

情報理論に基づいた高信頼かつ高効率なハンドサインの設計 Reliable and Efficient Hand Sign Design Based on Information Theory

高橋 時生[†] 内田 真人[†]
Tokio Takahashi Masato Uchida

1. はじめに

身体障害を持つ人々にとって、自分の能力を活かすことができる多様性のある社会が望まれている。そのような中で、社会的障壁を取り除き、障害を持つ人に対する合理的な配慮を行うことが期待されている[1]。身体障害は1級から7級まで程度が分かれており、級別の中でも障害の度合いが異なる[2]。このような身体障害の多様性に比べて、現状では障害を支援する方法の多様性は少なく、障害の度合いに応じた合理的な配慮を行えていない。例えば、言語障害を持つ人に対する配慮の手段として手話が使われているが、言語障害に加えて肢体が不自由な場合、一般に使われている手話の動作ができないことがある。そのような人に対しては一般に使われている手話は配慮の手段として適切ではない。つまり、身体の障害により通常の手話言語や音声言語によるコミュニケーションが困難な人の生活の質を向上するためには、各々の障害の特性に応じた円滑な意思疎通の手段が必要である。

このような場合の意思疎通の手段として、各自にとって可能な動作によって個別に設計された手話を使用することが考えられる。しかしその場合、各自が自由に手話を決めることになるため、一般に使われている手話とは異なるものとなり、相手にその意味が通じなくなってしまう。また、相手が一人一人異なる手話の意味を理解しようとするためには覚える手間がかかる。この問題を解決するためには、人間に代わってコンピュータシステムが手話を解釈することができればよい。すなわち、各自が自由に決めた手話の意味をコンピュータシステムが解釈できれば、相手に意思表示を行うことができる。よって、コンピュータシステムを利用したコミュニケーションの支援が合理的配慮として有効であり、このようなシステムの実現が重要である。

本論文では、コンピュータシステムを利用した障害者のコミュニケーション支援として、障害の度合いに応じたテーラーメイド式のハンドサイン識別システムの実現を目的とする。そして、各自が動作可能なハンドサインの中でどの組み合わせがコミュニケーションを行う上で高信頼かつ高効率であるのかを求める。ここでは、ハンドサインを事前に指定されたある文字を表す手の動きや形として定義する。障害によって一般に使われている手話が使えない人がいるため、各自が実現できる動作のみを用いて個別に設計された手話、すなわちハンドサインを個人ごとに決める必要がある。障害の度合いに応じたテーラーメイド式にすることで、既存のハンドサインに囚われることなく、自身の使えるハンドサインのある意思を伝えるものとして自由に紐づけることができる。つまり、同じ意思を伝えたいものでも人によってハンドサインが異なる場合がある。既存の手話とは表す意味が異なっても、個別に設計されたハンド

サインをコンピュータシステムが解釈することができれば、相手に意思を伝えることができる。

このとき、各自が動作可能なハンドサインの中で最もコミュニケーションが行いやすいものを個人ごとに決める必要がある。ここで、最もコミュニケーションが行いやすいものとは、コンピュータシステムで識別する上で、できるだけ誤りが少なく、相手に早く伝えることができるハンドサインのことである。つまり、高信頼かつ高効率なハンドサインが求められる。これを決めるためには、コンピュータシステムにおけるハンドサインの識別率を求める必要がある。そして、その識別率に基づいて各自の識別しやすいハンドサイン・識別しにくいハンドサインを考慮しながら最適なハンドサインの組み合わせを決めていかなければならない。ハンドサインの識別率が高いほどシステムの実現に適していると言えるが、本論文ではその識別率の向上を図るのではなく、得られたハンドサインの識別率に応じて高信頼かつ高効率である最適なハンドサインを導出する。

本論文の構成は以下の通りである。まず、次節において本論文に関連する研究について紹介する。次に第3節において本論文の提案手法について説明する。そして第4節では提案手法に対する評価実験とその結果について議論する。第5節で本論文のまとめを述べる。

2. 関連研究

ハンドサインの識別を目的とする研究として、カメラを使用したものが知られている[3]。この研究では、ニューラルネットワークを使用して静止画像からハンドサインを識別している。この手法では、1つのハンドサインに対して動き出す前・動いた後・その中間位置の3つの静止画像をあらかじめ切り出して識別する。そのため、実際のシステムではカメラから得られる映像から静止画をリアルタイムで切り出さなければならず、手の動きを追跡することが難しい。

ウェアラブル端末を使用してハンドサインの識別を試みる研究も行われている。ウェアラブル端末を使うことには、カメラのように間接的にデータを集めるのではなく、人間の動作データを直接集めることができるという利点がある。スマートグローブとよばれる手袋型のウェアラブル端末を使用した研究[4]では、グローブに指や手の甲の曲がり角度を測るセンサーや指の開閉を検出するホールセンサー、加速度センサーを取り付け、ハンドサインの識別をしている。筋電位から指文字を識別する研究[5]では、前腕に筋電位センサーを18チャンネル分取り付け、記録した筋電位を用いて手話の一部として利用されている指文字を識別している。識別にはニューラルネットワークを使用している。この手法では、筋電位を測定するためには筋肉の位置に合わせて前腕にテープ状の電極を貼り付けなければならず、簡単に測定することができない。

筋電位を用いてハンドサイン識別を行う研究としては、

[†] 早稲田大学基幹理工学研究科情報理工・情報通信専攻
Department of Computer Science and Communications
Engineering, Waseda University, Tokyo, Japan.

一般に市販されているウェアラブル端末 Myo を使用したものの[6, 7]がある。ウェアラブル端末 Myo については、第 3 節で詳しく述べる。Myo を使用したブラジル手話の識別に関する研究[6]では、Myo で測定した筋電位データを SVM を用いて識別している。Myo を使用したアメリカ手話の識別に関する研究[7]では、DTW (Dynamic Time Warping) を用いて手話の識別をしている。様々な識別アルゴリズムの識別率を比較した結果、DTW を用いた k 近傍法が最適であると結論づけられている。

このように、ハンドサインの識別に関する研究は数多く行われている。しかし、いずれの研究もアメリカ手話や指文字といった既存のハンドサインの識別率を導出し、その向上を目的としたものである。しかし、何らかの障害により既存のハンドサインを使うことができない人には、これらの研究成果を適用することができない。よって、身体の障害を持った人が動作可能なハンドサインを用いる必要がある。そして、その動作可能なハンドサインの中でコミュニケーションを行いやすいものを調べる必要がある。できるだけ誤りが少なく、相手に早く伝えることができるハンドサインであれば、コミュニケーションを行いやすい。そこで本論文では、高信頼かつ高効率なハンドサインの設計について検討する。

3. 提案手法

3.1 概要

本論文では、高信頼かつ高効率なハンドサインの最適な組み合わせを求めるために、情報理論の考え方を導入する。ハンドサインを識別する手段として、筋電位に着目する。本論文では、ウェアラブル端末を使用してハンドサイン動作中の筋電位データを測定する。 k 近傍法を採用して、得られた筋電位データを識別し、どのハンドサインが入力されたのかを識別することでハンドサインの識別率を求める。ハンドサインの識別率とは、「入力されたハンドサインがどのハンドサインとして識別されるか」を表す(条件付き)確率である。このような筋電位の測定結果からハンドサインを識別するコンピュータシステムを通信路とみなし、その入出力をハンドサインのラベルとする。そして、ハンドサインの識別率から定義される通信路行列から、それぞれのハンドサインの組み合わせにおいて通信路容量を導出する。導出した通信路容量をそれぞれ比較し、最大の通信路容量を与えるハンドサインの組み合わせを最適であるとする。通信路容量は、いくらかでも小さい復号誤り率で通信を行えるという条件のもとで、通信路を 1 回使うたびに送信できる最大ビット数(情報伝送速度)に等しい[8]。通信路における信頼性は複合誤り率、効率性は情報伝送速度により評価される。したがって、通信路容量を評価尺度とすれば信頼性と効率性を考慮できると言える。

以上のように、はじめはハンドサインの識別率を求めるために、ハンドサイン動作中の筋電位データを測定し、その識別を行う。そして、識別結果から得られた識別率に基づき、すべてのハンドサインの組み合わせについて通信路容量を導出し、それぞれを比較する。これらの提案手法の流れを図 1 に示す。

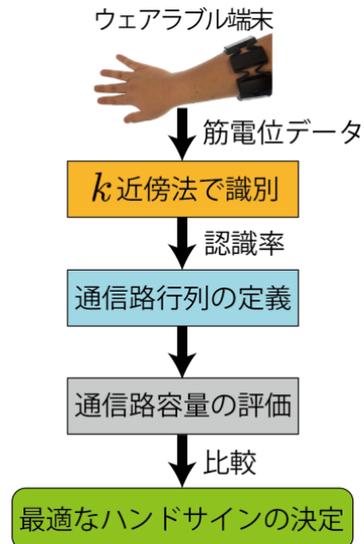


図 1 提案手法の流れ

3.2 Myo

ウェアラブル端末 Myo は Thalmic Labs 社によって開発されたジェスチャーコントロールアームバンドである。Myo は 8 チャンネル分の筋電位センサーを搭載している。コンピュータとは Bluetooth で接続し、腕に取り付けることでハンドサイン動作中の各センサーの値を送信することができる。

Myo による筋電位データの測定は、サンプリング周波数 200Hz で行われる。また、筋電位データは -128 から 127 の間の整数値に収まるように Myo によって変換される。

本論文では、Myo を筋電位データを測定する手段として用いる。大掛かりな装置を必要とせず腕に簡単に付けて使用できるのが Myo の利点である。また、無線通信によってデータを送信することができるため、人間が動作を行いやすく、場所も問われない。さらに、民生用として市販されているため入手しやすく、低コストでシステムを実現することができる。

3.3 筋電位データの識別と識別率の導出

3.3.1 k 近傍法

Myo を使用して測定した筋電位データの識別手法として、 k 近傍法を用いる。 k 近傍法に学習データとしてハンドサインの筋電位データとそれに対応するハンドサインのラベルを入力する。また、テストデータとして筋電位データを入力し、テストデータのハンドサインのラベルを出力として得る。この流れを図 2 に示す。

k 近傍法における予測には学習データとテストデータの乖離度を表す尺度が必要となる。本論文では、その尺度として DTW (Dynamic Time Warping) 距離を用いる。

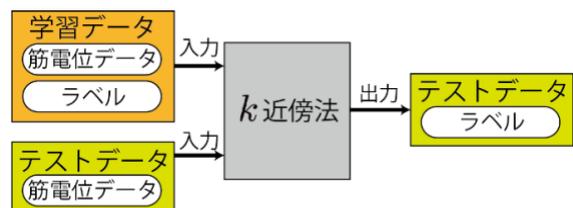


図 2 本論文における k 近傍法

3.3.2 ハンドサインの識別率の導出

本論文におけるハンドサインの識別率とは、「入力されたハンドサインがどのハンドサインとして識別されるか」を表す(条件付き)確率である。

Myo は 8 チャンネル分の筋電位データを一度に計測することができるため、1 つのハンドサインに対して 8 つの筋電位データが得られる。よって、 k 近傍法を用いて 1 チャンネルごとに筋電位データの識別を行うと、1 つのテストデータにつき 8 つの識別結果が得られる。これらの識別結果に基づいて識別率を導出する。まず、ラベルごとに識別された回数を求める。例として、使用するハンドサインのラベルが「あ、い、う、え、お」の 5 種類、 k 近傍法への入力として使用するテストデータのラベルが「あ」であるときのラベルごとの識別回数を図 3 に示す。多くの場合、図 3 に示すようにチャンネルによって識別結果が異なる。例えば、チャンネル 1 では「あ」と正しく識別するのにに対し、チャンネル 2 では入力ラベルとは異なるラベル「い」と誤って識別している。本研究では、ラベルごとの識別回数を識別結果の総数で割ることで識別率を求める。図 3 においては、ラベルごとの識別回数をそれぞれ 8 で割れば識別率を導出することができる。

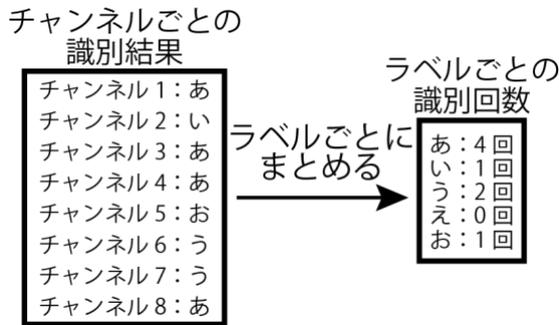


図 3 ラベルごとの識別回数の例

しかしながら、同じハンドサインであっても取得される筋電位データは測定ごとに変動する。そのため、学習データ・テストデータとして選んだデータによって識別結果が異なり、選んだデータに依存することになる。したがって、識別率を導出する際には、学習データ・テストデータの選び方を考慮する必要がある。以下では、その詳細な方法について説明する。

本論文では、それぞれのハンドサインにおいて学習データ・テストデータの比率を 8:2 とする。学習データ・テストデータの選び方は K 分割交差検証に基づいて決定し、 $K = 5$ とする。ここで、ハンドサインを M 種類使い、それぞれのハンドサインの筋電位データが N 個あるとする。このとき、すべての筋電位データは MN 個となる。

次に、それぞれのハンドサインにおける識別結果の総数を求める。前述のように、1 つのテストデータに対して 8 つの識別結果(チャンネル数分)が得られる。また、1 つの学習データ・テストデータの選び方につき、それぞれのハンドサインの筋電位データから $N/5$ 個のテストデータが選ばれる。よって、それぞれのハンドサインにおいて 1 つの選び方で $8N/5$ 個の識別結果が得られる。ゆえに、5 通りの選び方の識別結果を集計すると、それぞれのハンドサインにつき計 $8N$ 個の識別結果が得られる。また、ラベルごとの識別回数をすべての選び方について集計することで、

識別結果 $8N$ 個におけるラベルごとの識別回数を求めることができる。さらに、求めたラベルごとの識別回数を $8N$ で割ることで、識別率を求めることができる。

3.4 最適のハンドサインの組み合わせ

本論文では、ハンドサインの識別率から定義される通信路行列から、それぞれのハンドサインの組み合わせにおいて通信路容量を導出し、最大の通信路容量を与えるハンドサインの組み合わせを最適であるとする。以下では、この最適な組み合わせを導出する手順について説明する。

3.4.1 通信路行列

本論文では、筋電位の測定結果からハンドサインを識別するコンピュータシステムを通信路とみなす。この通信路の入出力はハンドサインのラベルである。しかし、この通信路は入力されたハンドサインのラベルと同じラベルが出力されるとは限らず、出力ラベルが入力ラベルによって確率的に変化するため、雑音のある通信路とみなすことができる。

ここで、通信路に入力するハンドサインのラベルを $A = \{a_1, a_2, \dots, a_m\}$ 、通信路から出力されるハンドサインのラベルを $B = \{b_1, b_2, \dots, b_n\}$ とし、 A を入力ラベル、 B を出力ラベルとよぶことにする。この通信路のモデルを図 4 に示す。この通信路は前述のように雑音のある通信路であるため、入力ラベル a_i を出力ラベル b_j として識別する識別率を条件付き確率とみなし、

$$p_{ij} = P(b_j | a_i) \quad (1)$$

と表すことにする。この p_{ij} を (i, j) 要素とする $m \times n$ の行列として、次のような通信路行列 T を定義することができる。

$$T = \begin{bmatrix} p_{11} & p_{12} & \dots & p_{1m} \\ p_{21} & p_{22} & \dots & p_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ p_{n1} & p_{n2} & \dots & p_{nm} \end{bmatrix} \quad (2)$$



図 4 通信路のモデル

3.4.2 通信路容量の導出と比較

通信路容量は、いくらでも小さい復号誤り率で通信を行えるという条件のもとで、通信路を 1 回使うたびに送信できる最大ビット数に等しい[8]。つまり、信頼性を確保した上で通信路を介して伝達することができる情報量の限界を表す。よって、通信路容量を評価尺度とすれば、信頼性と効率性の双方を考慮でき、高信頼かつ高効率である最適なハンドサインの組み合わせを求めることができる。ハンドサインの識別率(通信路行列)に基づき、ハンドサインの選ぶ個数・選ぶ種類のすべての組み合わせについて通信路容量を導出し、それらを比較して最大の通信路容量を与えるハンドサインの組み合わせを最適であるとする。

本論文ではハンドサインの選ぶ個数や選ぶ種類を変化させる中で、どの条件における通信路容量が最大であるのかを求める。ハンドサインの選ぶ個数や選ぶ種類が変化した場合における通信路容量の導出方法について、以下で説明する。

入力ラベル A 、出力ラベル B を持つ通信路の通信路容量 C は、

$$C = \max I(X; Y) \quad (3)$$

と表される。ここで、 $I(X; Y)$ は相互情報量である。すなわち、通信路容量 C は相互情報量 $I(X; Y)$ の最大値として定義される[8]。

ただし, (X, Y) は同時確率 $P(a, b) = P(a)P(b|a)$ にしたがう確率変数である. よって, 通信路容量 C を求めるためには, 相互情報量 $I(X; Y)$ を求めなければならない.

あるハンドサインの組み合わせにおける通信路容量を求めるときに, 通信路への入力として使うハンドサインのラベルの集合を A_1 , 使わないハンドサインのラベルの集合を A_2 とする. このとき, $A_1 \cup A_2 = A$, $A_1 \cap A_2 = \phi$ となる. このとき, 使わないハンドサインが入力されることはないことから, 以下が成り立つ.

$$P(X) = 0 \quad (X \in A_2) \quad (4)$$

ここで, 相互情報量 $I(X; Y)$ を次のように分解する.

$$\begin{aligned} I(X; Y) &= \sum_{(X, Y) \in A \times A} P(X, Y) \log \frac{P(X, Y)}{P(X)P(Y)} \quad (5) \\ &= \sum_{(X, Y) \in A_1 \times A_1} P(X, Y) \log \frac{P(X, Y)}{P(X)P(Y)} \\ &\quad + \sum_{(X, Y) \in A_1 \times A_2} P(X, Y) \log \frac{P(X, Y)}{P(X)P(Y)} \\ &\quad + \sum_{(X, Y) \in A_2 \times A} P(X, Y) \log \frac{P(X, Y)}{P(X)P(Y)} \quad (6) \end{aligned}$$

一方, 式(4)の制約条件を考慮すると, 式(6)の第 3 項は 0 となり無視できる. よって, 相互情報量 $I(X; Y)$ は,

$$\begin{aligned} I(X; Y) &= \sum_{(X, Y) \in A_1 \times A_1} P(X, Y) \log \frac{P(X, Y)}{P(X)P(Y)} \\ &\quad + \sum_{(X, Y) \in A_1 \times A_2} P(X, Y) \log \frac{P(X, Y)}{P(X)P(Y)} \quad (7) \end{aligned}$$

となる.

式(7)の第 2 項は使わないハンドサイン A_2 が出力された場合に相当する. この場合は, エラーが発生し, 通信が成立しなかったものとみなすことができる. エラーが発生した場合には, 再送等の対処が必要となる. このことから, 本論文においては式(7)の第 1 項のみを評価し, 使わないラベルの集合 A_2 についての総和部分(第 2 項)は相互情報量を求める際に考慮しない.

なお, 使うハンドサインに応じて通信路行列を個別に求める場合には, $P(Y|X) = 0 \quad (X \in A_1, Y \in A_2)$ から, 式(7)の第 2 項は必然的に 0 となり無視される. しかしこの場合は, すべてのハンドサインの組み合わせについて機械学習を行う必要があり, 莫大な計算量を必要とする. そこで本論文では, すべてのハンドサインを使用した 1 つの組み合わせについて機械学習を行い, 通信路行列を導出した. このとき, A_2 の識別率については前述のように通信が不成立であるとみなし, 通信路容量の評価から除外する.

4. 評価実験と考察

4.1 筋電位データの識別

3 名の被験者に対し, ハンドサインごとに筋電位を測定し, 得られた筋電位データを k 近傍法により識別した. そして, その識別結果からハンドサインの識別率を導出した. 詳細を以下に示す.

4.1.1 使用したハンドサイン

ハンドサインのサンプルとして, 手話の一部として実際に利用されている図 5 に示す 10 種類の指文字を使用した. これらは NHK オンライン[9]において定義された指文字であり, それぞれ「あ, い, う, え, お, か, き, く, け, こ」を表す. 本論文では, ハンドサインをあらかじめ定められた手の動きや形として定義する. 本論文では, 使用し

た 10 種類のハンドサインに対し, 既存の指文字と同様にそれぞれ「あ, い, う, え, お, か, き, く, け, こ」とラベルづけした. なお, このラベルづけは便宜的なものであり, それぞれのハンドサインを区別するために行った. よって, 実際のコンピュータシステムでは, これらのハンドサインをラベルづけした文字に対応させるとは限らない.



図 5 使用した 10 種類のハンドサイン[9]

4.1.2 筋電位データの測定

ウェアラブル端末 Myo を筋電位を測定するために使い, 右腕に装着した. 測定には, ホームポジションとして手を握る動作(じゃんけんのグー)を設けた. そのホームポジションを安静状態とみなし, ホームポジション→ハンドサイン→ホームポジションの順で動作を行い, その間の筋電位を測定した. ホームポジションを設けることにより, ハンドサイン動作時のみの筋電位とそれ以外のときの筋電位を区別しやすくした. 本論文では図 5 に示す 10 種類のハンドサインに対して筋電位の測定を 10 回ずつ行った. よって, 合計で 100 個の筋電位データを取得した.

4.1.3 k 近傍法を用いた筋電位データの識別

第 3 節で述べたように, 筋電位データを識別する手法として, k 近傍法を用いた.

1 つのハンドサインの 10 回の測定結果は, 学習データとテストデータの比率を 8:2 とし, 8 回を学習データ, 2 回をテストデータとして使用した. よって, 合計 100 個の筋電位データのうち, 80 個を学習データ, 20 個をテストデータとして k 近傍法に入力した. 入力として使用する学習データ・テストデータはハンドサインごとに K 分割交差検証 ($K = 5$) にしたがって決めた. なお, 学習データ・テストデータの取り出し方はランダムに決めた.

以上のように, K 分割交差検証 ($K = 5$) にしたがって 5 通りの学習データ・テストデータの選び方を求め, それぞれの選び方において k 近傍法を行った. Repeated K -fold cross-validation ($K = 5$) とよばれる手順で, 取り出すデータを変えながら 100 回行った. なお, $k = 3$ とした.

前述のように, 1 つのテストデータに対して 8 つの識別結果(チャンネル数分)が得られる. また, 1 つの学習データ・テストデータの選び方につき, それぞれのハンドサインの筋電位データから 2 つのテストデータが選ばれる. よって, それぞれのハンドサインにおいて 1 つの選び方で 16 個の識別結果が得られる. ゆえに, 5 通りの選び方の識別結果を集計すると, それぞれのハンドサインにおいて計 80 個の識別結果が得られる. さらに, これらの手順を取り出すデータを変えながら 100 回行うため, それぞれのハンドサインにおける識別結果は 8000 個得られる.

4.2 通信路容量の導出と比較

4.2.1 通信路行列の導出

本論文では、ハンドサインの識別率を通信路行列として定義する。被験者 A における通信路行列 T_A を式(8)に示す。通信路行列 T_A の各要素は式(1)の通りである。例えば、式(8)において「あ」を入力して「え」が出力される確率、すなわち「あ」が「え」と識別される確率は 0.243 となる。

$$T_A = \begin{bmatrix} 0.462 & 0.014 & 0.032 & 0.243 & 0.071 & 0.029 & 0.025 & 0.043 & 0.017 & 0.063 \\ 0.363 & 0.066 & 0.05 & 0.141 & 0.115 & 0.018 & 0.013 & 0.037 & 0.127 & 0.069 \\ 0.28 & 0.025 & 0.122 & 0.145 & 0.106 & 0.114 & 0.027 & 0.062 & 0.037 & 0.081 \\ 0.314 & 0.036 & 0.077 & 0.245 & 0.062 & 0.069 & 0.002 & 0.097 & 0.033 & 0.064 \\ 0.255 & 0.055 & 0.08 & 0.121 & 0.196 & 0.045 & 0.021 & 0.106 & 0.026 & 0.094 \\ 0.193 & 0.04 & 0.093 & 0.168 & 0.096 & 0.164 & 0.003 & 0.071 & 0.017 & 0.154 \\ 0.357 & 0.056 & 0.109 & 0.13 & 0.049 & 0.077 & 0.02 & 0.059 & 0.091 & 0.051 \\ 0.201 & 0.037 & 0.109 & 0.105 & 0.08 & 0.062 & 0.017 & 0.177 & 0.069 & 0.142 \\ 0.348 & 0.081 & 0.1 & 0.164 & 0.084 & 0.032 & 0.026 & 0.051 & 0.092 & 0.021 \\ 0.169 & 0.081 & 0.067 & 0.117 & 0.055 & 0.119 & 0.001 & 0.149 & 0.023 & 0.22 \end{bmatrix} \quad (8)$$

同様に、被験者 B・被験者 C における通信路行列 T_B , T_C をそれぞれ式(9), 式(10)に示す。

$$T_B = \begin{bmatrix} 0.329 & 0.024 & 0 & 0.01 & 0.199 & 0.083 & 0.186 & 0.125 & 0 & 0.043 \\ 0.184 & 0.029 & 0.008 & 0.035 & 0.206 & 0.096 & 0.209 & 0.028 & 0.195 & 0.011 \\ 0.128 & 0.054 & 0.064 & 0.051 & 0.322 & 0.031 & 0.184 & 0.012 & 0.127 & 0.028 \\ 0.081 & 0.179 & 0.014 & 0.139 & 0.196 & 0.019 & 0.153 & 0.026 & 0.183 & 0.012 \\ 0.123 & 0.049 & 0.011 & 0.025 & 0.481 & 0.081 & 0.08 & 0.026 & 0.079 & 0.045 \\ 0.143 & 0.02 & 0 & 0.021 & 0.208 & 0.173 & 0.339 & 0.015 & 0.023 & 0.059 \\ 0.065 & 0.022 & 0.002 & 0.04 & 0.223 & 0.093 & 0.379 & 0.05 & 0.011 & 0.114 \\ 0.014 & 0.011 & 0.002 & 0.078 & 0.128 & 0.068 & 0.256 & 0.198 & 0.081 & 0.163 \\ 0.043 & 0.036 & 0.017 & 0.072 & 0.18 & 0.112 & 0.199 & 0.012 & 0.307 & 0.024 \\ 0.01 & 0.013 & 0.004 & 0.009 & 0.176 & 0.114 & 0.268 & 0.162 & 0.027 & 0.218 \end{bmatrix} \quad (9)$$

$$T_C = \begin{bmatrix} 0.67 & 0.001 & 0.128 & 0.017 & 0.056 & 0.051 & 0.003 & 0.01 & 0.033 & 0.032 \\ 0.555 & 0.048 & 0.024 & 0.004 & 0.073 & 0.08 & 0.039 & 0.001 & 0.13 & 0.045 \\ 0.517 & 0.005 & 0.13 & 0.071 & 0.092 & 0.057 & 0.038 & 0.029 & 0.018 & 0.042 \\ 0.554 & 0.05 & 0.067 & 0.049 & 0.028 & 0.056 & 0.009 & 0.01 & 0.084 & 0.092 \\ 0.514 & 0.017 & 0.014 & 0.011 & 0.274 & 0.029 & 0.046 & 0.004 & 0.045 & 0.045 \\ 0.609 & 0.02 & 0.119 & 0.012 & 0.075 & 0.075 & 0.015 & 0.015 & 0.028 & 0.031 \\ 0.522 & 0.017 & 0.064 & 0.005 & 0.043 & 0.053 & 0.056 & 0 & 0.183 & 0.057 \\ 0.529 & 0.002 & 0.058 & 0.04 & 0.065 & 0.04 & 0.038 & 0.018 & 0.159 & 0.05 \\ 0.621 & 0.025 & 0.039 & 0.04 & 0.01 & 0.062 & 0.027 & 0.019 & 0.11 & 0.045 \\ 0.566 & 0.011 & 0.073 & 0.013 & 0.077 & 0.06 & 0.031 & 0.014 & 0.067 & 0.086 \end{bmatrix} \quad (10)$$

なお、式(8), 式(9), 式(10)における太字の値は、各入力ハンドサインに対して得られる出力ハンドサインの識別率の中で最も高いものである。

この評価実験において、入力ラベル A と出力ラベル B は、

$$\begin{aligned} A &= \{\text{あ, い, う, え, お, か, き, く, け, こ}\} \\ B &= \{\text{あ, い, う, え, お, か, き, く, け, こ}\} \end{aligned} \quad (11)$$

となり、各通信路行列の行・列はそれぞれ入力ラベル A ・出力ラベル B の要素に順に対応する。

4.2.2 通信路容量の導出と比較

式(8), 式(9), 式(10)に示す通信路行列に基づいて通信路容量を求めた。例えば、式(8)において入力に{あ, え, か}の

3 つのハンドサインを使用したときの通信路行列は、式(12)のようになる。これは式(8)から{あ, え, か}の3つのハンドサインに関する確率を抜き出したものであり、この通信路行列から通信路容量を導けばよい。

$$\begin{bmatrix} 0.462 & 0.243 & 0.029 \\ 0.314 & 0.245 & 0.069 \\ 0.193 & 0.168 & 0.164 \end{bmatrix} \quad (12)$$

同様の手順で、使用するハンドサインの種類・個数を変えながら通信路容量を導出した。

被験者 A・被験者 B・被験者 C におけるハンドサインの使用個数ごとの通信路容量の最大値を表 1 に示す。なお、通信路容量の計算結果は小数第 5 位で四捨五入した。

表 1 より、通信路容量が最大であるときのハンドサインの組み合わせは、被験者 A では「い, か, き, く, け, こ」の 6 種類、被験者 B では「あ, え, お, か, き, く, け, こ」の 8 種類、被験者 C では「え, お, か, き, く, け, こ」の 7 種類と求めることができた。

以上より、高効率かつ高信頼なハンドサインの組み合わせは被験者ごとに異なることが確かめられた。すなわち、被験者ごとに最適なハンドサインの設計を行えたと言える。

4.3 考察

表 1 より、被験者ごとに最適なハンドサインの組み合わせの種類・個数が異なっていることが分かる。また、通信路容量は必ずしもハンドサインの選ぶ個数が多いほど高いとは限らないことが分かる。最適なハンドサインについて 2 つの観点から考察する。

1 つ目は必ずしも識別率の高いハンドサインが最適なハンドサインに含まれないということである。例えば、被験者 A・C において、「あ」の識別率が高い。しかし、双方の最適なハンドサインには「あ」は含まれていない。これは、「あ」への誤識別率が高いことが要因であると考えられる。被験者 A では「こ」以外のハンドサインにおいて「あ」への誤識別率が最も高い。また、被験者 C では他のすべてのハンドサインにおいて「あ」への誤識別率が最も高い。一方で、被験者 A の別のハンドサインの個数においては「あ」が最適なハンドサインに含まれているものもある。これは「あ」の識別率が高いことが要因であると考えられる。これらは、被験者 B においても同様にみられる。被験者 B においては、「お」の識別率が最も高いが、他のハンドサインにおける誤識別率も高い。しかし、「お」は最適なハンドサインに含まれている。よって、他のハンドサインにおいて誤識別率が高いハンドサインであっても、識別率の高いハンドサインは最適なハンドサインを選ぶ上で無視できない存在であると言える。

2 つ目は、どの被験者の最適なハンドサインに「か, き, く, け, こ」の 5 種類が共通して含まれているということである。被験者 A においては、「か行」のハンドサイン同士は「あ行」のハンドサイン同士に比べて誤識別率が低いことが要因であると考えられる。また、被験者 B においては、「か行」の識別率の高さが要因であると考えられる。さらに、被験者 C においては、「あ」に次いで識別率の高い「お」に着目する。表 1 (被験者 C) より、ハンドサインの使用個数が 2・3 個のときのハンドサインには「お」「け」が含まれている。また、4・6 個のときは「え」「お」が含まれている。一方で 5 個のときは「お」は含まれておらず、「い」「う」が含まれている。これより、誤識別率を考慮すると「お」と「け」、「え」と「お」は同時に選ばれやすく、「お」と「い」「う」は同時に選ばれにくいと考えられる。よって、最適なハンドサインに「か行」が含まれたのは「お」と同時に選ばれにくいハンドサインが取り除かれた結果で

表 1 ハンドサインの使用個数ごとの通信路容量の最大値 (被験者 A・B・C)

ハンドサインの使用個数	被験者A		被験者B		被験者C	
	通信路容量の最大値	ハンドサインの種類	通信路容量の最大値	ハンドサインの種類	通信路容量の最大値	ハンドサインの種類
2	0.0861	け, こ	0.2458	あ, け	0.1148	お, け
3	0.0974	か, け, こ	0.2477	あ, け, こ	0.1148	お, け, こ
4	0.1	あ, か, き, こ	0.225	あ, え, く, け	0.1105	え, お, き, こ
5	0.1025	あ, か, き, く, こ	0.2402	あ, え, か, く, け	0.1162	い, う, え, く, け
6	0.1037	い, か, き, く, け, こ	0.2467	あ, え, か, く, け, こ	0.1178	え, お, か, き, け, こ
7	0.0717	あ, う, え, お, か, く, け	0.2478	あ, え, か, き, く, け, こ	0.1187	え, お, か, き, く, け, こ
8	0.0725	あ, う, え, お, か, く, け, こ	0.2479	あ, え, お, か, き, く, け, こ	0.1184	い, う, え, お, か, く, け, こ
9	0.0725	あ, う, え, お, か, き, く, け, こ	0.2002	あ, う, え, お, か, き, く, け, こ	0.1184	い, う, え, お, か, き, く, け, こ
10	0.084	あ, い, う, え, お, か, き, く, け, こ	0.1569	あ, い, う, え, お, か, き, く, け, こ	0.0989	あ, い, う, え, お, か, き, く, け, こ

あると考えられる。これらのことを踏まえると、被験者ごとの識別率の違いからそれぞれ異なる理由で「か, き, く, け, こ」の 5 種類が最適なハンドサインに含まれたと言える。

このように、被験者ごとの最適なハンドサインの組み合わせの選ばれる方は、識別率の高さも重要な要素である一方で、他のハンドサインにおける誤識別率も考慮されて決められている。よって、使用するハンドサインの個数を多くしたり、識別率の高いハンドサインを使用したりすることが必ずしも最適ではない。

5. まとめ

身体の障害により通常の手話言語や音声言語によるコミュニケーションが困難な人の生活の質を向上するためには、それぞれの障害の特性に応じた手段により、円滑な意思の疎通が図られることが重要である。本論文では、障害の度合いに応じたテーラーメイド式のハンドサイン識別システムの実現を目的とした。その上で、自身にとって動作可能な手指の伸展や屈曲により表現されるハンドサインによるコミュニケーションを高信頼かつ高効率に実現する手法を提案した。提案手法では、情報理論に基づき、手指の動作により発生する筋電位の測定結果からハンドサインを識別する過程を通信路とみなす。そして、その識別率から定義される通信路行列を用いて導出される通信路容量を評価尺度とすることで最適なハンドサインの組み合わせを求める。通信路容量は、信頼性を確保した上で通信路を介して伝達することができる情報量の限界を表しており、この値が最大のものが最適なハンドサインであると言える。

提案手法に対する評価実験は、3 人の被験者に対して行った。その結果、最適なハンドサインの組み合わせは被験者によって異なること、個数が多いほど良いとは限らないこと、識別率の高いハンドサインを使用することが必ずしも最適ではないことが明らかとなった。

一方で、通信路容量を尺度として使用した提案手法の

アプローチでは、ハンドサインの個数による使い勝手を評価することができない。使用するハンドサインの個数が多いと符号長は短くなるが、自身が覚える手間がかかる。逆に、使用するハンドサインの個数が少ないと符号長が長くなり、意思疎通を早く行えなくなる。よって、これらのトレードオフを踏まえた実用性に関する評価を検討する必要があるが、これは今後の課題とする。

謝辞

本研究の一部は、日本学術振興会における科学研究費補助金基盤研究(C)(課題番号 17K00135)による支援を受けている。ここに記し謝意を表す。

参考文献

- [1] 内閣府, “合理的配慮を知っていますか?”, http://www8.cao.go.jp/shougai/suishin/pdf/gouriteki_hairyo/print.pdf.
- [2] 厚生労働省, “身体障害者障害程度等級表”, <http://www.mhlw.go.jp/bunya/shougaihoken/shougai-shatechou/dl/toukyu.pdf>.
- [3] 平山亮, 船川政博, “ニューラルネットによる静止画像からの指文字認識”, 情報処理学会第 72 回全国大会 (2010).
- [4] Thshar Chouhan, Ankit Panse, Anvesh Kumar Voona, S.M. Sameer, “Smart Glove With Gesture Recognition Ability For The Hearing And Speech Impaired”, IEEE Global Humanitarian Technology Conference (2014).
- [5] 平山亮, 米山和也, 小池康晴, “表面筋電信号からの指文字認識手法”, 情報処理学会第 73 回全国大会 (2011).
- [6] João Gabriel Abreu, João Marcelo Teixeira, Lucas Silva Figueiredo, Veronica Teichrieb, “Evaluating Sign Language Recognition Using the Myo Armband”, Symposium on Virtual and Augmented Reality (2016).
- [7] Jackson Taylor, “Real-time translation of American Sign Language using wearable technology”, <http://scholarship.richmond.edu/honors-theses/928/> (2016).
- [8] 植松友彦, “イラストで学ぶ情報理論の考え方”, 講談社 (2012).
- [9] NHK オンライン, “指文字一覧 | NHK 手話 CG”, <http://cgi2.nhk.or.jp/signlanguage/syllabary.cgi>.