H-017

個別閾値を用いたオンライン署名認証システム
Online signature verification system using individual threshold

山田 太一† 平川 豊†
Taichi Yamada Yutaka Hirakawa

1. はじめに

現在、個人認証を行うために生体特徴を用いたバイオメトリクスが注目されている。バイオメトリクスには大きく 2 種類に分類でき、指紋や静脈のような身体的特徴を用いたものと、署名や音声といった行動的特徴を用いたものがある。本研究では署名時の行動的特徴を用いて個人認証を行う手法の研究を行った。また本研究では高精度の署名認証システムの構築を目的とし、認証精度の向上を図るため入力データが本人かどうかの判定の際に用いる閾値を個別に用意する手法を提案する。しかし本稿では個別閾値を用いた手法の有効性を示すまでであり、高精度の署名認証システムの構築までは至っていない。

2. SVC データベース

システム評価には First International Signature Verification Competition (SVC)[1]で使われたオンライン署名データを用いた。この署名データは、2004 年香港にて開催されたバイオメトリクスの国際会議、International Conference on Biometric Authentication にて行われた初めてのオンライン署名認証のコンペティション SVC[1] のために収集されたオンライン署名データである。

公開されている署名数は 40 名分で、一人当たり 20 の本人署名と 20 の訓練偽造署名が登録されている。一つのファイルには、ペン座標(x,y)、筆圧、ペン方位・仰角、ペンアップダウンの時系列データが取得されている。
本稿におけるデータ、数値等はこのデータベースを元に算出している。

3. 処理の流れ・方法

署名認証の基本構成は図 1 に示す。

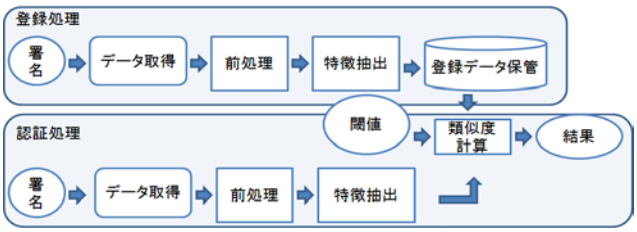

図 1 署名認証の基本構成

3.1 前処理・特徴抽出

本研究において前処理は座標 x,y についてのみ式 1,2 の方法で正規化を行った。

特徴抽出のフェーズでは座標 x,y、筆圧 P、仰角 Al、方位 Az の 5 つのパラメータのほかそれぞれの変化量（Vx,Vy、⊿P、⊿Al、⊿Az の計 10 のパラメータ）を特徴とした。

$$x^*(t) = \frac{x(t) - x_{min}}{x_{max} - x_{min}} \quad (1) \qquad y^*(t) = \frac{y(t) - y_{min}}{y_{max} - y_{min}} \quad (2)$$

3.2 類似度計算アルゴリズム

データの長さの異なる 2 つの波形の比較には伸縮関数を用いて距離を計算する必要があり、そのアルゴリズムとして DP マッチングを用いた[2]。 2 つのデータ A,B の長さを i,j とすると式(3)により類似度計算を行った。

初期値 $d(0,0)=0$

漸化式

$$d(i,j)=\min\begin{cases} d(i-1,j-1) + dist(A_i,B_j) \\ d(i-1,j) + dist(A_i,B_j) \\ d(i,j-1) + dist(A_i,B_j) \end{cases} \quad (3)$$

3.3 共通閾値

判定の際に用いる閾値は上記方法で本人同士の署名の類似度 7600 (40×20C2)のサンプル、本人と偽造署名の類似度 16000 (40×20×20)のサンプルを求め、本人拒否率(FAR)と他人受入率(FRR)が等しくなる等価エラー率(EER)が得られる値を閾値として各パラメータに設定する。

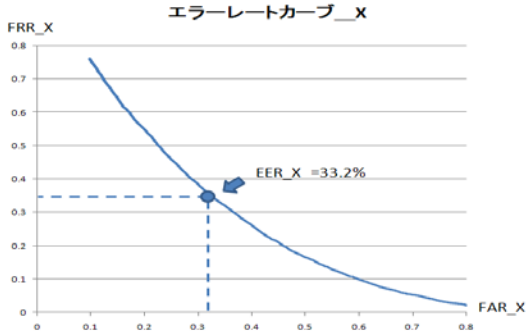

図 2 共通閾値を決めるエラーレートカーブ

3.4 個別閾値

個別に閾値を設定することは認証精度の面で有効であると言われている[3]。しかし高精度の個別閾値を得るには登録処理の際、多くのリファレンスデータが必要となりユーザの負担が大きくなり実用的でない。そこで少数のリファレンスデータから個別の閾値を設定する下記のような方法を提案する。

個人ごとに各パラメータの安定度は様々である。(図 2) このことを利用し共通閾値をもとに安定度が高い特徴は閾値を狭く、不安定な特徴は閾値を広く設定することで個別の閾値を設定する。ここで安定度とはリファレンスデータ同士の類似度の分散をもとに登録データ 40 人中での標準偏差をとり偏差値化したものをいう。個別閾値の計算方法を式 4、5 に示す。

$$個別閾値 = 共通閾値 \times (1-\alpha) \quad (4)$$

$$\alpha = \beta \times (Log(50 - HENSA値)/100) \quad (5)$$

※ βは補整係数 本稿ではβ＝8

† 芝浦工業大学 SHIBAURA INSTITUTE OF TECHNOLOGY

4. 予備実験

個人ごとに各パラメータの安定度が様々であるかどうか調べる予備実験を行った。SVC データベースから 40 人分の各パラメータにおいて一人当たり 190(20C2)の類似度の分散を 40 人の中で比較した。パラメータごとに単位が異なるので偏差値化することで同一の尺度でそれぞれの項目を比較したものを図 3 に示す。

分散が小さいほど安定した特徴といえるのでここでは偏差値が低いものほど安定度が高い。

図 2 から各パラメータの安定度には個人差があることがわかる。この安定度の個人差を利用すれば、個別の閾値を設定できると考えられる。

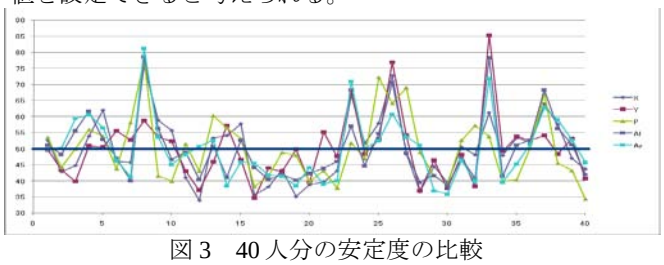

図 3　40 人分の安定度の比較

5. 実験

本人同士の署名の類似度 7600 のサンプル、本人と偽造署名の類似度 16000 のサンプルから共通閾値が得られた。その時の等価エラー率と式 4,5 で得た個別閾値を用いた場合の本人排他率・他人受入率の比較を行った。（Ⅰ）

またリファレンスデータ数 20 個の場合と 5 個の場合のエラーレートの比較も行った。（Ⅱ）

5.1 結果

(Ⅰ),(Ⅱ)の実験結果を表 1，2 に示す。

共通閾値の場合　　個別閾値の場合

	閾値	EER	FAR	FRR
座標 X	0.054697	33.5%	16.4%	15.6%
座標 Y	0.064931	33.5%	16.5%	20.0%
筆圧 P	203.6433	27.7%	14.5%	14.2%
仰角 AL	86.00829	28.7%	14.3%	16.1%
方位 Azi	41.05293	30.2%	12.4%	21.4%
速度V	96.00259	42.3%	27.9%	26.7%
速度Y	125.9727	40.9%	19.6%	30.1%
⊿P	23.18922	52.3%	34.0%	25.9%
⊿Al	10.18648	46.5%	31.6%	34.3%
⊿Azi	5.051176	48.5%	38.0%	28.8%

表 1　共通閾値と個別閾値のエラーレート比較

サンプル数	20	5	20	5
	FAR	FAR	FRR	FRR
座標 X	16.4%	22.8%	15.6%	20.6%
座標 Y	16.5%	20.4%	20.0%	17.7%
筆圧 P	14.5%	25.6%	14.2%	12.9%
仰角 AL	14.3%	19.9%	16.1%	24.5%
方位 Azi	12.4%	17.1%	21.4%	19.3%
速度V	27.9%	24.7%	26.7%	37.4%
速度Y	19.6%	20.5%	30.1%	33.8%
⊿P	34.0%	29.3%	25.9%	40.7%
⊿Al	31.6%	37.5%	34.3%	43.1%
⊿Azi	38.0%	36.8%	28.8%	35.3%

表 2　署名数 20 個と 5 個の場合のエラーレート比較

6. 評価・考察

実験(Ⅰ)の結果から個別閾値を用いた場合の認証精度はがどのパラメータにおいても上昇していることがわかる。このことから個別に閾値を設ける手法の有効性が言えそうである。特徴については 5 つの基本パラメータは識別に有効そうであるが、それぞれの変化量は個別閾値を用いても良くない結果となった。また共通閾値での EER が松本氏[2]の報告に比べ極端に精度が低い。つまり特徴の抽出方法や距離計算のアルゴリズムまたはプログラムが適切でない可能性が考えられる。基本となる共通閾値での EER を上昇させることで、より個別の閾値を用いた場合での有効性が現れるのではないかと考えられる。

個別閾値は式 4，5 の方法で設定したが、より有効な設定方法があるか検討が必要である。

実験(Ⅱ)の結果からリファレンスデータ数が少なくなると個別閾値の有効性も薄れてしまう結果となったが、共通閾値の EER と比べると、認証精度は向上している。このことから少数リファレンスデータでの個別閾値の設定を実現したと考えられる。

7. 課題

今回の実験では個別閾値を設けることの有効性を示すことが目的であったため、特徴抽出では簡単な処理しか施していない。今後、より有効と思われる特徴抽出の処理を加える必要がある。

また、今回出したエラーレートは一つのパラメータを用いて本人かどうかの判定をした場合での認証精度であり、個別閾値を用いた場合でも精度は十分であるとは言い難い。そのため複数の特徴パラメータを用いて、有効な特徴選択や特徴融合を行いより高精度を実現するアルゴリズムを構築する必要がある。

今回実験には SVC 署名データを用いたが有効な特徴はデータベースによって多少異なる[2]ので複数のデータベースを用いて実験することが望ましい。

8. まとめ

本稿では署名認証システムにおいて判定の際に用いる閾値を少数の署名から個別閾値を用意する手法について説明を行なった。そして SVC 署名データベースを利用し署名から取得される特徴について個々に認証精度について評価実験を行った。

実験結果から提案手法の有効性を示すことができたが、十分な認証精度を得るにはまだ多くの課題が残る結果となった。

参考文献

[1]Dit-Yan Yeung, Hong Chang, Yimin Xiong, Susan George, Remanujan Kashi, Takashi Matsumoto, and Gerhard Rigoll, 2004. SVC. First International Signature Verification Competition. Springer Lecture Notes in Computer Science,3702, 16-22.

[2]松本大吾, “オンライン署名の動的特徴量を用いた個人認証手法に関する研究”, 早稲田大学大学院理工学研究科 博 士 論 文 (2005)

[3]Anil K. Jain , , Friederike D. Griess and Scott D. Connell , "On-line signature verification" Department of Computer Science and Engineering, Michigan State University, 3115 Engineering Building, East Lansing, MI 48824, USA Available online 15 January 2002.