

指文字入力インタフェース「Fingual」の開発

福島大志[†] 宮崎文夫[†] 西川 敦^{††}

我々は「Fingual」という非接触文字入力インタフェースを新しく開発した。Fingualは磁石と磁気センサを取り付けたグローブを使ったインタフェースであり、ユーザはグローブを手に装着し手話の一種である指文字を形作ることで文字入力を行える。Fingualの特徴は、指文字一つ一つの形状に対応する磁場（データセット）をユーザに合わせて予め取得しておくことで、逆問題などの複雑な計算を必要とせずに文字認識を行えるところにある。本稿ではFingualを用いた文字入力実験を行うことで、Fingualの有用性について述べる。

The development of a noncontact letter input interface “Fingual”

TAISHI FUKUSHIMA,[†] FUMIO MIYAZAKI[†] and ATSUSHI NISHIKAWA^{††}

We have newly developed a noncontact letter input interface called “Fingual”. Fingual uses a glove mounted with inexpensive and small magnetic sensors. Using the glove, users can input letters to form the finger alphabets, a kind of sign language. Though Fingual needs to collect magnetic data for each letter, its feature is the ability to recognize letters without the complicated calculations such as inverse problems. This paper shows results of the recognition experiments, and shows the utility of the proposed system “Fingual”.

1. はじめに

今日、PCや携帯電話の普及には目覚ましいものがあり、我々の生活には不可欠なものとなっている。これらの機器では文字入力インタフェースとしてたいていはキーボードが利用されている。しかし、誰もが初めからキーボードを上手く利用できたわけではなく、練習しながら徐々に体得していくのが普通である。特に高齢者はキーボードの操作を苦手とする人が多く、パソコンによる文字入力を敬遠しがちである。また、携帯電話の場合、入力キーが少ないがために携帯電話による文字入力を面倒に思う人も多い。そこで我々は小さな子供や高齢者であっても直感的に文字入力を行えるような指文字インタフェース「Fingual」を開発した（Fingualは「finger」と「lingual」を組み合わせた造語である）。指文字は手話の一種であり、一つ一つの文字を表現したい時に用いられる。

我々は日常の会話において、言葉と同時に手の動きや形状などを用いたジェスチャを交えることが多い。ジェスチャは言葉では表せないことを表現したりコミュニケーションをサポートしてくれる。また、手話のようにジェスチャのみでコミュニケーションを行うことも可能であり、非常に多くの意思を表現することができる。ジェスチャは人が自然に行っている動作であるため、ジェスチャを利用することで自然な入力インタフェースを構築できるのではないかと考えられる。

ジェスチャを認識する研究は数多く行われているが大別すると「カメラを利用するタイプ¹⁾」と「センサなどの機器を装着するタイプ²⁾」、「カメラとセンサの両方を用いるタイプ³⁾」の3つに分けられる。「カメラを利用するタイプ」の場合、ユーザは機器を装着する必要はないが、精度のよいカメラを使うとなると高価になってしまう。「センサなどの機器を装着するタイプ」の場合、ユーザが機器を装着するという煩わしさがあるものの比較的安価に構築できる。「カメラとセンサの両方を用いるタイプ」は前記2つの手法より自由度が大きいというメリットがあるが、両者の短所も合わせ持ってしまう。それぞれのタイプで一長一短があるものの、本研究では、コスト面や携帯電話との親和性を優先し、磁石と磁気センサを取り付けたグローブを用いる方法を採用した。

[†] 大阪大学大学院基礎工学研究科機能創成専攻
Department of Mechanical Science and Bioengineering,
Graduate School of Engineering Science, Osaka University

^{††} 信州大学繊維学部応用生物学系バイオエンジニアリング課程
Bioengineering Course, Division of Applied Biology,
Faculty of Textile Science and Technology, Shinshu University

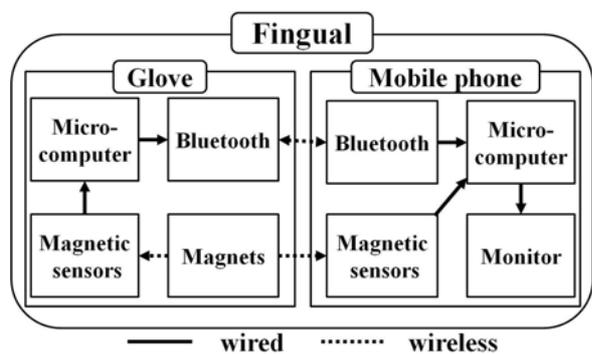


図 1 Fingual の構成
Fig.1 Configuration of Fingual

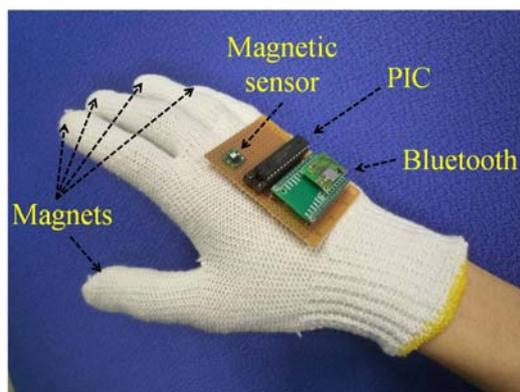


図 2 Fingual グローブ
Fig.2 Fingual glove

2. 方 法

2.1 システム構成

Fingual に必要な機器を以下に示す.

- (1) 3 軸磁気センサ
- (2) ネオジム磁石 (角型 15 × 3 × 3[mm], 表面磁束密度 320[mT])
- (3) Bluetooth
- (4) グローブ
- (5) シングルチップマイクロコンピュータ (PIC16F876A)
- (6) 携帯電話 (本稿では PC (CPU Intel Core2Duo 2.80[GHz], メモリ 4[GB]) を用いた)
- (7) 電池 (5V)

Fingual は図 1 に示すように磁石の磁場を磁気センサで検知し, Bluetooth を使い携帯電話や PC と交信することでモニタに文字を表示させるシステムである. 近年では多くの携帯電話や PC に Bluetooth が内蔵されており, 携帯電話にはさらに磁気センサが内蔵されているものも多い. Fingual は携帯電話のこれらの機能をそのまま利用する. 図 2 は指文字入力に使うグローブである. 各指先に磁石が一つずつ合計 5 個, 磁気センサが手の甲と手首の一つずつ合計 2 個付いている. このセンサが各指文字を表現した時の磁場を計測する. 図 1 の携帯電話側の磁気センサは認識された指文字を決定 (入力) する為に用いられ, Fingual はグローブを携帯電話に近づけることで文字を決定 (入力) することができる. なお, 本稿では Bluetooth による無線通信ではなく有線にて行った実験結果を述べる.

2.2 対象とする指文字

本研究で対象とする指文字は小文字のアルファベット 24 文字である (図 3). アルファベットの「j」と「z」は表現に動きを伴うものとなっており, Fingual を応用すれば表現可能であるが本稿では対象外とした.

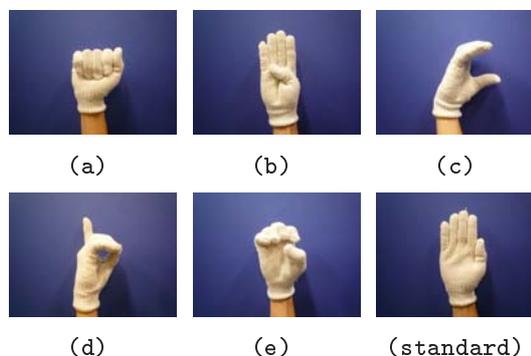


図 3 指文字の例
Fig.3 Examples of finger alphabets

図 3 の「standard」を指文字の基準姿勢と定義した. 基準姿勢は環境磁場の影響を排除する為に利用し詳細は次節で説明する.

2.3 文字認識手法 A

初めに, ユーザはグローブを装着し各指文字を表した時のそれぞれの磁場を登録する ($B_1 \sim B_n$). この事前登録した磁場と指文字との組み合わせをデータセット (表 1) と定義する. 今, ユーザが任意の指文字を作った時の磁場を A とする. この指文字を判別する為に A とデータセットの磁場とを比較する. A とデータセットの磁場 $B_1 \sim B_n$ をそれぞれ式 (1) に代入して, それぞれのデータナンバーで誤差の 2 乗和 $f_1 \sim f_n$ を求める.

$$f_i = \|A - B_i\|^2 \quad (1)$$

$f_1 \sim f_n$ の中で最も値の小さかったものと同じデータナンバーの文字を認識した指文字とする. なお, 3 軸磁気センサを 2 つ用いるので A, B は 6 次元ベクトルである.

本研究は磁場を使ったインタフェースであるため, 地磁気などの環境磁場の影響を受けやすい. そこで, 本研究では図 3 の (standard) のように表される手の

表 1 データセット

Table 1 Dataset

Data number	Magnetic field	Letter
1	B_1	文字 1
2	B_2	文字 2
⋮	⋮	⋮
n	B_n	文字 n

形状を基準姿勢とし、各指文字の磁場データはこの基準姿勢の磁場との差分を登録する。登録後に実際に入力を行う際も、最初に基準姿勢をとり磁場を記録することで表現したい指文字の磁場との差分を計測することができる。こうすることで、データセットを登録する際とそのデータセットを使って入力を行う際に環境が変化しても、同じ条件での磁場の比較を行うことが可能となる。

2.4 文字認識手法 B

文字認識手法 A は誤差の 2 乗和 (f 値) がもっとも小さいデータナンバーの文字を入力された指文字とするものである。しかし、適した指文字が必ずしも f 値が最小のデータナンバーのものであるという確証はなく、 f 値が 2 番目や 3 番目に小さいデータナンバーなどに対応する文字である可能性もある。例えば表 2 のようになる時を考える。これは複数のデータセットを用いた場合に、 f 値の最も小さいデータナンバーから順位づけを行った時の一例である。なお、文字の添字はどのデータセットから取得した文字かを示している。この時、文字認識手法 A では f 値が最小である「a」が認識された。しかしこの場合、2 位 3 位に「b」という文字が認識されており、正しい答えが「b」である可能性も高い。そこで、図 4 のようなガウス関数を用いた f 値の順位を考慮した重みづけをし、最も信頼性の高い文字が認識されるようにする。

一般にガウス関数は式 (2) のように表される。

$$y = a \exp\left(-\frac{(x-b)^2}{2c^2}\right) \quad (2)$$

本研究では $a = 1, b = 0, c = 0.075$ とし、式 (3) のように分子に式 (1) の f 値を代入し、 y をその順位のポイント p とする。

$$p = \exp\left(-\frac{f}{2c^2}\right) \quad (3)$$

本稿では f 値の小さいものから 3 位までのデータナンバーを考慮し、各順位でポイント p を計算する。もし、3 位以内に同じ文字があればポイントの和を計算する。そして、最も和の大きい文字を認識された文字とする。つまり、表 2 の場合、2 位 3 位である「b」のポイントの和が 1 位である「a」のポイントを上回れ

表 2 認識されたデータセットの例

Table 2 Example of the recognized dataset

f 値順位	1 位	2 位	3 位
対応する文字	a_3	b_1	b_2

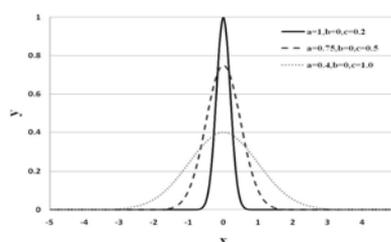


図 4 ガウス関数

Fig. 4 Gaussian function

ば文字「b」が認識されることになる。

3. 実験

3.1 被験者自身のデータセットを用いた実験

3.1.1 実験内容

被験者 A, B, C の 3 名に対してアルファベット 24 文字 (入力に動きを伴う「j」と「z」を除く) の認識率実験を行った。実験の初めに各被験者自身のデータセットを 3 セット合計 72 個 (24 文字 \times 3) 登録した。被験者により入力された文字がリアルタイムに表示されるモニタを用意し、被験者はモニタを見ながら手形状を調整することができる。モニタには被験者が入力すべき文字 (ターゲット) も同時に表示される。ターゲットと同じ文字を被験者が入力すると、ターゲットとなる文字が次の文字に変更される。今回、ターゲットはアルファベット順に変わるものとする。被験者にはアルファベット 24 文字 \times 10 セット (240 文字) の入力を行ってもらい、認識率とターゲット 1 文字当たりの入力にかかる時間を計測した。また、ターゲットが表示されて 10 秒経過しても認識されない場合を認識エラーとした。なお、本項では認識手法 A を用いた。

3.1.2 結果と考察

表 3 に本実験の結果を示す。全ての被験者において、ほぼ 100[%] の認識率を示していることがわかる。また、指文字一つ当たりの入力時間に関しても 1[s] 前後しかかかっておらずスムーズな入力が可能であることがわかる。被験者 3 名で入力時間に若干の差があるが、提案システムはいかに自分の登録した指の形状を記憶しているかが肝要であり、最も結果の良かった被験者 C は指文字形状の再現が上手かったとも言える。また、本実験の被験者は指文字を日常で使用していない人物であるので指文字の形状を思い出す時間の差に

表 3 被験者自身のデータセットを用いた時の認識率と入力時間
Table 3 Recognition rate and input time
when subjects used their own dataset

	認識率 [%]	1 文字当たりの入力時間 [s]
被験者 A	98.3	1.27
被験者 B	98.8	1.10
被験者 C	100.0	0.78

よる影響も考えられる。しかし、どの被験者においても良い結果が得られており本手法の入力インタフェースとしての有用性を示すことができたと思われる。

3.2 他人のデータセットを用いた実験

3.2.1 実験内容

前項の実験では参照するデータセットとして被験者自身の登録したものを使っていた。本項では他人のデータセットを利用した場合にどのような認識率を示すかを検証した。被験者は前節と同じ A, B, C であり、被験者とは別の協力者 D のデータセット (3 セット) を用いて前節と同じ内容の実験を行った。

3.2.2 結果と考察

表 4 に本実験の認識率を示す。他人のデータセットを用いた場合は被験者自身のデータセットを用いた場合と比べ認識率は落ちている。これは、被験者によって手の大きさや指文字を作る際の指の微妙な曲げ具合などで作り出す磁場に違いが生まれてしまうこと原因だと思われる。また、認識手法 B を使って f 値による順位を考慮した方が認識率は良くなっており認識手法 B の方が有効であることがわかる。表 5 は本実験の 1 文字当たりの入力時間である。いずれの手法でも被験者本人のデータセットを使った時よりも時間かかってしまっている。これは、被験者が指形状を微調整しながら協力者 D の登録した指形状による磁場を見つけて出すことに時間がかかるのが原因だと考えられる。また、入力時間においても認識手法 A よりも認識手法 B の方が有効であることがわかる。表 6 は 24 文字のうち 1 度でも誤認識があった文字が何種類あったかを示している。これを見ると、認識手法 B の方が誤認識される種類が減っておりこれが入力時間に影響しているものと考えられる。認識手法 B を用いることで、指文字と指文字の区別がより明確になるということも言える。しかし、認識手法 B を用いてもいくつかの特定の文字だけは他人のデータセットを用いた場合認識できないことがわかった。

4. ま と め

本研究は新しく開発した指文字インタフェース「Fingual」の有用性を示すものである。Fingual は事

表 4 他人のデータセットを用いた時の認識率
Table 4 Recognition rate
when subjects used another person's dataset

	認識手法 A [%]	認識手法 B [%]
被験者 A	77.9	88.8
被験者 B	78.3	88.8
被験者 C	91.3	96.7

表 5 他人のデータセットを用いた時の
1 文字当たりの入力時間

Table 5 Input time per 1 letter
when subjects used another person's dataset

	認識手法 A [s]	認識手法 B [s]
被験者 A	2.71	1.53
被験者 B	2.16	1.77
被験者 C	1.42	1.19

表 6 他人のデータセットを用いた時の
24 文字のうち 1 度でも誤認識があった文字の種類数

Table 6 The number of letters out of 24 which are falsely
recognized at least once when subjects used
another person's dataset

	認識手法 A [種類]	認識手法 B [種類]
被験者 A	13	9
被験者 B	12	4
被験者 C	6	4

前に各指文字ごとの磁場情報を登録することで使用できる。一度登録すれば磁石の位置がずれたりすることが無ければ再度登録する必要はない。また、指文字形状をユーザ自身で変更してオリジナルの指文字を作ることでもできる。本稿では 2 つの手法を用いて、被験者本人のデータセットを用いた場合と他人のデータセットを用いた場合の認識率の違いを示し、Fingual の有用性を示した。Fingual は現状ではデータセットの磁場と対になるものとして文字を利用しているが、声や音を登録することも可能である。今後は Fingual の無線化に取り組み、さらなる応用を考えていきたい。

参 考 文 献

- 1) 上田悦子, 松本吉央, 今井正和, 小笠原司. 多視点シルエット画像を用いた手の形状推定. 情報処理学会研究報告, CVIM, コンピュータビジョンとイメージメディア 2001(66), pp.25-31, 2001.
- 2) 塚田浩二, 安村通晃. Ubi-Finger: モバイル指向ジェスチャ入力デバイスの研究. 情報処理学会論文誌 43(12), pp.3675-3684, 2002.
- 3) 渡辺賢, 岩井儀雄, 八木康史, 谷内田正彦. カラーグローブを用いた指文字の認識. 電子情報通信学会論文誌, D-II, Vol.J80-D-II, No.10, pp.2713-2722, 1997.