1K2-3

視線情報データマイニングを用いた個人認証システムに向けての基礎研究

Research of Personal Authentication System using Gaze Data Mining

八木 大介 Daisuke YAGI
清水 惇司 Atsushi SHIMIZU
松本 一教 Kazunori MATSUMOTO

神奈川工科大学大学院 工学研究科 情報工学専攻
Course of Information and Computer Sciences,
Graduate School of Engineering,
Kanagawa Institute of Technology

This paper presents fundamental considerations on a new personal authentication method based on eye tracking data. Several previous studies show eye tracking data reflect personal properties and can become a tool for the authentication. In most of such the studies, they collect eye tracking data over texts in English. Then the tracking data depend on knowledge about the texts and English. This means the tracking data over texts primarily has a difficulty in a use for the authentication. To avoid this disadvantage, we in this study use tracking data over simple diagrams instead of texts. We extract features from the raw tracking data, and apply machine learning algorithms over them. We demonstrate with experimental results the proposed method is applicable for the personal authentication.

1. まえがき

本研究の目標は，汎用性の高い機器を用いた個人認証を実現することである．現在，個人が所有するパソコン・スマートフォンなどの情報端末の普及率は 75%を超えており，1 人が 2,3 台保有していることも珍しくない．日常で情報端末を使用する機会が増えるにつれ，個人認証が必要となる場面が増えていると言える．一方，ユーザーが持つ個人認証に対する意識は低くなりがちである．例えば Web ブラウザにパスワードを記憶させる，パスワードの使い回しが一般的になるなど，個人認証が行われる意義は認知されておらず，なりすましなどの危険が高くなっている．そのため，より利用者の負担が少ない認証方式が求められている．利用者の負担が少ない方式として，指紋・静脈など個人の身体的特徴を用いた認証も普及が進められている．しかし身体的特徴を用いた認証は特徴を抽出するために専用機器が必要となり，機器管理者への負担が大きい．また，タイピングの得手不得手を用いて個人の癖を検出する方法も開発されている[Rick 90]．そこで本研究では，視線検出装置を用いた認証を開発する．その方法として単純な動作を行う際に出現する個人の癖をデータマイニングの理論を用いて分析する．分析した癖を用いて，知識の影響を受けない視線認証手法の開発，及び精度検証を行う．

2. 使用機材・先行研究

今回の研究では個人認証の方法として注視点データを利用する．注視点データとは被験者が画像・映像媒体などを確認する際に，最も注目が集まった点を時系列的にまとめたものである．最も注目した点であるために，単純な視界，視点とは異なる．今回の実験では Tobii 社製の注視点検出装置(Tobii Pro X2-30)を使用して注視点を計測した． 注視点を用いた個人認証の手法として，画面上にキーボードなどを表示し，注視によって単語入力を行わせる方法がある．しかし，この方式では単純にキーボードの代わりに注視点検出装置を用いているだけであるため，被験者がパスワードを覚える手間があるという本質的な問題は解決できない．また，被験者に英文読解やタイピングを行わせ，その際の注視点運動を用いて分析した単語注視順番や注視点のブレ，注視時間を用いて認証を行う手法がある．この方式では被験者は自然な動作をするだけで認証が行えるために，記憶的な負荷が少なく，模倣も起きにくいという利点がある．しかし，英文読解のパターンは個人の英語読解能力による影響が強いと考えられる．このため，長期間にわたって特徴を維持できず，認証には不適切である．また，この方式では英語に関する知識がないものは認証を行うことが出来ない．そこで本項では上記二つの問題を改善するために，知識を必要としない測定対象で，なおかつ汎用性があるものを利用する．その対象として，単純な図形な角を次々に注視する，という動作を測定対象とした個人認証方式について検討する．

連絡先：松本一教，神奈川工科大学大学院 情報工学専攻，
matumoto@ic.kanagawa-it.ac.jp

3. 分析実験

3.1 実験目的

視線情報を用いた認証システムを確立するために，被験者に注視動作を行わせる測定対象および分類に使用する特徴量についての検討を行う．今回作成する認証システムでは，個人が注視する際に発生する癖を元とした認証を行う．そのために癖が多く出現しやすい測定対象を選定することが必要となる．一方，癖が出現しやすい測定対象であっても，環境設定時や測定時に混乱を招くような測定対象は望ましくない．そのため，今回は一般的な正多角形を複数用意し，個々の測定対象での分類精度を計ることで使用する測定対象を選定した．正多角形のならば，単純な数値を用いて要素が表現できるために設定が簡易であり，また被験者が移動方向で混乱をすることも少ない．
今回は三角～六角の正多角形を三種のサイズで用意した．以後，サイズを小，中，大の三種で区別する．それぞれ直径 80, 120, 160mm の円に内接する大きさである．
また，使用する特徴量は安定して癖が出現しやすいものを選定する必要がある．癖が安定しない特徴量は認証エラーを招くため，精度が高くなる特徴量を選択しなければならない．そこで今回は，考えられるだけの特徴量を用意し，実際に個人分類を行う分類器を作成する．作成した分類器は交差検証法を用いて

精度検証を行う．精度が高くなる対象で使用される特徴量を用いて使用特徴量を決定する．今回の実験で想定する特徴量のリストを表 1 に示す．また，表 1 の固有名詞を具体的に図示したものを図 2 に示す．なお，表 1 の測定対象は全ての角及び辺について算出を行う．よって，分類に使用する特徴量の数は 8 ×角数となる．

表 1 使用する特徴量のリスト

番号	測定対象	特徴量の説明
1	辺	元画像の辺の角度と実際に移動した角度の差
2	辺	注視目標に対する角度と実際に移動した角度の差
3	辺	元画像の辺の長さと実際に移動した距離の差
4	辺	注視目標への距離と実際に移動した距離の差
5	角	注視目標に対する絶対角度
6	角	注視目標に対するX座標の差
7	角	注視目標に対するY座標の差
8	角	注視目標に対する絶対距離

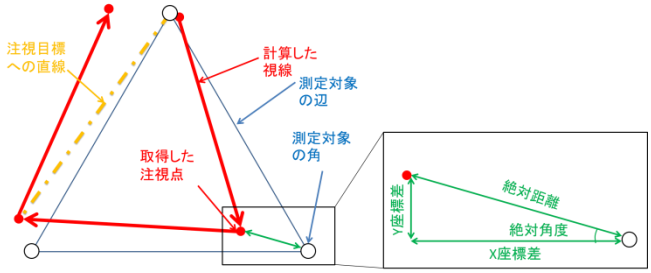

図 2 固有名詞の図示

表 1 の特徴量の内，個人の特徴が出現するものを判定する手法として，今回は決定木による分類を利用する．決定木による分類は，特徴量の値や属性などを用いて分類することで分析モデルを作成する手法である．決定木を作成する際は一つの特徴量のみを基準として分割を繰り返し，分類器を作成する．決定木の作成方法の例を図 3 に示す．

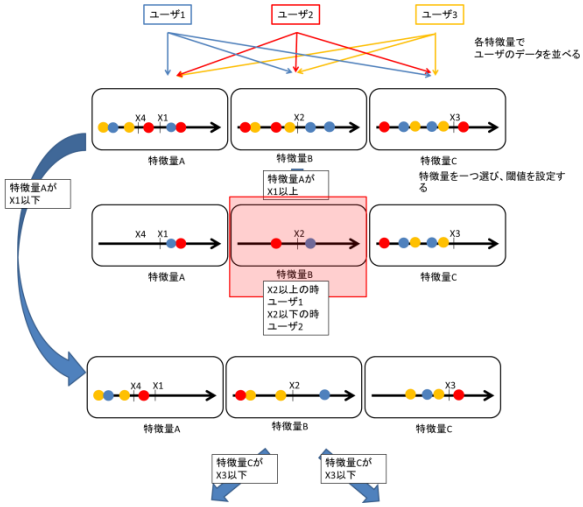

図 3 決定木の作成方法

図 3 のように決定木は特徴量の組み合わせによる効果を考えず，個々の特徴量で基準となる値を考え，分割を繰り返すことで作成される．そのため，分類の経過がわかりやすく重要な特徴量を判定する必要があるとき特に有用であるといえる．一方，未知データに対する分類精度自体は低くなる傾向がある．そのため，個人の特徴が出現する点を判定するのに向くが，実際に使用する分類器としては精度が劣る．以上の点から，今回は決定木を用いて個人の特徴が出現する特徴量を判別するが，実際の分類には異なるアルゴリズムを使用する．

3.2 実験手順

今回は，12 種の測定対象をランダムに並び替え，被験者に提示することで実験を行った．並び替えた画像セットを 10 個用意し，各被験者に提示した．被験者はそれぞれ提示された画像の角を時計回りに注視した．注視動作終了後，簡単なキーボード操作により次の画像が表示されるようにした．注視動作時間には制限をもたせず，また各画像セット間では自由に休憩を取れるようにした．分析実験での被験者は 20 代学生 8 名であった．実験終了後，得られた注視点データから表 1 の特徴量を算出し，各被験者名をクラスとする決定木を J48 アルゴリズムにて作成した．交差検証法を用いて決定木の精度を計算し，各測定対象での最大精度を検証した．

3.3 実験結果

分析実験で得た，各測定対象での最大精度を表 2 に示す．

表 2 各測定対象での最大分類精度[%]

	サイズ		
角数	小	中	大
3	46.66	47.45	57.04
4	52.63	61.40	41.38
5	32.20	53.45	34.00
6	62.50	52.11	36.84

表 2 の通り，分析実験全体での最大精度は小サイズ六角形の 62.50%であった．このことから，認証システムでは小サイズ六角形を使用する．また，被験者数は 8 名であるため，ランダムな分類では 12.50%となる．すなわち，どの測定対象を用いても，単純な分類よりも精度が高くなっているため，視線の癖を用いた個人識別が可能であるといえる．最大精度となった小サイズ六角形の決定木を図 4 に示す．図 4 にて分類基準として使用されている数字は，表 1 の通し番号に対応する． 本稿では詳細を省くが，表 2 の赤色で表示されている精度上位三種の測定対象における決定木を用いて特徴量の考察，選別を行った．

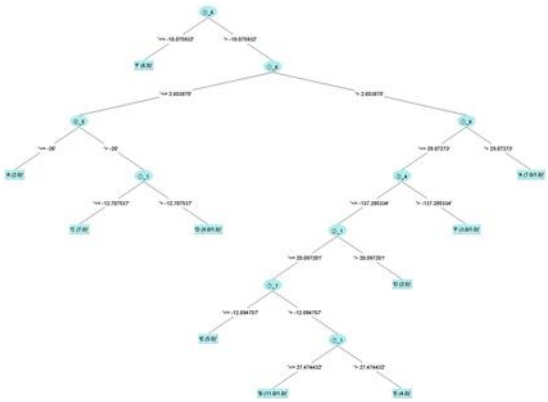

図 4 分析実験にて最大精度となった決定木

図 4 に示した通り，1，3，7 が多く使用されている．他 2 種の測定対象でも同様の傾向が見られた．このことから，今回の分析実験では元画像の辺に対する角度，長さの差，及び注視目標に対する Y 座標差に個人の特徴が出現することが判明した．

4. 認証システム

4.1 注視点の抽出

本システムでは各角周辺に出現する注視点を元に特徴量を計算するため，各注視点がどの角に属するか分類を行う必要がある．また，使用する特徴量は各注視点の位置関係によるものが存在するために，出現した注視点のうち，特定の点を除いて削減を行う必要がある．そこで本システムでは，図 5 のように注視点の所属の判定，及び削減を行うものとした．各角を中心とした円を従属範囲とし，各従属範囲で最初と最後に出現した注視点を除いて削減を行った．また，この円はユーザー入力によって直径を自在に変動できるものとした．

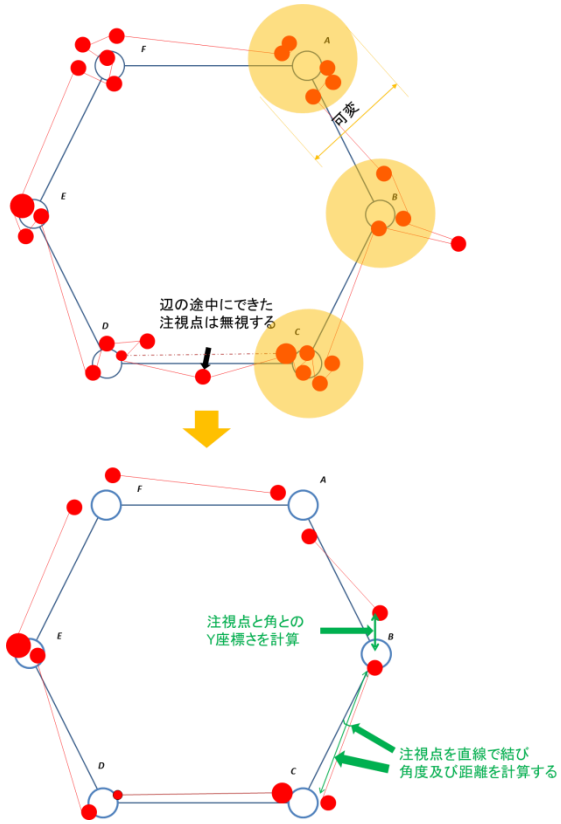

図 5 注視対象となる角の判定

図 5 のように本システムでは元となる図形の辺上に出現した線のみを分析対象とする．

4.2 認証アルゴリズム

今回作成したシステムは，まずアイトラッカーを用いて取得した注視点データを.tsv ファイルとして出力する Tobii Studio．出力した.tsv ファイルを元に注視点情報を読み取り，特徴量の計算を行う JavaGUI，GUI から入力された特徴量を元に決定木を作成・個人分類を行う Weka の三種ソフトウェアによって成り立つ．今回作成したシステムでの，ユーザーが視線情報を取得してから実際に分類を行うまでの流れを図 6 に示す．

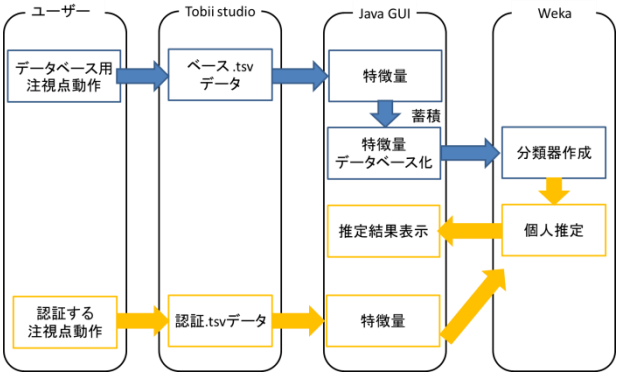

図 6 インタフェースデザイン

今回作成したシステムでは，不正認証に対応するために複数の分類器を用いて認証を行う．単一の分類器を用いた認証では認証の基準が単純になるために，模倣による不正認証が容易になる．そこで今回は複数分類器の判定結果を累計し，最も多くの分類器で判定されたユーザーを求めることで認証を行う．しかし，単純な多数決では判定精度が低い分類器の影響が強くなる．そのため，判定に適していない分類器を識別して，その影響が低くなるように調整を行う必要がある．今回は，幾度か個人分類テストを行うことで各分類器の判定精度を計算し，ユーザーに提示する機能を構築した．さらに，ユーザーが個々の分類器を使用するか，しないかについて自由に設定できる機能を作成した．これらの機能を用いることで，判定に適していない分類器を削減し，システム全体の精度向上を計ることが出来る．まず，ユーザーはシステムを数十回稼働させ，システムが判定したユーザー名と，実際に入力したユーザー名の成否の照合を繰り返し行う．システムは照合した結果を元に各分類器の判定確率を計算する．システムが各分類器の判定確率を計算する流れを図 7 に示す．

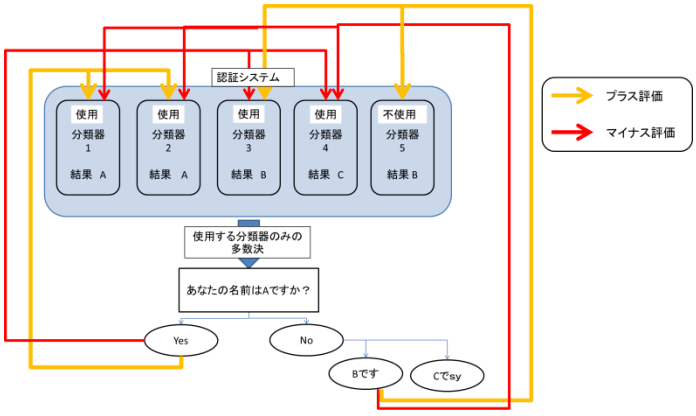

図 7 判定確率評価の流れ

今回作成したシステムにて，初期設定した分類アルゴリズムを表 3 に示す．

表 3 初期設定での分類アルゴリズム

番号	分類器名	weka ファンクション名
1	単純ベイズ分類器	weka.classifiers.bayes.bayesNet
2	サポートベクターマシン	weka.classifiers.functions.livsvm
3	パーセプトロン	weka.classifiers.functions.MultilayerPerceptron
4	k近傍法	weka.classifiers.lazy.IBk
5	決定木	weka.classifiers.trees.j48
6	ランダムフォレスト	weka.classifiers.trees.RandomForest
7	k-means	weka.clusterers.SimpleKmeans
8	ナイーブベイズ分類器	weka.classifiers.bayes.NaiveBayes
9	ナイーブベイズ決定木	weka.classifiers.trees.NBTree
10	ディシジョンスタンプ	weka.classifiers.trees.DecisionStump

なお，この分類アルゴリズムは Weka ファンクション内に存在するアルゴリズムであれば，自由に追加することが出来る．

4.3 精度評価

今回作成した個人認証システムの精度評価実験を行った．測定対象は分析実験にて最大精度が出現した，小サイズ六角形を用いた．被験者は分析実験と同様の視線運動を複数回計測し，そのうちシステムが処理できる 15 個の視線運動を得た．得た 15 個の視線運動データの内 10 個を分類基準データとして，システムに入力，表 3 の各分類アルゴリズムを作成した．アルゴリズムに入力し，実際に個人判定を行うデータは，残り 5 個の視線運動データを用いた．精度評価実験では，20 代学生 10 人の視線データを使用した．まず 50 個の個人判定データからランダムに1つ選んだ視線データを入力，判定する動作を繰り返し行なった．この動作を 10 回繰り返すごとに，分類精度 40%以下の分類器を判定から除外した．以上の操作を 5 回繰り返した後，全ての判定データを入力することでシステムの精度を測定した．測定結果を表 4 に示す．

表 4 判定したユーザーと入力の対応表

		システムが判定したユーザー									
		1	2	3	4	5	6	7	8	9	10
入力したユーザー	1	4	0	0	0	0	0	1	0	0	0
	2	0	4	0	0	0	0	0	0	0	1
	3	0	1	4	0	0	0	0	0	0	0
	4	0	0	0	5	0	0	0	0	0	0
	5	0	0	0	0	5	0	0	0	0	0
	6	0	0	1	0	0	3	0	0	0	1
	7	0	0	0	0	1	0	4	0	0	0
	8	0	0	0	0	0	0	0	5	0	0
	9	0	0	0	0	0	0	0	0	5	0
	10	0	1	0	0	0	1	0	1	0	2

表 4 では，入力したユーザーとシステムが判定したユーザーを数値で表している．個々の数値は判定が行われた回数を示している．そのため，入力したユーザーとシステムが判定したユーザーが一致している時の数値が高いほど正しい識別ができているといえる．表 4 の結果より，82%の精度で認証が行われたことがわかる．個々のユーザーに目を向けると，ユーザー6 や 10 のように精度が低いユーザーが存在する．一方，ユーザー6 やユーザー8 等のように高精度となるユーザーが存在する．このことから今回作成したシステムではユーザーによって精度のブレが大きくなることが判明した．また，最終的に使用された分類アルゴリズム，及び各アルゴリズムの判定確率を表 5 に示す．

表 5 各アルゴリズムの判定確率

番号	分類器名	認証精度[%]
3	パーセプトロン	78.00
4	k近傍法	60.00
5	決定木	74.00
6	ランダムフォレスト	78.00
9	ナイーブベイズ決定木	72.00

表 5 の通り，全てのアルゴリズムはシステム全体の正答率よりも低い精度を示している．このことより，複数アルゴリズムを使用することで判定エラーが起きにくい認証が実現できたことがわかる．また，精度が高くなる分類器はパーセプトロンやランダムフォレストなど，ノイズとなる特徴量があったとしても対応できる分類器であった．このことから，表 4 で得られたユーザーごとの認証率のズレも合わせて，今回使用した特徴量は全てのユーザー分類に対して有用なものでなく，ユーザーによっては分類ノイズとなることが考えられる．

5. むすび

今回の実験では先行研究に見られるような自由視線運動パターンを用いる分類手法でなく，個人が注目するとき，視線を動かすときに発生する癖を用いた分析を行った．その方式として，単純図形の角を定められた順番に注視するという注視点動作を対象とした，また，分類器の精度計測，選別することで，複数分類器を用いた模倣が起きにくいシステムを開発した．また，システム全体の認証精度は 82%となった．今後の課題として，今回作成したシステムの長期利用が挙げられる．今回は数日中に計測した視線データを元に分析を行ったが，月単位で間隔を開けた際の精度などを検証する必要がある．また，今回分類アルゴリズムは主に一般的なものを用いたが，アルゴリズムの精査を行うことでシステム分類精度を安定させる必要がある．

今後の展望としては，なりすまし等の防止のために，常に個人を識別するシステムの開発があげられる．そのために今回作成したシステムのように単純図形を対象とするものだけではなく，より複雑な測定対象に適応させ E-Learning などでパソコンを使用している最中，常に個人判定を行う方式を考える．

参考文献

[Eberz 15] Simon Eberz, Kasper B. Rasmussen, Vincent Lenders, Ivan Martinovic : Preventing Lunchtime Attacks: Fighting Insider Threats With Eye Movement Biometrics, (2015)

[Deravi 11] Farzin Deravi, Shivanand P. Guness, Gaze Trajectory as a Biometric Modality, Biosignals, pp335-341, (2011)

[Kumar 07] Manu Kumar, Tal Garfinkel, Dan Boneh, Terry Winograd : Reducing Shoulder-surfing by Using Gaze-based Password Entry, In Proceedings of the 3rd symposium on Usable privacy and security:pp.13-19, (2007)

[Rick 90] Joyce Rick, Gopal Gupta : Identity authentication based on keystroke latencies. Communications of the ACM, 33.2, pp168-176, (1990)

[Silver 08] Daniel L. Silver, Adam J. Biggs : Keystroke and Eye-Tracking Biometrics for User Identification, International Conference on Artificial Intelligence, pp344-348, (2008)

[Soumen 14] Roy Soumen, Utpal Roy, D. D. Sinha : Enhanced knowledge-based user authentication technique via keystroke dynamics, Int. J. Eng. Sci. Invention (IJESI) 3.9, pp41-48, (2014)