

# How to Control Acceptance Threshold for Biometric Signatures with Different Confidence Values?

Yasushi Makihara, Md. Altab Hossain, Yasushi Yagi  
*The Institute of Scientific and Industrial Research*  
*Osaka University*  
*Ibaraki, Osaka, Japan*

Email: {makihara, hossain, yagi}@am.sanken.osaka-u.ac.jp

**Abstract**—In the biometric verification, authentication is given when a distance of biometric signatures between enrollment and test phases is less than an acceptance threshold, and the performance is usually evaluated by a so-called Receiver Operating Characteristics (ROC) curve expressing a tradeoff between False Rejection Rate (FRR) and False Acceptance Rate (FAR). On the other hand, it is also well known that the performance is significantly affected by the situation differences between enrollment and test phases. This paper describes a method to adaptively control an acceptance threshold with quality measures derived from situation differences so as to optimize the ROC curve. We show that the optimal evolution of the adaptive threshold in the domain of the distance and quality measure is equivalent to a constant evolution in the domain of the error gradient defined as a ratio of a total error rate to a total acceptance rate. An experiment with simulation data demonstrates that the proposed method outperforms the previous methods, particularly under a lower FAR or FRR tolerance condition.

**Keywords**—acceptance threshold control; ROC curve; biometrics; quality measure

## I. INTRODUCTION

Biometrics-based verification has recently gained considerable attention due to wide applications including access control systems, smart log-in systems, and surveillance systems. In the verification process, a distance of a user's biometric signatures between enrollment and test phases is measured at first, and then authentication is given to the user if the distance is less than the predefined acceptance threshold. The verification performance is usually measured by False Rejection Rate (FRR) of the same person and False Acceptance Rate (FAR) of imposters. FRR and FAR depends on the acceptance threshold and their tradeoff curve is called Receiver Operating Characteristics (ROC) curve [1].

In addition, quality measures are introduced to estimate the verification performance in many biometrics areas[2][3][4][5], which are typically obtained as signal to noise ratio in speech recognition or degrees of situation differences between enrollment and test phases. The quality measures play important roles in fusion of multiple samples[6], multiple features[7], or multimodal biometrics[8][9], score normalization[10].

Moreover, Kryszczuk et al.[11] propose a method to find an optimal decision boundary in the evidence space composed of a score and the quality measure and to adaptively control the acceptance threshold based on the decision boundary to improve the verification performance. Because the method provides a single decision boundary as a result, the performance is given as not a whole ROC curve but a pair of FAR and FRR which is expressed in just one point on the ROC curve. Strictly speaking, though a ROC curve can be generated by evolving decision boundaries based on a certain criteria (e.g. signed distances to the decision boundary), the generated ROC curve is not guaranteed to be optimal.

Therefore, we propose a method of adaptive acceptance threshold control in terms of ROC curve optimization in the presence of scores with different quality measures. For this purpose, we define a new measure *error gradient* as a increasing ratio of total error samples and total acceptance samples and show that constant acceptance threshold in the error gradient domain realizes the ROC curve optimization.

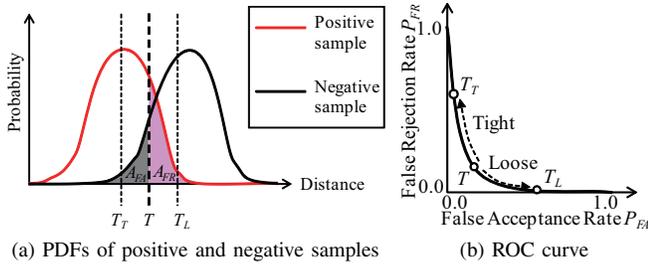
## II. ADAPTIVE ACCEPTANCE THRESHOLD CONTROL

### A. ROC curve

In usual biometric verification systems, a kind of distance between biometric signatures is exploited as metrics for verification. Naturally, it is desirable that the distance is smaller for the same person trial (call it positive sample), and is larger for imposter trial (call it negative sample).

Here, let's assume the Probability Distribution Functions (PDFs) of the distances for the positive samples and negative samples are given, and that an acceptance threshold  $T$  for verification is set as shown in Fig. 1(a). In this case, the area  $A_{FR}$  and  $A_{FA}$  correspond to False Rejection Rate (FRR) of the same person and False Acceptance Rate (FAR) of the imposter respectively. If the acceptance threshold is moved to tight direction  $T_T (< T)$  as shown in Fig. 1(a), FRR gets larger while FAR gets smaller, and if it is moved to loose direction  $T_L (> T)$  as shown in Fig. 1(a), FRR gets smaller while FAR gets larger.

Therefore, by continuously evolving the acceptance threshold from the tightest value (e.g. minus infinity) to the loosest value (e.g. plus infinity), a tradeoff curve of



(a) PDFs of positive and negative samples (b) ROC curve  
Figure 1. Examples of PDFs of positive and negative samples and ROC curve

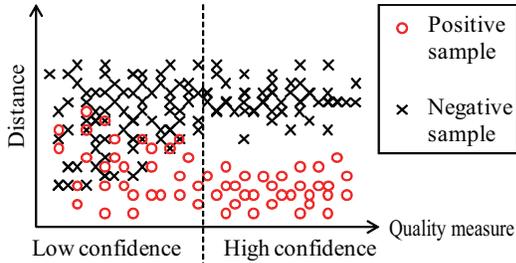


Figure 2. Distribution of distance and quality measure

FRR and FAR is obtained as well-known Receiver Operating Characteristics (ROC) curve[1] as shown in Fig. 1(b).

### B. Simplified example

In this subsection, we give a simplified example to help an intuitive understanding of adaptive acceptance threshold control. First, let's assume that positive and negative samples are distributed in the distance and quality measure domain as shown in Fig. 2. For simplicity, we roughly classify into two quality measures in this subsection: high quality measures with well discriminable samples and low quality measures with relative confusing samples. Then, let's assume that PDFs and ROC curve for each quality measure is given as shown in Fig. 3 and let upper bounds of acceptance thresholds where no false alarm exist in low and high quality measures be  $T_{NoFA}^L$  and  $T_{NoFA}^H (> T_{NoFA}^L)$  respectively.

Next, it is assumed that acceptance thresholds of both high and low quality measures have reached  $T_{NoFA}^L$ , and that they are subsequently evolved to  $T_{NoFA}^H$ . In this case, because FAR in the low quality measure increases to some extent (gray area in Fig. 3(a)) while FAR in the high quality measure does not increase at all, it is obviously better to evolve only the acceptance threshold of the high quality measure while stopping the acceptance threshold of the low quality measure.

### C. Adaptive acceptance threshold control

In this subsection, a method of general adaptive acceptance threshold control is described. First, we define PDFs of distance  $t$  for positive and negative samples as  $p^P(t)$  and  $p^N(t)$  respectively. If the acceptance threshold is set to  $T$ ,

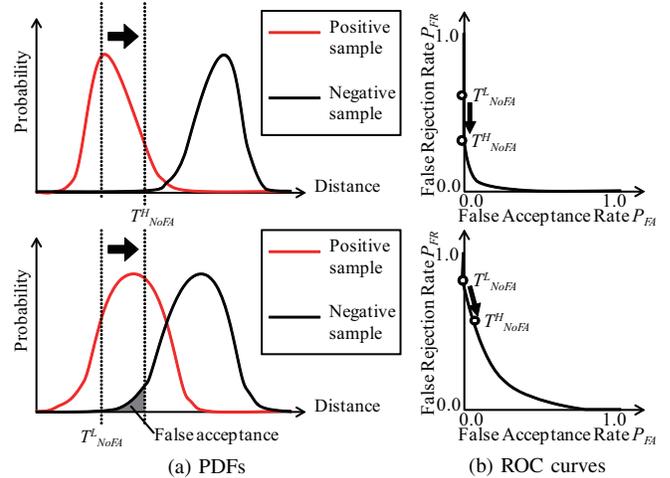


Figure 3. PDFs and ROC curves for high (top) and low (bottom) quality measure.

the FAR  $R_{FA}(T)$  and FRR  $R_{FR}(T)$  are defined as

$$R_{FA}(T) = \int_{-\infty}^T p^N(t) dt \quad (1)$$

$$R_{FR}(T) = 1 - \int_{-\infty}^T p^P(t) dt. \quad (2)$$

Then, we subsequently define new variables: error rate  $R_E(T)$  and acceptance rate  $R_A(T)$  as

$$R_E(T) = R_{FA}(T) + R_{FR}(T) \quad (3)$$

$$R_A(T) = R_{FA}(T) + (1 - R_{FR}(T)). \quad (4)$$

Next, we consider to error gradient of the error rate  $R_E(T)$  to the acceptance rate  $R_A(T)$  as

$$g(T) = \frac{dR_E(T)}{dR_A(T)} = \frac{p^N(T) - p^P(T)}{p^N(T) + p^P(T)}. \quad (5)$$

Here, a lower error gradient (e.g.  $g(T) \simeq -1.0$ ) means that the most of accepted samples are positive samples, and a higher error gradient (e.g.  $g(T) \simeq 1.0$ ) means that the most of accepted samples are negative samples. In addition, a middle error gradient (e.g.  $g(T) \simeq 0.0$ ) means that the positive and negative samples in the accepted samples are balanced, that is, the more confusing area.

Next, in order to realize minimal pairs of FAR and FRR, we equalize the error gradient  $g(T)$  on a controlled acceptance threshold line for each quality measure, in other words, contour line of the gradient is a set of controlled acceptance threshold lines. Moreover, once the original distance  $t$  is mapped to the error gradient  $g(t)$ , the constant acceptance threshold line in the gradient domain is the same meaning of the contour line of the error gradient in the original distance domain. Note that the optimality of the method is guaranteed under the condition where the error gradient  $g(T)$  is semi-monotone increase.

#### D. Implementation

In this section, a method to implementation of the adaptive acceptance threshold control is addressed. First, assuming that  $N_P$  positive training samples and  $N_N$  negative samples are given and that  $i$ th positive and negative samples are composed of pairs of distance and quality measure  $(t_i^P, c_i^P)$  and  $(t_i^N, c_i^N)$  respectively. Then, the distance  $t$  and the quality measure  $c$  is quantized respectively by appropriate lower and upper bound and quantization step as

$$t_j = t_{min} + j s_t, j \in \mathbb{Z}, 0 \leq j \leq (t_{max} - t_{min}) / s_t \quad (6)$$

$$c_k = c_{min} + k s_c, k \in \mathbb{Z}, 0 \leq k \leq (c_{max} - c_{min}) / s_c \quad (7)$$

Next, a weight  $w_{i,k}^P$  of  $i$ th positive sample is calculated for  $k$ th quality measure control point as

$$w_{i,k}^P = \max(1.0 - |c_i^P - c_k| / s_c, 0) \quad (8)$$

Subsequently, the positive samples PDF  $p_{j,k}$  of distance at  $j$ th distance control point and  $k$ th quality measure control point is calculated by Gaussian kernel-based non-parametric PDF estimation as

$$p_{j,k}^P = \frac{1}{Z_k} \sum_i^{N_P} w_{i,k}^P \exp\left(-\frac{(t_i^P - t_j)^2}{2\sigma^2}\right), \quad (9)$$

where  $Z_k$  is a distribution function so as to normalize the PDF as  $\sum_j p_{j,k}^P s_t = 1$ , and  $\sigma$  is a pre-determined Gaussian kernel size. After the negative samples PDF  $p_{j,k}^N$  is calculated in the same way, the error gradient  $g_{j,k}$  at control point  $(t_j, c_k)$  is calculated by eq. (5). To satisfy the optimality, the error gradient should be semi-monotone increase for each quality measure control point  $c_k$ , the optimal approximation of the error gradient  $\mathbf{g}^* = \{g_{j,k}^*\}$  is estimated by the following minimization scheme

$$\mathbf{g}^* = \arg \min_{\hat{\mathbf{g}}} S(\hat{\mathbf{g}}) \quad (10)$$

$$S(\hat{\mathbf{g}}) = \sum_{k,j} \{(\hat{g}_{j,k} - g_{j,k})^2 + \alpha(\hat{g}_{j,k} - \hat{g}_{j-1,k})^2\} \quad (11)$$

$$\text{subject to } \hat{g}_{j-1,k} \geq \hat{g}_{j,k}, \quad (12)$$

where  $\alpha$  is a coefficient for regularization term. The above optimization is efficiently solved by the convex quadratic programming. Once the error gradients at all the control points are obtained (call it an error gradient map), an arbitrary pair of a distance and a quality measure is mapped to the error gradient domain by linear interpolation of the error gradient map and conventional constant acceptance threshold control is applied to them in the error gradient domain.

Note that the error gradient map can be applied not only to the training samples but also to test samples in case of where the tendency between the training samples and test samples are similar.

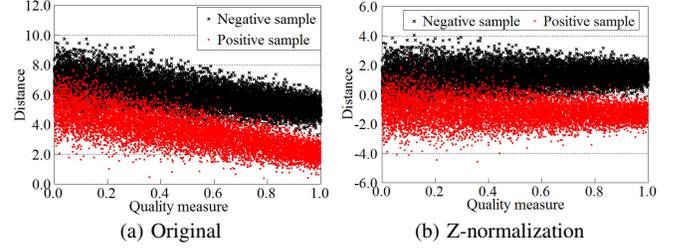


Figure 4. Distance and quality measure distribution of training samples

### III. EXPERIMENTS

#### A. Data set and parameter setting

Experiments on simulation data were made to show the effectiveness of the proposed method. First, we set PDFs where positive and negative samples are drawn for a training set and a test set. A PDF of the quality measure  $c$  commonly used in positive and negative samples is set to a uniform distribution in the domain  $0 \leq c \leq 1$ . Then, PDFs of distances of the positive and negative samples are set to Gaussian Distributions  $\mathcal{N}(\mu^P(c), \sigma^P(c))$  and  $\mathcal{N}(\mu^N(c), \sigma^N(c))$  respectively. Basically, these two PDFs are decided so as to more confusing as the quality measure  $c$  gets low, we set means and standard deviations of the PDFs as  $\mu^P(c) = 5.0 - 3.0c$ ,  $\mu^N(c) = 7.0 - 2.0c$ ,  $\sigma^P(c) = 1.0 - 0.5c$ , and  $\sigma^N(c) = 1.0 - 0.5c$  respectively.

Then, a set of 10,000 positive and negative samples each are drawn from the PDFs for a training set as shown in Fig. 4(a). From the figure, it is obviously seen that the two PDFs are more discriminable in the high quality measure domain (e.g.  $c = 1.0$ ) while that they are more confusing in the low quality measure domain (e.g.  $c = 0.0$ ). In the same way, another set of 10,000 positive and negative samples are drawn for a test set used in the performance evaluation.

In addition, in order to compare the proposed method with the z-normalization[1], the distance is converted by subtracting a total mean  $\mu(c)$  and by dividing it by a total standard deviation  $\sigma(c)$  for each quality measure  $c$  as shown in Fig. 4(b). Note that the mean and the standard deviation are equal to 0 and 1 respectively after the z-normalization.

Next, we set the parameters used in the experiments. In the PDF estimation process, the Gaussian kernel size and control point intervals are set to  $\sigma = 0.3$ ,  $s_t = 0.01$ , and  $s_c = 0.1$  respectively. The domains of distance are set to  $t_{min} = 0.0, t_{max} = 10.0$  for the original set, and to  $t_{min} = -0.5, t_{max} = 5.0$  for the z-normalization set. In the estimation of the semi-monotonically increasing error gradient, the regularization coefficient is set to  $\alpha = 1.0$ .

#### B. Performance evaluation

The error gradients are estimated from the training set and resultant contours of the error gradients are shown in Fig. 5(a). From the figure, we can see that the contours are dense in the high quality measure area and vice versa, which derives from better and worse discrimination capability in

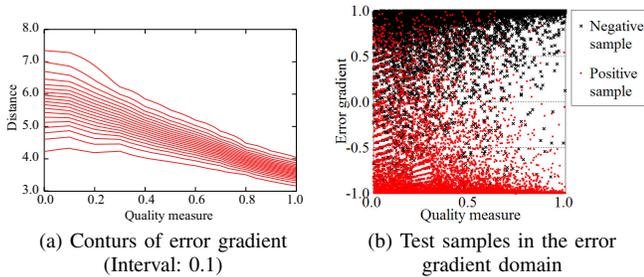


Figure 5. Error gradient

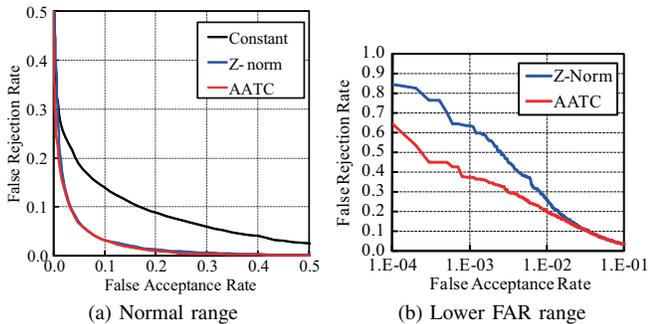


Figure 6. ROC curves

each quality measure. Then, based on the trained error gradient map, the test samples are converted into the error gradient domain as shown in Fig. 5(b). We can see that the samples around the balanced area of error gradient (e.g.  $g = 0.0$ ), that is, more confusing area, are sparse in the high quality measure area and vice versa.

Finally, the verification performance for the test set is evaluated with the ROC curve as shown in Fig. 6. As a result, the proposed Adaptive Acceptance Threshold Control (AATC) outperforms a conventional constant acceptance threshold control (Constant). On the other hand, although the z-normalization (Z-norm) seems to be competitive to the proposed method in a normal range, the proposed method outperforms in lower FAR range (Fig. 6(b)) and lower FRR range (graph is omitted due to page limitation). This indicates that the proposed method is particularly effective in case where a tolerance of either FAR or FRR is very tight.

#### IV. DISCUSSION

Application targets of the proposed method are not limited to biometric verification but also two-class classification problems in the broad sense like detection problem, for example, classification of human and non-human in human detection problem. In such a case, a kind of quality measures can be still measured based on occluding areas. Thus, experiments on real data on the two-class classification problems are remained as one of future directions.

On the other hand, the proposed method assumes that the distributions of distance and quality measures are consistent in the training and test sets, and therefore the optimality is not guaranteed in case where the distributions are inconsistent. Hence, another future work is investigation and

enhancement of generalization capability of the proposed method.

#### V. CONCLUSION

We proposed a method to adaptively control an acceptance threshold with quality measures for biometric verification. The optimal evolution of the adaptive threshold in the domain of the distance and quality measure is formulated as a set of contours of the error gradient defined as a ratio of a total error rate to a total acceptance rate. An experiment with simulation data demonstrates that the proposed method outperforms the previous methods in terms of the ROC curve, particularly under a lower FAR or FRR tolerance condition.

#### ACKNOWLEDGMENT

This work was supported by Grant-in-Aid for Scientific Research(S) 21220003.

#### REFERENCES

- [1] P. Phillips, H. Moon, S. Rizvi, and P. Rauss, "The feret evaluation methodology for face-recognition algorithms," *Trans. of Pattern Analysis and Machine Intelligence*, vol. 22, no. 10, pp. 1090–1104, 2000.
- [2] A. Harriero, D. Ramos, J. Gonzalez-Rodriguez, and J. Fierrez, "Analysis of the utility of classical and novel speech quality measures for speaker verification," in *ICB '09: Proceedings of the Third International Conference on Advances in Biometrics*. Berlin, Heidelberg: Springer-Verlag, 2009, pp. 434–442.
- [3] F. Alonso Fernandez, F. Roli, G. Marcialis, J. Fierrez, and J. Ortega Garcia, "Comparison of fingerprint quality measures using an optical and a capacitive sensor," in *BTAS07*, 2007, pp. 1–6.
- [4] S. Muller, O. Henniger, and T. U. D. Darmstadt, "Evaluating the biometric sample quality of handwritten signatures," in *In 2nd International Conference on Biometrics*, 2007.
- [5] E. Krichen, S. Garcia-Salicetti, and B. Dorizzi, "A new probabilistic iris quality measure for comprehensive noise detection," in *Biometrics: Theory, Applications, and Systems, 2007. BTAS 2007. First IEEE International Conference on*, 27–29 2007, pp. 1–6.
- [6] H. Sellaheewa and S. Jassim, "Illumination and expression invariant face recognition: Toward sample quality-based adaptive fusion," in *BTAS08*, 2008, pp. 1–6.
- [7] M. Gales and S. Young, "Robust continuous speech recognition using parallel model combination," *IEEE Transactions on Speech and Audio Processing*, vol. 4, no. 5, pp. 352–359, 1996.
- [8] J. Fierrez-Aguilar, "Adapted fusion schemes for multimodal biometric authentication," Ph.D. dissertation, Universidad Politecnica de Madrid, 2006.
- [9] S. Bengio, C. Marcel, S. Marcel, and J. Mariethoz, "Confidence measures for multimodal identity verification," *Information Fusion*, vol. 3, no. 4, pp. 267–276, 2002.
- [10] N. Poh, J. Kittler, and T. Bourlai, "Improving biometric device interoperability by likelihood ratio-based quality dependent score normalization," in *BTAS07*, 2007, pp. 1–5.
- [11] K. Kryszczuk and A. Drygajlo, "Improving classification with class-independent quality measures: Q-stack in face verification," in *In 2nd International Conference on Biometrics, Seoul, South Korea*, 2007.