

ユーザの入力の特徴を用いた個人識別

5 T - 5

石川 一郎[†] 河上 伸一^{††} 勅使河原 可海^{††}
 日本電子開発[†] 創価大学工学部^{††}

1. はじめに

今日、ネットワーク社会の進展に伴い、ネットワークアクセスに対して、確実に、簡易な個人認証の方法が求められている。

本論文は、コンピュータシステムのログインにIDとパスワードを用いる代わりに、ユーザの手書き図形の特徴を利用する試みについてのものである。手書き文字入力やサインによる筆跡照合の研究はすでに広く行われている。しかしこれらは、言語毎に様々なシステム上の制約を受けることが考えられる。このため本研究では、言語の違いに制約を受けない「図形」というシンボルに着目し、オンライン手書き図形認識による筆跡照合を個人識別に用いることを提案する。

2. 個人識別

「識別」と「認証」の関わりについては、「識別」は「認証」の前段階である。「識別」によって「認証」すべき対象を特定し、しかるのちに「認証」を行うからである。したがって、両者は互いに独立した処理であるが深い関連があるといえる。

3. 認証システムの提案

利用者のログイン要求があったときに、システムは鍵となる図形列の入力を求める。利用者はあらかじめ登録しておいた図形列を入力し、システムは事前登録されていた図形列と照合し、本人であると判断すればログインを認める。本人でないと判断された場合にはログインを拒否する。

このように図形列は二つの情報を持つ。一つはそれぞれの図形が持つ描画の特徴である。この特徴によって筆跡の照合を行う。もう一つは図形列の構成要素である。特徴が一致していても、異なる図形列を入力した場合には認証しない。システムはこの二つが同一であると判断した場合にのみ、本人であると認証する。

但し、現時点では開発の第一段階として図形列の長さは1とし、入力される図形は三角形に限定する。

4 システムの概要

システムの概略図を図1に示す。認証処理の中心となる認証エンジン部は、利用者の情報を格納しておくデータベースに結ばれている。

入力デバイスは描画動作すべてを数値化し、次か

ら次へとシステムに向けて送る。一時メモリは、入力デバイスから次々と流れてくる情報をたくわえて、必要に応じて前処理部に送り出す。前処理エンジンでは、連続していた描画データを図形ごとに切り分け、認証エンジンに渡す。認証エンジンでは、図形列の一致・不一致、図形ごとの特徴をデータベースと照合し、認証の可否を決する。

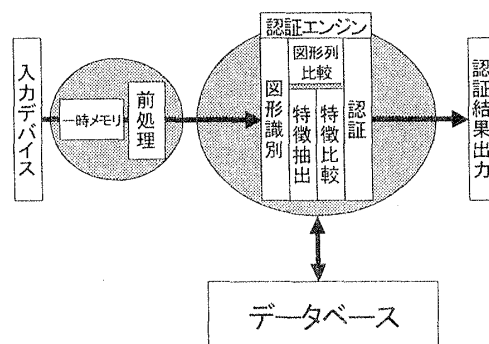


図1 システム概略図

5 描画データからの個人の特徴抽出

ポインタの軌跡を意味する描画データは、画面上の座標と時刻の3次元の座標を持つ。このデータから個人を識別できそうな特徴を設定し、どのようにして抽出するかを考える。今回は、個人の特徴が表れそうなものとして、形の歪みとストロークの速度変化の抽出方法について考えた。

5.1 形の歪み

手書き図形に現れる筆跡の描画の形状特徴は、各辺の歪み、角の鋭さ、面積などが考えられる。これらの特徴を測定するために、描画データに基づいて三角形の頂点を決め、頂点を結ぶ。このようにしてできた三角形を基準三角形と名づける。ここでは特に、辺の歪みに着目することにした。描画データが基準三角形に対し、どれだけ歪んでいるかを測定し特徴量とする。歪みを表す数値として、凹凸の個数、凹凸の大きさを考えた。

さらに、特徴値として、底辺をはさむ2つの角度を考えた。

5.2 ストロークの速度変化

ストローク(手書き軌跡)を、3つの辺ごとに分け、各辺ごとにストロークの速さの最大値、平均速度、さらに速さの変化量を加速度とみて最大加速度、最大減速度を特徴量とする。

さらに、各辺のストロークの軌跡が全ストロークにおいて占める割合を特徴量とする。

User Identification using Characteristics of User's inputted Drawing
 Ichiro Ishikawa*, Shinichi Kawakami**, Yoshimi Teshigawara**
 kawa@exch.nsl.t.soka.ac.jp

*Nippon Electronics Development Co. Ltd.

**Faculty of Engineering, Soka University

6 特徴の識別

特徴の識別に、マハラノビス距離を用いる。マハラノビス距離は、距離が等しければ確立密度が等しくなるように定義された距離である。

パターン Y が、 i 個の特徴を持つとき、その成分を、 Y_i とし、その分散を λ_i で表すと、平均 $(\mu_1, \mu_2, \dots, \mu_i)$ から、点 (Y_1, Y_2, \dots, Y_i) までのマハラノビス距離は、次のようになる。

$$D^2 = \frac{(Y_1 - \mu_1)^2}{\lambda_1} + \frac{(Y_2 - \mu_2)^2}{\lambda_2} + \dots + \frac{(Y_i - \mu_i)^2}{\lambda_i}$$

6.1 特徴データからの判別分析アルゴリズム

前項で述べたマハラノビス平方距離を用いて、特徴データが誰のものであるかを判別する。1対1照合と、1対 n 照合の2種類を考慮することができる。

(1) 1対1照合

誰かのものである1つの描画データを、あらかじめ登録しておいた1人分の特徴データに対して照合し、描画した者が登録された本人であるかどうかを判別する。登録されている特徴データの平均と分散を用いて、20個の描画データから20個のマハラノビス距離をそれぞれ求め、そのうちの最大値を、 D_{\max} とし、筆者 X による描画データのマハラノビス距離 D_X が、

$D_X \leq D_{\max}$ である場合に、本人であると判別する。

また、マハラノビス距離 $(D_1, D_2, \dots, D_{20})$ のうち、大きいほうから何番目の値を判別に用いるかによって、本人誤認率と他人許容率の変動する。1対1照合における結果を、表1に示す。

本人誤認率:

本人の描画データを本人と認識できない割合

本人正答率:

本人の描画データを本人と認識でき割合

(注) (本人誤認率) = 1 - (本人正答率) である。

他人許容率:

本人以外の描画データを本人の描画データであると誤認する割合

表1 本人誤認率と他人許容率の関係

本人誤認率	30%	20%	10%	5%	0%
平均(%)	8.452	12.86	19.64	24.17	28.10

(2) 1対 n 照合

1つの描画データを、データベースに存在するすべての人数分の特徴データと照合し、その中の誰によるものであるかを判別する。

A が描いた図形の持つ特徴の母集団を G_A とし、 B が描いた図形の持つ特徴の母集団を G_B とする。この二人によって図形が描かれたとき、その図形が持

つ特徴によって G_A 、 G_B のどちらに属するものかを判別しなければならない。この図形の特徴 (Y_1, Y_2, \dots, Y_p) から G_A 、 G_B までのマハラノビス距離をそれぞれ D_A 、 D_B とする。この D_A 、 D_B の大小関係を見ることによって判別を行う。

たとえば筆者 X による描画データのマハラノビス距離が

$$D_A > D_B$$

の場合には、その描画は B によるものであるとされる。つまり筆者 X は B であると判別される。

同様に、 p 人の場合には D_1, D_2, \dots, D_p のなかで、最も小さな値を持つ母集団に属することになる。注意しなければならないのは、この方法では、ある筆者の本人誤認が即、他の登録者にとっての他人許容となることである。さらに、「登録していない利用者」でさえも必ず「登録済みの誰か」であると判別されてしまう。1対 n 照合における結果を表2に示す。

表2 本人誤認率と他人許容率

	本人誤認率	他人許容率
平均	17.86%	2.976%

7 本提案システムの評価

評価結果から、提案する方式に基づく1対1照合では、本人誤認率20%のときに他人許容率が11%となりかなり低い数値を達成できた。また、同じく1対 n 照合では、本人誤認率18%のときに他人許容率が3%に抑えられるという結果となった。現状の本人誤認率や他人許容率の数値は決して実用に耐えるものではないが、これらの結果は当初予想していたよりはるかに良好な結果となり、本研究の有用性と実用化に対する可能性を示しており、さらに改良を重ね実用に向けることが可能であると考えられる。

8 今後の課題

今後の課題として、以下のものがある。

- 1) 特徴の類似度と図形の見た目の類似度の関係
- 2) 特徴の重み付け
- 3) 学習機能の付加
- 4) 異なる環境下での評価

参考文献

- [1] 電子商取引実証推進協議会, “本人認証技術検討WG報告書—評価基準(第1版)—”, http://www.ecom.or.jp/about_wg/wg06/h9doc/wg06-list.htm, 1998.