Combination of signature verification techniques by SVM

Takashi Ito, Wataru Ohyama, Tetsushi Wakabayashi and Fumitaka Kimura
Division of Computer Science, Graduate School of Engineering
Mie University
Tsu, Mie, Japan

***Abstract*—This paper proposes a new SVM based technique for combining signature verification techniques using off-line features and on-line features. The off-line feature based technique employs gradient feature vector representing the shape of signature image, and the on-line feature based technique employs dynamic programming (DP) matching technique for time series data of the signatures. The final decision (verification) is performed by SVM based on output from those off-line and on-line techniques. In the evaluation test the proposed technique achieved 92.96% verification accuracy, which is 1.4% higher than the better accuracy obtained by the individual techniques. This result shows that combining multiple techniques by SVM improves signature verification accuracy significantly.**

***Keywords*-signature verification; gradient feature; HOG; SVM; DP;**

I. Introduction

Automatic signature verification, a behavioral biometrics, has been studied by many researchers[1,2,3,4]. It can be performed using the scanned signature image or using a tablet with a stylus. The former is called off-line verification and the later is called on-line verification. On-line verification with the availability of dynamic information such as stroke order, velocity, or local pressure has already had commercial applications.

Most of the techniques for the on-line signature verification are based on waveform analysis of time series data in terms of Dynamic Programming (DP) or Hidden Markov Model (HMM) [2,3,4]. However the problems such as accuracy improvement and reduction of required size for training sample are still remaining as the research topics. In particular, the shape of signature image (off-line features) is not directly employed for its verification while on-line features are focused on. The characteristic of an individual's signature can only be established using an appropriate number of signature specimens. Since human signature could vary overtime, too few samples will increase the false rejection rate of genuine signatures while too many samples will be labor intensive for the user.

Aiming to solve these problems this paper proposes a new SVM based technique for combining signature verification techniques using off-line features and on-line features.

The off-line feature based technique employs gradient feature vector representing the shape of signature image[7], and the on-line feature based technique employs dynamic programming (DP) matching technique for time series data of the signatures. By combining those two techniques, it is possible to verify the signature synthetically using both off-line and on-line features. The proposed technique can be applied to combine more techniques with other features to achieve higher performance.

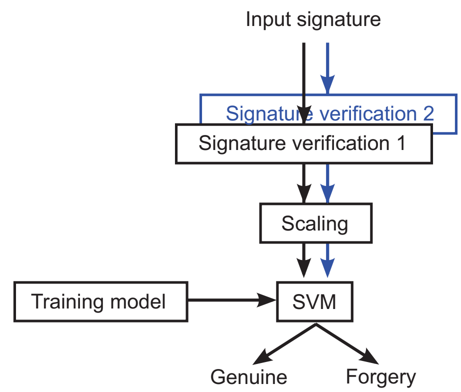


Figure 1. Block diagram of proposed signature verification system.

The rest of this paper is organized as follows: Section 2 describes the signature verification techniques based on the gradient features (2.1), DP matching (2.2) and the combining technique by SVM (2.3). Section 3 details the experimental settings, results and discussions. Finally, Section 4 concludes with future topics.

II. Signature verification technique

Fig. 1 shows the block diagram of combined signature verification system.

The proposed system consists of procedures of signature verification by individual verification techniques and the final decision (verification) performed by SVM based on output from those techniques.

A. Signature verification using gradient features

The gradient feature based technique consists of each steps for image generation, feature extraction, training and verification.

978-0-7695-4774-9/12 $26.00 © 2012 IEEE
DOI 10.1109/ICFHR.2012.192

(a) genuine

(b) forgery

Figure 2. Examples of generated images.

1) Image Generation: A signature image reflecting the velocity of pen movement as the gray scale is generated from a series of coordinates and time. Points with higher (lower) pen velocity are drawn with brighter (darker) grayscale.

Size normalization and smoothing are applied to the input data. The pen velocity is obtained as a quotient of the distance to the duration between adjacent points.

The brightness (grayscale) P_i is defined by

$$P_i = \frac{v_i - v_{min}}{v_{max} - v_{min}} \times P_{max}, \qquad (1)$$

where v_i is the velocity and v_{min}, v_{max} and P_{max} are minimum velocity, maximum velocity and the maximum brightness, respectively.

The values are set to P_{max} = 250, v_{min} = 0, v_{max} = 50,000 in the following experiments. The signature image is dilated towards 4-neighborhood by gray scale morphology operation so that the strokes have predefined width (7).

Fig. 2 shows examples of generated images for a genuine signature (a) and the forgery (b).

2) Feature Extraction: The gradient feature vector is extracted from the generated signature image[5], [6]. The gradient feature vector is composed of directional histogram of gradient of the image. Signature image is segmented into blocks and 576 dimensional feature vector is composed of the local directional histograms. The gradient feature extraction is performed as in the following steps:

Step 1: A 2×2 mean filtering is applied 5 times on the input image.

Step 2: The gray-scale image obtained in Step 1 is normalized so that the mean gray scale becomes zero with maximum value 1.

Step 3: The normalized image is initially segmented into 17(width) $\times$ 7(height) blocks. Compromising trade-off between accuracy and complexity, this block size is decided from the experiment.

Step 4: A Roberts filter is then applied on the image to obtain gradient image. The arc tangent of the gradient (direction of gradient) is initially quantized into 32 directions and the strength of the gradient is accumulated with each of the quantized direction. The strength of Gradient $f(x, y)$ is defined as follows:

$$f(x, y) = \sqrt{(\Delta u)^2 + (\Delta v)^2} \qquad (2)$$

and the direction of gradient $\theta(x, y)$ is:

$$\theta(x, y) = \tan^{-1} \frac{\Delta v}{\Delta u} \qquad (3)$$

where

$$\Delta u = g(x + 1, y + 1) - g(x, y) \qquad (4)$$

and

$$\Delta v = g(x + 1, y) - g(x, y + 1) \qquad (5)$$

and $g(x, y)$ is the gray level of (x, y) point.

Step 5: Histograms of the values of 32 quantized directions are computed in each of 17×7 blocks.

Step 6: Directional histogram of 17×7 blocks is down sampled into 9×4 blocks and 16 directions using Gaussian filters. Finally, a $9 \times 4 \times 16 = 576$ dimensional feature vector is obtained.

Figure 3 illustrates the gradient feature extraction. Figure 3(a), (b) and (c) show the mean filtered image (Step1), the block segmentation (Step 3) and the gradient image (Step 4), respectively. The direction and the strength are represented by the hue and the brightness in Figure 3(c), respectively.

3) Verification: Regularized Mahalanobis distance is calculated for the obtained feature vector. The Mahalanobis distance is defined by

$$g(\boldsymbol{X}) = (\boldsymbol{X} - \boldsymbol{M}_l)^T \boldsymbol{\Sigma}_w^{-1} (\boldsymbol{X} - \boldsymbol{M}_l), \qquad (6)$$

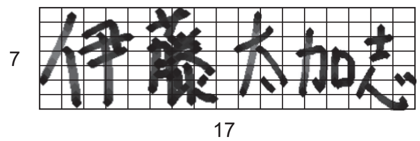
where $\boldsymbol{X}$ is the feature vector, $\boldsymbol{M}_l$ is the mean vector of the l-th writer and $\boldsymbol{\Sigma}_w$ is the pooled within-covariance matrix, respectively.

Generally the rank of $\boldsymbol{\Sigma}_w$ is less or equal to the total sample size minus the number of writers. Hence $\boldsymbol{\Sigma}_w$ will be singular if the feature dimension is greater than the rank. We regularize $\boldsymbol{\Sigma}_w$ to avoid the singularity problem as follows,

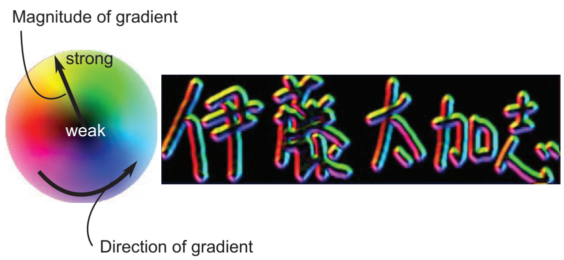
$$(1 - \alpha)\boldsymbol{\Sigma}_w + \alpha \frac{\mathrm{trace}\{\boldsymbol{\Sigma}_w\}}{n} \boldsymbol{I}, \qquad (7)$$

where $\boldsymbol{I}$ denote the identity matrix.

(a) Mean filtered image.



(b) Block segmentation.



(c) Gradient image.

Figure 3. Gradient feature extraction.

B. Signature verification using DP matching

The DP matching based technique consists of each step for data normalization, velocity calculation and the DP matching.

1) Data normalization: Since the size and position vary even in genuine signatures input signature is normalized in its size and position to reduce the variation. The total writing time that varies together with the signature size is also normalized to a fixed time to reduce the variation.

The data normalization is performed as in the following steps:

Step 1: Size of the signature is normalized so that the enclosing rectangular of the signature has pre-specified size.

Step 2: Coordinates are transformed by such translation that the centroid of the signature comes to the origin of the coordinate system.

Step 3: The total writing time is normalized to a fixed time.

2) Pen velocity calculation: The pen velocity v is obtained as a quotient of the distance to the duration t between adjacent points. The velocities of adjacent two points are averaged to reduce and smoothen the variation of the velocity due to observation errors.

Pen velocity is defined by

$$v_{xi} = \frac{\sqrt{(x_{i+1} - x_{i-1})^2}}{t_{i+1} - t_{i-1}}, v_{yi} = \frac{\sqrt{(y_{i+1} - y_{i-1})^2}}{t_{i+1} - t_{i-1}} \tag{8}$$

3) DP matching: DP matching is a technique of pattern matching based on dynamic programming, which evaluate similarity between two sequences of data with different length.

The dissimilarity $g(i, j)$ is recursively defined by

$$\begin{aligned} g(0,0) = d(0,0) = (x_0 - x'_0)^2 + (y_0 - y'_0)^2 \\ +\lambda(v_{x0} - v'_{x0})^2 + \lambda(v_{y0} - v'_{y0})^2 \\ +\mu(z_0 - z'_0)^2 + \nu(t_0 - t'_0)^2 \end{aligned} \tag{9}$$

$$g(i,j) = \min \begin{cases} g(i-1,j) + d(i,j)\, i = 1 \sim n \\ g(i-1,j-1) + d(i,j)\, j = 1 \sim m \\ g(i,j-1) + d(i,j) \end{cases} \tag{10}$$

$$\begin{aligned} d(i,j) = (x_i - x'_j)^2 + (y_i - y'_j)^2 + \lambda(v_{xi} - v'_{xj})^2 \\ +\lambda(v_{yi} - v'_{yj})^2 + \mu(z_i - z'_j)^2 + \nu(t_i - t'_j)^2 \end{aligned} \tag{11}$$

The smaller the dissimilarity is the higher the similarity between two data.

Variables $x_i, y_i, z_i, v_{xi}, v_{yi}, t_i, (x'_j, y'_j, z'_j, v'_{xi}, v'_{yi}, t'_j)$ are the i-th or j-th coordinates, velocity and time for test (for learning) and λ, μ, ν are the weighting factors. Coordinates z, z' take values 0 for pen up, and 1 for pen down. t, t' are the elapsed time from the starting point of the signature (seconds).

C. Combining technique by SVM

SVM is a learning algorithm that classifies the input to two classes.

The regularized Mahalanobis distance and the dissimilarity are calculated by techniques describe in (2.1), (2.2) . These outputs are rescaled so that they range from -1 to 1 for the learning sample. The rescaled outputs and the final decision (genuine or forgery) for the learning sample are used to train the combining SVM.

Used kernel of the SVM is Gaussian.

III. Experiment

A. Data Acquisition

Signature data of 44 individuals were collected by a tablet PC. All individuals signed after some practice to get along with the writing equipment. While the writer of the genuine signature is not allowed to see his/her earlier signatures, the forger make the signature seeing the genuine signature.

The genuine signatures were collected in four days. Total of 42 signatures/individual, 15 signatures in the first day and 9 signatures/day in the following three days were collected. Total of 36 forgeries/signature, 9 forgeries/forger by 4 forgers, were collected. Total samples consists of 1,848 genuine signatures (42 specimens/individual) and 1,584 skilled forgeries (36 speciments/forgery) . Up to six signatures/user collected in the first day are used for training individual techniques. The rest of signatures are used for test. The

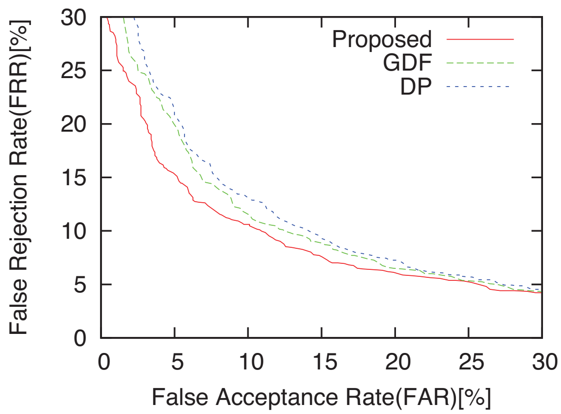


Figure 4. DET curve when three signatures/user are used for training.

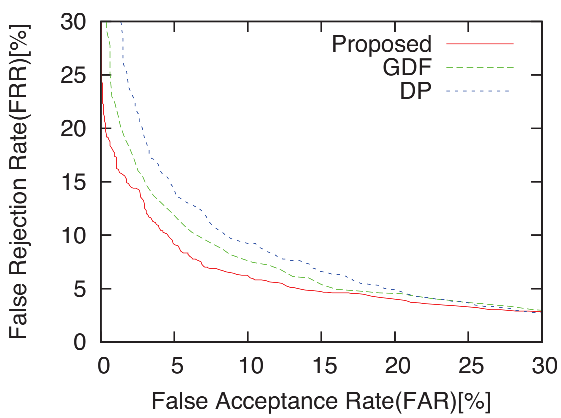


Figure 5. DET curve when six signatures/user are used for training.

combining SVM is trained and tested by two fold cross validation for the test data.

B. Verification Experiment

In order to show the effectiveness of the proposed method the verification accuracies of the individual techniques as well as the SVM combined technique are evaluated by verification test.

Figure 4 shows the DET curve when three signatures/user are used for training individual verifiers and Figure 5 shows the DET curves when six signatures/user are used.

Table 1 shows the verification accuracies with even FRR(False rejection rate) and FAR(False accept rate).

When three signatures/user are used for training the proposed combined technique achieved 89.65% verification accuracy, which is 0.5% higher than the better accuracy obtained by the individual techniques. When six signatures/user are used for training the proposed combined technique achieved 92.96% verification accuracy, which is 1.4% higher than the better accuracy obtained by the individual techniques.

Table I
EXPERIMENTAL RESULTS.

Signatures/user	GDF	DP	Proposed
Three	89.20%	88.23%	89.65%
Six	91.51%	90.59%	92.96%

IV. CONCLUSIONS AND FUTURE WORK

This paper proposed a new SVM based technique for combining signature verification techniques using off-line features and on-line features. The result of evaluation test showed that combining multiple techniques by SVM improves signature verification accuracy significantly.

Following studies are remaining as future research topics: (1) Combining three or more techniques of signature verification based on different features and different algorithms to further improve the verification accuracy, (2) application to signature verification of other script than Japanese and application to accuracy improvement of off line signature verification and (3) testing the proposed method on a public dataset.

REFERENCES

[1] R. Plamondon and G. Lorette, Automatic signature verification and writer identification-The state of the art. Pattern Recognition, Vol.22, No.2, pp.107-131, 1989.

[2] P. Zhao, A. Higashi and Y. Sato, On-Line Signature Verification by Adaptively Weighted DP Matching. IEICE Trans.Inf.&Syst., Vol.E79-D, No.5, pp.535-541, 1996.

[3] Y. Komiyama, T. Ohishi and T. Matsumoto, A Pen Input On-Line Signature Verifier Integrating Position, Pressure and Inclination Trajectories. IEICE Trans.Inf.&Syst., Vol.E84-D, No.7, pp.833-838, 2001.

[4] J. Fierrez, J. Ortega-Garcia, D Ramos and J. Gonzalez-Rodriguez, HMM-based on-line signature verification: Feature extraction and signature modeling. Pattern Recognition Letters, Vol.28, No.16, pp.2325-2334, 2007.

[5] T. Wakabayashi, S. Tsuruoka, F. Kimura and Y. Miyake, Increasing the feature size in handwritten numeral recognition to improve accuracy. Systems and Computers in Japan, Vol.26, pp.2046-2053, 1995.

[6] M. Shi, Y. Fujisawa, T.Wakabayashi and F. Kimura, Handwritten numeral recognition using gradient and curvature of gray scale images. Pattern Recognition, Vol.35, pp.2051-2059, 2000.

[7] Yumiko Kawazoe, Wataru Ohyama, Tetsushi Wakabayashi and Fumitaka Kimura, Improvement of On-line Signature Verification Based on Gradient Features. ICFHR2010, pp.410-414, 2010.