

Title	動作を伴う指文字および連続した指文字の認識システム
Author(s)	後藤, 岳志
Citation	
Issue Date	1997-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/1060
Rights	
Description	Supervisor:堀口 進, 情報科学研究科, 修士

修 士 論 文

動作を伴う指文字および連続した指文字の認識システム

指導教官 堀口 進 教授

北陸先端科学技術大学院大学
情報科学研究科 情報システム学専攻
マルチメディア統合システム講座

510048 後藤 岳志

1997 年 2 月 14 日

目 次

1	序論	1
1.1	研究の背景と目的	1
1.2	本論文の構成	2
2	手形状を用いたインターフェース	3
2.1	はじめに	3
2.2	直接操作環境	3
2.3	コミュニケーション支援	4
2.3.1	手話	4
2.3.2	指文字	4
2.4	入力装置の概要	6
2.4.1	CyberGlove	6
2.4.2	FASTRAK	9
3	全ベイズ識別法による連続指文字認識	10
3.1	はじめに	10
3.2	全ベイズ識別法	10
3.2.1	手形状・向き・動きの測定	10
3.2.2	ベイズ識別法	12
3.2.3	入力データ	13
3.2.4	標準辞書	13
3.3	全ベイズ識別法による連続指文字動作の認識実験	14
3.3.1	全ベイズ識別法の連続指文字認識	14

3.3.2	実験結果	15
3.3.3	全ベイズ識別法の問題点	19
3.4	まとめ	19
4	ベイズ・距離併用法による連続指文字認識	21
4.1	はじめに	21
4.2	動作を伴う指文字の特徴	21
4.3	ベイズ・距離併用法による指文字認識	22
4.3.1	手形状の認識	25
4.3.2	向きの認識	25
4.3.3	動きの認識	27
4.3.4	切り出し手法	27
4.4	ベイズ・距離併用法による連続指文字動作の認識実験	30
4.4.1	実験結果	30
4.4.2	比較・検討	32
4.5	まとめ	33
5	指文字学習支援システム	34
5.1	はじめに	34
5.2	従来の学習支援システム	34
5.3	Cyber Touch を用いた指文字学習支援システム	35
5.3.1	Cyber Touch	35
5.3.2	指文字学習支援システムの概要	35
5.3.3	評価実験	42
5.4	まとめ	42
6	結論	45
6.1	まとめ	45
6.2	今後の課題	46
	謝辞	47
	参考文献	47

目 次

2.1	指文字	5
2.2	CyberGlove の外観	6
2.3	システム構成	8
3.1	手形状入力装置の関節角測定点	11
3.2	連続指文字認識の流れ	15
3.3	全ベイズ識別法のフローチャート	16
3.4	切り出し例	17
4.1	手形状・向き・動きによる指文字の分類	23
4.2	連続指文字認識の流れ	25
4.3	手形状・向き・動き認識のフローチャート	26
4.4	切り出し手法のフローチャート	28
4.5	切り出し手法の例	29
5.1	CyberTouch の概観	36
5.2	指文字学習支援システム概念図	37
5.3	出力される指文字の例	38
5.4	出力例 (1)	39
5.5	出力例 (2)	40
5.6	出力例 (3)	41

表 目 次

2.1	Sensor Data Byte Ordering	7
3.1	実験に用いる単語	14
3.2	$N = 5, 6, 7, 8$ のときの認識率 (全ベイズ識別法)	17
3.3	実験結果 (全ベイズ識別法)	18
3.4	静的指文字の実験結果 (全ベイズ識別法)	18
3.5	動作を伴う指文字の実験結果 (全ベイズ識別法)	18
3.6	動作を伴う指文字の脱字の分類	18
4.1	動作を伴う指文字の条件	24
4.2	$N = 5, 6, 7, 8$ のときの認識率 (ベイズ・距離併用法)	31
4.3	実験結果 (ベイズ・距離併用法)	31
4.4	静的指文字の実験結果 (ベイズ・距離併用法)	31
4.5	動作を伴う指文字の実験結果 (ベイズ・距離併用法)	31
4.6	誤字・脱字の分類	32
4.7	正しく認識される単語数 (全ベイズ識別法)	33
4.8	正しく認識された単語数 (ベイズ・距離併用法)	33
5.1	関節角の測定点と振動させる指の対応	37
5.2	学習前の認識結果	43
5.3	学習後の認識結果	43

第 1 章

序論

1.1 研究の背景と目的

近年、聴覚障害者の社会進出が進むにつれ、聴覚障害者と健聴者とのコミュニケーションを必要とする機会が増えてきている。聴覚障害者とのコミュニケーション手段として手話が一般的であるが、手話通訳者が少なく健聴者とのコミュニケーションが十分に行なわれていない。また、手話を修得するにはかなりの時間がかかるため、実用レベルに到達するには何年も要するのが普通である。このような背景のもと、手話通訳を機械に代行させる手話認識、指文字認識の研究が数多く行なわれている。

手話認識の研究では、佐川らの研究 [1][2] がある。佐川らは、データグローブから入力した手話パターンのパターン圧縮を行い、標準パターンと照合する圧縮連続 DP 照合を提案している。この方式は、620 語を用いた大語彙手話単語認識実験を行なった結果、98.7% の高い認識率が得られている。しかし、認識システムは高価であり、また認識に用いた実験データは個人データであるため、認識システムを利用できる人は限定されるという問題点がある。

聴覚障害者の社会参加を支援する上で、より多くの人々が手話を習得することが望まれる。そのため、寺内らは手話の学習を支援する手話学習システムの研究 [3] を行なっている。寺内らは、電子化辞書を利用し、手話の自学自習を支援するための学習システムを検討している。このシステムでは、電子化辞書のアニメーション部、および実録画像を活用し手話学習を進めて行く。しかし、アニメーションや実録画像では、手の動きはわかりやすいものの、手が顔の近くにあると手形状がわかりにくいという問題点がある。

本研究の目的は、手話の基本である指文字の学習支援システムを構築することである。本研究では、Cyber touch (バイブレータ機能のある手形状入力装置) を使用する。この装置を使用することにより、間違っただけの手形状のときはバイブレータを振動させることが可能となり、直接ユーザの手に間違いを知らせることができ、正しい手形状が学習できる。正しい手形状を提示しているか確認するため、指文字の認識を行なう必要がある。指文字 (76 文字) は、手の静的形状で意味の定まる静的指文字 (41 文字) と、手の形状と動きで意味の定まる動作を伴う指文字 (30 文字) に分類される。従来の指文字認識の研究では、高橋らの手の関節角をコード化して認識を行なう手法 [5] や内田らのニューラルネットを適用させた手法 [6]、萩原らの骨格モデル法 [7]、関節角の統計的特徴量を用いる方法などが提案されている。特に、萩原らの統計的特徴量を用いる方法は高い認識率が得られている。しかし、これらは静的指文字と動作を伴う指文字の一部 (手の形状が他の文字と区別できる 5 文字) を対象としており、手形状のみでは区別できない動作を伴う指文字 30 文字を対象としていない。指文字を使用する時は、名前や外来語を示す時であり静的指文字と動作を伴う指文字の連続で示される。よって、両方の指文字を含む連続した指文字動作の認識手法が必要である。

本論文では、連続認識の手法として、手形状・向き・動きの全てをベイズ識別法で識別する全ベイズ識別法を提案する。また、全ベイズ識別法の次元数をへらしたベイズ・距離併用法を提案し有用性を示す。さらに、指文字認識を応用させた指文字学習支援システムを構築し、その有効性を確かめるための評価実験を行なう。

1.2 本論文の構成

本論文の構成は次のとおりである。第 2 章では指文字、本研究で使用する入力装置について述べる。第 3 章では、手形状入力装置、三次元位置計測装置による計測データからのベイズ識別による連続認識手法、問題点について述べる。第 4 章では、ベイズ識別法の計算量の問題点を解決する手法として、全ベイズ識別法の次元数を減らしたベイズ・距離併用法を提案し、有用性について述べる。第 5 章では、指文字認識を応用させた指文字学習支援システムと、このシステムの有効性を確かめる評価実験の結果について述べる。第 6 章は結論である。

第 2 章

手形状を用いたインターフェース

2.1 はじめに

手形状を用いたインターフェースとして、現在までに、仮想現実 (VR : Virtual Reality) 環境での操作環境、ポインティングデバイス、コミュニケーション支援などの研究が報告されている [8]。この様な新しいコミュニケーション技術の導入の背景には、ハードウェア技術の急速な進展による新しいデバイスの出現が大きな役割を担っていると言える。

本章では、こうした新しいデバイスを利用することで可能となる、手振りをを用いたインターフェースについて議論する。また、本研究で用いる手形状入力装置について説明する。

2.2 直接操作環境

仮想現実環境の大きな特徴の一つは、操作する人間と計算機の作り出す人工的な三次元画像空間とのインタラクションが可能なことである。特に、手はこの様な仮想環境内を直接操作することが出来る。ATR 通信システム研究所の岸野らは自分の手を仮想的な 3 次元空間と融合させて対象操作を行なう環境を作った [9]。この環境では液晶シャッターで左右眼に異なる画像を提示し立体対象を作り、データグローブを装着した手で対象物を自由に操ることが出来る。この様な直接操作は自分の手による直接的な操作であって、違和感なく操作性が高まる利点がある。直接操作は、現実の力学世界により近いシミュレーションを行なえば行なうほど、本物に近い操作が可能となる。例えば、「仮想物体を持ち上げ

る」という操作に対してより現実性を高めようとする、物体の持ち上がる条件を静力学的に求めること（指と物体の接触判定の計算や、摩擦力によりその物体が持ち上がるか否かを計算する）が考えられる。しかし、この様な計算をリアルタイムで実現するには、現在ではまだ計算機的能力が不足している。これは仮想空間を表現する一画面を作成する単位時間ごとの計算時間が長くなり、画面の動きがぎこちなくなることを意味し、逆に現実感が失われることになる。従って、簡単なルール（「手が物体と接触して」おり、かつ「手が握る動作をしている」など）を適用せざるを得ない。このことは、手の形状自体に意味を持たせる必要性を示している。

2.3 コミュニケーション支援

2.3.1 手話

聴覚障害や発話障害者の積極的な社会進出に伴い、障害者と健常者との間のスムーズなコミュニケーションを必要とする機会が社会生活の様々な場面で多くなりつつある。聴覚、発話障害者のコミュニケーション手段は手話が主流である。しかし、健常者で手話の出来る人は限られており、コミュニケーションを行うには手話通訳者を必要とすることが多い。しかし、手話通訳者の不足等により、通訳を受けられないといった場合がある。また、一般に手話通訳者は通訳を必要としている人とは何ら関わりの無い第三者であるため、プライバシーの問題も発生してくる。そのため、通訳に代わるヒューマンコミュニケーション支援システムが望まれてきた。しかし、手話表現は複雑な手形状を用いているため、画像処理のみによる手話認識は困難であった。手形状入力装置を用いることにより、手話を認識するシステムを作り出すことができ、障害者自身が手話や指文字を熟知していれば、手話を知らない相手とのスムーズなコミュニケーション支援が可能となる。

2.3.2 指文字

指文字は手話法において、手話語がない場合や、新しい用語、単語などを表現する時、助詞を補間する場合などに用いる。一般に指文字は手話語と組み合わせて用いられる。日本語の指文字 [10][11] は、手話表現の一つであり、五十音を表す基本形の指文字は、‘あ’～‘ん’の 46 形状からなる。濁音 20 文字は、対応する静音の手形状を右に動かすことで表現す

る。ぱ行の半濁音 5 文字は、対応する静音の手形状を上にかぶることによって表現する。拗音（や、ゆ、よ）、促音（っ）、‘を’ は対応する静音の手形状を手前に動かすことによって表現する。‘の’、‘り’、‘ん’ は向きを変えて表現し、‘も’ は手形状を変えて表現する。図 2.1 に日本の指文字五十音の形状を示す。手話や指文字のように、手の形状自体が意味を持つような場合、その形状の意味を対応付ける標準辞書が必要となる。

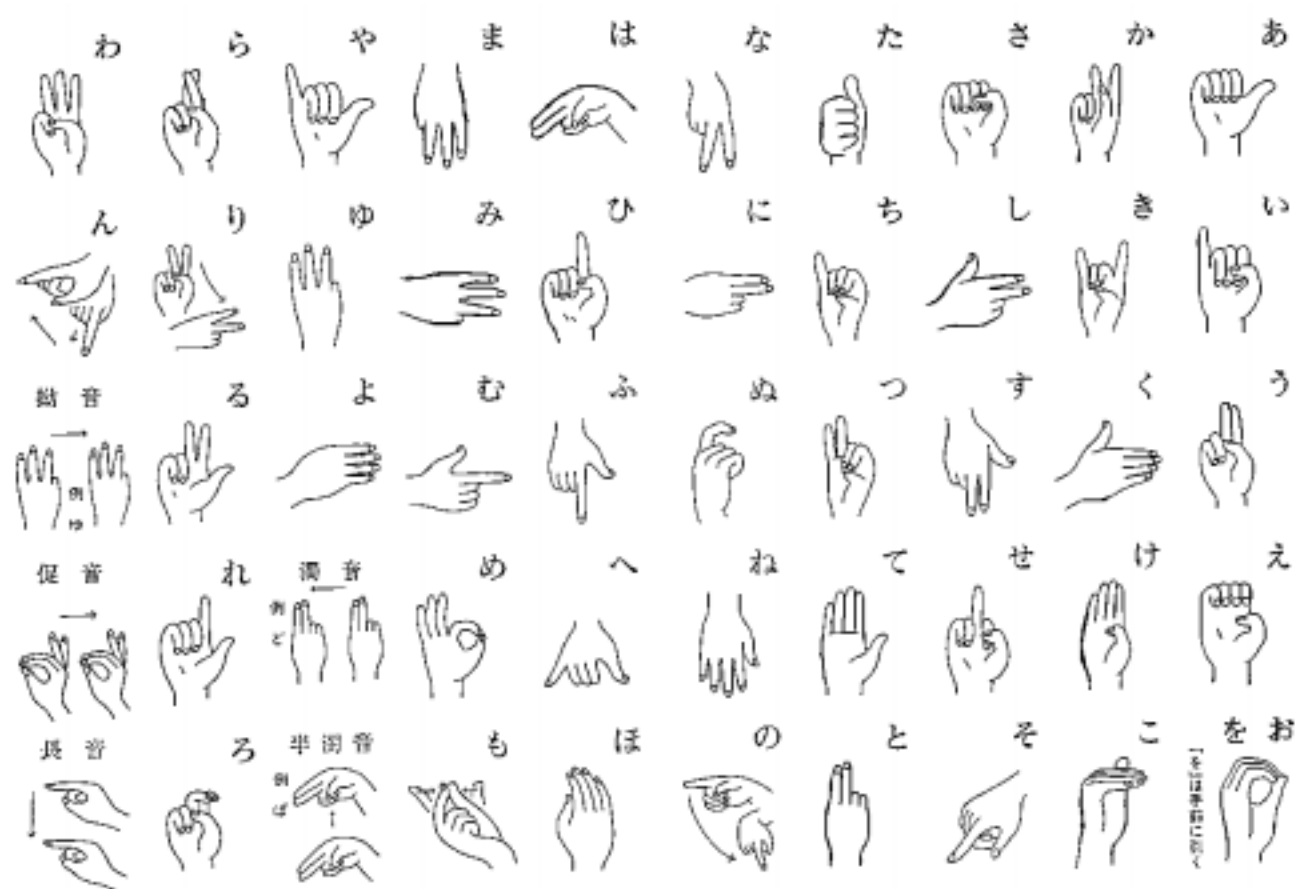


図 2.1: 指文字

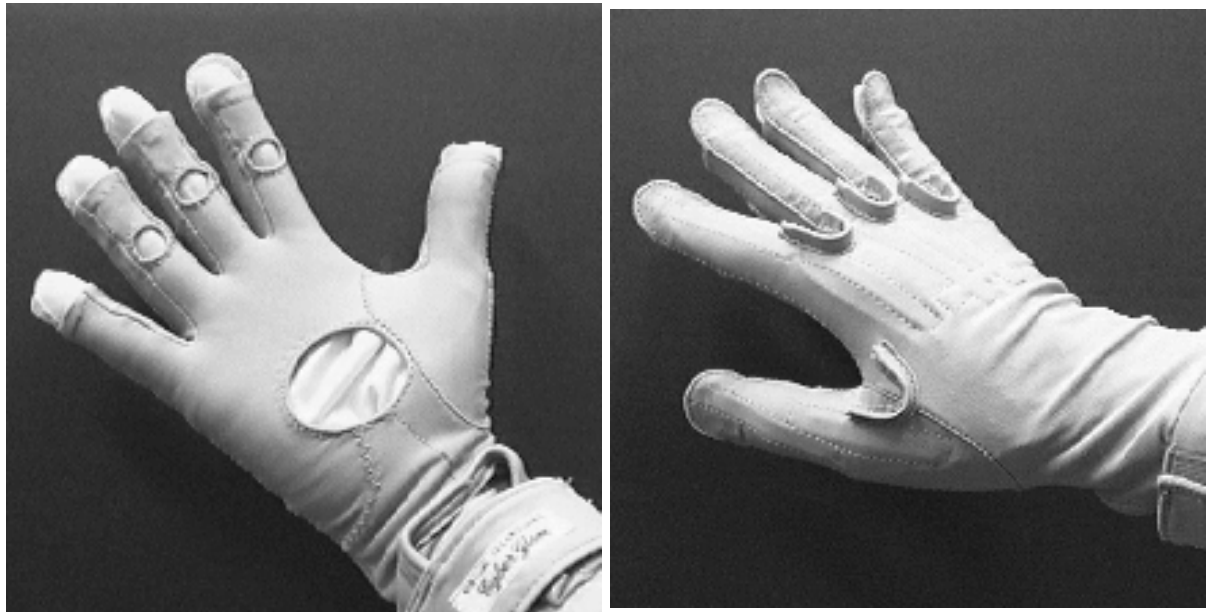


図 2.2: CyberGlove の外観

2.4 入力装置の概要

2.4.1 CyberGlove

本研究では、Virtual Technologies 社製の CyberGlove という手形状入力装置を用いる。CyberGlove は手指の屈曲角を測定する手袋状のもので、手に直接装着して用いる。外観を図 2.2 に示す。

この装置は 18 自由度 (表 2.1) を有し、人の複雑な手形状を精密に測定できる。得られる測定データも線形であり、複雑な変換操作を必要としない。

センサー類は、手と手指の関節部や他の主要な部分を覆うように取り付けられている。各指には、屈曲を測定するセンサーが 2 個ずつ付いており、親指には手の掌上を小指方向にどれだけ外転したかを測定するセンサーがある。各指間のセンサーはグローブの表面に貼り付けられた U 字型のもので、指間の外転角を測定する。これらのセンサーの出力は指の曲げ角度に対して線形であるため、複雑な変換を必要としない。出力からの線形変換は、 $\text{gain}(\text{slope})$ と $\text{offset}(\text{y-intercept})$ の 2 個のパラメータによりソフトウェア上で行なう。

データの入出力は、図 2.3 に示すように CGIU (CyberGlove Interface Unit) を通して

表 2.1: Sensor Data Byte Ordering

1. thumb rotation/tmc (angle of thumb rotating across palm)
2. thumb mcp (joint where the thumb meets the palm)
3. thumb ip (outer thumb joint)
4. thumb abduction (angle between thumb and index finger)
5. index mcp (joint where the index meets the palm)
6. index pip (joint second from finger tip)
7. middle mcp
8. middle pip
9. middle-index abd'n (angle between ring and middle fingers)
10. ring mcp
11. ring pip
12. ring-middle abduction (angle between pinkie and ring finger)
13. pinkie mcp
14. pinkie pip
15. pinkie-ring abduction (angle between pinkie and ring finger)
16. palm arch (causes pinkie to rotate across palm)
17. wrist pitch
18. wrist yaw

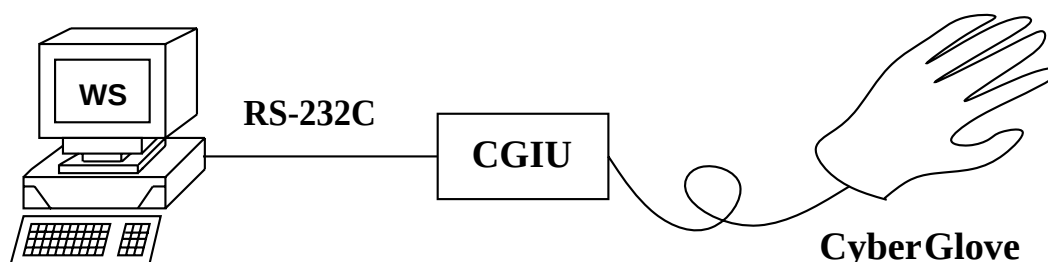


図 2.3: システム構成

Silicon Graphics 社製のワークステーションに接続して行なう。この CGIU には CyberGlove の信号を増幅および量子化する回路が含まれている。各センサーの信号は CGIU の中で増幅され、8bit に量子化され出力される。

キャリブレーション

手形状入力装置の性質上、操作者ごとの手の動作特性の違い、あるいは装着時におけるセンサーのずれなどをに対処するため、被験者は計測前にキャリブレーションを行なう必要がある。具体的なキャリブレーションは以下の手順で行なう。

1. 5 指全部を反らす。
2. 親指以外の指を曲げる。
3. 親指だけを曲げる。
4. 5 指全部を横で密着させる。
5. 5 指全部を外側に開く。
6. 親指と小指を接触させる。

これによって、被検者個人の、5 指全て関節角、および各指間の外転センサー出力の最大値、最小値が得られる。これらのセンサーから出力される値を C_{max_i} 、 C_{min_i} 、と表わす ($i = 0, 1, 2, \dots, 14$)。

屈伸率

各関節の測定値を正規化し、関節角を指が曲ってる割合を屈伸率 (bend ratio) を用いて表す。

入力値を $input$ とした時、各関節の屈伸率 $bend_i$ は次式で求めることができる。

$$bend_i = \frac{input_i - Cmin_i}{Cmax_i - Cmin_i} \quad (i = 0, 1, 2, \dots, 14) \quad (2.1)$$

以後、手形状入力装置から得られる測定データは屈伸率で表現する。

2.4.2 FASTRAK

FASTRAK とは、三次元位置情報を測定するためのセンサーである。得られるデータは三次元座標 (x,y,z) と回転座標 (yaw,pich,roll) である。送られてくるデータは float 型の精度で ASCII 形式で送られる。三次元座標のデータを連続で得ることにより、手の甲の動きを計測可能である。

第 3 章

全ベイズ識別法による連続指文字認識

3.1 はじめに

静的指文字は手形状・向きを認識の際の特徴量として用いることで認識が可能であった [7]。しかし、動作を伴う指文字を認識するには、手形状・向きの他に動きの変化の特徴量が必要である。本章では、手形状入力装置および FASTRAK から得られるデータの統計的特徴を用いた認識手法について述べる。そして、認識実験を行ない認識手法の評価を行なう。

3.2 全ベイズ識別法

3.2.1 手形状・向き・動きの測定

手形状の測定

手形状の測定には、2.4.1節で述べた CyberGlove を使用する。本研究で使用する関節角測定点は図 3.1 に示す 15 箇所（測定点 0 は TR 軸に対する回転角を表す）と、手首に関する測定点 3 箇所の計 18 箇所である。

向き・動きの測定

向き・動きの測定には、2.4.2節で述べた FASTRAK を使用する。FASTRAK からは三次元位置座標 (x,y,z) 、 x,y,z 軸に関する回転角 $(yaw,pitch,roll)$ が得られる。向きは、回転角から

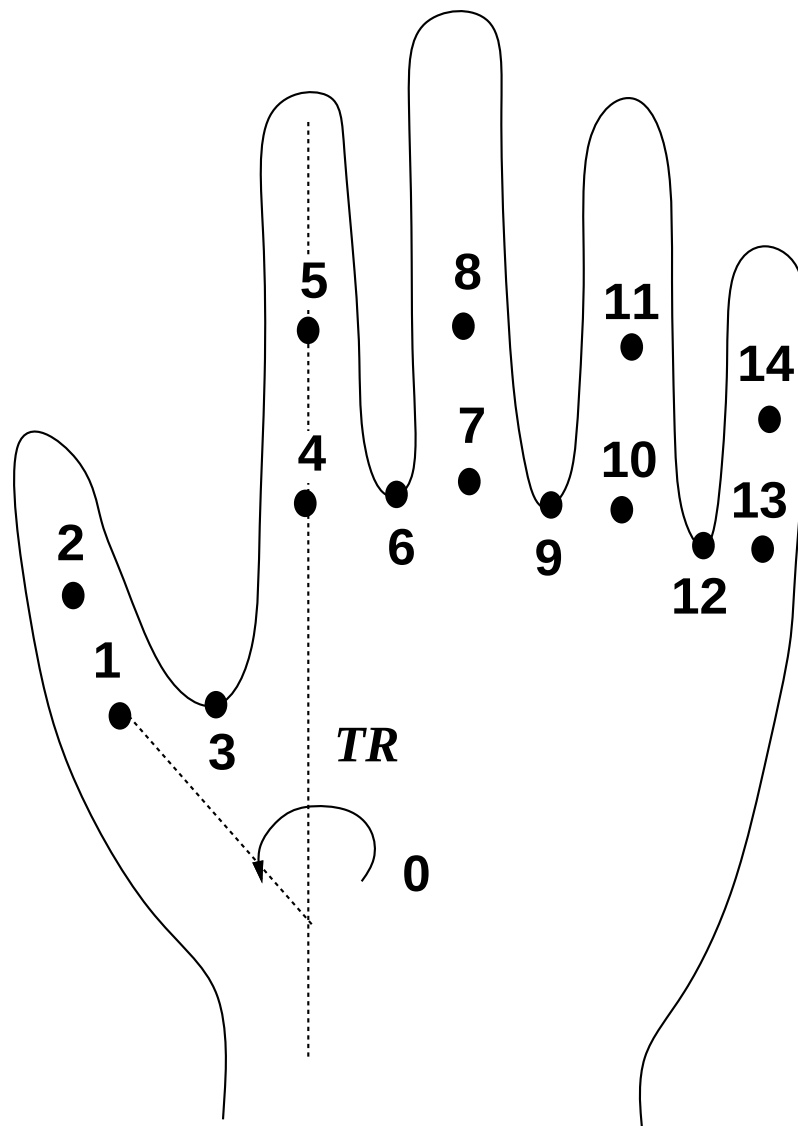


図 3.1: 手形状入力装置の関節角測定点

測定する。動きは三次元位置座標のデータを連続でとり、三次元位置の変化 ($\Delta x, \Delta y, \Delta z$) で示す。ここで、 $\Delta x, \Delta y, \Delta z$ は

$$\Delta x = x_i - x_{i-1}, \Delta y = y_i - y_{i-1}, \Delta z = z_i - z_{i-1} \quad (3.1)$$

である。 x_i, y_i, z_i は最初から i 番目のデータであることを示す。

3.2.2 ベイズ識別法

認識対象とするベクトルを $\mathbf{x}$ とする。 $\mathbf{x}$ の要素は各関節角の屈伸率 (18 次元)、回転角 (3 次元)、3 次元位置の変化 (3 次元) の計 24 次元である。入力された $\mathbf{x}$ から、その形状が指文字 l を表す場合を考える。 l を表す時、 $\mathbf{x}$ が正規分布確率を持っているとすると、まずカテゴリ条件付確率は、

$$P(\mathbf{x}|l) = \frac{1}{(2\pi)^{\frac{I}{2}} |\Sigma^{(l)}|^{\frac{1}{2}}} \cdot e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}^{(l)})\Sigma^{(l)-1}(\mathbf{x}-\boldsymbol{\mu}^{(l)})^T} \quad (3.2)$$

となる。ここで、 $\boldsymbol{\mu}^{(l)} : \mathbf{x}^{(l)}$ の平均ベクトル、 $\Sigma^{(l)} : \mathbf{x}^{(l)}$ の共分散行列、 I : ベクトル $\mathbf{x}^{(l)}$ の次元数、 l : 各文字パターンである。

一方、手の形状が $\mathbf{x}$ であるときに表す文字が l である条件付き確率は、ベイズの定理から、

$$P(l|\mathbf{x}) = \frac{P(l)P(\mathbf{x}|l)}{P(\mathbf{x})} \quad (l = 1, 2, \dots, L) \quad (3.3)$$

従って、手形状 $\mathbf{x}$ が入力された場合、その形状がどの文字を表すかは $P(l|\mathbf{x})$ が最大となる l を求めれば良い。

式 3.3 において分母の $P(\mathbf{x})$ はすべての l に関して同一であるから、 $P(l|\mathbf{x})$ の大小関係を判断する際には影響を与えない。従って、この項を除いた式、

$$g^{(l)}(\mathbf{x}) = P(l)P(\mathbf{x}|l) \quad (3.4)$$

の $g^{(l)}(\mathbf{x})$ の大小のみを判断すれば良い。

これに、式 3.2 を代入し、

$$g^{(l)}(\mathbf{x}) = P(l) \cdot \frac{1}{(2\pi)^{\frac{I}{2}} |\Sigma^{(l)}|^{\frac{1}{2}}} \cdot e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}^{(l)})\Sigma^{(l)-1}(\mathbf{x}-\boldsymbol{\mu}^{(l)})^T} \quad (3.5)$$

が得られる。さらに右辺の自然対数をとる。

$$\log P(l) - \frac{I}{2} \log 2\pi - \frac{1}{2} \log |\Sigma^{(l)}| - \frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}^{(l)})\Sigma^{(l)-1}(\mathbf{x} - \boldsymbol{\mu}^{(l)})^T \quad (3.6)$$

自然対数は単調増加関数であるから、やはり値の大小関係を変更しない。また第二項は定数ゆえ、問題なく省略可能である。以上より正規分布確率の場合に対する識別関数は次のようになる。

$$g^{(l)}(x) = -\frac{1}{2}(x - \mu^{(l)})\Sigma^{(l)-1}(x - \mu^{(l)})^T - \frac{1}{2}\log|\Sigma^{(l)}| + \log P(l) \quad (3.7)$$

このとき、各文字の事前生起確率は同じとすれば、最終的な識別関数は、以下の式で表される。

$$g^{(l)}(x) = -\frac{1}{2}(x - \mu^{(l)})\Sigma^{(l)-1}(x - \mu^{(l)})^T - \frac{1}{2}\log|\Sigma^{(l)}| \quad (l = 1, 2, \dots, L) \quad (3.8)$$

これに基づき、識別は、 $g^{l_0}(x) = \text{Max } g^{(l)}(x) \quad (1 \leq l \leq L)$ である l_0 を求めれば良い。

標準辞書には、 $\mu^{(l)}$ 、 $\sigma^{(l)}$ 、 $\Sigma^{(l)-1}$ および、 $\log|\Sigma^{(l)}|$ が必要である。

3.2.3 入力データ

実験では、あらかじめ 1 人の被験者に 50 人の哲学者の名前 (表 3.1) を 2 回提示してもらい、32Hz でサンプリングされた連続データを入力データ (データ 1、データ 2) として用いる。被験者は指文字経験者である。この入力データは濁音、半濁音、拗音、促音、長音、‘の’、‘り’、‘ん’ の動作を伴う指文字を多く含んでいる。そのため、これらの単語を正しく認識することにより、動作を伴う指文字を含めた指文字の連続認識が可能であると思われる。

3.2.4 標準辞書

手形状入力装置を装着した被験者に対して、10 人の被験者から、‘あ’～‘ん’ の 45 文字を一組として、1 人について 10 組のデータを採取する (1 文字につき 100 個のデータ)。被験者は指文字経験者、および、指文字の知識は有するが、日常生活において手話・指文字の使用経験は無い健常者である。この 100 組のデータを手形状の辞書作成に用いる。

向き・動きは個人差があるため、入力データを辞書作成に用いる。入力データで同じ文字を提示している区間を切り出し、その区間の回転角、三次元位置の変化のデータにより向き・動きの辞書を作成する。ただし、‘ん’ に対しては下向きの動きと上向きの動きがあるため、下向き用と上向き用の 2 つのデータを用意する。実験では動作を伴う指文字を含めた 77 個 (‘ん’ は 2 個) の辞書を使用する。

表 3.1: 実験に用いる単語

ミル	ルカーチ	デカルト	カッシラー	アンセルムス
キケロ	プラトン	サルトル	エウリゼナ	エックハルト
ゼノン	オッカム	パスカル	クザーヌス	キルケゴール
カント	ベーコン	ホッブス	ソシュール	アダムスミス
ロック	スピノザ	ヒューム	リッケルト	ブレンターノ
タレス	ベンサム	アベラール	マキャベリ	ライプニッツ
バルト	ヘーゲル	ポアンカレ	フッサール	ハイデッガー
デリダ	ブロッホ	ベルグソン	ピタゴラス	アリストテレス
マルクス	マルセル	ヤスパース	エピキュロス	モンテスキュー
デューイ	フーコー	ソクラテス	パルメニデス	アルチュセール

3.3 全ベイズ識別法による連続指文字動作の認識実験

3.3.1 全ベイズ識別法の連続指文字認識

図 3.2 に全ベイズ識別法の連続指文字認識の流れを示す。まず、Cyberglove から 18 次元の関節角データ、および FASTRAK から回転角 (3 次元)、三次元位置 (3 次元) のデータを得る。三次元位置は前のデータとの差をとり、変化量 ($\Delta x, \Delta y, \Delta z$) を計算する。次に、この 24 次元の入力に対してベイズ識別法の識別関数 (式 3.8) を 77 個について計算する。そして、最も識別関数の値が高いものを文字候補とし、同じ文字候補の続いた回数 (ren とする) を違う文字候補が現れるまで数える。同じ文字候補の続いた回数が N 回以上のとき、文字として出力する。ただし、‘ん’ に関しては ‘ん’ の下向きが N_d 回続いた後、‘ん’ の上向きが N_u 回続いた時、‘ん’ を出力するものとする。図 3.3 に切り出し法のフローチャートを示す。ここで、chr は認識された文字候補、moji は連続回数を数えるために chr と比較する文字列、ren は連続回数を示す。切り出しの例を図 3.4 に示す。

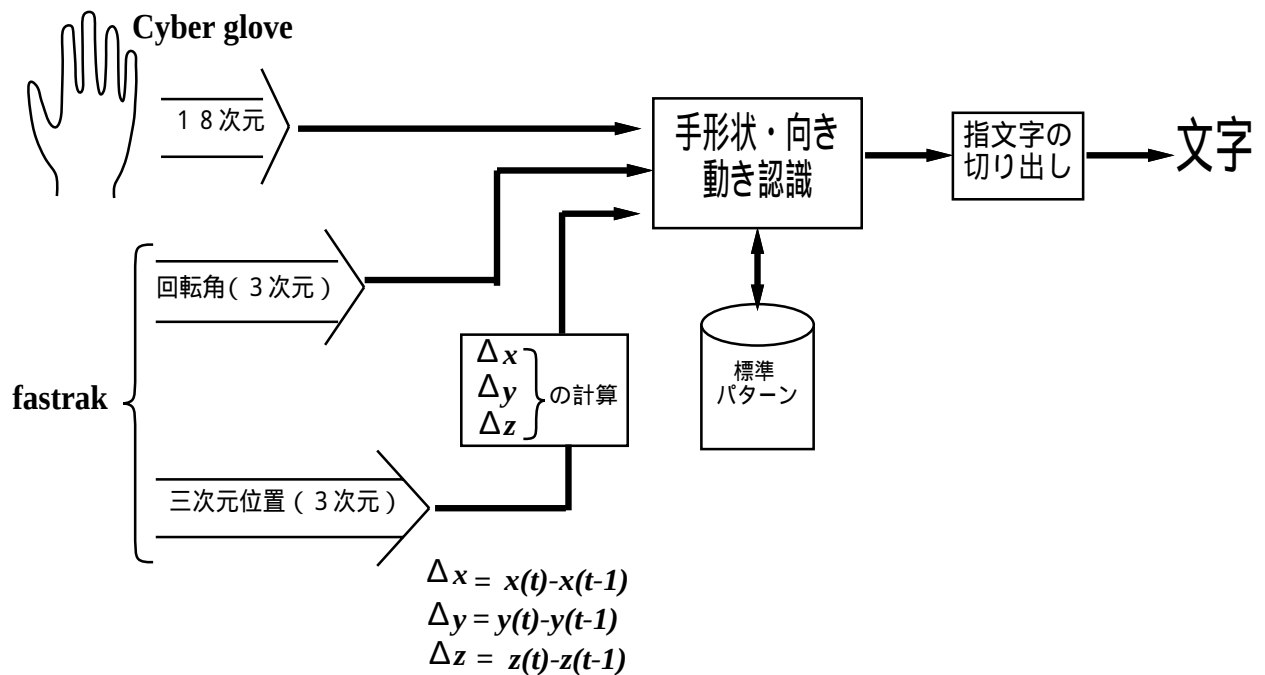


図 3.2: 連続指文字認識の流れ

3.3.2 実験結果

入力データは 50 人の哲学者の名前を用いる。文字数は 231 文字を 2 組、合計 462 文字である。適切な N を設定するため、 $N = 5, 6, 7, 8$ と変化させ認識実験を行なった。 $N = 5, 6, 7, 8$ のときの認識率を表 3.2 に示す。 $N_d = 3$ 、 $N_u = 2$ とした。これは、入力データで‘ん’の文字の認識結果を参照したところ、‘ん’の下向きの連続回数の最小が 3 回、‘ん’の上向きの連続回数が 2 回であったためである。ここで、誤字数は切り出された文字が提示した指文字と違う文字の数であり、脱字数は文字候補の続いた回数が N 未満であることを示す。表 3.2 を見てわかるように、 $N = 6$ のとき認識率が最高になる。誤字数は N が増えるにしたがい減っていくものの、脱字数は増えていく。これは、 N が増えるにしたがい誤字であったものが脱字になっていることを示す。

$N = 6$ での実験結果を表 3.3、静的指文字だけの実験結果を表 3.4、動作を伴う指文字の実験結果を表 3.5 に示す。静的指文字の認識率が 96.50%、動作を伴う指文字の認識率が 92.05%、全体で 95.45% と高い認識率が得られる。実験結果を見てみると、誤字数は少ないものの脱字数が多い。以下では、誤字・脱字の原因について述べる。

N=6のとき

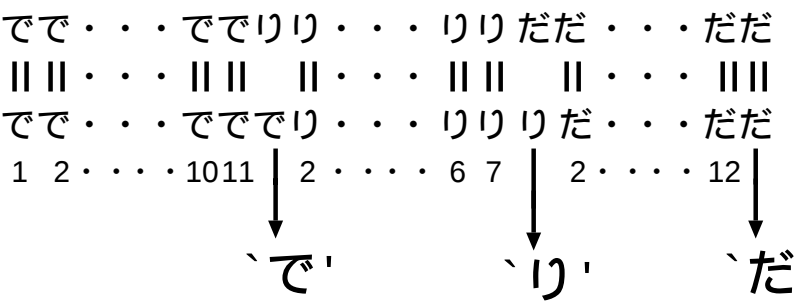


図 3.4: 切り出し例

表 3.2: $N = 5, 6, 7, 8$ のときの認識率 (全ベイズ識別法)

N	誤字数	脱字数	認識率 (%)
5	10	19	93.72
6	2	21	95.02
7	2	41	90.69
8	2	52	88.31

表 3.3: 実験結果 (全ベイズ識別法)

	文字数	誤字数	脱字数	認識率 (%)
データ 1	231	1	9	95.67
データ 2	231	1	12	94.37
合計	462	2	21	95.02

表 3.4: 静的指文字の実験結果 (全ベイズ識別法)

	文字数	誤字	脱字	認識率 (%)
データ 1	143	0	6	95.80
データ 2	143	0	4	97.20
合計	286	0	10	96.50

表 3.5: 動作を伴う指文字の実験結果 (全ベイズ識別法)

	文字数	誤字数	脱字数	認識率 (%)
データ 1	88	1	3	95.45
データ 2	88	1	8	89.77
合計	176	2	11	92.61

表 3.6: 動作を伴う指文字の脱字の分類

	脱字数	長音	‘の’	‘ん’	濁音	半濁音	拗音
データ 1	3	3	0	0	0	0	0
データ 2	8	3	1	1	1	1	1
合計	11	6	1	1	1	1	1

静的指文字の脱字 10 文字は、似たような指文字に誤認識され切り出されなかったものである。例えば、‘ス’と‘リ’、‘タ’と‘ア’、‘レ’と‘ヒ’などである。動作を伴う指文字のデータ 1 の誤字は、デリダの‘デ’が‘テ’と誤認識されたものである。これは、動きがなめらかであったため、静的指文字と識別したためである。データ 2 の誤字は、ブレンターノの‘一ノ’が‘ン’と誤認識されたものである。これは、‘一’から‘ノ’へ移動するときの動きが‘ん’と似ているため、誤認識したためである。

動作を伴う指文字の脱字の分類を表 3.6 に示す。長音の脱字のうち 4 文字は、長音の提示区間が短かったために切り出されなかったものである。残り 2 文字は、長音の提示区間に‘ん’の下向きと認識されているのが含まれていたために切り出されなかったものである。‘の’の脱字は、‘ん’の下向きに誤認識されて切り出されなかったものである。‘ん’の脱字は、‘ん’の下向きのあと‘ん’の上向きが続かなかったためのものである。濁音 1 文字、拗音 1 文字は提示区間が短かったために切り出されなかったものである。半濁音 1 文字は、エピキュロスの‘ピ’で‘ん’の上向きと誤認識されたものである。長音、‘の’、‘ん’の脱字は 11 文字中 8 文字である。この 3 文字は同一形状で動きも似ているため、それぞれ誤認識される場合が多い。また、提示区間が短かかったために切り出されなかったものは、11 文字中 6 文字であり、濁音、半濁音などの動作を伴う指文字に応じて、切り出しの回数を変える必要があると思われる。

3.3.3 全ベイズ識別法の問題点

この手法は静的指文字、動作を伴う指文字を同等に扱うことができ、切り出し法も簡単であるという利点がある。しかし、文字数、入力される次元数が増加することにより、計算量が非常に多くなり、認識に時間がかかってしまう。また、文字数が多くなると、似た形状、似た動きが多くなり、識別が困難となる。そのため、同一の文字が長く続かなくなり、脱字数が多くなる。

3.4 まとめ

本章では、動作を伴う指文字を含めた連続指文字動作を認識するための手法として全ベイズ識別法を提案し、認識実験を行なった。その結果 $N = 6$ で 94.37% の高い認識率が得られた。動作を伴う指文字の認識率も 92.05% と高く、全ベイズ識別法は連続指文字認識

に有効な手法であると言える。しかし、指文字学習支援システムへの応用を考えると、計算量が少ないことが望まれる。よって、次元数を削減した認識手法が必要がある。また、動作を伴う指文字の種類(濁音、長音など)により、提示している区間が異なるため、それぞれにあった切り出し回数が必要であると思われる。

第 4 章

ベイズ・距離併用法による連続指文字認識

4.1 はじめに

指文字動作を連続で認識する際、リアルタイムな認識が必要となる。全ベイズ識別法では、手形状、向き、動きの 24 次元を入力し、ベイズ識別法の識別関数を計算する認識手法であった。そのため辞書サイズ、計算時間とも膨大なものとなる。指文字学習支援システムへの応用を考えると、認識の高速化が必要となる。そこで本章では、計算量を削減するため全ベイズ識別法の次元数を減らしたベイズ・距離併用法を提案する。この手法では、動きを認識するため動作を伴う指文字の特徴を考慮しなければならない。そこで、FASTRAK から得られる三次元位置のデータをもとに動作を伴う指文字の特徴をとらえる。そして、提案した認識手法の有効性を検討するため、全ベイズ識別法との比較、検討を行なう。

4.2 動作を伴う指文字の特徴

ベイズ・距離併用法は手形状のみをベイズ識別法で認識する手法である。よって、動きを認識するには FASTRAK から入力される三次元位置のデータより、静的指文字、動作を伴う指文字の特徴をとらえることが必要である。

図 4.1 に手形状・向き・動きによる指文字の分類を示す。指文字は手形状により 24 に分類され、向き・動きにより 1 文字に決まる。よって、手形状のみ、あるいは手形状・向きで文字候補をしばり、動きの特徴をとらえ 1 文字に特定すれば良い。そこで、濁音、半濁

音、拗音、促音、長音、‘の’、‘ん’の動作を伴う指文字の特徴を FASTRAK から得られる三次元位置のデータをもとに抽出する。‘も’、‘り’は動作を伴う指文字であるが、手形状・向きで識別可能であるため、動きの特徴は考慮しない。

表 4.1 に動作を伴う指文字の特徴を示す。この特徴は、50 人の哲学者の名前のデータをもとにしたものである。濁音 (1)、半濁音 (1) は相手に対して掌を向ける指文字、濁音 (2)、半濁音 (2) は相手に対して甲を向ける指文字である。2 つに分類したのは掌を向けるか、甲を向けるかによって FASTRAK から得られる三次元位置の変化の方向が異なるためである。また、- は特徴がないことを示す。‘ん’ は下向きと上向きの動きがあるため、両方の特徴を記す。動作を伴う指文字を決めるには特徴の続いた回数、特徴が続いている区間での移動距離を考慮すれば良い。例えば、濁音 (1) の手形状で $\Delta x > 0$ 、 $\Delta y > 0$ が最低 5 回続き、この特徴が続いている区間での x 軸の移動距離が 5.02、 y 軸の移動距離が 1.95 以上であることが濁音の条件である。表 4.1 の移動距離は、特徴の始まったところの三次元位置 (bg_x, bg_y, bg_z) と終わったときの三次元位置 (end_x, end_y, end_z) の差分の最小値 ($\min(end_x - bg_x), \min(end_y - bg_y), \min(end_z - bg_z)$) である。

長音、‘の’、‘ん’の下は手形状が同一であり、動きの特徴も似ている。長音は x, y, z の変化量は全て負、‘の’は x, y のみが負、‘ん’の下向きの動きは y のみが負である。‘の’の z の変化量、‘ん’の x, z の変化量は正にも負にもなるため、この 3 つの識別は難しい。

4.3 ベイズ・距離併用法による指文字認識

連続認識の流れを図 4.2 に、手形状・向き・動き認識のフローチャートを図 4.3 に示す。まず、Cyberglove から得られる 18 次元の関節角データより、手形状を認識する。手形状認識により、同一の手形状 (表 4.1 の 1 ~ 24 の手形状) に候補をしぼる。そして、三次元位置の変化量の計算を行なう。手形状認識で得られた結果が、表 4.1 の 1 ~ 6 であれば文字候補が決定する。8 ~ 14 であれば、手形状と動きにより、文字候補が決定する。15, 22 では動作を伴う指文字はそれぞれ ‘だ’、‘ぶ’ と ‘ぶ’ であり、他は静的指文字である。よって、濁音、半濁音 (2) の両方の特徴である x の差分をみて、動作を伴う指文字と静的指文字を識別し、文字候補を決定する。 $\Delta x > 0.2$ にしたのは、‘だ’、‘ぶ’、‘ぶ’ を提示している区間での $\min\{\Delta x\}$ が 0.21 であるためである。それ以外の場合は、向き認識を行ない文字候補を決定する。この文字候補を chr とする。

動きの認識では、chr と動きの特徴より濁音、半濁音、拗音、促音、長音の特徴を識別

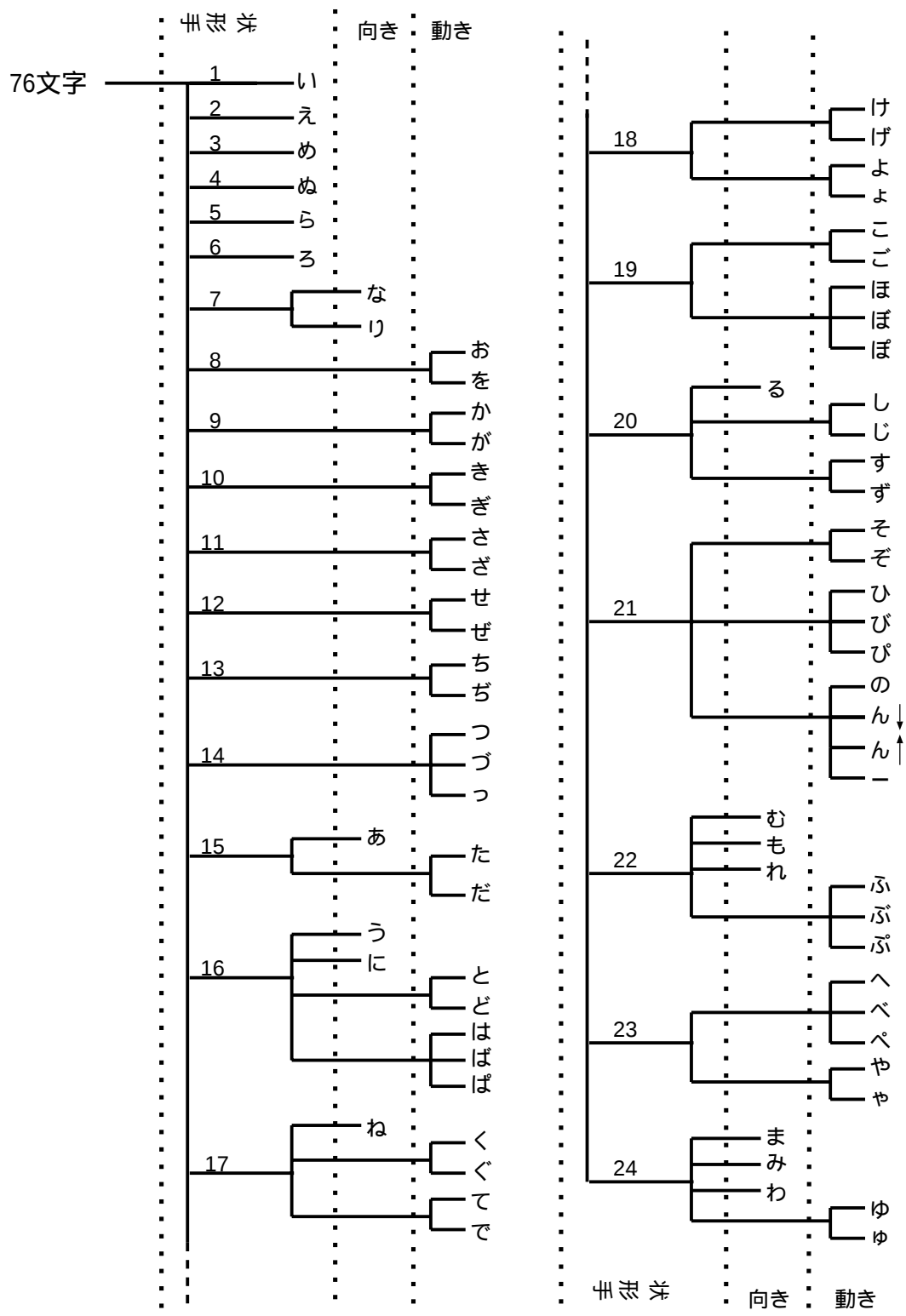


図 4.1: 手形状・向き・動きによる指文字の分類

表 4.1: 動作を伴う指文字の条件

	特徴	連続回数	移動距離
濁音 (1)	$\Delta x > 0$ $\Delta y > 0$ $-$	5 回	5.02 1.95 $-$
濁音 (2)	$\Delta x > 0$ $\Delta y < 0$ $-$	5 回	4.21 -0.85 $-$
半濁音 (1)	$-$ $\Delta y > 0$ $\Delta z > 0$	5 回	$-$ 6.06 2.86
半濁音 (2)	$\Delta x > 0$ $-$ $\Delta z > 0$	5 回	8.41 $-$ 5.96
拗音、促音	$\Delta x > 0$ $\Delta y < 0$ $\Delta z > 0$	4 回	1.91 -3.35 2.49
長音	$\Delta x < 0$ $\Delta y < 0$ $\Delta z < 0$	4 回	-2.53 -4.71 -3.34
の	$\Delta x < 0$ $\Delta y < 0$ $-$	4 回	-5.08 -9.06 $-$
ん (下向)	$-$ $\Delta y < 0$ $-$	4 回	$-$ -2.15 $-$
ん (上向)	$-$ $\Delta y > 0$ $-$	3 回	$-$ 3.31 $-$

濁音 (1) ... か、き、け、こ、さ、せ、そ、た行、は、ひ、ほ

濁音 (2) ... く、し、す、ふ、へ

半濁音 (1) ... は、ひ、ほ

半濁音 (2) ... ふ、へ

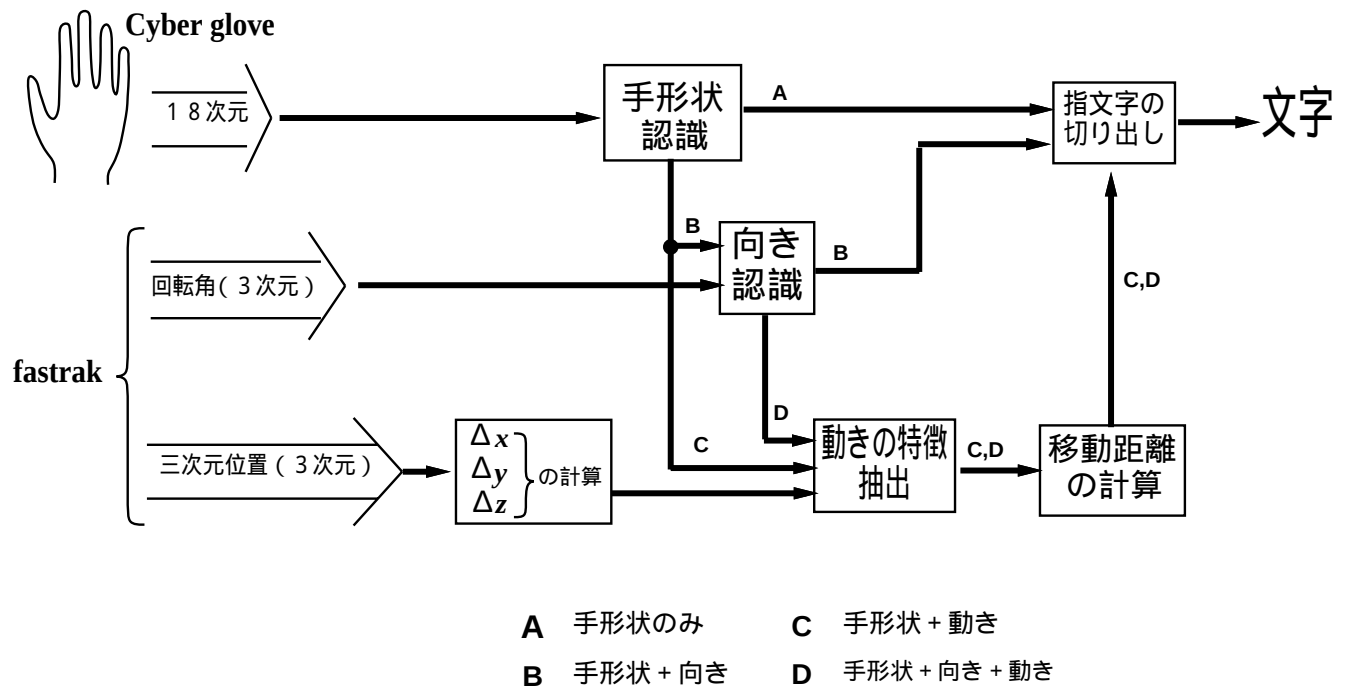


図 4.2: 連続指文字認識の流れ

する。

手形状・向き・動き認識で得られた chr と動きの特徴をもとに文字の切り出しを行なう。

以下では、手形状、向き、動きの認識、および切り出し手法について述べる。

4.3.1 手形状の認識

手形状認識では、Cyberglove から得られる 18 次元のデータに対して、ベイズ識別法の識別関数 (式 3.8) を計算する。そして、識別関数の値が最も高いものを手形状として出力する。これにより、24 形状のうち 1 形状に選ばれる (図 4.1)。

4.3.2 向きの認識

手形状のみ、手形状+動きでは識別できない文字候補に対して、向きの認識を行なう。向き認識の辞書には、 x, y, z 軸に関する回転角の平均値 ($x_{\mu}^{(l)}, y_{\mu}^{(l)}, z_{\mu}^{(l)}$) を用いる。ここで、 $l = \text{あ, い, } \dots, \text{ん}$ である。向きの認識は手形状認識で選ばれた 1 形状に対して距離計

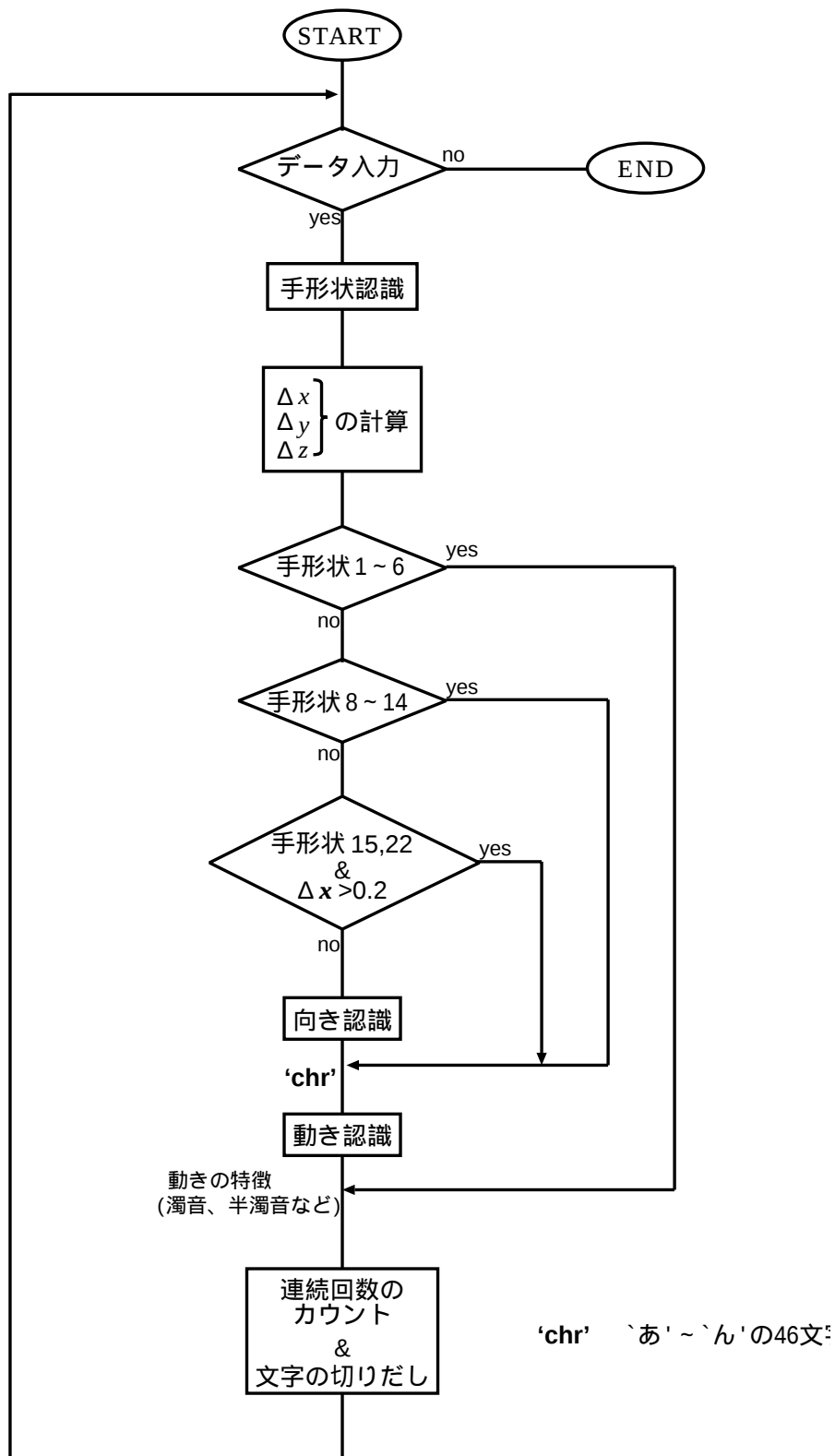


図 4.3: 手形状・向き・動き認識のフローチャート

算を行なう。入力される回転角データ $(x_{roll}, y_{roll}, z_{roll})$ とする。同一の手形状をもつ文字候補に対して、

$$DIST^{(l)} = |x_{roll} - x_{\mu}^{(l)}| + |y_{roll} - y_{\mu}^{(l)}| + |z_{roll} - z_{\mu}^{(l)}| \quad (4.1)$$

の計算を行ない、 $\min\{DIST^{(l)}\}$ の l を認識結果として出力する。

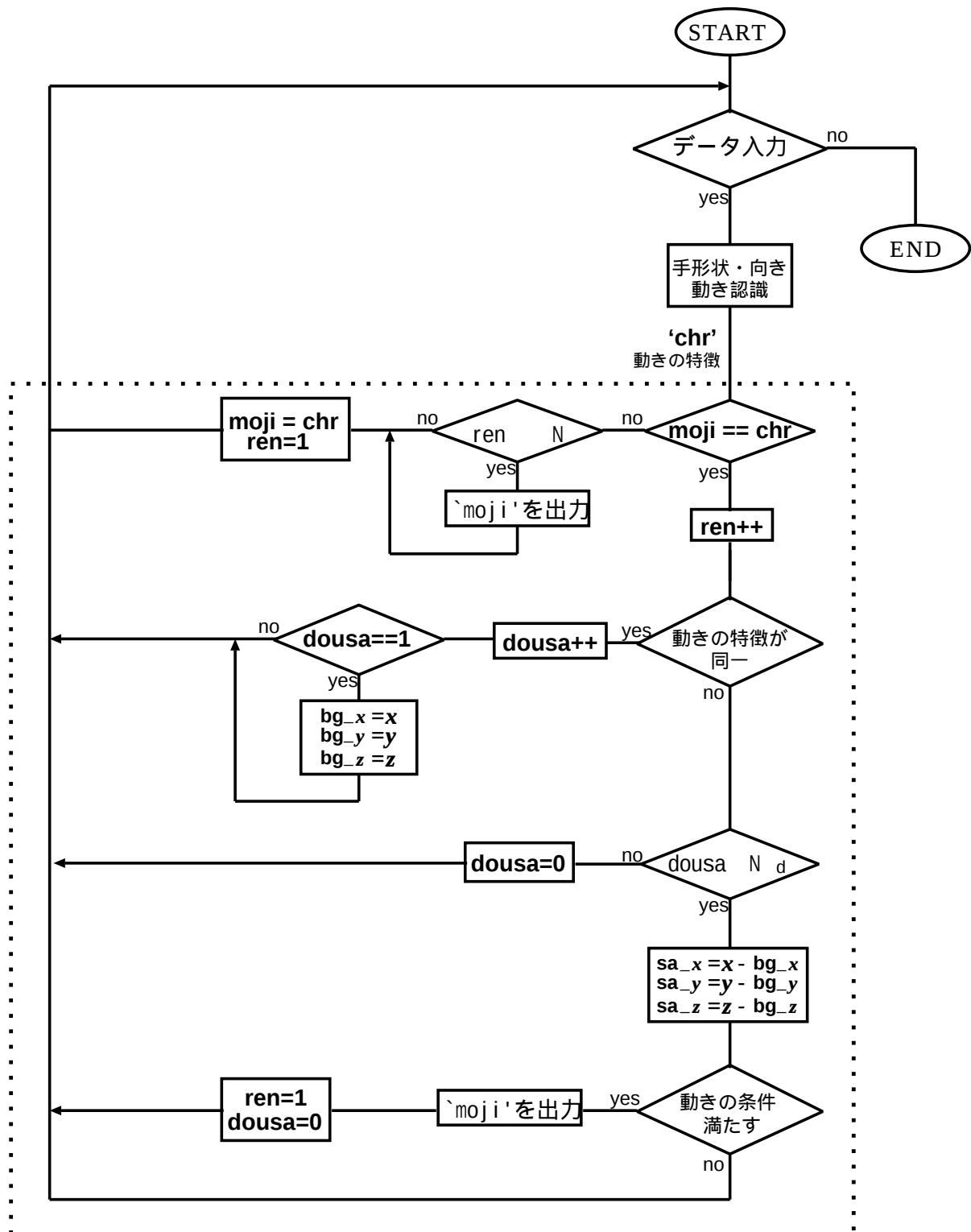
4.3.3 動きの認識

動き認識では、chr と三次元位置の変化量 $(\Delta x, \Delta y, \Delta z)$ より、濁音、半濁音などの動作の伴う指文字の特徴を識別する。表 4.1 の特徴がないときは静的指文字と認識する。手形状・向き・動きの認識により、chr と動作を伴う指文字の特徴が得られ、これをもとに切り出しを行なう。

4.3.4 切り出し手法

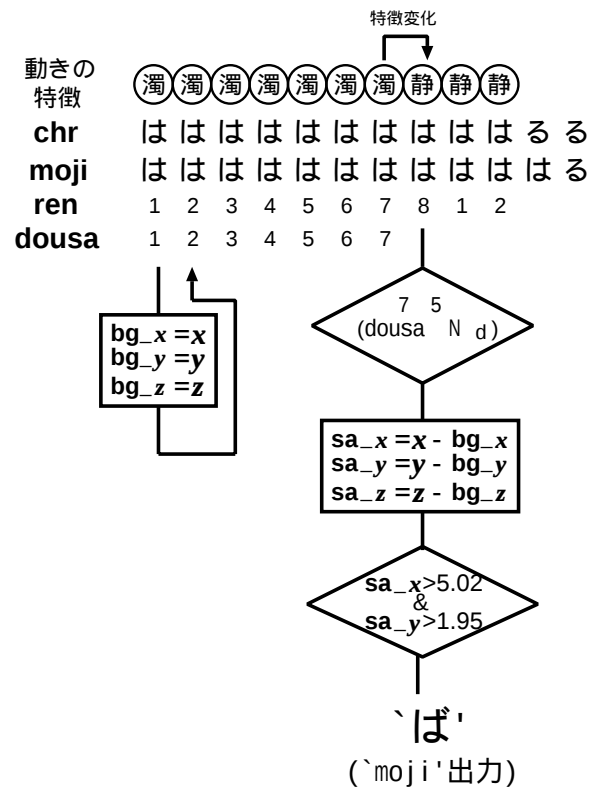
切り出し手法を図 4.4 に示す。手形状のみ、あるいは手形状・向きの認識より、文字候補 chr が、動きの認識より動きの特徴が出力される。moji は同一の文字候補の続いた回数を数えるために、chr と比較する文字である。まず、同じ文字候補の続いた回数 (これを ren とする) を数える。動きの特徴を満たす (表 4.1 の特徴参照) ときは、続いた回数 (これを dousa とする) を数える。 $dousa = 1$ のときは、指文字動作区間内に動いた距離を計算するため、三次元位置を格納しておく (これを bg_x, bg_y, bg_z とする)。そして、同一の特徴が続かなくなったとき dousa と表 4.1 の連続回数 (これを N_d とする) とを比較する。 $dousa \geq N_d$ のときは動作区間内に動いた距離 (これを sa_x, sa_y, sa_z とする) を計算し、その移動距離が表 4.1 の移動距離を満たすとき、動作を伴う指文字と認識し、相当する文字を出力する。そうでないときは、 $dousa = 0$ とする。

切り出しの例を図 4.5 に示す。手形状・向きの認識結果として chr が ‘は’、動きの認識結果として濁音の特徴があらわれる。そして、‘は’ の続いた回数と濁音の特徴が続いた回数を違う特徴が出力するまで数える。違う特徴が出力されたら、それまで濁音の続いた回数と表 4.1 の連続回数を比較する。連続回数が条件を満たしたら、三次元位置の移動距離を計算する。その移動距離が条件を満たしたら、動作を伴う指文字とみなし、‘ば’ を出力する。



文字の切り出し

図 4.4: 切り出し手法のフローチャート



① 濁音の特徴

$$\begin{pmatrix} \Delta x > 0 \\ \Delta y > 0 \end{pmatrix}$$

図 4.5: 切り出し手法の例

4.4 ベイズ・距離併用法による連続指文字動作の認識実験

4.4.1 実験結果

入力データは 50 人の哲学者の名前を用いる。文字数は 231 文字を 2 組、合計 462 文字である。適切な N を設定するため、 $N = 5, 6, 7, 8$ と変化させ認識実験を行なった。 $N = 5, 6, 7, 8$ のときの認識率を表 4.2 に示す。 $N = 5$ では指文字提示区間の文字だけでなく、次の指文字へ移行する区間の指文字も認識されてしまい誤字数が多くなる。 $N = 6$ で誤字数が安定していることから、次の指文字へ移行している区間で同じ指文字が続く回数は 5 回程度であると考えられる。

$N = 6$ のときの実験結果を表 4.3、静的指文字のみの実験結果を表 4.4、動作を伴う指文字のみの実験結果を表 4.5 に示す。静的指文字の認識率が 95.10%、動作を伴う指文字の認識率が 93.18%、全体で 94.37% と高い認識率が得られている。

動作を伴う指文字の誤字・脱字の分類を表 4.6 に示す。全誤字数 10 文字中 5 文字は、長音、‘ん’ の誤字である。‘ん’ の誤字 (4 文字) は、長音への誤認識が 2 文字、‘の’ が 1 文字、静的指文字の‘そ’ が 1 文字である。長音の誤字 (1 文字) は、動きが少なく静的指文字の‘そ’ と認識されたものである。濁音の誤字 2 文字は、向きが誤認識されたものである。残り 3 文字は動きが少なく静的指文字と認識されたものである。脱字の 2 文字は、ハイデッカーの‘デ’ とモンテスキューの‘ユ’である。‘デ’ は提示区間の途中まで同一形状の‘く’ と認識されていたため、切り出されなかったものである。‘ユ’ は途中まで同一形状の‘マ’ と認識されていたため、切り出されなかったものである。

静的指文字の誤字 5 文字は、全て向きの誤認識である。‘す’ が‘し’ と認識されているのが 3 文字、残りは‘に’ が‘と’、‘て’ が‘く’ となるものである。脱字の 8 文字中 6 文字は、連続回数が 4, 5 回であったため切り出されなかった文字である。脱字の残り 2 文字はベルグソンの‘そ’ であり、次の文字の‘ん’ と同一の手形状であるため、‘ん’ と認識されたものである。

表 4.2: $N = 5, 6, 7, 8$ のときの認識率 (ベイズ・距離併用法)

N	誤字数	脱字数	認識率 (%)
5	23	7	93.51
6	15	11	94.37
7	15	14	93.72
8	15	21	92.21

表 4.3: 実験結果 (ベイズ・距離併用法)

	文字数	誤字数	脱字数	認識率 (%)
データ 1	231	8	7	93.51
データ 2	231	7	4	95.24
合計	462	15	11	94.37

表 4.4: 静的指文字の実験結果 (ベイズ・距離併用法)

	文字数	誤字数	脱字数	認識率 (%)
データ 1	143	2	5	95.10
データ 2	143	3	4	95.10
合計	286	5	9	95.10

表 4.5: 動作を伴う指文字の実験結果 (ベイズ・距離併用法)

	文字数	誤字数	脱字数	認識率 (%)
データ 1	88	6	2	90.91
データ 2	88	4	0	95.45
合計	176	10	2	93.18

表 4.6: 誤字・脱字の分類

	誤字数	‘ん’	‘長音’	濁音	脱字数	濁音	拗音
データ 1	6	2	1	3	2	1	1
データ 2	4	2	0	2	0	0	0
合計	10	4	1	5	2	1	1

4.4.2 比較・検討

全ベイズ識別法とベイズ距離併用法との比較を行なう。誤字・脱字数をみると、全ベイズ識別法では脱字数が、ベイズ・距離併用法では誤字数が多くなっている。これは、全ベイズ識別法では 76 文字から 1 文字を選択することにより、長音、‘の’、‘ん’などの似た形状、向き、動きをもつ文字の認識が困難なためである。これに対しベイズ・距離併用法では、手形状認識で大きく文字候補がしぼられるため、脱字する可能性は少くなる。よって、脱字より誤字が多くなる。

次に認識率について比較する。全ベイズ識別法の認識率は 95.02%、ベイズ・距離併用法では 94.37%と高い認識率が得られている。正しく認識された単語数をみると、全ベイズ識別法では 78 単語、ベイズ・距離併用法では 77 単語と 2 つの手法にあまり変化はない。また、動作を伴う指文字の認識率は全ベイズ識別法では 92.61%、ベイズ・距離併用法では 93.18%であり、どちらも動作を伴う指文字の認識に有効である。しかし、全ベイズ識別法の脱字の 11 文字中 8 文字、ベイズ・距離併用法の誤字の 10 文字中 5 文字が長音、‘の’、‘ん’である。これらの文字は、同一形状で動きの方向も似ているため識別が困難である。

計算量について比較する。計算量は全ベイズ識別法では 24 次元の行列計算を行なうため、非常に多くなる。これに対し、ベイズ・距離併用法では手形状のみの 18 次元に対しベイズ識別法を行なうため、計算量が削減できる。指文字や手話の学習支援システムを構築するには、計算量が少なく、高い認識率の認識手法が必要となる。よって、ベイズ・距離併用法が学習支援システムに有効であると考えられる。

表 4.7: 正しく認識される単語数 (全ベイズ識別法)

	単語数	正しく認識された単語数
データ 1	50	41
データ 2	50	37
合計	100	78

表 4.8: 正しく認識された単語数 (ベイズ・距離併用法)

	単語数	正しく認識された単語数
データ 1	50	37
データ 2	50	40
合計	100	77

4.5 まとめ

本章では、全ベイズ識別法の次元数を削減したベイズ・距離併用法を提案した。全ベイズ識別法では、24 次元に対しベイズ識別法の識別関数を計算するため、非常に計算量が多くなる。手形状・向き・動きの全ての特徴量を用いなければ識別できない文字は 42 文字である。残りの 34 文字は、手形状のみ (6 文字)、手形状+向き (12 文字)、手形状+動き (15 文字) で認識できるものである。よって、手形状・向き・動きの全てを考慮する全ベイズ識別法は無駄な計算が多くなるといえる。これに対しベイズ・距離併用法では最も重要な特徴量である手形状 (18 次元) のみに対して識別関数を計算し、向きは距離計算、動きは三次元位置の変化からの特徴の抽出により認識を行なう。向き認識、手形状認識の必要のないものに対しては認識は行なわない。よって、全ベイズ識別法と比較して計算量が大幅に軽減された。また、認識率も大きな低下はなく 94.37%であった。そのため、ベイズ・距離併用法は計算量、認識率を考慮しても有効な認識手法であるといえる。

第 5 章

指文字学習支援システム

5.1 はじめに

本章では、指文字認識を応用させた指文字学習支援システムについて述べる。5.2節では、従来の学習支援システムとその問題点について述べる。5.3節では、従来法の問題点を解決する CyberTouch を用いた指文字学習システムの概要、および学習システムの有効性を確認するために行なった評価実験の結果について述べる。

5.2 従来の学習支援システム

従来の学習手法はグループ学習型と自習型の 2 種類に大別される。その中で、自習型教材としてパーソナルコンピュータなどの情報機器の普及に伴い CD-ROM を利用した辞書や学習教材が多く市販されるようになってきた。これらは、実画像やアニメーションを用いたものである。アニメーションは、登録されている音声表記記号から自動生成されるため、データ量を削減でき、ユーザが任意のアニメーションで再生することが可能である。また、実画像を用いたものは、データ量は大きくなるものの、非手指動作などを学習する上で適している。しかし、これらはいずれも、顔の色と手の色が似ているため、手形状がわかりにくくなるという問題がある。このような問題を解決するためには、提示した手形状が正しいかどうかを識別するセンサが必要である。

5.3 Cyber Touch を用いた指文字学習支援システム

5.3.1 Cyber Touch

指文字学習支援システムを構築するため本研究では、Virtual Technologies 社製の Cyber Touch を使用する。CyberTouch はバイブレータ機能の有する Cyber Glove である。バイブレータは各指に 1 本と掌に 1 本、合計 6 本おのこの指の後ろと、掌に取り付けられている。バイブレータの振動する周波数は 0 ~ 255 の間で設定できる。CyberTouch の使用例として、仮想空間上で仮想物体に触った時にバイブレータが振動させることにより、触った感触を手に伝えたりすることが挙げられる。CyberTouch の概観を図 5.1 に示す。

5.3.2 指文字学習支援システムの概要

図 5.2 に本研究で作成した指文字学習支援システムの構成を示す。まず、‘あ’ ~ ‘ん’ の中から 1 文字を入力する。支援システムには、図 5.3 のような指文字が格納されている。そうすれば、画面には入力された指文字の図と学習者が提示している手形状がアニメーションで表示される (図 5.4)。手形状が間違っている時、間違っている指のバイブレータを振動させると同時に、画面では振動している指が赤くなって表示される。図 1 では、中指、薬指、小指が振動している。よって、学習者はどの指が原因で間違っているか視覚的および触覚的に知ることができる。学習者が指文字に近づけていくと、振動する指が少なくなる (図 5.5)。そして、正しく認識された時全ての指が振動なくなり、正しく学習されたことがわかる (図 5.6)。

入力された文字と学習者が提示した指文字が同一であるか確かめるため、指文字認識を行なう。認識の結果、入力された文字と同一ならばバイブレータは振動しない。違う文字と認識されたとき、バイブレータを振動させる。振動させる条件は、 $|x_i - \mu_i^{(l)}| > n\sigma_i^{(l)}$ ($i = 1, \dots, 18$) のとき振動させることとする。ここで、 x_i は提示している手形状の関節角データ、 $\mu_i^{(l)}$ 、 $\sigma_i^{(l)}$ は入力された文字 ($l = \text{あ}, \dots, \text{ん}$) の関節角データの平均値、標準偏差を示す。 n は最初は $n = 3$ に設定し、振動しなければデクリメントする。表 5.1 に、関節角の測定点と振動させる指の対応を示す。関節角の測定点は図 3.1 に対応している。



図 5.1: CyberTouch の概観

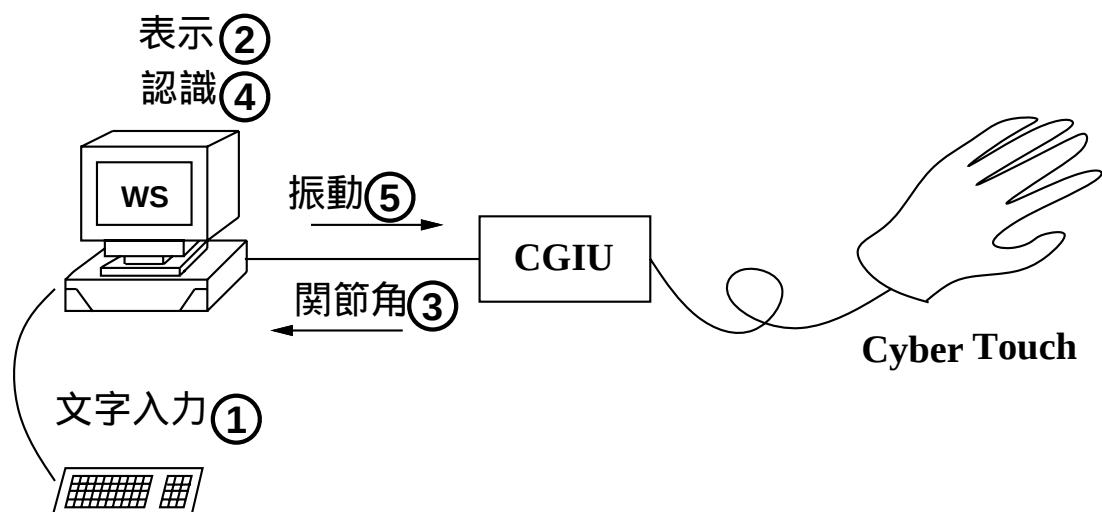


図 5.2: 指文字学習支援システムの概念図

表 5.1: 関節角の測定点と振動させる指の対応

測定点	振動させる指
0 1 2 3	親指
4 5	人差し指
6	人差し指、中指
7 8	中指
9	中指、薬指
10 11	薬指
12	薬指、小指
13 14	小指

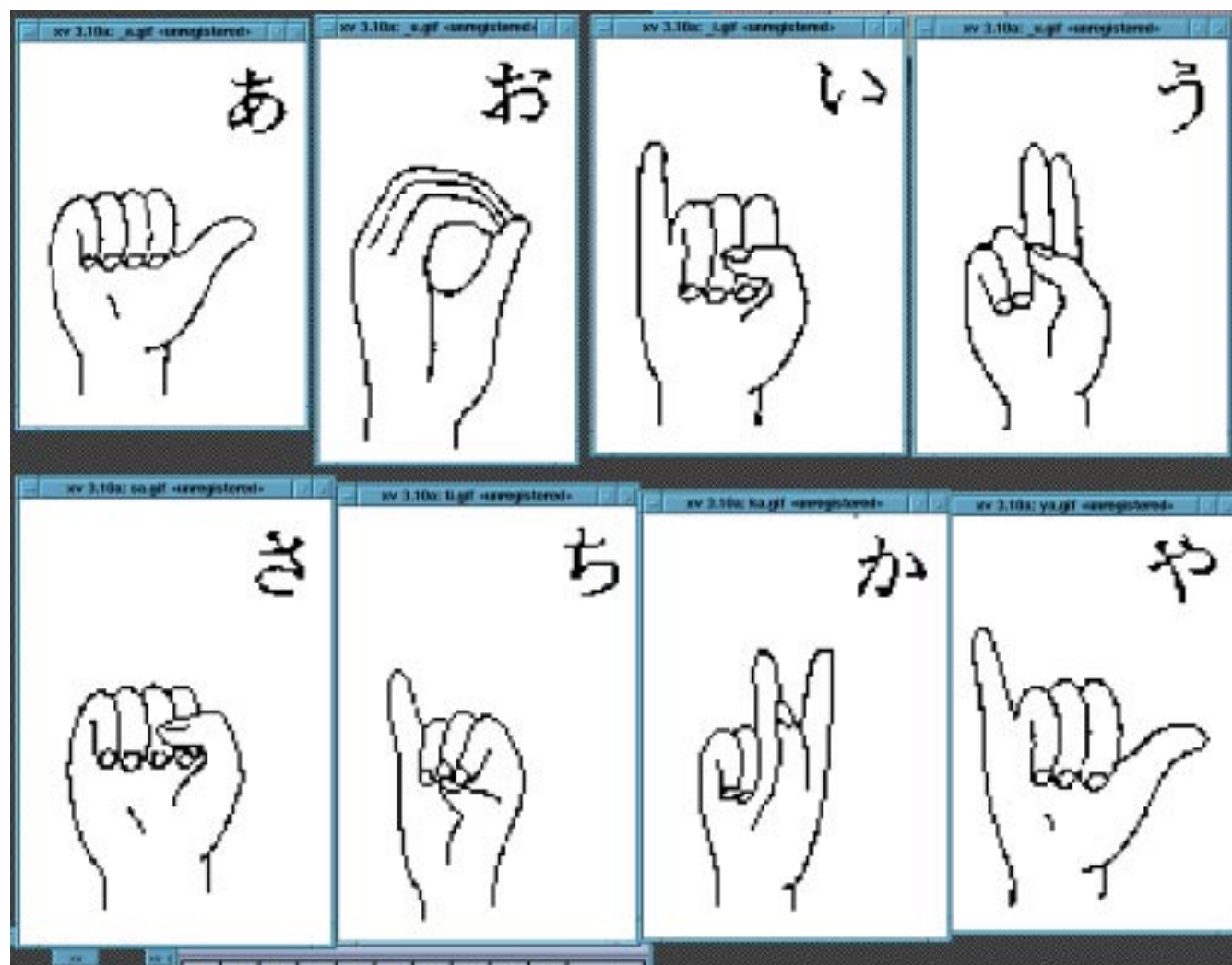


図 5.3: 出力される指文字の例

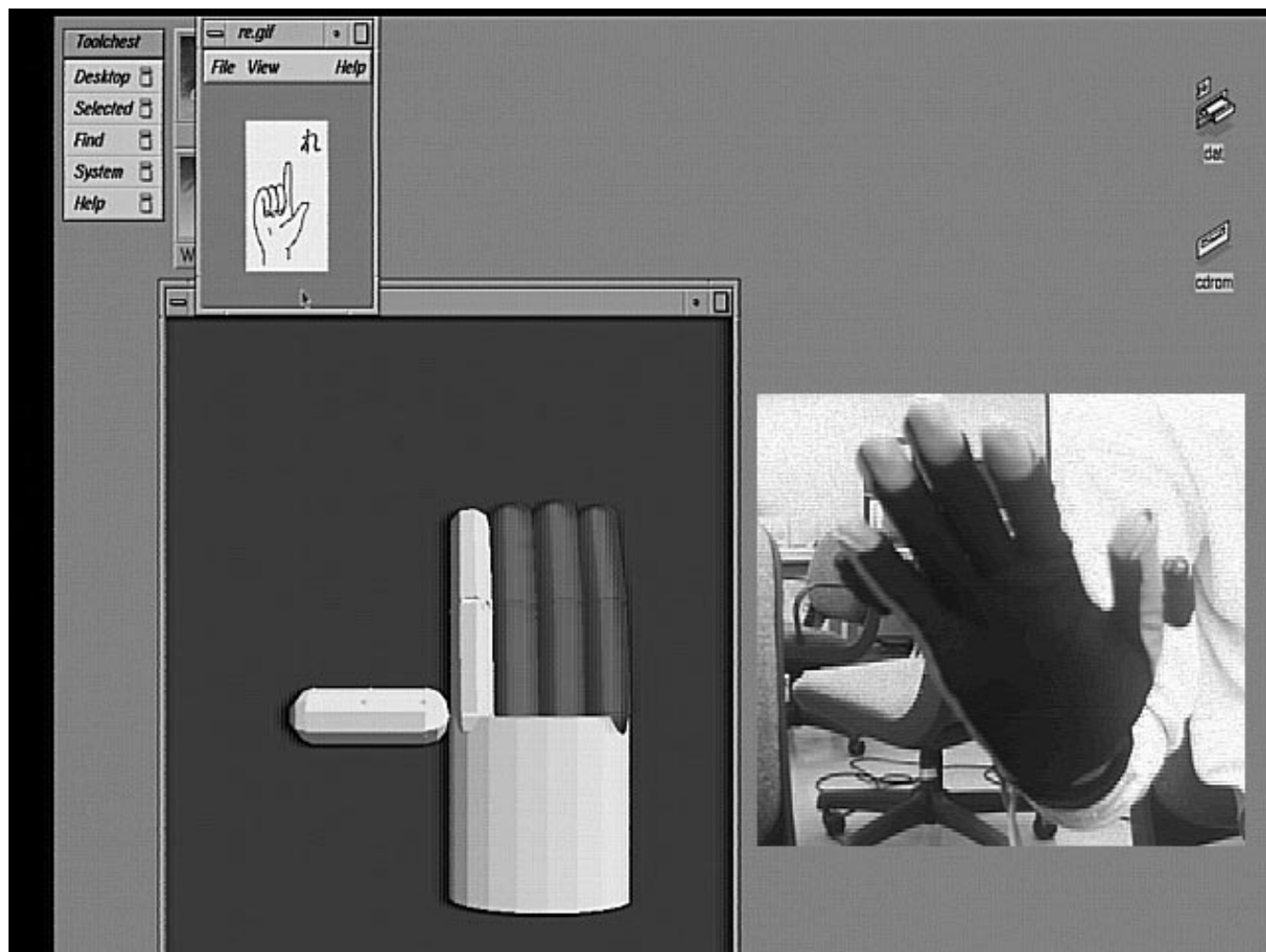


图 5.4: 出力例 (1)

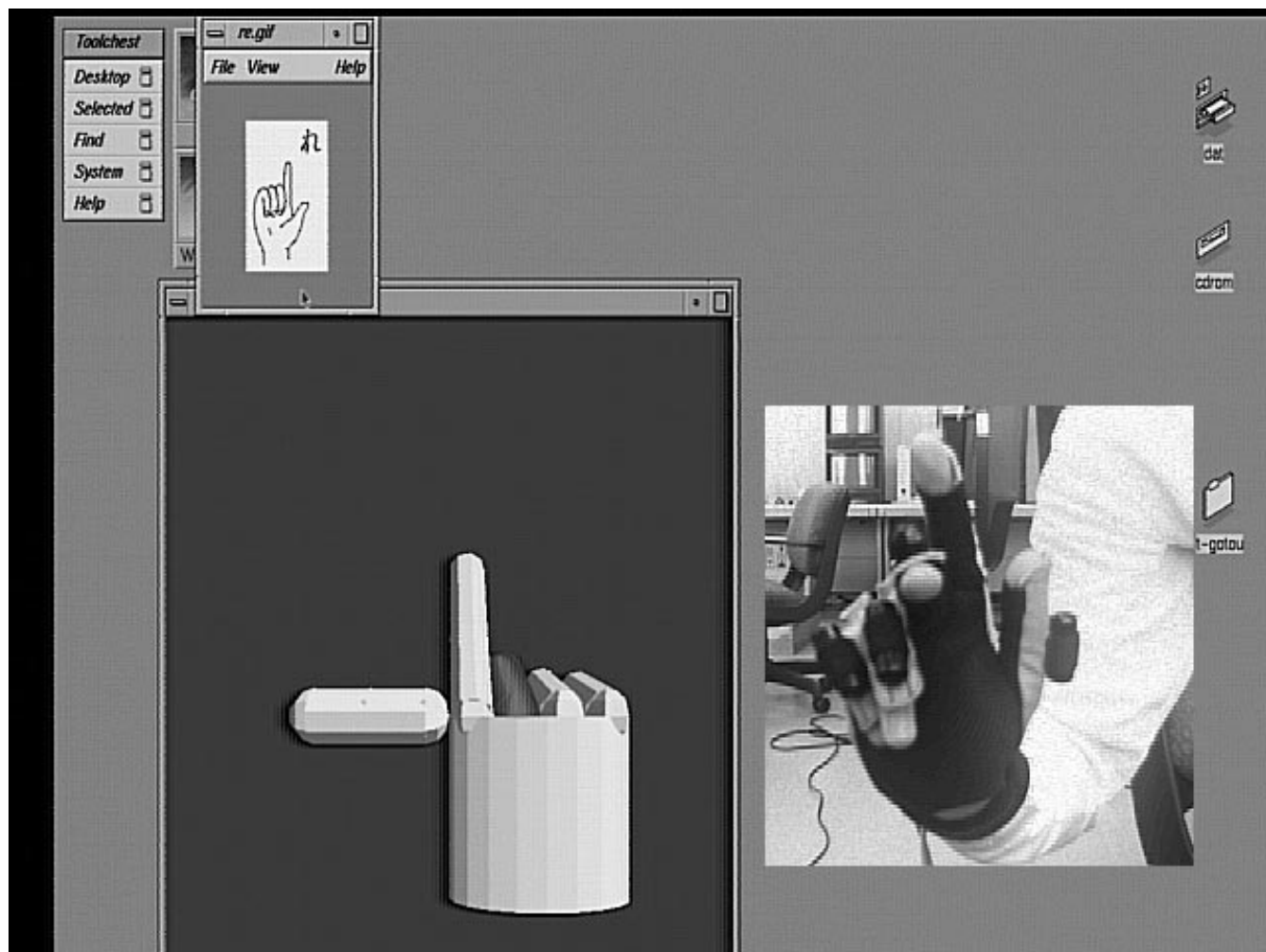


图 5.5: 出力例 (2)

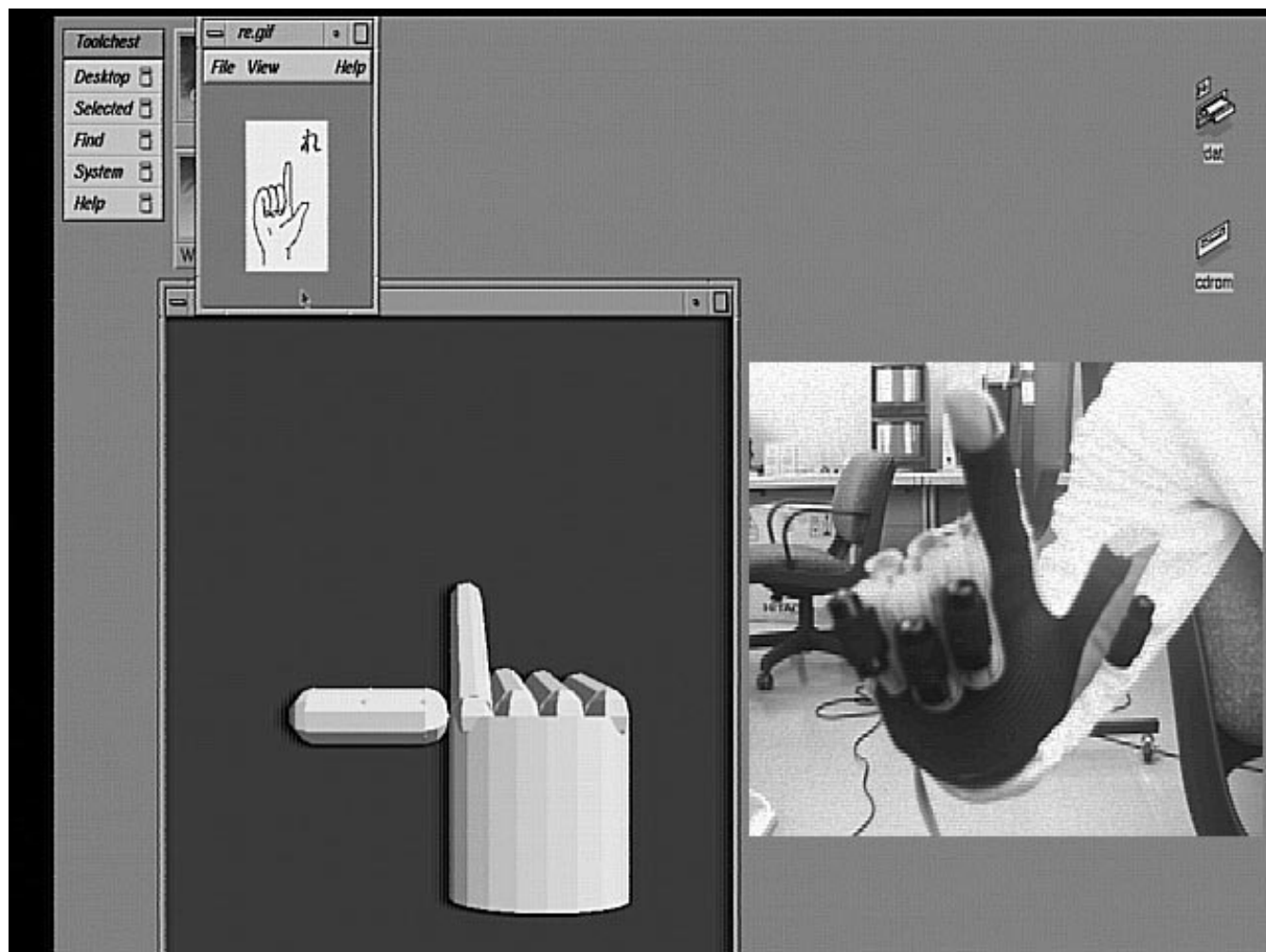


图 5.6: 出力例 (3)

5.3.3 評価実験

作成した指文字学習支援システムの有効性を確かめるための評価実験を行なう。評価実験の手順は以下の通りである。

- (1) 指文字の図を見て指文字を提示してもらい、その時の関節角データを採取する。
- (2) 1週間以上の期間をおき、指文字学習支援システムを使用したあと、指文字を提示してもらい関節角データを採取する。
- (3) (1) と (2) の関節角データをもとに指文字認識を行ない、正しく提示できたかどうかを比較する。

実験では10人分の指文字データを使用する。この10人は日常生活において手話・指文字の使用経験はない健聴者であり、標準辞書を作成した際の被験者とは違う人である。

指文字データは手形状を重視し、指文字を示す時に使う24の手形状データのみを採取する。しかし、‘ら’の手形状はCyberTouchのバイブレータが原因で正しく提示できないため除外する。

(1) での指文字データの認識結果を表5.2、(2) での指文字データの認識結果を表5.3に示す。(1) の結果より、指文字の図のみの情報では‘ち’や‘つ’の図のみではわかりにくい指文字があり、正しく提示されていないことがわかる。また、‘る’、‘ひ’、‘あ’なども正しく提示されていない。一方、(2) では間違いがバイブレータにより指に伝わるため、ほとんどの指文字が正しく学習され、正しく提示されていることがわかる。また、‘り’、‘か’、‘さ’などは学習前とあまり変化はない。しかし、提示している指文字は見た目にはあきらかにそれぞれの指文字を示している。よって、これらが正しく認識されていない原因は、標準辞書のデータと今回実験に協力してもらった人の手の特性が大きく違っていたためであると思われる。

5.4 まとめ

本章では、従来の学習支援システムの問題点を解決した指文字学習支援システムの概要について述べた。また、指文字学習支援システムの評価実験を行なった。その結果、このシステムを使用することによりほとんどの指文字が正しく提示でき、指文字学習支援シス

表 5.2: 学習前の認識結果

間違った個数	間違った指文字 (認識された文字)
7 個	り (る、ぬ、う)、ち (い)
6 個	る (う、か)
5 個	つ (い、や)
4 個	さ (え)、ひ (ぬ)、あ (え、こ、さ)
3 個	お (え、ぬ、や)、き (ひ、ぬ、ち)、や (え、い)、れ (ひ)
2 個	か (り、ろ)、ぬ (ひ、ろ)、う (か)
1 個	せ (い)、け (こ)、ろ (ぬ)、こ (え)、て (こ)

表 5.3: 学習後の認識結果

間違った個数	間違った指文字 (認識された文字)
5 個	り (う、ろ)
3 個	さ (え、こ)
2 個	あ (こ)、か (こ、ろ)
1 個	け (こ)、せ (ろ)、つ (や)

テムの有効性を確かめられた。しかし一部の指文字では、見た目にはあきらかにその指文字を提示しているのに、認識結果は違った文字を示しているものもある。このような見た目と認識システムの違いを改善することが今後の課題である。

第 6 章

結論

6.1 まとめ

最近、多くのメディアで手話が取り上げられ、人々の手話への関心が高まってきており、さまざまな形態の手話教材が市販されている。電子化された手話辞書や手話学習ソフトウェアは、CD-ROM の特性を利用した検索機能、文字や画像を同時に扱えるマルチメディア機能を生かしたシステムである。しかし、これらのシステムは動きはわかるものの、重要な情報である手形状がわかりにくく、正しい手形状が学習できない可能性がある。よって、間違った手形状を提示しているときは、間違いを指摘するシステムが必要である。

そこで、本研究では手形状を正しく認識するために、手話の基本である指文字の学習支援システムを構築した。このシステムは、間違った手形状のときはバイブレータにより指を振動させ、間違いを直接知らせることが可能である。また、このシステムの有効性を確かめるため評価実験を行なった。その結果、指文字の図だけでは正しく学習できなかった指文字が、システムの使用により正しく学習できることが確かめられた。

提示した指文字が正しいことを確認するためには、指文字認識を行なう必要がある。そこで、動作を伴う指文字を含めた連続な指文字動作の認識手法として、全ベイズ識別法を提案し、認識実験を行なった。その結果、95.02%の認識率が得られた。しかし、手話の学習支援システムへの応用を考えると計算量が少ないことが望ましい。そこで、全ベイズ識別法の次元数を減らしたベイズ・距離併用法を提案し、認識実験を行なった。その結果、認識率は 94.37%であり、大きな低下もなく、計算量を削減できた。

指文字は複雑な手形状を有する。よって、指文字が正しく認識できることは複雑な手形状が認識可能であり、手話認識に応用できると思われる。このシステムは、手話学習支援システムの基礎的なものであり、これを応用させることにより、多くの手話の知らない人が正しい手話を学習できると思われる。

6.2 今後の課題

指文字の長音、‘の’、‘ん’は同一の手形状で、動きの方向も似ている。このような、似た手形状、向き、動きを識別する認識手法を検討することが必要である。また、指文字学習支援システムを応用させた手話学習支援システムを構築することが今後の課題である。

謝辞

本研究を進めるにあたり、終始熱心な御指導、御鞭撻をいただいた北陸先端科学技術大学院大学 堀口 進 教授、阿部 亨 助教授に深く感謝いたします。

サブテーマで熱心に御指導をいただいた 石原 哉 助教授に感謝いたします。

日頃より有意義な御教示、検討、議論をしてくださった沼田一成助手に感謝いたします。

指文字データの収集に御協力頂いた奥村研究室の徳田晶晃さんに感謝いたします。

指文字学習支援システムの作成の際、御協力頂いた堀口研究室の堀井洋君、鈴木信勝君に感謝いたします。

さらに、日頃よりお世話になった堀口・阿部研究室の皆様に厚く御礼申し上げます。

参考文献

- [1] 佐川、酒匂、大平、崎山、阿部：“圧縮連続 DP 照合を用いた手話認識方式”，電子情報通信学会論文誌, Vol.J77-D-II, No.4, pp.753-763, Apr.1994.
- [2] E.Ohira, H.Sagawa, T.Sakiyama, M.Ohki：“A Segmentation Method for Sign Language Recognition”, IEICE TRANS.INF. & SYST., VOL E78-D, NO.1 Jan.1995.
- [3] 寺内 美奈, 長嶋 祐二：“手話学習のためのマルチメディア型教材開発”，日本手話学会 第 22 回全国大会予稿集, pp.19-22, Sep.1996.
- [4] 福本、間瀬、末永：“画像処理を用いた指示動作検出の実験システム” 電子情報通信学会春期全国大会論文集、 1-251, 1991.
- [5] 高橋 友一, 岸野 文郎：“手振り認識方法とその応用”，電子情報通信学会論文誌, Vol.J73-D-II, No.12, pp1985-1992, Dec.1990.
- [6] 内田 雅文, 石川 潔, 井出 英人：“手形状認識と手話への応用”，電気学会論文誌, Vol.114-C, No.10, pp.995-999, Oct.1994.
- [7] 萩原 芳彦, 堀口 進：“指手骨格モデルを用いた指文字認識”，日本手話学会 第 21 回全国大会予稿集, pp.32-35, Sep.1995.
- [8] “ノンバーバルインターフェース”，電子情報通信学会編, オーム社, 1994.
- [9] 竹村, 伴野, 岸野, 小林：“大型 3 次元表示装置を用いた仮想操作環境の試作”，第 20 回画像工学コンファレンス論文集, pp281-284, 1989.
- [10] “手話法辞典”，栃木県聾学校, 栃木県聴覚障害者協会, 1978.
- [11] 入江 督 編：“手話辞典”，大阪ダイヤモンド工業株式会社, 1993.

- [12] 伊藤, 竹村 : “手話入門”, 廣濟堂, 1982.
- [13] 中野, 伊藤, 松本 : “手話への招待”, 福村出版, 1977.
- [14] “*CyberGlove*TM User's manual”, Virtual Technologies, 1993.
- [15] “3SPACE USER'S MANUAL”, POLHEMUS, 1993.
- [16] 崎山、大平、佐川、大木、池田 : “リアルタイム手話アニメーションの合成方法”, 電子情報通信学会論文誌, Vol.J79-D-II, No.2, pp.182-190, Feb.1996.

研究業績

- [1] 動作を伴う指文字を含む連続指文字認識
後藤岳志 , 阿部亨、堀口進
電気関係学会北陸支部連合大会 F-54,pp369.1996.