RL-005

日本語自由文書入力におけるキーストローク認証

Keystroke Dynamics in Japanese Free Text Typing

平岡 佑基* 石井昌樹† 佐村敏治‡ 西村治彦§

Yuki Hiraoka Masaki Ishii Toshiharu Samura Haruhiko Nishimura

1. はじめに

近年の情報化社会の発展にともない、コンピュータへの不正アクセスが急増しているなかで、新しい認証技術として生体認証 (バイオメトリクス) が注目を集めつつある。生体認証は、人間が保有する生体的な特徴を利用した認証技術であり、認証のキーとなる情報の忘失、盗難、偽造などの心配が少ないため安全性が非常に高い。

キーストロークダイナミクスは生体認証の 1 つであり、キーボードから入力するときのキーストロークデータに存在する固有のパターンを利用した認証である。ほかの生体認証と比べ、キーボードのほかに特別な装置を必要としないという長所がある。

キーストロークダイナミクスにおけるこれまでの研究の多くは、利用者がログインするときの認証を対象としていた[1–5]。本人のみが知る知識 (パスワード) だけでなく、同時にキーストロークダイナミクスを用いて認証するというものである。最近では、パスワードのような定型語ではなく、全く異なった文書を入力しても個人の特徴が捉えられるような、非定型文書におけるキーストロークダイナミクスの研究が行われている[6–22]。しかし、多くの実験はあらかじめ決まっている文書を入力し (以下、指定文書入力という)、被験者はどうしてもタイピングすることを意識してしまう。

そこで本研究では、被験者が自由に文書を入力 (以下、自由文書入力という) したときのキーストロークダイナミクスを対象とする。更に日本語文を対象として日本語文に特化した特徴量を提案して認証を行う。そして被験者がメールやレポートなどの日本語文書を自由に入力している際に、バックグラウンドでキーストロークデータを収集し、それを解析することで認証が行えるシステムを開発する。実験は 25 名の被験者を対象に行い、日本語自由文書入力におけるキーストロークダイナミクスの認証率を調べる。さらに、指定文書入力におけるキーストロークダイナミクスとの比較を行う。

2. キーストロークダイナミクス認証システム

本章では開発したキーストロークダイナミクス認証システムについて説明する (図 1)。本システムには、2 つのモードが存在する。1 つはプロファイル登録モードであり、もう 1 つが認証モードである。

プロファイル登録モードは、図 1 の実線で示す流れである。キーボードからの入力データをキーストロークデータとして収集し、それが日本語文であるかどうかを判定する。日本語文であると判定されれば、キーストロークデータから特徴量を抽出し、その特徴量をプロファイルとして登録者データベースに格納する。

認証モードは、図 1 の破線で示す流れである。プロファイル登録モードと同様に収集したキーストロークデータから特徴量を抽出し、入力者のプロファイルを作成する。この入力者のプロファイルと、事前に登録者データベースに登録されているプロファイルとを比較することにより、入力者の認証を行う。そして、入力者が登録者であると認証された場合は、入力者のプロファイルを登録者データベースに格納し、登録者でないとみなされれば、システム管理者へ通報したり、当該入力者を強制的にログアウトさせたりするなどの措置を行う。

本システムを実現するには、次節に詳述する 5 機能が必要である。

2·1 キーストロークデータ収集機能

キーボードからの入力から、3 種のキーストロークデータを生成する。キーストロークデータの例を図 2 に示す。第 1 フィールドのデータは、入力されたキーの種類である。第 2 フィールドのデータは、キーが押されたのか (press)、それとも離されたのか (release) を表す情報である。第 3 フィールドのデータは、キーストロークイベントが発生したときの時刻である。これは、システム時間 (UNIX 時間) を用いてミリ秒の単位で取得する。

2·2 日本語判定機能

日本語判定機能とは、収集したキーストロークデータが、日本語文であるかどうかを判定する機能である。日本語文を判定する方法として、ひらがなの多くが子音・母音ペアが多いことに着目する[21]。子音・母音ペアとは、子音が入力された直後に母音が入力される場合の文字列であり、ka や si、mu などが該当する。そこで、入力された文字数のうち子音・母音ペアが占める割合を示す CVP(Consonant Vowel Pair) 率を導入する。入力された文書の文字数を N、その中に含まれる子音・母音ペアの数を N_{CVP} としたとき、CVP 率 r_{CVP} を式 (1) のように定義する。

$$r_{\mathrm{CVP}} = \frac{2 \cdot N_{\mathrm{CVP}}}{N} \tag{1}$$

そして、アルファベット 50 文字分のキーストロークデータが収集されたところで CVP 率を計算し、CVP 率しきい値以上の場合、日本語文であると判定する。

ここで、CVP 率しきい値を決定するために、6 言語 (日本語文書 A，日本語文書 B，英語文書，フランス文語書，ドイツ語文書，C 言語文書) について検証文書をそれぞれ 10 文書

* 明石工業高等専門学校専攻科
† (株)NTT ネオメイト
‡ 明石工業高等専門学校電気情報工学科
§ 兵庫県立大学応用情報科学研究科

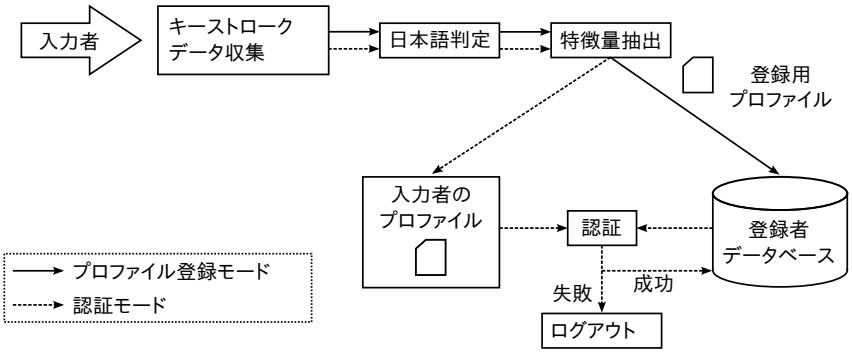

図 1　システムアーキテクチャ

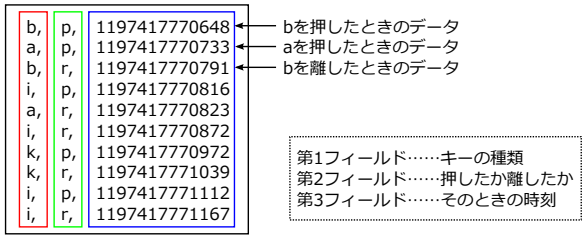

図 2　キーストロークデータの例

表 1　各言語における CVP 率

言語	CVP 率
日本語文書 A	0.834±0.0131
日本語文書 B	0.714±0.0192
英語文書	0.465±0.0204
フランス語文書	0.447±0.0217
ドイツ語文書	0.476±0.0253
C 言語文書	0.377±0.0828

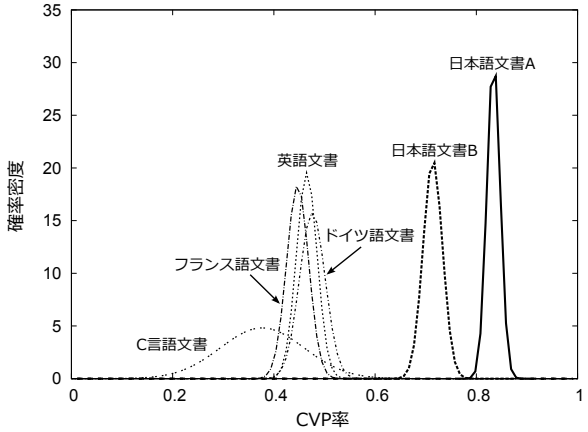

図 3　CVP 率の確率密度関数

ずつ用意し、CVP 率の平均値と標準偏差を計算した。結果を表 1 に示す。なお、日本語文書 A は書籍等から抜粋した文語的なものであり、日本語文書 B は電子掲示板等で用いられる口語的なものである。

さらに、正規分布に従う仮定すると、各言語における CVP 率の確率密度関数は図 3 のようになる。平均値 μ と標準偏差 σ の正規分布に従う確率密度関数では、ある観測値が $\mu \pm 3\sigma$ の範囲に入る確率は 99.7% である (3σ の法則)。日本語文書 A よりも CVP 率が低い日本語文書 B について、3σ の範囲に適用すると、0.657〜0.771 となり、日本語文書 B の次に CVP 率が高いドイツ語文書では、3σ の範囲は 0.401〜0.551 となる。以上の議論により、本研究では CVP 率のしきい値を 0.65 と設定する。

2·3　特徴量抽出機能

本節では、キーストロークデータからどのような特徴量を抽出するのかについて説明する [15,16,19–22]。キーストロークダイナミクスでは、あるキーを押す (press)、あるいは離す (release) というキーストロークイベントが発生したとき、それらの間の時間を特徴量として用いる。図 4 左に示すように、1 文字のキーストローク (1 文字打鍵) の場合は、ある 1 つのキーを押してから離すまでの時間を特徴量とする。この時間を押下時間 (*1pr*) と言い、その平均値 (*1pr.ave*) と標準偏差 (*1pr.sd*) が 1 文字打鍵における特徴量となる。特徴量抽出の対象となるキーは全てのアルファベットキーである。

2 文字のキーストローク (2 文字打鍵) の場合は、連続する 2 つのキーを押し離しするときのキーストロークイベント間の時間を考える。図 4 右に示すように、キーストロークイベント間の時間には 6 種類があるが、これらの時間の平均値 (*2xx.ave*) を 2 文字打鍵における特徴量とする。標準偏差についても特徴量の候補として考えられるが本研究では扱わない。標準偏差を特徴量とすると認証率を低下させてしまうことが、先行研究により示されていることによる [15,16,19,22]。特徴量抽出の対象となる 2 つのキーは、日本語文の特徴である子音・母音ペアと nn(ん) とする。

2·4　プロファイル登録機能

プロファイルは、あらかじめ設定した文字数 (50、100、200、300、400、500) ごとに作成する。これは、入力文字数ごとの認証率を調べるためである。また、1 人の被験者につき 1 つだけではなく、5 つ以上のプロファイルを登録する。

ここで、プロファイルを作成するとき、抽出した特徴量に次のような処理を行う。1 文字打鍵による特徴量において、

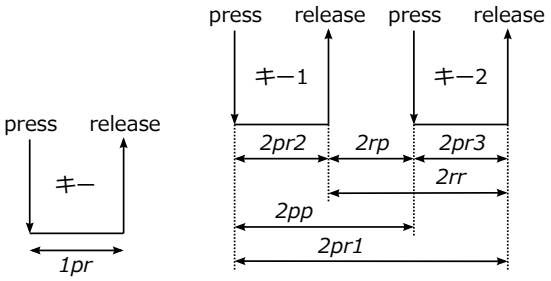

図 4　1 文字 (左) と 2 文字 (右) の打鍵測度

全てのアルファベットキーについて特徴量抽出を行うと、あまり使われなかったキー (c や v など) が 1 回や 2 回だけ出現することがある。このとき、出現回数の少ないキーが特徴を捉えきれず、認証率を低下させてしまう可能性がある。本研究では先行研究 [19] から、3 回以上出現するキーを対象とする。本システムにおいても 3 回以上出現しないキーはプロファイルから除外する。2 文字打鍵における特徴量についても同じ処理を行う。

さらに、ある特徴量を x とすると、式 (2) のように 0〜1 の範囲に標準化しておく。x' が標準化された特徴量であり、$x_{\min}$ と $x_{\max}$ は全てのプロファイルから得られた x の最小値と最大値である。

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{2}$$

2·5　認証機能

認証機能は、入力者のプロファイルと登録者データベースに事前に登録されているプロファイルとを比較し、入力者の認証を行う機能である。本システムでは、佐村らが提案している重みつきユークリッド距離 (WED) 法 [15,16]、Gunetti らが提案している Array Disorder(AD) 法 [8]、これら 2 つの手法を組み合わせたハイブリッド (HB) 法 [19,20] の 3 つの認証手法を使用する。本節では、これらの認証手法について解説する。

2·5·1　重みつきユークリッド距離 (WED) 法

重みつきユークリッド距離 (Weighted Euclidean Distance：WED) 法とは、入力プロファイルと登録プロファイルとの重みつきユークリッド距離を計算することにより、入力プロファイルの所有者を決定する認証手法である。入力プロファイルを docIN、登録者 A の 1 つめのプロファイルを docA1 と表すと、docIN と docA1 との WED(docIN,docA1) は式 (3) のようになる。

$$\mathrm{WED}(\mathrm{docIN}, \mathrm{docA1}) = \sqrt{\frac{1}{m}\sum_{\alpha=1}^{m}\frac{1}{n_\alpha}\sum_{i=1}^{n_\alpha}(k_{\alpha(i)} - r_{\alpha(i)})^2} \tag{3}$$

ただし、α は特徴量を表し、*1pr.ave*、*1pr.sd*、*2rr.ave* などに対応する。m は特徴量の種類数であり、本システムでは 8 種類である。次に $\alpha(i)$ は、docIN と docA1 のどちらにも存在する文字、すなわち比較する文字のうち、i 番目の文字における特徴量 α の値を表し、n_α はその種類数である。また、$k_{\alpha(i)}$ が docIN における $\alpha(i)$、$r_{\alpha(i)}$ が docA1 における $\alpha(i)$ を示す。ここで、docIN と docA1 のどちらかが持たない文字は欠損文字として扱い比較を行わない。なお、式 (3) の値は 0〜1 の範囲に規格化される。

2·5·2　Array Disorder(AD) 法

Array Disorder(AD) 法とは、入力プロファイルと登録プロファイルとの不揃度を計算することにより、入力プロファイルの所有者を決定する認証手法である。不揃度とは、特徴量の値を昇順に並び替えて文字を順位づけしたとき、入力プロファイルと登録プロファイルとの文字の並びがどれだけずれているかを表したものである。docIN と docA1 との不揃度 AD(docIN,docA1) を式 (4) および式 (5) に示す。

$$\mathrm{AD}(\mathrm{docIN}, \mathrm{docA1}) = \frac{1}{m}\sum_{\alpha=1}^{m}\frac{1}{\omega(n_\alpha)}\sum_{i=1}^{n_\alpha} \mid rk_{\alpha(i)} - rr_{\alpha(i)} \mid \tag{4}$$

$$\omega(n_\alpha) = \begin{cases} \frac{n_\alpha^2}{2} & (n_\alpha\text{：偶数}) \\ \frac{n_\alpha^2 - 1}{2} & (n_\alpha\text{：奇数}) \end{cases} \tag{5}$$

ただし、α、m、n_α は前節で述べた変数である。$rk_{\alpha(i)}$ は、docIN において特徴量 α での i 番目の文字の順位を示し、$rr_{\alpha(i)}$ は docA1 における順位である。WED 法と同様に、docIN と docA1 のどちらかが持たない文字は欠損文字と扱う。また、式 (4) の値に対しても 0〜1 の範囲に規格化される。

2·5·3　ハイブリッド (HB) 法

WED 法は絶対的な距離を用いており、AD 法は相対的な距離を用いているため、これらは異なる性質を持つ認証手法である。そこで、2 つの手法を組み合わせることによって、さらなる認証率の向上を図るのがハイブリッド (HB) 法である。HB 法では、2 つの手法で計算された距離を足し合わせて認証を行う。

3. 評価方法

本章では、キーストロークダイナミクスの評価方法について述べる。認証手順により、1 対 N 方式と 1 対 1 方式に分類できる。

3·1　1 対 N 方式

1 対 N 方式とは、入力プロファイルの所有者が、N 人のプロファイル登録者のうち誰なのかを正しく識別できるかどうかで評価を行う方式である。1 対 N 方式では、入力プロファイルは全ての登録プロファイルと比較される。そこで、入力プロファイル docIN と、登録者 A のプロファイル docA1〜docA5 との距離 d(docIN,docA1)〜d(docIN,docA5) が、認証手法により求められる。そして、式 (6) のように入力プロファイルと登録者 A との平均距離 md(IN,A) を求める。

$$\begin{aligned}\mathrm{md}(\mathrm{IN}, \mathrm{A}) = [&\mathrm{d}(\mathrm{docIN}, \mathrm{docA1}) + \mathrm{d}(\mathrm{docIN}, \mathrm{docA2}) + \\ &\cdots + \mathrm{d}(\mathrm{docIN}, \mathrm{docA5})]/5\end{aligned} \tag{6}$$

この平均距離を全ての登録者について算出し、最近傍決定則により入力プロファイルの所有者を決定する。1 対 N 方式では、評価指標として識別成功率で評価する。識別成功率の計算には leave-one-out クロスバリデーション法を採用する。まず、全ての登録プロファイルから 1 つのプロファイ

ルを取り出し、それをテストファイルとする。次に、テストファイルと全ての登録者との平均距離を求め、最近傍決定則により認証成功 (識別成功) か認証失敗 (識別失敗) かを判断する。そして、この一連の操作を全ての登録プロファイルがテストファイルとなるまで行い、(識別成功数/全テストファイル数)×100 により認証率 (識別成功率) を求める。

3·2 1 対 1 方式

1 対 1 方式とは、入力者の宣言を正しく認証できるかどうかで評価する方式である [8]。仮に入力者が登録者 A であると宣言したとする。まず、入力プロファイル docIN と登録者 A との平均距離 md(IN,A) は他の登録者との平均距離より最短距離とする。もし md(IN,A) より距離の短い登録者が見つかれば「入力者は登録者 A でない」と判断する。次に、登録者 A のプロファイル (docA1〜docA5) から、式 (7) のような登録者 A の平均距離 m(A) を求める。

$$\mathrm{m(A)} = [\mathrm{d(docA1, docA2)} + \mathrm{d(docA1, docA3)} + \cdots + \mathrm{d(docA4, docA5)}]/10 \quad (7)$$

m(A) を 1 対 1 方式におけるしきい値とし、md(IN,A) が、m(A) よりも小さい場合は、「入力者は登録者 A である」と判定する (式 (8))。

$$\mathrm{md(IN, A)} < \mathrm{m(A)} \quad (8)$$

一方、md(IN,A) が md(IN,X) よりも大きい場合は、式 (9) を満たせば「入力者は登録者 A である」とし、それ以外は「入力者は登録者 A でない」と判定する。

$$\mathrm{md(IN, A)} - \mathrm{m(A)} < \mathrm{md(IN, X)} - \mathrm{md(IN, A)} \quad (9)$$

ここで、md(IN,X) は md(IN,A) の次に小さい平均距離 (登録者を X とする) を示す。

1 対 1 方式では、評価指標として本人拒否率 (FRR：False Rejection Rate) と他人受入率 (FAR：False Acceptance Rate) が用いられる。

N 人の登録者が 5 個のプロファイルをもっているとき、本人拒否率の計算は次の方法で行う。まず、leave-one-out クロスバリデーション法に従い、全ての登録プロファイルから 1 つを取り出して、それをテストファイルとする。次に、上記の条件を満たせば認証成功 (本人受入) とし、満たさなければ認証失敗 (本人拒否) とする。そして、これを全ての登録プロファイルがテストファイルとなるまで行い、(本人拒否数/全テストファイル数)×100 により本人拒否率を求める。全テストファイル数は $\mathrm{N} \times 5$ である。

一方、他人受入率の計算は次の方法で行う。他人受入率を計算するときには、入力プロファイル docIN は外部からの侵入者が入力したものとして扱う。そこで、全ての登録プロファイルから 1 つを取り出して、それをテストファイルとしたとき、このテストファイルを侵入者とするために、テストファイルの所有者の他のプロファイルを登録プロファイルから一時的に削除する。そして、残りのプロファイルを対象に順次比較を行い、上記の条件を満たせば認証失敗 (他人受入) とし、満たさなければ認証成功 (他人拒否) とする。この操作を全ての登録プロファイルがテストファイルとなるまで行い、(他人受入数/全テストファイル数)×100 により他人受入率を求める。ここで、1 つのテストファイルは $(\mathrm{N} - 1)$ 人のふりをして不正侵入を試みるため、全テストファイル数は $\mathrm{N} \times 5 \times (\mathrm{N} - 1)$ となる。

4. 実験結果

本研究では 2 つの実験を行った。1 つ目の実験は 25 名の被験者を対象に行い、自由文書入力におけるキーストロークダイナミクスの認証率を入力文字数ごとに調べた。1 対 N 方式で評価を行った結果を図 5 に示す。まず、入力文字数の増加にともない認証率は向上している。また、WED 法と AD 法を組み合わせた HB 法では、いずれの入力文字数でもほかの 2 つの認証手法と同等以上の認証率が得られており、異なる認証手法を組み合わせることにより認証率を向上させられることが示された。

入力文字数の増加にともない認証率が向上する原因について考察する。図 6 に、1 文字打鍵と 2 文字打鍵の特徴量における比較文字種数の平均 $\overline{n_1}$、$\overline{n_2}$ を示す。入力文字数が多くなると、$\overline{n_1}$、$\overline{n_2}$ ともに増加することがわかる。すなわち、入力文字数が増加すると、n_α が増加して多くの特徴量を用いることができるため認証率が向上すると考えられる。

続いて、HB 法により 1 対 1 方式で評価を行った結果について、本人拒否率 (FRR) と他人受入率 (FAR) に対する入力

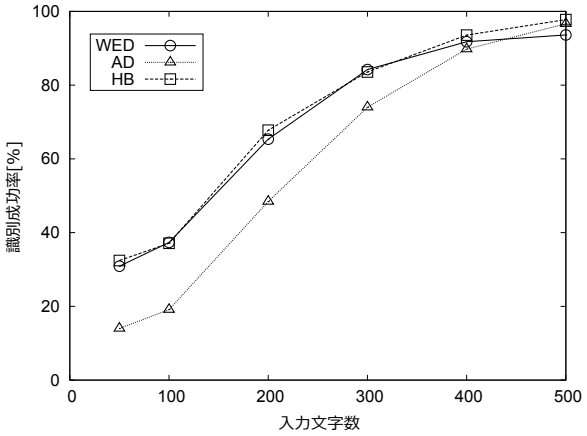

図 5 1 対 N 方式による認証率の入力文字数依存性 (自由文書入力)

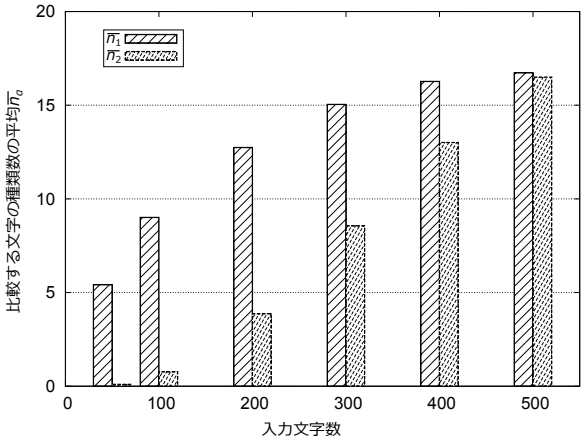

図 6 平均比較文字種数 $\overline{n_1}$、$\overline{n_2}$ の入力文字数依存性

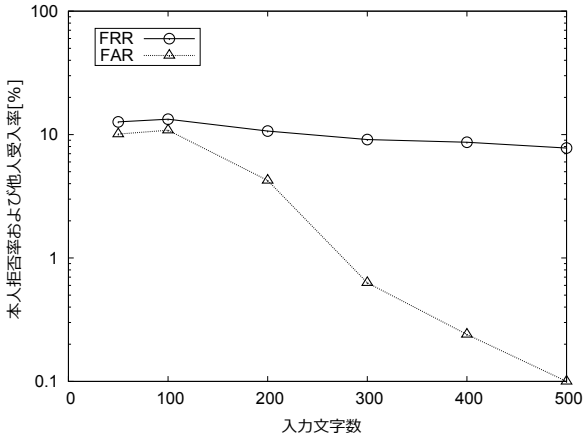

図 7　1 対 1 方式による認証率 (FRR,FAR) の入力文字数依存性 (自由文書入力)

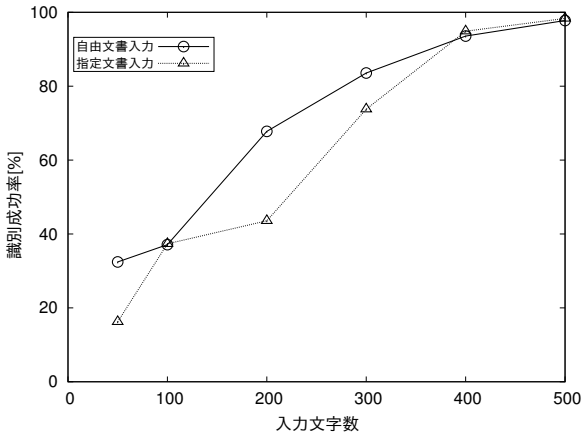

図 8　1 対 N 方式による認証率の入力文字数依存性 (自由文書入力および指定文書入力)

文字数への依存性を図 7 に示す。本人拒否率 (FRR) は入力文字数の増加に依存しない (8%〜12%) ことがわかる。一方、他人受入率 (FAR) は入力文字数の増加にともない大幅に向上 (FAR は減少) していることがわかる。

2 つめの実験として、指定文書入力における認証実験を被験者 25 名を対象に行い、実験 1 の結果と比較した。HB 法による 1 対 N 方式での結果を図 8 に、1 対 1 方式 (他人受入率) での結果を図 9 に示す。入力文字数が 50 文字から 300 文字までは、1 対 N 方式においても 1 対 1 方式においても自由文書入力での認証率が若干高くなっているが、入力文字数が 400 文字や 500 文字になると大差がないことがわかる。つまり、指定文書入力と自由文書入力のどちらでも、400 文字や 500 文字程度入力すれば、同程度の認証率が得られることがわかる。これは、指定文書入力での実験による先行研究[15,16,19,20,22] の結果が、自由文書入力にも適用できることを示唆している。また、自由文書入力のほうが入力文字数の変化に対して安定である。理由として、自由文書入力は本人が考えた文字を入力するため、同じフレーズや語が早い段階で登場し、それが比較文字種数の向上に役立つためではないかと考えられる。

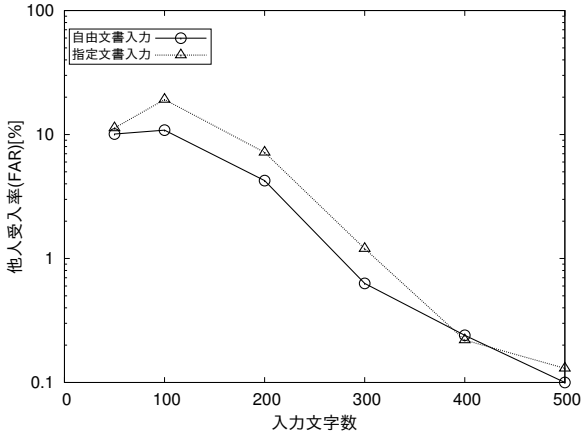

図 9　1 対 1 方式による他人受入率の入力文字数依存性 (自由文書入力および指定文書入力)

5. おわりに

本研究では、被験者が自由に文書を入力するときのキーストロークデータを対象とした、自由文書入力におけるキーストロークダイナミクスを解析した。実験は 25 名の被験者を対象に行い、自由文書入力におけるキーストロークダイナミクスの特性を解析した。

結果として、次のことが明らかになった。(1) 入力文字数の増加にともない識別成功率と他人受入率が向上するが、本人拒否率は大きな変化が見られなかった。(2)WED 法と AD 法を組み合わせた HB 法で、単独の方法よりも高い認証率を得た。(3) 入力文字数が 400 文字や 500 文字になると大きな差はない。以上の実験結果は、先行研究における指定文書入力での結果が自由文書入力にも適用できることを示唆している。

本研究では日本語入力におけるキーストローク認証における基本的な特徴として子音・母音に表れるアルファベット文字に着目し、日本語判定機能についてもこれらに登場するアルファベット文字のみ扱っている。一方、その他のキー（スペースキー、リターンキー等）や拗促音入力については取り扱っていない。これらのキーや入力についても認証率を向上させる特徴量となる可能性があるので今後の研究で検討していきたい。

謝辞

本研究の遂行に際して明石工業高等専門学校電気情報工学科卒業生の赤井優真氏、及び認証実験に協力していただいた皆様に深く感謝する。

参考文献

1) R. Joyce and G. Gupta: User authorization based on keystroke latencies; *Communications of the ACM*, 33-2, pp.168-176 (1990)
2) 粕川, 森, 小松, 赤池, 角田：打鍵データに基づく個人認証システムと改良, 情報処理学会論文誌, 33-5, pp.728-735 (1992)
3) 佐村, 高岡, 柴田, 西野, 小高, 小倉：打鍵データの特性を生かした個人認証システム, 福井工業大学研究紀要, 第二部, No. 29, pp.305-312 (1999)

4) 小谷, 法岡, 堀井：テンキーパネルを用いた打鍵認証システムの構築と評価，ヒューマンインタフェース学会論文誌，7-1，pp.149-156 (2005)
5) 山村，鈴木，ラシキア：打鍵署名を利用したパスワード認証の強化について，情報処理学会研究報告，CSEC，2009-20，pp.79-84 (2009)
6) F. Monrose and A. Rubin: Keystroke dynamics as a biometric for authentication, Future Generation Computer Systems, 16, pp.351-359 (2000)
7) F. Bergadano, D. Gunetti and C. Picardi: User authentication through keystroke dynamics, ACM Trans. Inf. Syst. Secur., 5-4, pp.367-397 (2002)
8) D. Gunetti and C. Picardi: Keystroke Analysis of Free Text,ACM Trans. Inf. Syst. Secur., 8-3, pp.312-347 (2005)
9) M. Curtin, C.C. Tappert, M. Villani, G. Ngo, J. Simone: H. St. Fort, and S. Cha, Keystroke Biometric Recognition on Long-Text Input: A Feasibility Study, Proc. Int. Workshop Sci Comp/Comp Stat (IWSCCS 2006), (2006)
10) 倉橋，田中，小松： スコアの累積値を用いたキーストロークダイナミクスによる個人認証手段，電子情報通信学会技術研究報告，IE，106-244，pp.35-40 (2006)
11) 石川, 岡, 加藤: キーストロークデータを対象としたコンテクスト抽出手法の提案, 電子情報通信学会技術研究報告, CSEC, 2006-81，pp.419-426 (2006)
12) 佐村，西村：キー押下時間に着目したキーストロークダイナミックス解析, 第 5 回情報科学技術フォーラム (FIT2006) 講演論文集, pp.459-462 (2006)
13) 片岡，宮本，青木，泉： キーストロークの統計情報を利用した個人認証手法の提案，電子情報通信学会技術研究報告，SEC，107-140，pp.23-30 (2007)
14) 宮崎，赤堀: 異種特徴量複合型キーボード・ダイナミクスによる個人認証，日本 e-Learning 学会論文誌，44-2，pp.80-89 (2008)
15) 佐村，西村: 非定型な日本語文入力におけるキーストロークダイナミクス識別，システム制御情報学会論文誌，22-4，pp.145-153 (2009)
16) T.Samura, H.Nishimura: Keystroke Timing Analysis for Individual Identification in Japanese Free Text Typing, Proc. ICCAS-SICE 2009, pp.3166-3170 (2009)
17) C.C. Tappert, M. Villani, and S. Cha: Keystroke Biometric Identification and Authentication on Long-Text Input, Behavioral Biometrics for Human Identification: Intelligent Applications, IGI global, pp.342-367 (2009)
18) 櫻井, 宮本, 青木, 岩田, 汐崎：日本語非定型文のキーストローク特性に着目した個人識別，電気学会研究会資料. IS, 情報システム研究会, 2010(21), pp. 19-24 (2010)
19) 佐村，西村: 日本語非定型文入力のキーストロークデータに基づく個人識別ハイブリッドモデル，計測自動制御学会論文集，Vol.46，No.11．pp.676-684 (2010)
20) 佐村，西村:非定型文入力のキーストロークダイナミクスのおける 1 対 1 認証，計測自動制御学会 SSI2010 講演論文集，2F2-2 (2010)
21) 石井，佐村，西村：自由文書入力のキーストロークダイナミクスによる持続的認証システム，計測自動制御学会システム・情報部門学術講演会 2010(SSI) 講演論文集，2F3-1 (2010)
22) T. Samura and H. Nishimura, Keystroke Dynamics for Individual Identification in Japanese Free Text Typing, SICE JCMSI, Vol. 4, No. 2, pp.172-176 (2011)