

[1] M. Turk and A. Pentland, "Face Recognition Using Eigenfaces," *Proc. IEEE Conf. Comput. Vision Pattern Recognition*, 1991, pp. 568-591.

[2] P. Belhumeur, J.P. Hespanha, and D.J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, 1997, pp. 711-720.

[3] P.J. Phillips, "Support Vector Machines Applied to Face Recognition," *Proc. Conf. Advances in Neural Inf. Process. Syst. II*, 1999, pp. 803-809.

[4] I. Choi, B. Jun, and D. Kim, "Facial Disguise Discrimination Using RMCT + ADABOOST," *Proc. POSTECH-KYUTECH Joint Workshop on Neuroinformatics*, Kitakyushu, Japan, 2006.

[5] B. Froba and A. Ernst, "Face Detection with the Modified Census Transform," *6th IEEE Int. Conf. Auto. Face Gesture Recognition*, 2004, pp. 91-96.

[6] W.H. Yun et al., "Robust Face Recognition Using the Modified Census Transform," *Int. Symp. Commun. Inf. Technol.*, 2007, pp.

749-752.
[7] Y. Freund and R.E. Schapire, "A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting," *8th Annual Conf. Computational Learning Theory: EuroCOLT*, 1995, pp. 23-37.
[8] P. Viola and M. Jones, "Robust Real-Time Face Detection," *Int. J. Computer Vision*, vol. 57, no. 2, 2004, pp. 137-154.
[9] J. Lattin, J.D. Carroll, and P.E. Green, *Analyzing Multivariate Data*, Pacific Grove, Ca., USA: Thomson Learning, 2003.
[10] B. Jun and D. Kim, "Robust Real-Time Face Detection Using Face Certainty Map," *Proc. 2nd Int. Conf. Biometrics*, vol. 4642, 2007, pp. 29-38.
[11] H. Baskoro, J. Kim, and C. Kim, "Mean-Shift Object Tracking with Discrete and Real AdaBoost Techniques," *ETRI J.*, vol. 31, no. 3, 2009, pp. 282-291.
[12] H. Kim and W. Kim, "Eye Detection in Facial Images Using Zernike Moments with SVM," *ETRI J.*, vol. 30, no. 2, 2008, pp. 335-337.

Woo-han Yun received the BS in electronic and electrical engineering from Sungkyunkwan University, Suwon, Korea, in 2004, and the MS in computer science and engineering from Pohang University of Science and Technology (POSTECH), Pohang, Korea, in 2006. He has been a research scientist at ETRI since 2006. His current research interests include object recognition, tracking, and human-robot interaction.

DoHyung Kim received the PhD from Pusan National University in 2009. He has been a research scientist at ETRI since 2002. His research interests are in the area of human-robot interaction, computer vision and pattern recognition, and especially vision-based user recognition.

Ho-Sub Yoon received the BS and MS in computer science from Soongsil University, Seoul, Korea, in 1989 and 1991, respectively. He received the PhD in image processing from KAIST, Daejeon, in 2003. He joined KIST/SERI in 1991 and transferred to ETRI in 1999. His major research interests include HRI, image processing, audio processing, and pattern recognition.

Jaeyeon Lee received his PhD from Tokai University, Japan, in 1996. He has been a research scientist at ETRI since 1986. His research interests include robotics, pattern recognition, and computer vision.