

ペンを持つ手の動き情報を用いた筆記者認証

Writer Authentication Using Motion Information of Hand Holding a Pen

田中 緑†

中井 満†

Midori Tanaka Mitsuru Nakai

1. はじめに

個人の癖などの行動的特徴を用いた認証方法にオンライン署名認証がある。オンライン署名は、筆跡以外にも筆点の移動速度や筆圧などの情報を利用できる [1]。また、これらに加わる情報として署名時のペンの持ち方が提案されている [2]。本発表では、ペンを持つ手の時間変化パターンを用いた認証法について報告する。

2. ペンを持つ手の動き情報の取得

2.1 データ収集

本研究では、手の形状とその動きを測定するために Kinect の深度センサを用いた。4 人の被験者 A, B, C, D にペンを持ってもらい、手を親指側から撮影した。手の領域検出を容易にするために、Kinect から約 80 cm の位置に印を付け、そこにペン先を置いた状態から筆記を開始した。筆記動作はすべて「丸を描く」ものとし、筆記速度や丸の大きさは自由とした。測定は 30fps で 100 フレームとし、その間に筆記してもらい、筆記していないときは静止するように指示した。これを 1 人につき 5 回行い、4 人から計 20 サンプル収集した。また被験者 A については、後日、5 日間に分けて 10 サンプル追加収集した。4 人の被験者のペンの持ち方を図 1 に示す。

3. ペンの持ち方による筆記者認証原理

3.1 前処理

撮影した動画像から筆記中の手を空間的・時間的に切り出す。まず、フレーム毎に手の関心領域 (ROI) 画像を次の手順で抽出する。640 x 480 ピクセルの深度画像の座標系から世界座標系に変換し、3D 点群データにする。ここで水平方向を x 、垂直方向を y 、深度方向を z とする。点群のクラスタリングにより、手の置かれた距離 ($z = 80$ cm 近傍) にある最大クラスタを選ぶ。そのクラスタからペンを把持する手の領域とし

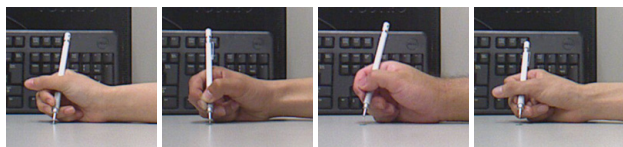


図 1: 4 人の被験者のペンの持ち方



図 2: ペンを持つ手の時間変化

て、 x 端から 15 cm まで、および y 端から 15 cm までを ROI 点群として切り出す。収集したサンプルでは ROI 点群はおおよそ 4000 点くらいの大きさとなった。これを画像座標系に変換し、ROI 画像とする。

また ROI 点群の重心位置の移動が筆跡に類似することから、移動量が閾値を超える連続した区間を筆記区間とする。今回は目視で区間決定したところ、最小で 23 フレーム、最大で 47 フレームとなった。あるデータの書き始めの 1 フレーム目、筆記中の 11 フレーム目、書き終わりの 36 フレーム目の手形状を図 2 に示す。同じ被験者でも筆記動作中に手形状が変化していることが分かる。

3.2 距離計算

認証時には、登録動画像 (フレーム数 I) と入力動画像 (フレーム数 J) の距離 \bar{D} を計算し、 \bar{D} が閾値以下のときに本人と認証とする。登録動画像の任意の 1 フレーム i ($1 \leq i \leq I$) と入力動画像の任意の 1 フレーム j ($1 \leq j \leq J$) の距離 $d_{(i,j)}$ の計算手順を述べる。2 つのフレームの ROI の重なる面積が最も大きくなるように画像を重ねる位置補正を行ってから計算を行う。図 3 に被験者 A, B の ROI およびその 2 値化画像 (シルエット) を示す。2 値化画像の排他的論理和を取った領域は手が重なっていない領域を示し、この面積をシルエット距離 d_1 とする。2 値化画像の論理積を取った領域は手が重なっている領域を示し、この領域の深度差 (絶対値) の総和を深度距離 d_2 とする。 d_1, d_2 は、単位がピクセルと mm で異なるため、それぞれあらかじめ決められた係数によりスケール変換する。

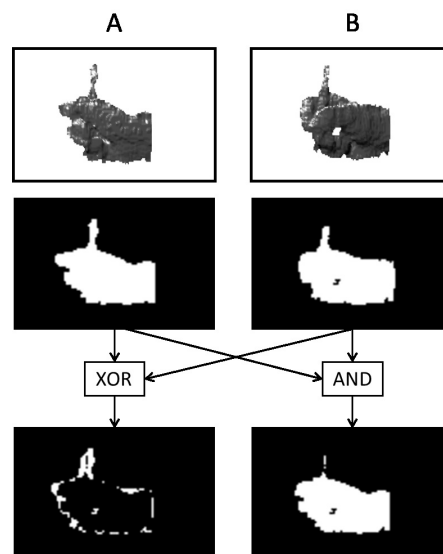


図 3: 筆記者間の距離計算

†富山県立大学, Toyama Prefectural University

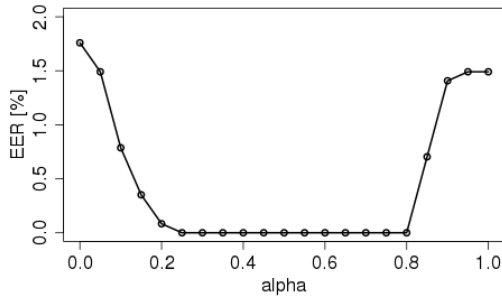


図 4: 動画像における重みと EER の関係

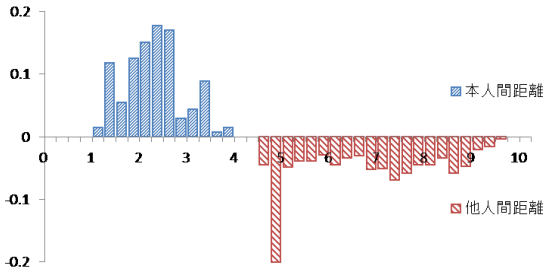
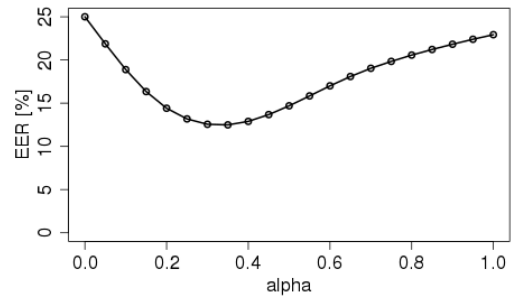
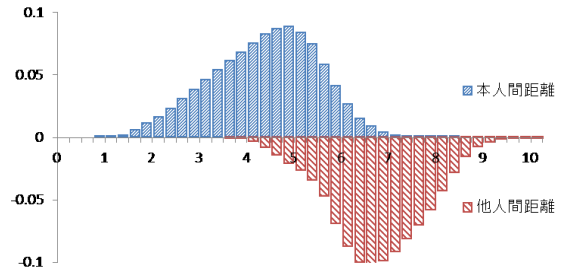
図 5: 動画像における距離分布の割合 ($\alpha = 0.65$)

図 6: 静止画像における重みと EER の関係

図 7: 静止画像における距離分布の割合 ($\alpha = 0.35$)

このとき、距離 $d_{(i,j)}$ は結合重み α を用いて

$$d_{(i,j)} = (1 - \alpha) \times d_1 + \alpha \times d_2$$

とする。動画像間距離 \bar{D} は、フレーム間の距離の総和が最小になるように DP マッチングで対応付けて算出する。登録動画像の $1 \sim i$ ($i \leq I$) フレームと、入力動画像の $1 \sim j$ ($j \leq J$) フレームの累積距離 $D_{(i,j)}$ は以下の漸化式で求める。

$$D_{(i,j)} = d_{(i,j)} + \min \begin{cases} D_{(i-1,j)} \\ D_{(i-1,j-1)} \\ D_{(i-1,j-2)} \end{cases}$$

すなわち、登録動画像の各フレームによく一致するフレームが認証時に時間順に入力されているかチェックしている。登録フレームあたりの平均距離 $\bar{D} = D_{(I,J)}/I$ を求め、認証する。

4. 筆記者認証実験

4.1 動画像による認証精度の調査

30 サンプルから任意の 2 つを選ぶとき、同一被験者となるのは 270 組、異なる被験者となるのは 600 組である。それぞれの組について、前者は本人間距離、後者は他人間距離として計算する。EER は等誤り率であり、本人拒否率と他人受入率が等しくなる閾値での誤り率を指す。

本実験では、重み α 毎に EER の調査を行った。図 4 に重み α と EER の関係を示す。シルエット距離のみを用いた場合 ($\alpha = 0$ のとき) は 1.8%，深度距離のみを用いた場合 ($\alpha = 1$ のとき) は 1.5% であった。 α が 0.25～0.8 の場合には 0% となった。 $\alpha = 0.65$ のときの本人間距離と他人間距離の分布の割合を図 5 に示す。このとき 2 つの距離分布に重なりは無く、本人間距離の最大値と他人間距離の最小値の差が最も大きくなった。

4.2 静止画像による認証精度の調査

比較として、動画像ではなく任意の 1 フレームを登録・照合した場合の実験を行った。登録動画像と入力動画像から任意の 1 フレームを選ぶとき、同一被験者の組合せは 420262 組、異なる被験者の組合せは 756558 組である。図 6 に重み α と EER の関係を示す。動画像による認証と同様に、 α を 0 から増やすと EER が低くなり、1 に近付くと再び EER が増加する。このことから、シルエット距離のみを用いるよりも、深度距離を併せて用いた方が認証の精度が上がるのがわかる。 α が 0.35 のとき EER は最も低く 12.5% であった。このときの本人間距離と他人間距離の分布の割合を図 7 に示す。図 7 では 2 つの距離分布が広い範囲で重なっている。静止画像の場合、手形状が筆記中に変わるので選択したフレームによっては本人でも大きく異なる場合がある。また、他人でも偶然近くなる場合があり、認証精度の低下に繋がったと考えられる。

5. まとめ

ペンを持つ手の動き情報を用いた筆記者認証について検討した。動画像間の距離計算において、DP マッチングを用いて時系列の対応付けを行う手法を提案し、少人数の認証実験ではあるが EER 0.0% を達成した。また、距離計算には手のシルエット距離だけではなく深度距離も加えた方が認証精度が良く、最適な深度距離の重みは 0.35 であった。今後は被験者および登録データ数を増やして評価を行う予定である。

謝辞 本研究は JSPS 科研費 17K00275 の助成を受けて行った。
参考文献

- [1] 山中 晋爾, 浜本 隆之, 半谷 精一郎, “3 種類の筆記情報を用いる動的な署名照合,” 2000 年映情学年次大, 2000.
- [2] 村松 大吾, 橋本 侑樹, 小方 博之, “ペン持ち方特徴を用いたバイOMETリック個人認証,” 情処学 CVIM 研報, 2011-177, 2011.