

修 士 論 文

線形予測誤差を用いた  
オンライン署名照合

平成 23 年度修了

三重大学大学院工学研究科

博士前期課程 情報工学専攻

中西 真一郎

# はじめに

近年の情報化社会の発展により多様化する犯罪を防ぐため、企業や個人でのセキュリティ対策が必須となっている。セキュリティには様々なものがあるが、その中でも当人かどうかを判断する個人認証が極めて重要であり、これにより多くの犯罪を防ぐことが可能となる。

個人認証は大きく3つに分けられる。印鑑やカード、鍵などによる所有物認証、暗証番号やパスワードなどによる知識認証、指紋や虹彩、顔、手書き署名などによる生体認証である。これらの中で生体認証はバイオメトリクスと呼ばれ、近年特に注目されている。所有物認証は偽造や盗難の危険性があり、知識認証は忘却の可能性もある。それらに比べ、本人の身体的特徴や行動的特徴を利用したバイオメトリクスは盗難や偽造の心配が少なく、記憶する必要もない。

本研究ではバイオメトリクスの1つである手書きオンライン署名照合について研究を行っている。手書きオンライン署名照合は従来社会的に受け入れられている認証方法であり、指紋などのバイオメトリクスに比べ利用者の受容性が高い。また、身体的特徴を用いる個人認証では登録データが盗難されればデータの変更は不可能であるが、手書きオンライン署名では登録データを変更することも可能である。このような特徴から、手書き署名による個人認証はバイオメトリクスの中でも重要な技術の1つであると言える。

オンライン署名照合は、既に書かれた署名には利用できないが、筆速や筆圧など多くの情報が得られる。そして、署名時にペンやタブレットなどの専用機器を必要とするが、近年はiPadなどに代表されるようにコンピュータの標準入力装置の利用環境が整ってきていることから、実用化が進められている。しかし、他の個人認証に比べ認識率において劣るため、さらなる精度向上が求められている。

そこで本研究では、線形予測誤差を署名照合に利用し、手書き署名照合における線形予測分析の有用性を検証した。線形予測とは、離散信号の将来の値をそれまでの標本群の線型写像として予測する数学的操作であり、音響信号処理および音声処理の分野でも波形解析手法としてよく用いられる。デジタル信号処理では、線形予測法を線形予測符号(LPC)と呼び、デジタルフィルタのサブセットと見ることができる。この線形予測分析の考え方

の一部である線形予測誤差を利用する。

本実験はタブレット PC により取得した 19 人分の署名，合計 1425 個を用いる．学習用データとして真筆署名を 1 人につき 3 個ずつ，合計 57 個を使用．評価用データとして，真筆署名を 1 人につき 36 個ずつ，合計 684 個，偽筆署名を 1 人につき他人 4 人が書いたものを各 9 個ずつ，合計 684 個を使用．各人の署名から得られる時系列データ，署名の形 ( $x$  座標,  $y$  座標)，ペンの接触・非接触情報，筆速，筆記時間を用いて算出された線形予測誤差を利用することで，最高 81.50% の照合率を得た．

今後の課題として，より多くの署名データでの実験，本手法の精度向上が挙げられる．さらに，本手法と従来手法とを組み合わせることにより，さらなる署名照合の精度向上が期待できる．

# 目次

はじめに	i
第 1 章 署名照合	1
1.1 手書き署名照合	3
1.2 本研究の目的	4
1.3 本論文の構成	4
第 2 章 署名データと提案手法	5
2.1 手書き署名データ	5
2.2 データの正規化	7
2.3 照合手順	8
2.4 線形予測誤差	11
2.5 真偽判定	16
第 3 章 実験	17
3.1 実験データ	18
3.2 照合率の定義	19
3.3 実験 1	21
3.4 実験 2	23
3.5 実験 3	25
3.6 実験 4	27
3.7 実験 5	29
3.8 実験 6	31
3.9 考察	33
第 4 章 まとめと今後の課題	34
4.1 まとめ	34
4.2 今後の課題	34

---

付録 A	実験データとプログラム	35
付録 B	プログラムソースリスト	37
謝辞		39
参考文献		40

# 第 1 章

## 署名照合

本人かどうか確認することは，クレジットカードの利用，特定場所の入退室など社会生活のいたるところで必要であり，個人情報の保護や犯罪防止のためにも，重要な問題である．本人認証手法として，以下のものがある．

- 本人が持つ知識による認証：パスワード，暗証番号．
- 本人の所有物による認証：IC カード，印鑑．
- 本人の身体的特徴や行動的特徴による認証：署名，指紋，音声．

本人しか知り得ない，本人しか所有していないものによる認証は，他人に盗まれたり，なくしたりすることにより，他人に悪用される危険性がある．それらに比べ，本人の身体的特徴や行動的特徴を用いた認証（バイオメトリクス）は，盗難，紛失，忘れる等の心配が非常に小さい．つまり他人による代行が困難であることから，近年特に注目されている．

バイオメトリクスは大きく二つに分類できる．一つは指紋，虹彩，顔，静脈などの身体的特徴であり，もう一つは声紋，署名などの行動的特徴である．主なバイオメトリクスとその特徴を表 1.1 に示す [1]．ここで普遍性とは誰もが持っている特徴であること，唯一性とは本人以外は同じ特徴を持たないこと，永続性とは時間の経過とともに変化しないことである．また，拒否率とは本人を間違って拒否する誤りであり，受入率とは他人を間違って受け入れる誤りである．表を見ると，指紋を用いた認証が良いように思われるが，犯罪捜査などに利用されていることから，登録において心理的な抵抗感があると言える [2]．

このように，バイオメトリクスにはそれぞれ一長一短があり，要求されるセキュリティレベルや利用方法に応じた選択が必要となる．本研究で取り上げる署名は，以下のような特徴がある．

- クレジットカード利用時など，従来より社会的に受け入れられている個人認証方法

表 1.1: バイオメトリクスの比較

情報	特徴量	普遍性	唯一性	永続性	コスト	拒否率 (%)	受入率 (%)	データ量 (Byte)
指紋	手指の指紋特等量	◎	◎	◎	◎	1.0	0.01	250
掌形	手の大きさ・長さ・比率	◎	○	○	△	0.1	0.1	10
顔	顔輪郭・目・鼻の形・配置	◎	△	△	○	5	5	2,000
虹彩	虹彩の模様	◎	◎	◎	◎	10	$10^{-6}$	200
静脈	手指の静脈特徴点	◎	○	○	△	1.0	0.01	500
声紋	話者の音声特徴	◎	△	△	◎	10	10	1,500
署名	字体・書き順・筆速	◎	△	△	○	5	5	1,000

であり、他のバイオメトリクスに比べて利用者の受容性が高い。

- 行動的特徴による個人認証であるため、本人の意志確認を伴う。
- 永続性の高い身体的特徴の登録データとは違い、署名は万が一、流出や漏洩が生じても登録データそのものを変更することが可能である。
- 電子文書への貼り付けが可能である。

このような特徴を持つことから、署名による個人認証はバイオメトリクスの中でも重要な技術の一つである [3]。

## 1.1 手書き署名照合

署名照合は、手書きの署名を用いて真筆（本人が書いた本物の署名）であるか偽筆（他人が書いた偽物の署名）であるかを判断し、署名した人が登録されている本人であるかどうかを判定する個人認証の1つである。認証に手書き署名を用いるため、記憶する必要や紛失の心配がない。また、全く同じ署名を書くことはありえないため、過去に書かれたものと同じ署名は破棄するようにすれば、署名データの盗難も防ぐことができる。万が一盗まれたとしても、登録署名の変更が可能であるため漏洩対策の面でも有効である。

### 1.1.1 オンライン情報とオフライン情報

署名照合方式には、オンライン署名照合とオフライン署名照合があり、それぞれの特徴は以下の通りである。

- オンライン署名照合
  - － 照合に用いる特徴：筆跡、筆速、筆圧、筆順。
  - － 専用機器を用いるため、使用環境の制約がある。
  - － 既に紙などに書かれた署名（書類や小切手など）は照合できない。
- オフライン署名照合
  - － 照合に用いる特徴：筆跡。
  - － 専用機器を用いないため、使用環境の制約が小さい。
  - － 過去の署名も照合できる。

オフライン署名照合とは、既に書かれた署名から得られる筆跡など形態情報を利用したものである。この場合、文字を特定しない筆者証明として古くから筆跡鑑定があり、脅迫状や遺言状などにおける特定筆者の証明を行う技術である。

一方、オンライン署名照合とは、タブレットなどの専用機器より取得できる筆順、筆圧、筆速など署名をしているときの筆記運動情報を利用するものである。この場合、専用機器としてタブレットや電子ペンが必要になる。最近ではタブレット PC、iPad のようなタブレット端末、スマートフォン、任天堂 DS のようなゲーム機など、ペン入力さまざまな場面で利用可能となってきた。これらの普及により、実装もより容易になってきており、比較的低コストで実装可能である。

また、オフライン情報よりも得られる情報が多く、筆速やペンの傾きなどは、他人が容易に真似ることが困難であると考えられる。



現実においてどちらの情報が利用できるかは状況次第であるが、オンライン情報の取得環境が整いつつあることと、得られる情報の多さより、本研究ではオンライン署名照合を研究対象とする。

### 1.1.2 従来研究

オンライン署名照合に関して、署名文化のある欧米で研究が盛んに行われている [4]。同様に、日本語署名を対象にした研究も数多く発表されている [5]。そのほとんどが、筆跡の動的情報であるペンの座標、筆速、筆圧、傾きなどの時系列データに対して DP (Dynamic Programming) マッチングを行う手法 [6][7][8]、セグメントごとの曲率や向きなどの局所的特徴の時系列を隠れマルコフモデルにより照合を行う手法 [9] など、筆跡を時間関数と捉えて照合するものが主流である [3]。

最近では、濃度こう配特徴によりオンラインとオフライン双方の特徴を総合的に用いることで照合を行う手法 [10]、ユーザ共通の Fusion モデルにより複数の距離を組み合わせることで照合を行う手法 [11]、動的情報の時間変化を離散ウェーブレット変換によりサブバンド分解した信号を特徴とし、適応信号処理技術を応用し照合する方法 [12] が提案されている。

## 1.2 本研究の目的

1.1.1 で述べたことから、署名照合においては今後オンライン照合の発展が期待されている。そこで本研究では、オンライン情報から  $x, y$  座標データ、ペンの接触・非接触情報（ペンアップデータ）、時刻データを利用し、波形解析手法である線形予測分析（線形予測誤差）による署名照合の有用性を検証することで、新たな視点から署名照合精度を向上させる手法を提案する。ここで、ペンアップデータとは、ペンがタブレット PC に接触しているかどうかを表すデータであり、時刻データとは、署名開始からの経過時間のデータである。

さらに、将来的に他の署名照合手法と併用することで、精度の向上が見込まれ、これにより署名照合による個人認証はさらに大きな可能性を持つこととなる。

## 1.3 本論文の構成

第 2 章では、処理の流れ、各処理の詳細について説明する。第 3 章では、実験方法、結果、考察を述べる。第 4 章では、研究のまとめと今後の課題について述べる。

## 第 2 章

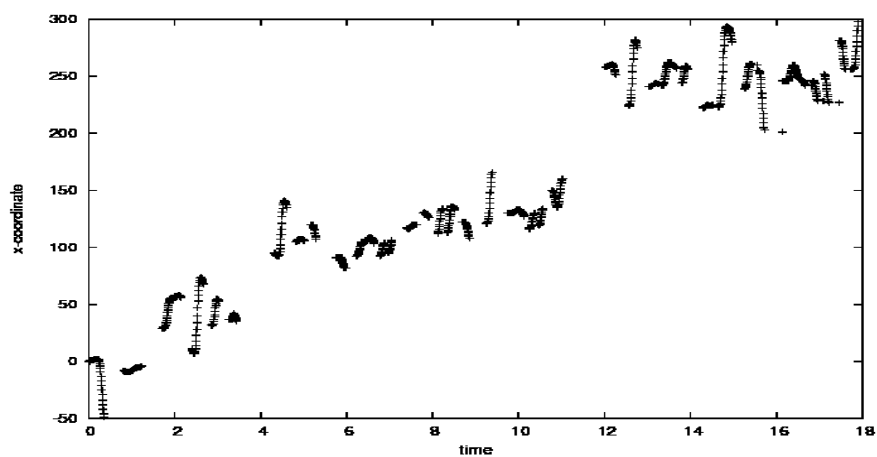
# 署名データと提案手法

本章では，本研究における署名データと署名照合の概要を以下に示す．

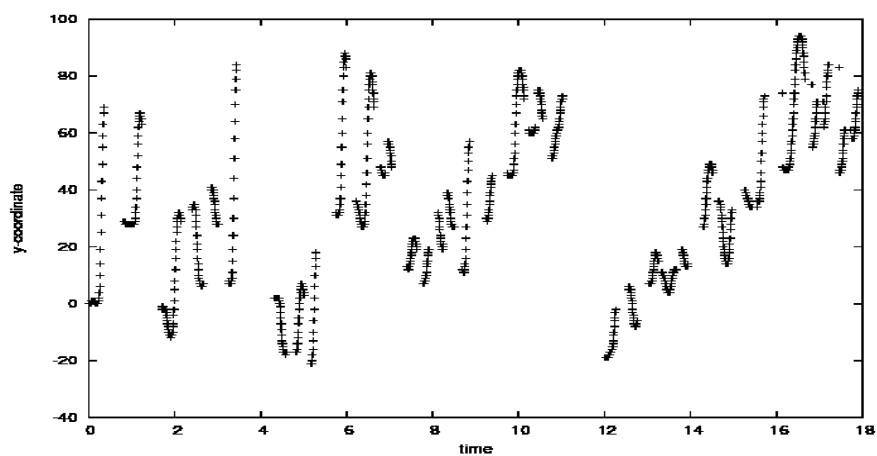
### 2.1 手書き署名データ

本実験では，タブレット PC を用いて  $600 \times 300$  pixel の枠内に書かれた手書き署名データを用いる．時系列点座標データは 5 種類の観測値を持ち，時刻  $t$  における時系列点座標データは  $P_i(x_i, y_i, z_i, v_i, t_i)$  と表すことができる． $P_t$  における観測値  $x_i, y_i, z_i, v_i$  は，それぞれ時刻  $t_i$  における  $x, y, z$  座標，筆速である．なお， $z$  座標とはペンがタブレット PC と接触しているか，していないかを 2 値で示すデータである．1 署名につき，平均約 1000 点の時系列点座標データで構成されている．

時系列点座標データの一例を図 2.1 に示す．(a)  $x$  座標，(b)  $y$  座標中の縦軸と横軸は，それぞれ座標位置と時間とを示し，(c) 筆速  $v$  中の縦軸と横軸は，それぞれ速度と時間とを示す．



(a) x 座標



(b) y 座標

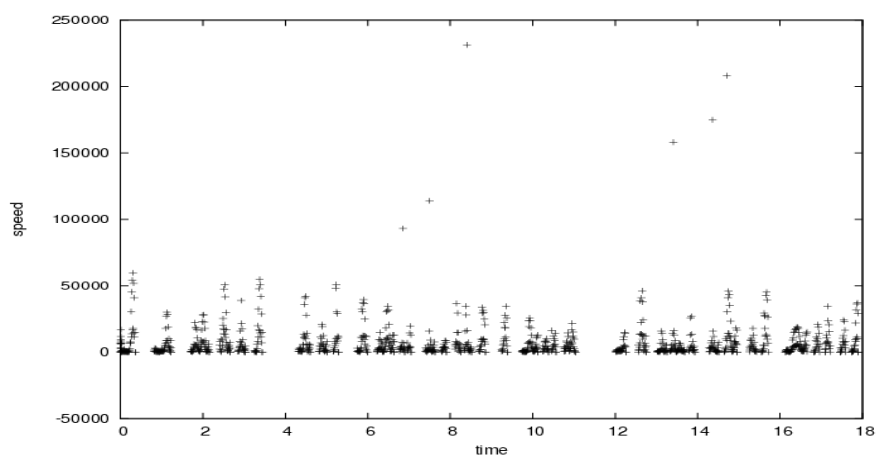
(c) 筆速  $v$ 

図 2.1: 時系列点座標データの例

## 2.2 データの正規化

### 2.2.1 全実験データに行う正規化

筆記開始位置  $x, y$  座標と開始時刻  $t$  とを学習用データと評価用データで合わせる．具体的には，筆記開始時の  $x, y$  座標と開始時刻  $t$  を 0 となるように時系列点座標データを変換する．時系列点座標データのサンプリング点において 2 点目以降は 1 点目との差を取り，その値を正規化データとする．

### 2.2.2 観測値 $x, y$ 座標の変化量を用いる正規化

2.2.1 で示した正規化に加え，さらに観測値  $x, y$  座標に連続するサンプリングデータ点の変化量（差分） $\Delta x, \Delta y$  を用いたデータも作成する．時系列点座標データ  $P_i(x_i, y_i, z_i, v_i, t_i)$  と  $P_{i+1}(x_{i+1}, y_{i+1}, z_{i+1}, v_{i+1}, t_{i+1})$  から， $\Delta x_i = x_{i+1} - x_i$ ,  $\Delta y_i = y_{i+1} - y_i$  により観測値  $x, y$  座標の変化量  $\Delta x_i, \Delta y_i$  を算出し， $P'_i(\Delta x_i, \Delta y_i, z_i, v_i, t_i)$  を得る．

## 2.3 照合手順

本提案手法は、以下の 3 種類のデータに対して実験を行っている。

1. 対象署名から得られる時系列データ（署名 1 個分）
2. 対象署名から得られる時系列データの観測値  $x, y$  座標 に変化量を用いたもの（署名 1 個分）
3. 対象署名から得られる時系列データを分割したもの（分割署名）

以下にそれら 3 パターンにおける処理の流れを記述する。

本提案手法の処理の概要を 2.3.1 に、分割署名を用いた場合の詳細を 2.3.2 に示す。

### 2.3.1 処理の概要

時系列データに対する処理の流れを示す。

1. タブレット PC から得られた署名の時系列点座標データを取得する。
2. 時系列点座標データに、開始点の座標と時刻を 0 とする正規化を行う。
3. 時系列点座標データの連続したサンプリング点を用いて  $N$  次元の特徴ベクトル群に変換する。（ここまでの処理は学習用、評価用署名共に同じ）
4. 各人の学習用データである真筆署名 3 個から得られた  $N$  次元特徴ベクトル群に対して共分散行列を求める。
5. 学習用データの共分散行列を用いて、固有ベクトルを求める。
6. 学習用データの固有ベクトルより求めた最小 2 乗超平面と、各評価用署名データの特徴ベクトルとのそれぞれの線形予測誤差（投影距離）を計算する。
7. 各評価用署名の時系列点座標データのサンプリング数分算出された線形予測誤差（投影距離）から平均値を対象署名データ毎に求める。
8. 真筆を偽筆と間違える割合、偽筆を真筆と間違える割合、これらが等しくなるようにしきい値を設定する。
9. 評価用署名データから得た平均投影距離がしきい値より小さければ真筆、大きければ偽筆と判定する。

なお、手順 3 までは学習用、評価用データ共に同じ処理を行う。

### 2.3.2 分割署名と投票の処理詳細

本実験では，分割署名（対象署名から得られる時系列点座標データを均等に分割したもの）を用いる（図 2.2）．学習用，評価用共に分割署名のそれぞれを 1 つの署名として 2.3.1 と同様に線形予測誤差を求める処理を行い，真偽判定を行う．その後，各分割署名に票を与え，投票により分割元の対象署名の真偽を判定する．つまり，分割署名判定結果の真筆票が偽筆票より多ければ対象署名は真筆，分割署名の偽筆票が真筆票より多ければ対象署名は偽筆と判定する．

さらに，この各分割署名に与える票数は，以下の 2 通りの方法を用いる．

- 1 分割署名につき 1 票を与える．
- 各分割署名から算出された平均投影距離値としきい値との差の絶対値をそれぞれに重みとして票を与える．
  - － 例えば，ある分割署名の平均投影距離値がしきい値より 100 小さかった場合，その分割署名には真筆票 100 票が与えられる．

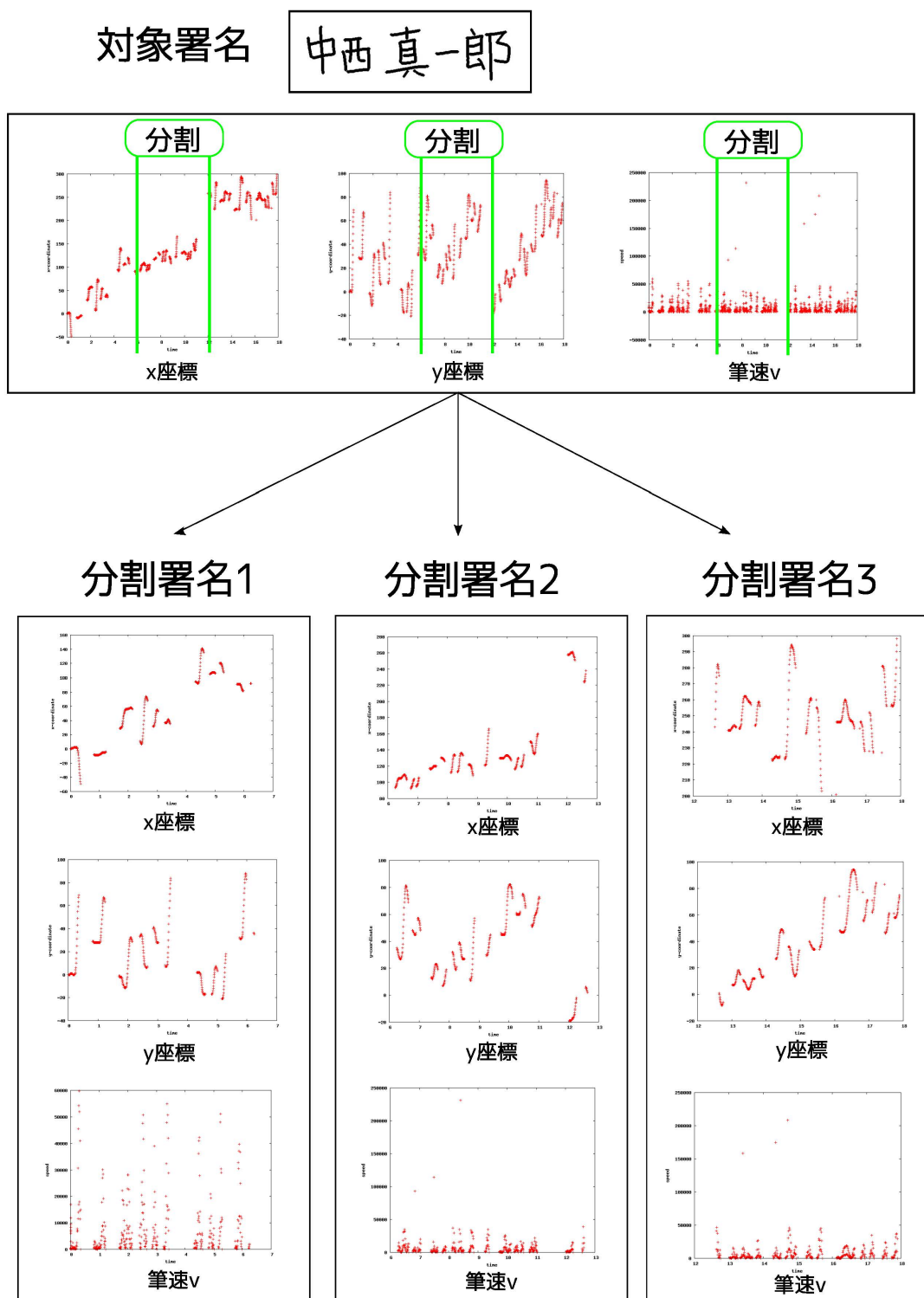


図 2.2: 対象署名 (3 分割) の概念図

## 2.4 線形予測誤差

線形予測分析とは、離散信号の将来の値をそれまでの標本群の線型写像として予測する数学的操作であり、音響信号処理および音声処理の分野でよく用いられる手法である。その中でも本研究では、線形予測符号 (LPC) により算出される線形予測誤差の考え方を利用する。

この線形予測誤差の手順（本研究で用いた部分）の概要を以下に示す。

1. 座標空間 (本研究では、特徴ベクトル空間) 上に  $k$  次元の時系列データ（提案手法における学習用データ）がある。  
(ここで  $k$  は任意の数であり、提案手法では後述する 2.4.3 内の  $k$  がここでの  $k$  に相当する.)
2. それらを近似し、超平面（あるいは、直線や平面）を得る。
3.  $N$  次元 ( $N > k$ ) の入力データ  $a$ （提案手法における評価用データ）がある。
4. この入力データ  $a$  から超平面に対して垂線を降ろし（投影）、超平面と交わる点を  $a'$  とする。この  $a'$  が  $a$  の線形予測符号となる。
5. この入力データから線形予測符号への投影距離が線形予測誤差となる。

本研究では、入力と線形予測符号との誤差、つまり超平面への投影距離である線形予測誤差を用いて照合を行う。

この手法の簡単な例を図 2.3 に示す。この図 2.3 では、3 次元座標空間上で行っている。



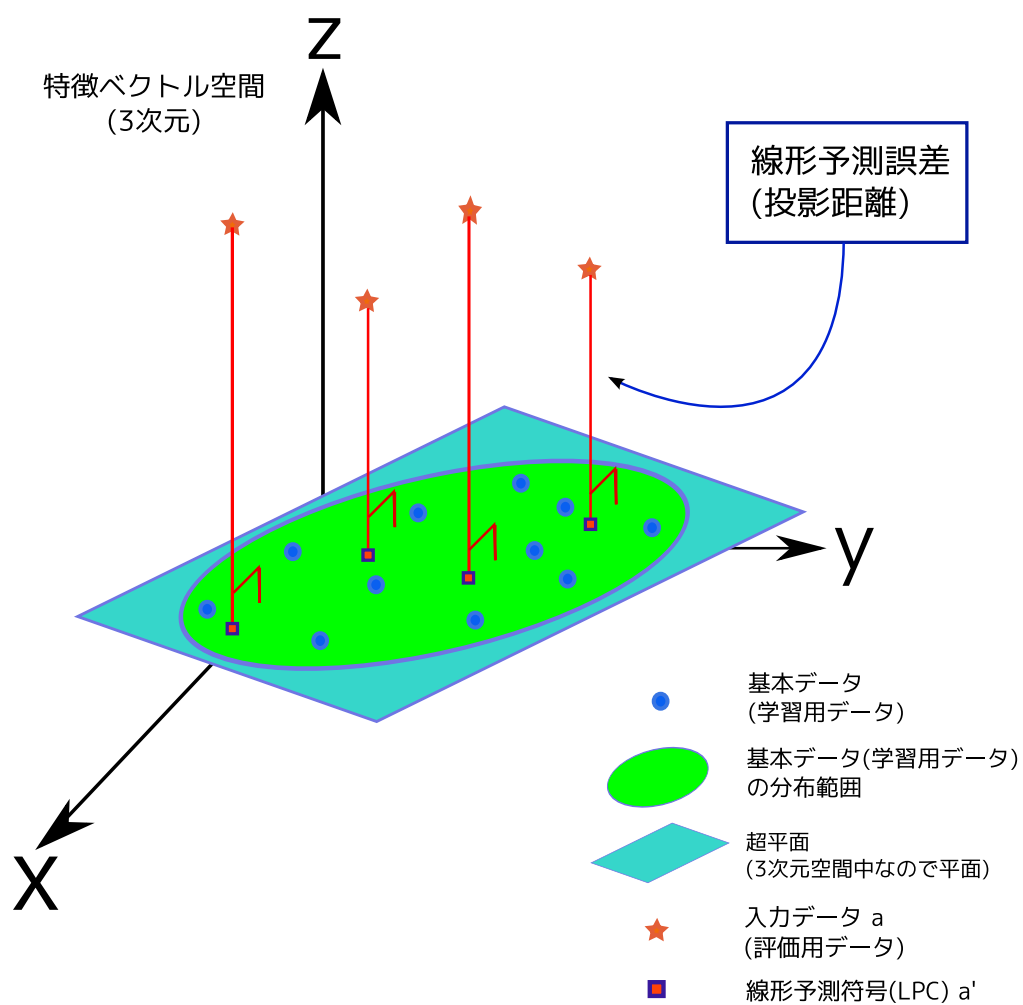


図 2.3: 線形予測誤差の概念図

### 2.4.1 線形予測誤差から利用する考え方

本研究では、線形予測誤差の考え方を利用する。その手法を以下に示す。

まず、1つの特徴ベクトルを作成するのに、取得された時系列データ何点分を用いるのか、5種類ある観測値中、何種類を特徴ベクトル作成に使用するのか、この2点を決める。

時系列データを  $A$  点分、観測値を 5 種類中  $B$  種類を使用する場合、 $A \times B$  次元の特徴ベクトルを作成する事になる。

例えば、署名の時系列データが  $D$  点ある場合  $P_1, P_2, \dots, P_D$ ，まず  $P_1$  から  $A$  個を順番に用いて 1 つの多次元データ  $(P_1, P_2, \dots, P_A)$  を生成する。次に  $P_2$  から順番に  $A$  個のデータを用いて多次元データ  $(P_2, P_3, \dots, P_{A+1})$  を生成する。そして、それぞれの  $P$  に  $B$  種類の観測値データ（パラメータ）を持たせる。この操作を繰り返し行い、 $D - A + 1$  個の  $N = A \times B$  次元データを得る。

具体例を次節で示す。

### 2.4.2 特徴ベクトル作成の具体例

この手法の簡単な例を以下に示す。

$(P_1, P_2, P_3, P_4, P_5)$  という 5 点で構成される時系列データがある場合、これらを  $A = 3$  点毎に用い、 $(P_1, P_2, P_3), (P_2, P_3, P_4), (P_3, P_4, P_5)$  と 3 つの 3 次元データができる。これらに観測値  $(x, y)$  の  $A = 2$  種類を持たせた場合、 $(x_1, y_1, x_2, y_2, x_3, y_3)$ ， $(x_2, y_2, x_3, y_3, x_4, y_4)$ ， $(x_3, y_3, x_4, y_4, x_5, y_5)$ ，となり、 $N = A \times B = 3 \times 2 = 6$  次元データ 3 個を得る。（参考：図 2.4）

1. 5点の時系列データからなる署名

$(P_1, P_2, P_3, P_4, P_5)$

- 各点を持つ5つの観測値
1. タブレット上のx座標
  2. タブレット上のy座標
  3. ペンのタブレットへの接触・非接触  $z$
  4. 筆速  $v$
  5. 経過時間  $t$

2. サンプルングデータ数 $A$ ごとにグループ化し, 特徴ベクトル $X_i$ を求める ( $A = 3$ の場合)

$X_1 = (P_1, P_2, P_3), \quad X_2 = (P_2, P_3, P_4), \quad X_3 = (P_3, P_4, P_5)$

3.  $P_i = (x_i, y_i)$ の場合

ここで  $A \times B = N$  とする

$$X_1 = \begin{pmatrix} x_1 \\ y_1 \\ x_2 \\ y_2 \\ x_3 \\ y_3 \end{pmatrix}, \quad X_2 = \begin{pmatrix} x_2 \\ y_2 \\ x_3 \\ y_3 \\ x_4 \\ y_4 \end{pmatrix}, \quad X_3 = \begin{pmatrix} x_3 \\ y_3 \\ x_4 \\ y_4 \\ x_5 \\ y_5 \end{pmatrix}$$

図 2.4: 特徴ベクトル作成の概要

### 2.4.3 投影距離の計算方法

投影距離  $d^2$  は以下の式で計算する.

$$d^2 = \|X - M\|^2 - \sum_{i=1}^k \{\phi_i^T (X - M)\}^2. \quad (2.1)$$

ここで,  $d^2$ : 投影距離

$X$ : 評価用  $N$  次元特徴ベクトル

$M$ : 学習用  $N$  次元特徴ベクトルの平均ベクトル

$\|X - M\|^2$ : ユークリッド距離

$k$ : 超平面の次元数

$\phi_i$ : 学習用  $N$  次元特徴ベクトルの共分散行列の固有ベクトル

とする.

実験を行う場合,  $N$  次元特徴ベクトルで投影距離を計算する際に,  $k$  の値の設定可能範囲は 0 から  $N - 1$  までとする. なぜなら,  $k = N$  とした場合, 投影距離は常に 0 となり投影距離で評価する本実験では必要がないためである.

### 2.4.4 投影距離計算における $k$ の値を変化させた場合の例

信号データを 4 次元データとして扱う場合,

- $k = 0$  の場合, 投影距離 = ユークリッド距離,
- $k = 1$  の場合, 投影距離 = 直線までの距離,
- $k = 2$  の場合, 投影距離 = 平面までの距離,
- $k = 3$  の場合, 投影距離 = 超平面までの距離,
- $k = 4$  の場合, 投影距離 = 0,

となる.

## 2.5 真偽判定

線形予測誤差（投影距離）の計算結果により，一定のしきい値を設定し，真筆か偽筆かの判定を行う．各署名から算出された線形予測誤差がしきい値より小さければ真筆，大きければ偽筆という照合結果となる．

しきい値の決定法については，ミニマックス誤り確率を使用する．つまり，他人受入れ率（FAR:False Acceptance Rate）と本人拒否率（FRR:False Rejection Rate）が等しくなるしきい値を設定することになる．（他人受入れ率と本人拒否率については 3.2 にて詳しく触れる．）

## 第 3 章

# 実験

先に記述した線形予測誤差を用いたオンライン署名照合が有用であるか，評価実験を行う．

本実験では，

- 特徴ベクトル作成時に用いるデータ点の数  $A$
- 投影距離算出式内の  $k$  の値

の 2 点を変化させることで，照合率はどのような結果になるのかを調べた．

1.  $A$  の値は，10 点, 20 点, 50 点, 100 点の 4 パターンを検証する．
2.  $k$  の値は， $N$  次元特徴ベクトルを用いる場合に， $N$  を 100% とし，0% から 100% まで 10% 毎に設定する．（0%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%）（ただし，100% の場合だけ  $N - 1$  とする．先に記述したように  $k = N$  とすると投影距離が 0 となるからである．）

## 3.1 実験データ

タブレット PC を用いて収集した，研究室の研究生 19 人分の署名データを使用する．学習用署名データ，評価用署名データの内訳を以下に示す．筆記者には署名筆記前にペンタブレットの使用方法に慣れてもらい，自然な状態で筆記ができるように筆記者本人の名前を何度か練習してもらい，その後，本実験データの取得を行っている．これにより，ペンタブレットでの筆記に慣れていないが為に生じる署名情報の変動を小さくする．偽筆者は，利用者が署名を登録するのと同時期に書いた真筆署名を参考に，偽筆を作成する．ここで，偽筆者は真筆のサンプル（オフライン署名）のみを見て偽筆を作成しており，真筆が書かれた場に居合わせているわけではない．

なお，本実験は予備実験の結果より， $z$  座標（ペンとタブレット PC の接触・非接触データ）を 10 と 0 の 2 値で示す．

取得した実験データの内訳を以下に示す．

### 1. 学習用署名データの内訳

- 真筆署名：57 個 (19 人分 × 3 個)

### 2. 評価用署名データの内訳

- 真筆署名：684 個 (19 人分 × 36 個)
- 偽筆署名：684 個 (19 人分 × 偽筆者 4 人 × 9 個)

真筆署名は 4 日間に分けて取得している．1 日目に連続で 6 個取得し，その後，時間を空けずに 9 個署名してもらう．次に，2，3，4 日目にはそれぞれ連続で 9 個の署名をしてもらう．このようにし，一人当たり 36 個の署名を集めた．この内，1 日目に連続して取得した 6 個の署名のうち 3 個を，各人の学習用署名データとする．偽筆署名は，1 個の対象署名に対し，偽筆者 4 人から連続で 9 個の署名を取得することで，対象署名 1 個につき偽筆 36 個を収集した．

## 3.2 照合率の定義

署名照合には二つの誤り，他人受入れ率（FAR:False Acceptance Rate）と本人拒否率（FRR:False Rejection Rate）が存在する．本人拒否率は，本人の真筆署名に対し偽筆と誤って判定する割合であり，他人受入れ率は，偽筆署名に対し真筆と誤って判定する割合である．他人受入れ率と本人拒否率はトレードオフの関係にある（図 3.1 参照）．どちらが低い方がよいかは用途によって異なるため，本研究では， $FAR = FRR$  となったときの照合率を示す．

照合率  $R$ ， $FAR$ ， $FRR$  を次のように定義する．

$$\text{照合率 } R = \frac{a + d}{a + b + c + d} \times 100(\%)$$

$$FAR = \frac{c}{c + d} \times 100(\%)$$

$$FRR = \frac{b}{a + b} \times 100(\%)$$

$a$  : 真筆に対して正しく真筆と判定した数

$b$  : 真筆に対して誤って偽筆と判定した数

$c$  : 偽筆に対して誤って真筆と判定した数

$d$  : 偽筆に対して正しく偽筆と判定した数

表 3.1: 判定数の内訳（ $a, b, c, d$  はデータ数）

		システム	
		真	偽
正解	真	a	b
	偽	c	d



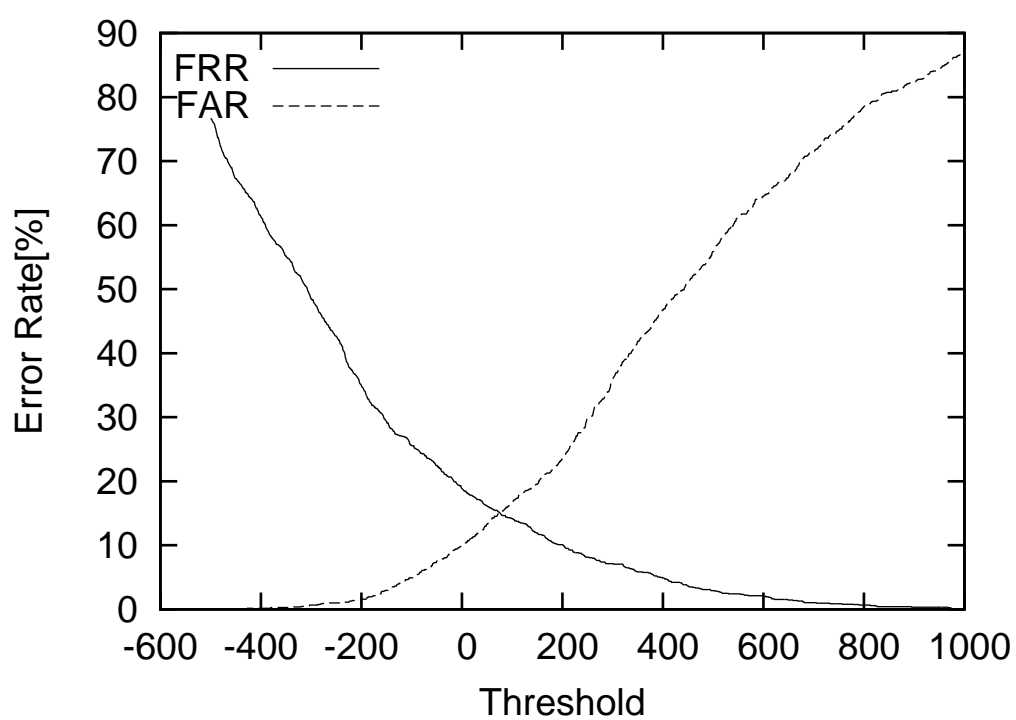


図 3.1: FAR と FRR の関係

### 3.3 実験 1

この実験は以下の条件にて行う。

- 対象署名から得られる時系列データ（署名 1 個分）を用いる。

ここでは予備実験により，特徴ベクトル作成時に用いるサンプリングデータ点の数  $A = 50$  とした実験結果を示す． $A = 10, 20$  の場合，全体的に  $A = 50$  の場合より照合率が低い結果となった。

#### 3.3.1 結果

実験結果を図 3.2 に示す。

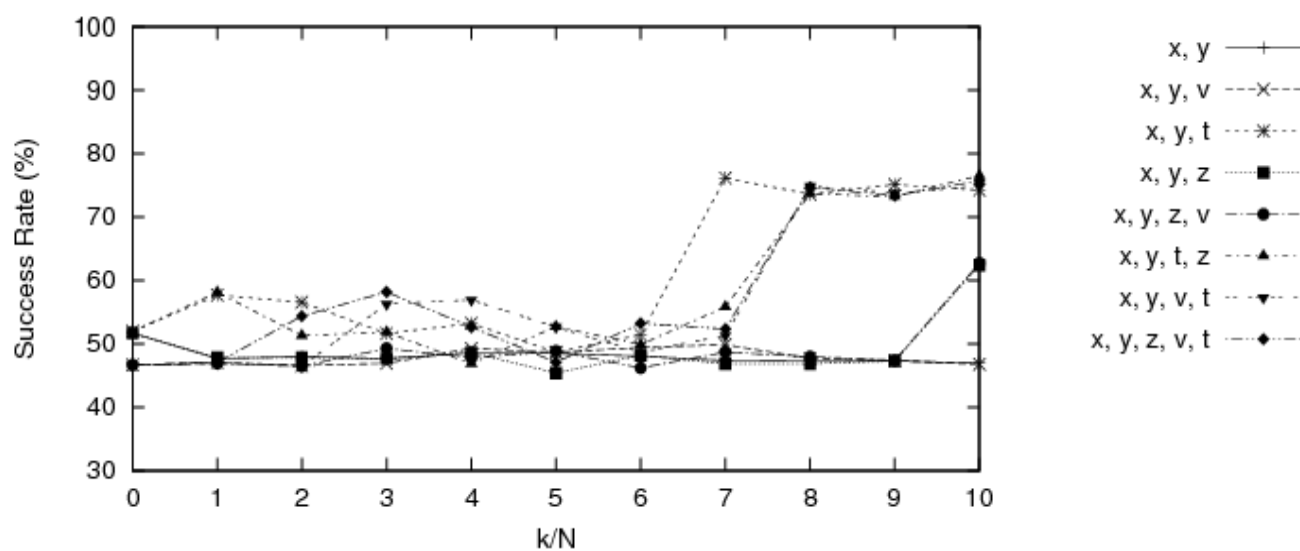


図 3.2: 照合実験結果

観測値  $(x, y, z, t)$  を用いて特徴ベクトルを作成し，投影距離計算時に  $k = N - 1$  と設定した場合に，最大照合率 76.53% という結果が得られた。

全体的に見ると，観測値に筆記時間  $t$  を用いた場合に高い照合率を得ていることがわかる。反対に，観測値  $t$  以外， $x, y, z, v$  だけを用いた場合は低い照合率を得ている。

## 3.4 実験 2

この実験は以下の条件にて行う。

- 分割署名から得られる時系列データを用いる。
  - 分割数は 3 とする。
- 1 分割署名につき 1 票を与えて，多数決により対象署名の真偽を判定する。

ここでは予備実験により，特徴ベクトル作成時に用いるサンプリングデータ点の数  $A = 50$  とした実験結果を示す。  $A = 10, 20$  の場合，全体的に  $A = 50$  の場合より照合率が低い結果となった。

### 3.4.1 結果

実験結果を図 3.3 に示す。

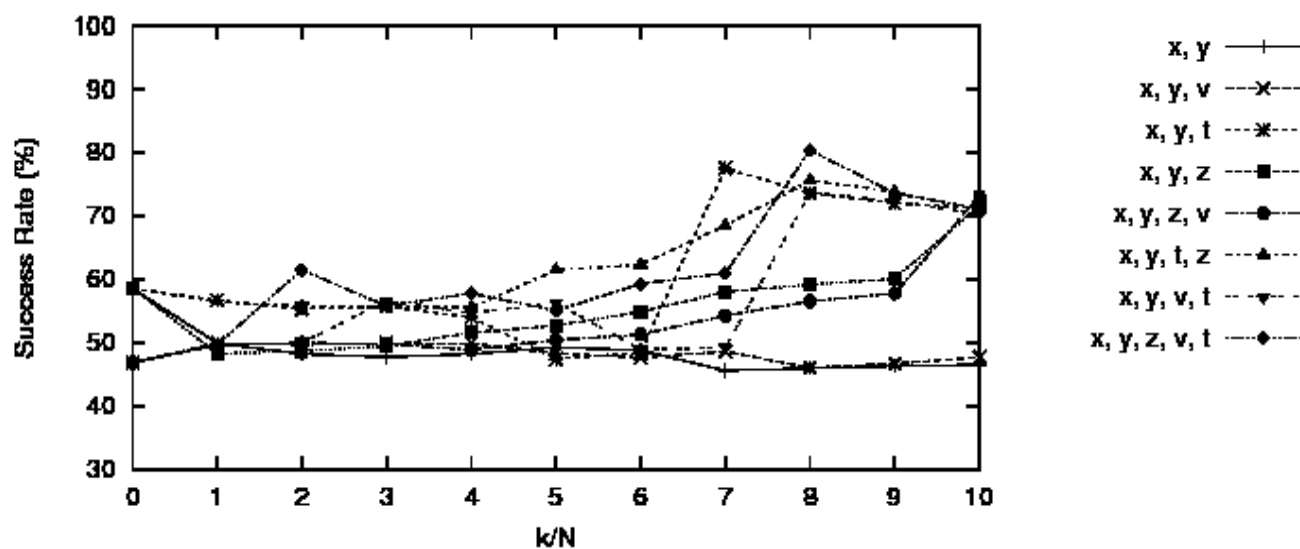


図 3.3: 照合実験結果

観測値  $(x, y, z, v, t)$  を用いて特徴ベクトルを作成し，投影距離計算時に  $k = 0.8N$  と設定した場合に，最大照合率 80.40% という結果が得られた。

観測値  $(x, y, z, v, t)$  と  $(x, y, t)$  を用いた場合に，他の組み合わせより比較的高い照合率を得ている。全体的に見ると，観測値に筆記時間  $t$  を用いた場合に高い照合率を得ていることもわかる。反対に，観測値  $(x, y)$  と  $(x, y, v)$  を用いた場合は低い照合率を得ている。

## 3.5 実験 3

この実験は以下の条件にて行う。

- 分割署名から得られる時系列データを用いる。
  - 分割数は 5 とする。
- 1 分割署名につき 1 票を与えて、多数決により対象署名の真偽を判定する。

ここでは予備実験により、特徴ベクトル作成時に用いるサンプリングデータ点の数  $A = 50$  とした実験結果を示す。  $A = 10, 20$  の場合、全体的に  $A = 50$  の場合より照合率が低い結果となった。

### 3.5.1 結果

実験結果を図 3.4 に示す。

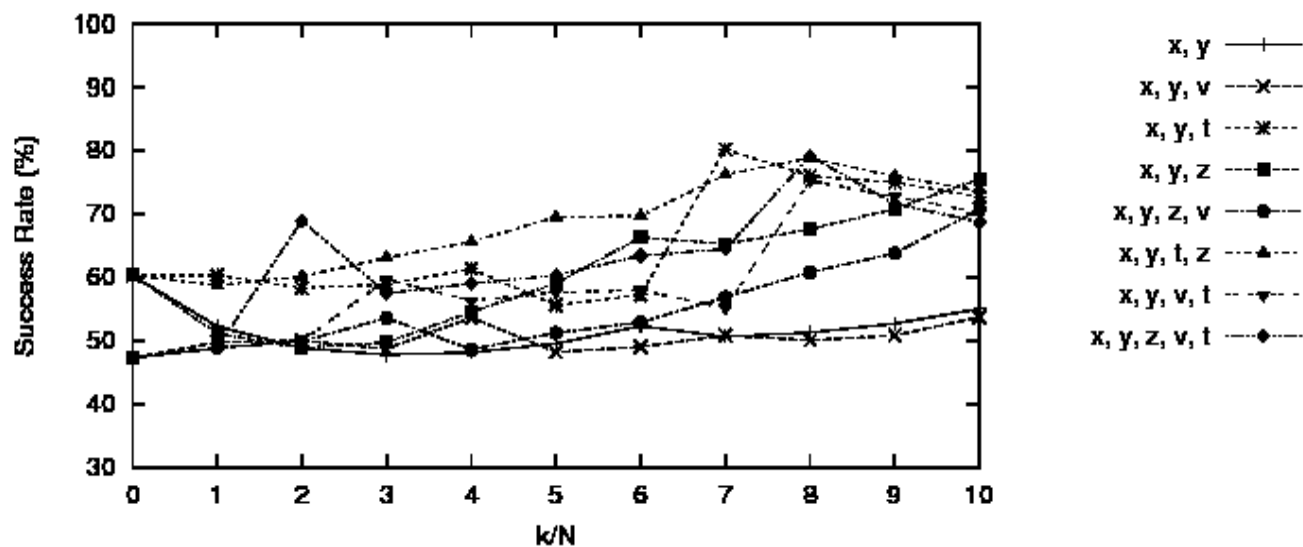


図 3.4: 照合実験結果

観測値  $(x, y, t)$  を用いて特徴ベクトルを作成し，投影距離計算時に  $k = 0.7N$  と設定した場合に，最大照合率 80.40% という結果が得られた。

観測値  $(x, y, z, v, t)$ ,  $(x, y, z, t)$ ,  $(x, y, t)$  を用いた場合に，他の組み合わせより比較的高い照合率を得ている．全体的に見ると，観測値に筆記時間  $t$  を用いた場合に高い照合率を得ていることもわかる．反対に，観測値  $(x, y)$  と  $(x, y, v)$  を用いた場合は低い照合率を得ている．

## 3.6 実験 4

この実験は以下の条件にて行う。

- 分割署名から得られる時系列データを用いる。
  - 分割数は 3 とする。
- 各分割署名から算出された投影距離としきい値との差の絶対値をそれぞれに重みとして票を与える。

ここでは予備実験により，特徴ベクトル作成時に用いるサンプリングデータ点の数  $A = 50$  とした実験結果を示す。  $A = 10, 20$  の場合，全体的に  $A = 50$  の場合より照合率が低い結果となった。

### 3.6.1 結果

実験結果を図 3.5 に示す。

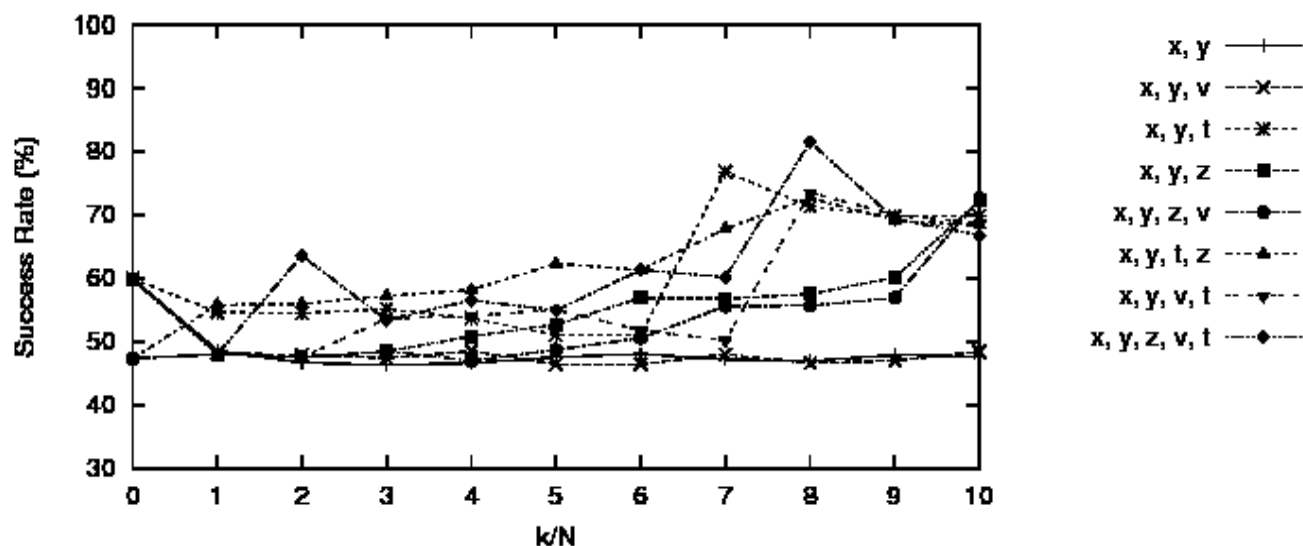


図 3.5: 照合実験結果



観測値  $(x, y, z, v, t)$  を用いて特徴ベクトルを作成し，投影距離計算時に  $k = 0.8N$  と設定した場合に，最大照合率 81.50% という結果が得られた。

観測値  $(x, y, z, v, t)$  と  $(x, y, t)$  を用いた場合に，他の組み合わせより比較的高い照合率を得ている。全体的に見ると，観測値に筆記時間  $t$  を用いた場合に高い照合率を得ていることもわかる。反対に，観測値  $(x, y)$  と  $(x, y, v)$  を用いた場合は低い照合率を得ている。

## 3.7 実験 5

この実験は以下の条件にて行う。

- 分割署名から得られる時系列データを用いる。
  - 分割数は 5 とする。
- 各分割署名から算出された投影距離としきい値との差の絶対値をそれぞれに重みとして票を与える。

ここでは予備実験により、特徴ベクトル作成時に用いるサンプリングデータ点の数  $A = 50$  とした実験結果を示す。 $A = 10, 20$  の場合、全体的に  $A = 50$  の場合より照合率が低い結果となった。

### 3.7.1 結果

実験結果を図 3.6 に示す。

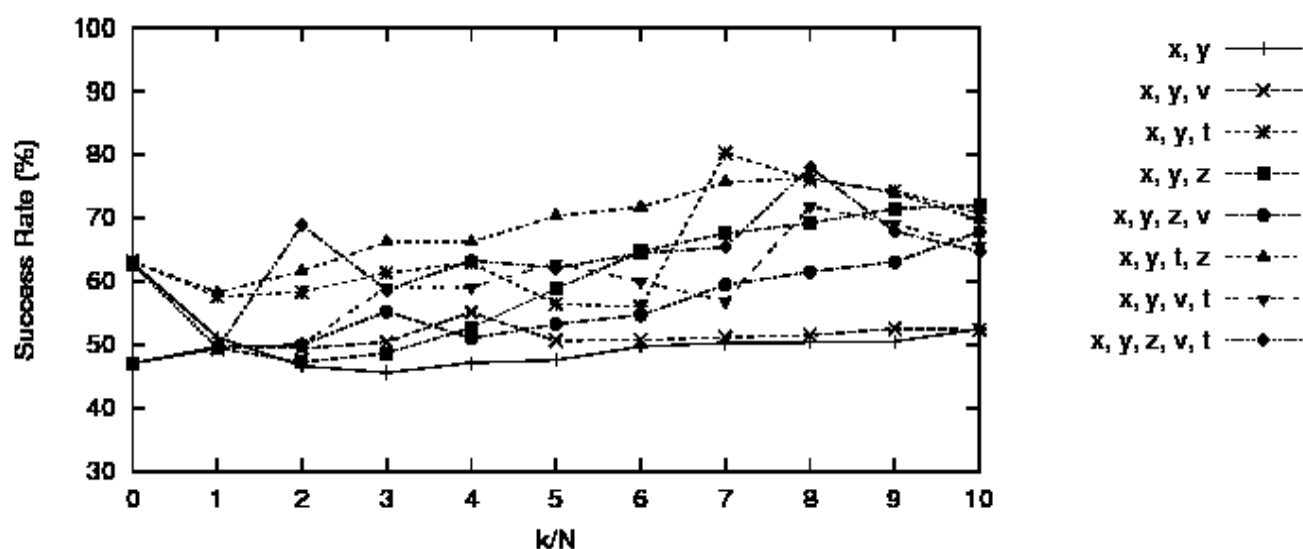


図 3.6: 照合実験結果

観測値  $(x, y, t)$  を用いて特徴ベクトルを作成し，投影距離計算時に  $k = 0.7N$  と設定した場合に，最大照合率 80.26% という結果が得られた。

観測値  $(x, y, z, v, t)$ ,  $(x, y, z, t)$ ,  $(x, y, t)$  を用いた場合に，他の組み合わせより比較的高い照合率を得ている．全体的に見ると，観測値に筆記時間  $t$  を用いた場合に高い照合率を得ていることもわかる．反対に，観測値  $(x, y)$  と  $(x, y, v)$  を用いた場合は低い照合率を得ている．

## 3.8 実験 6

この実験は以下の条件にて行う．

- 時系列データの観測値  $x, y$  座標 にサンプリングデータの変化量（差分）を用いる．

ここでは予備実験により，特徴ベクトル作成時に用いるサンプリングデータ点の数  $A = 100$  とした実験結果を示す． $A = 10, 20, 50$  の場合，全体的に  $A = 100$  の場合より照合率が低い結果となった．

### 3.8.1 結果

実験結果を図 3.7 に示す．

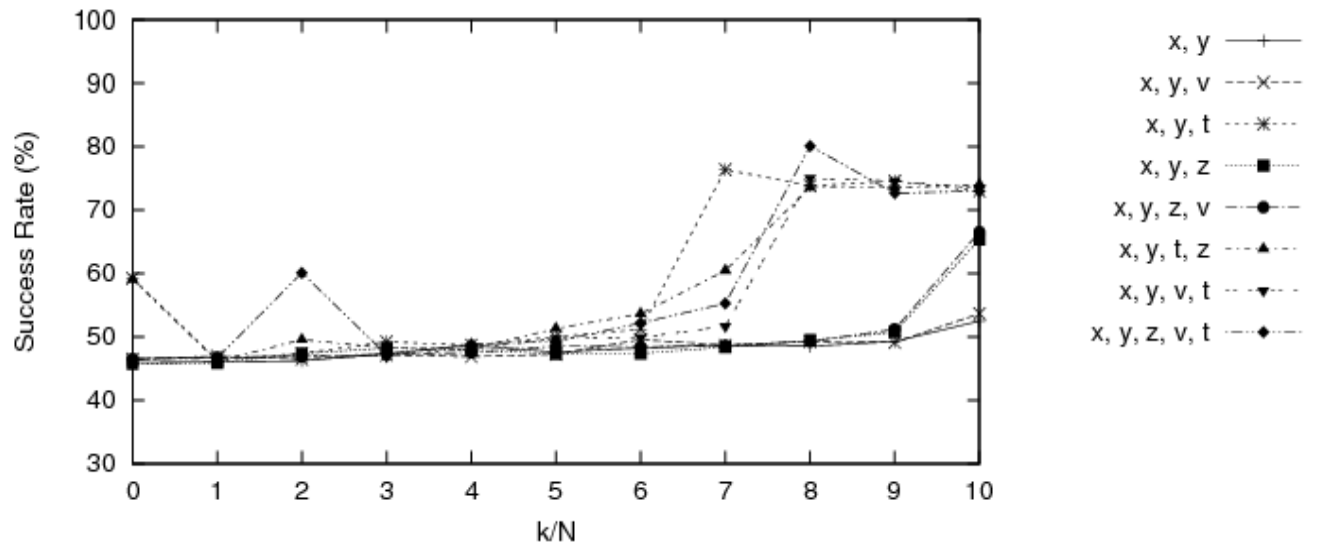


図 3.7: 照合実験結果

観測値  $(x, y, z, v, t)$  を用いて特徴ベクトルを作成し，投影距離計算時に  $k = 0.8N$  と設定した場合に，最大照合率 80.11% という結果が得られた．

観測値  $(x, y, z, v, t)$ ， $(x, y, t)$  を用いた場合に，他の組み合わせより比較的高い照合率を得ている．全体的に見ると，観測値に筆記時間  $t$  を用いた場合に高い照合率を得ていることもわかる．

## 3.9 考察

本報告で行った実験では、タブレット PC から得られた署名データを分割し、それら分割署名の各時系列点座標データを 50 点毎に使用し特徴ベクトルを作成、特徴ベクトルに使用する観測値を  $(x, y, z, v, t)$  の 5 種類とした場合に、照合率が 81.50% という結果が得られ、これが最も高い照合率となった。

この結果から、署名データは 1 枚の対象署名を分割したものを使用し、それら分割署名の真偽判定結果による投票をもって対象署名を分類する手法がより高い照合率を示すことから、署名の局所的な部分に個人の特徴が現れることが推測される。

観測値  $x, y$  座標に差分を用いた場合も照合率に大きな変化は見られなかった。このことから、オンライン情報（時系列座標データ）において、 $x, y$  座標は各データの数値の大小ではなく、署名の形としてのデータが重要であると考えられる。

また、特徴ベクトル作成において観測値を  $x, y$  座標だけでなく、他のパラメータ、特に経過時間  $t$  を用いた場合、線形予測誤差（投影距離）計算内の  $k$  の値によらず、照合率が全体的に高い傾向にある。このことから、特徴ベクトル作成に使用する観測値の種類は、 $x, y$  座標だけでなく他のパラメータも用いる方がよいと考えられる。つまり、照合には各人の分割署名を使用した投票形式で真偽判定を行い、特徴ベクトル作成に利用する観測値の種類は多い、かつ経過時間  $t$  を用いたほうが、より個人の特徴が現れるものと推測される。

さらに、線形予測誤差（投影距離）算出における  $k$  の値を  $0.7N$ 、もしくは  $0.8N$  に設定する、つまり特徴ベクトルを高次元の超平面で近似し、適度に分散の小さい固有ベクトルに投影することにより、照合率が高くなる傾向が全体的に見られる。

従来研究である、古市ら [6]、川添ら [10] の 2 手法との比較を行う。同じ実験データを用いた比較実験の結果、古市らの DP マッチングを用いた手法は最高照合率 91.75%、川添らの濃度こう配特徴を用いた手法は最高照合率 92% となり、本手法の最高照合率 81.50% という結果は依然改善の余地が残されている。しかし、これらの従来研究は 1 個の対象署名全体から特徴を抽出し分類を行っているのに対し、本手法では 1 個の対象署名を分割し、局所的な特徴を抽出、それらを統合して真偽を分類しているので、特徴抽出の対象が署名全体と局所という違いがある。将来的に、この 2 種類の手法を融合させることで、さらなる照合率の精度向上が期待できる。

## 第 4 章

# まとめと今後の課題

### 4.1 まとめ

本研究では，線形予測誤差を用いたオンライン署名照合の有用性について検証を行った．署名データの分割や正規化，ペンアップデータ，筆速，経過時間も用いることで照合率が高くなることがわかった．実験結果から， $x, y$  座標値，筆速データ，ペンアップデータ，時刻データを用いることで照合率 81.50% が得られた．このことより，本手法は有用であるとわかった．

### 4.2 今後の課題

今後の課題として以下のことが挙げられる．

1. より多くのデータで評価実験，比較を行うことで，実験結果の信頼性を高める．
2. 署名のサイズを全て統一するなど，さらに有用な正規化の方法はないか調査する．
3. 用いる観測値の種類，特徴ベクトル作成に用いるサンプリングデータ点の数， $k$  の値の最適な組み合わせを調査する．
4. 投影距離の代わりに改良投影距離を用いた場合，どのように結果が変わってくるのか検証する．
5. しきい値の最適な設定方法について検討する（各人毎に固有のしきい値 1 つずつ設定するのか，もしくは全体で共通のしきい値 1 つ設定するのか）．
6. 線形予測誤差（投影距離）算出時の  $k$  の値を  $0.7N$ ， $0.8N$  と設定すると高い照合率を得るという結果から，本実験では  $k$  は 10% 区切りに設定したが， $0.7N$ ， $0.8N$  付近でももう少し細かく区切り実験し，照合率にどのような変化が現れるか検証する．
7. より低次元の超平面で近似できる特徴抽出手法を検討する．
8. 他の照合手法と融合させることで，照合率の精度向上を図る．

## 付録 A

# 実験データとプログラム

研究に使用したプログラムやデータはすべて

~/nakanisi/study/

内のディレクトリ以下に格納されている。各ディレクトリとファイルについて以下に示す。詳細は、各ディレクトリ内の *README* を参照のこと。

```
| -- study
|
| # 2007 年度の署名データ 19 人分
|--True
| # 真筆のテキストファイル用ディレクトリ
|--True_convert1~56
| # 各種真筆のテキストファイル（データ正規化後）用ディレクトリ
|--False
| # 偽筆のテキストファイル用ディレクトリ
|--False_convvert1~56
| # 各種偽筆のテキストファイル（データ正規化後）用ディレクトリ
|--Learn
| # 真筆（学習用）のテキストファイル用ディレクトリ
|--Learn_convert1~56
| # 各種真筆（学習用）のテキストファイル（データ正規化後）用ディレクトリ
|--normalize-
| # 各種正規化用プログラム
```



```
| --program  
| # 各種プログラム
```

## 付録 B

# プログラムソースリスト

作成したプログラムを以下のディレクトリ

~/nakanisi/study/program

に置く．すべてのプログラムファイル，コンパイル，実行形式，シェルスクリプトに関して詳細はディレクトリ内の *README* 参照のこと．

```
| --type1  
| # 観測値 x 種類を用いて特徴ベクトルを作成した場合の投影距離を計算し，指定した  
新しいファイルに保存する．  
|  
| --type2  
| # 観測値 x, y を用いて多次元データを作成した場合の投影距離を計算し，指定した新  
しいファイルに保存する．  
|  
| --type3  
| # 観測値 x, y, z を用いて多次元データを作成した場合の投影距離を計算し，指定し  
た新しいファイルに保存する．  
|  
| --type4  
| # 観測値 x, y, z, v を用いて多次元データを作成した場合の投影距離を計算し，指定  
した新しいファイルに保存する．  
|  
| --type5
```

```
| # 観測値 x,y,z,v,t を用いて多次元データを作成した場合の投影距離を計算し, 指定した新しいファイルに保存する.
```

```
|
```

```
| --type6
```

```
| # 観測値 x,y,v,t を用いて多次元データを作成した場合の投影距離を計算し, 指定した新しいファイルに保存する.
```

```
|
```

```
| --type7
```

```
| # 観測値 x,y,z,v を用いて多次元データを作成した場合の投影距離を計算し, 指定した新しいファイルに保存する.
```

```
|
```

```
| --type8
```

```
| # 観測値 x,y,v を用いて多次元データを作成した場合の投影距離を計算し, 指定した新しいファイルに保存する.
```

```
|
```

```
| --shikiichi3
```

```
| # しきい値, 照合率を算出し, 結果をファイルに保存する.
```

```
|
```

```
| --memo.txt
```

```
| # 研究に関するメモ
```

```
|
```

# 謝辞

本研究に関し，研究に対する数多くの適切な助言，御指導を頂いた三重大学大学院工学研究科情報工学専攻，教授木村文隆先生に深く感謝致します．研究に関する多くの知識や研究分野における様々な技術を御教授下さった准教授若林哲史先生に深く感謝致します．資料作成からプレゼンテーションのコツなどを御指導頂きました助教大山航先生に深く感謝致します．専門的な観点から深く幅広いアイデアや知識を御教授下さった三重大学名誉教授三宅康二先生に深く感謝致します．

また，日頃色々とお世話になった田中みゆき事務，研究に対し多くの有益なアドバイスや御指導頂いた先輩方，共に多くの壁を乗り越えてきた後輩達，署名データ収集に御協力頂いた皆様に感謝致します．さらに，研究室での生活を支え，楽しい学生生活を与えて下さった学科の同期，ヒューマンインタフェース研究室の皆様に深く感謝致します．

最後になりましたが，長きに渡る私の学生生活を支えてくれた祖母，両親，叔母，弟に今一度の感謝を表して，本論文の結びと致します．

## 参考文献

- [1] 瀬戸洋一. バイオメトリックセキュリティ認証技術の動向と展望 (<特集>バイオメトリック認証システム). IPSJ magazine, Vol.47, No.6, pp.571-576, 2006.
- [2] (社) 日本自動認識システム協会. よくわかるバイオメトリックスの基礎. オーム社, 2005.
- [3] 中村善一. 日本語筆跡に現れる個人性の抽出とオンライン筆者照合に関する研究. 奈良先端科学技術大学院大学博士論文, 2008.
- [4] R.Plamondon and G.Lorette. Automatic signature verification and writer identification-The state of the art. Pattern Recognition, Vol.22, No.2, pp.107-131, 1989.
- [5] 吉村ミツ, 吉村功. 筆者認識の現段階と今後の動向. 信学技報, PRMU96-48, pp.81-90, 1996.
- [6] 古市唯, 大山航, 若林哲史, 木村文隆: “DP マッチングを用いた手書き署名照合の高精度化”: 平成 22 年度電気関係学会東海支部連合大会 (8 月 30-31 日)(N2-2)
- [7] P.Zhao, A.Higashi, and Y.Sato. On-Line Signature Verification by Adaptively Weighted DP Matching. IEICE Trans.Inf.&Syst., Vol.E79-D, No.5, pp.535-541, 1996.
- [8] Y.Komiyama, T.Ohishi, and T.Matsumoto. A Pen Input On-Line Signature Verifier Integrating Position, Pressure and Inclination Trajectories. IEICE Trans.Inf.&Syst., Vol.E84-D, No.7, pp.833-838, 2001.
- [9] J.Fierrez, J.Ortega-Garcia, D Ramos and J.Gonzalez-Rodriguez. HMM-based on-line signature verification: Feature extraction and signature modeling. Pattern Recognition Letters, Vol.28, No.16, pp.2325-2334, 2007.
- [10] 川添 由美子, 大山 航, 若林 哲史, 木村 文隆: “濃度こう配特徴を用いたオンライン署名照合の高精度化”, 電学論 C, Vol. 130, No. 12, pp.2142-2149, 2010
- [11] 村松大吾, 本郷保範, 松本隆. ユーザ共通 Fusion モデルを用いたオンライン署名認証. 信学論, Vol.J90-D, No.2, pp.450-459, 2007.
- [12] 中西功, 西口直登, 伊藤良生, 副井裕. DWT によるサブバンド分解と適応信号処理を用いたオンライン署名照合. 信学論, Vol.J87-A, No.6, pp.805-815, 2004.