

# 筋肉の非活動部位を入力モダリティとして用いた ハンドジェスチャ認識手法の提案

御手洗 彰<sup>1,a)</sup> 棟方 渚<sup>1</sup>

**概要：**本稿ではユーザが物を把持しながら入力可能な手法としてハンドジェスチャ列認識手法を提案する。具体的には、複数回のハンドジェスチャを1入力として受け付けて識別することで、把持方法に依存せず高精度に識別可能な手法の開発を試みる。本研究では、ユーザの把持に使用されていない筋肉の筋電位を「筋肉の非活動部位」という新たな入力モダリティとして用いることで、提案手法の実現を試みた。具体的には、筋肉の非活動部位を入力モダリティとしたハンドジェスチャ列認識手法のプロトタイプを試作して入力識別精度の評価を行なった。さらに、ユーザ実験として把持分類学に基づく6種の把持条件において、同一の認識モデルを用いて、ハンドジェスチャ列の分類精度を評価する実験を行った。その結果、提案手法では93.13%の精度で基本的な把持方法におけるハンドジェスチャ列の分類が可能であることがわかった。最後にユーザ実験で得られた知見と実応用に向けた検討事項について考察する。

**キーワード：**入力モダリティ、ハンドジェスチャ列、筋電図、ジェスチャインタラクションデザイン

SHO MITARAI<sup>1,a)</sup> NAGISA MUNEKATA<sup>1</sup>

## 1. はじめに

人間とコンピュータのインタラクションは、近年ますます多様化してきている。ウェアラブル技術や仮想現実技術の進歩により、画面を見ずにコンピュータから情報を得たり、コントローラデバイス（マウスやスマートフォン）を把持せずに、ハンドジェスチャでコンピュータを操作したりすることが可能となってきている。HCI研究においては手や足、目、身体など人間の様々な身体部位を入力モダリティとした入力手法の提案がされている。

一方で、このような入力インタラクションの大きな問題点として、入力モダリティが占有されている状況では使用できない点が挙げられる。例えば、料理をしながらハンドジェスチャを行うことや、人と会話しながら音声入力を行うことは困難である。これは、入力モダリティである手や声帯がユーザの他の活動（料理や会話）によって占有されてしまうためである。従って、このような状況におけるシステムとのインタラクションには、日常生活で占有されない入力モダリティが必要となる。本研究では、日常生活で

使われない入力モダリティとして、「筋肉の非活動部位」における筋電位に着目した。例として、図1では、バッグを把持する際に前腕にかかる力の強さと、バッグを把持しながらハンドジェスチャを行った際に前腕にかかる力の強さを比較している（力の強さについて図では筋電位を指標とした百分率の値で示している）。図の青斜線部に着目すると、バッグの把持に使われている力は100%ではなく、前腕部の筋肉には非活動領域が存在することを示している。そのため、ユーザがバッグを把持しながらハンドジェスチャを行うと、バッグの把持に使用されていない筋肉、つまり筋肉の非活動部位が活動し、ハンドジェスチャをしていないときより強い力が発揮される（図1赤斜線部）。本研究ではハンドジェスチャによって発揮される余力を「筋肉の非活動部位」における入力モダリティとして定義する。筋肉の非活動部位における入力モダリティを用いることで、日常生活で使われる様々な「モノ」を把持し、手の入力モダリティが占有された状況下におけるハンドジェスチャ入力手法の実現を試みる。

## 2. 関連研究

モノを把持しながら行うジェスチャ入力は、ユーザの主

<sup>1</sup> 京都産業大学

<sup>a)</sup> i2086123@cc.kyoto-su.ac.jp

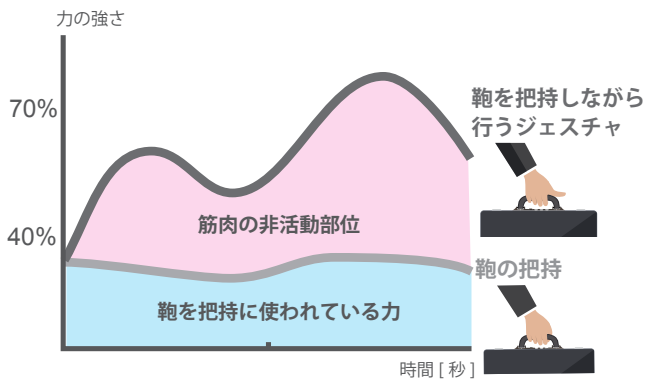


図1 バッグを把持している際の筋電位と、バッグを把持しながらハンドジェスチャを行った際の筋電位の比較（赤斜線部が筋肉の非活動部位を示す）．縦軸は筋電位を指標とした力の強さを百分率で表した値、横軸は時間（秒）を表す．

要なタスクを中断せずにコンピュータへの入力を可能とする手法として検討されている．これらの研究では、主に、モノを把持する際にどのようなジェスチャがユーザにとって好ましいかを調査している．Wolfら[15]は、モノを把持している最中に、どのようなジェスチャがインタラクション手法として適しているかを調査した．Sharmaら[14]は、モノを把持しながら行うマイクロジェスチャの体系的な調査を行った．具体的には、Schlesinger[13]の古典的な分類法に基づいて、6つの異なる把持タイプにどのようなマイクロジェスチャが適しているかを調査するための誘発調査を実施した．これらの研究の結果、ユーザの主要なタスクを妨げないコンピュータとのインタラクション手法が明らかになってきた．一方で、モノを把持しながら行うジェスチャを認識するための技術に関する調査は、まだ初期段階にある．

ハンドジェスチャの認識技術に関して様々な手法が研究されている．手は身体部位の中でも細やかな制御が可能である．そのため、自然な手の動作でコンピュータを操作できるようにすることは、インタラクション研究の課題となっている[6]．これまでの研究では、手首の曲げ伸ばし[7],[8],[9]や手話のような手の形を認識する手法[5],[10]やデバイスが開発されてきた．また、ハンドジェスチャによるインタラクションの社会的受容性を考慮して、指のジェスチャなどを用いた微小な動作による入力を実現する研究が行われている[1],[3]．

しかし、これらの技術は、ハンズフリーの状況を対象としたジェスチャ入力手法であり、モノを把持するような状況で適用することはできない．それに対して、Saponasら[12]は筋電センサを用いてモノを把持しているときのハンドジェスチャ認識システムを実現した．彼らの研究では、旅行用マグカップとノートパソコン用バッグの2つの把持条件で指のピンチジェスチャを認識する技術を開発し、高い精度でジェスチャの分類が可能であることを示した．しかし、実用を考えた場合、Saponasらの手法の有用性は限

定的である．その理由として、彼らのジェスチャ認識技術は対象物を把持した状態のハンドジェスチャを学習する必要があるためである．つまり、我々が日常生活で把持している全てのモノを学習する必要があるためである．これを解決するためには、把持の対象物に依存しない形で入力の学習を行う必要がある．

そこで、本研究では、モノを把持しながら行うハンドジェスチャ認識の課題として、ハンズフリーの状況の入力のみで機械学習により分類モデルを作成し、モノを把持した状況における入力の分類を試みる．そのためのアプローチとして、筋肉の非活動部位を入力モダリティとして用いた入力認識手法として、ハンドジェスチャ列認識手法を提案する．

### 3. ハンドジェスチャ列認識手法

我々の提案手法であるハンドジェスチャ列認識手法では、入力モダリティとして人間の筋肉の非活動部位における筋電位を用いる．既存の入力モダリティは、ユーザの日常の活動においてある身体部位が使用されているかによって、入力モダリティとして使用可能か不可能かが明確に分かれる．例として、歩行中にはフットジェスチャは使用できない、車の運転中にはハンドジェスチャは使用できない、などの状況が挙げられる．我々の目的は、ある入力モダリティが使用可能かを2値ではなく、連続値として捉えることで、入力モダリティがユーザの活動によってある程度占有されていたとしても、入力可能であることを示すことである．例えば、ユーザがバッグを手で把持している状況を考える．この時の筋肉の活動を筋電位として図1で表している．図のようにユーザはバッグの把持に100%の力を使っていないため、まだ力を加える余力を残している．本研究ではこの余力をユーザがモノを把持したり、手を使った活動をしたりする際に筋肉が非活動となっている部位、つまり筋肉の非活動部位であると定義し、これを入力モダリティとして利用する．

本研究では筋肉の非活動部位を入力モダリティとして利用するための入力手法として、ハンドジェスチャ列認識手法を提案する．本章では、提案手法の実装要件について説明し、最後に本研究で開発した提案手法のプロトタイプについて述べる．

#### 3.1 ハンドジェスチャ列

本研究の対象とする筋肉の非活動部位を入力モダリティとして用いる場合、入力モダリティはユーザに何らかの動作のために使用されていることが前提であるため、誤認識に対する堅牢性が重要な要素として挙げられる．日常におけるユーザの動作はジェスチャやセンサを用いた入力手法において、意図しない動作が入力と認識されることで、誤

認識の原因となる。様々な手法において誤認識を防ぐための試みがなされている [11] が Fukabori ら [2] は、ユーザの動作によっては誤認識が発生することが多く、適切に対処することが難しいと述べている。

その中で、誤認識に対する堅牢性を高めている入力手法として、Whack Gestures[4] が挙げられる。Whack Gestures は 2 つの Whack ジェスチャを入力するジェスチャの間に挟む形式で 3 つのハンドジェスチャによって構成されるハンドジェスチャ列を入力とすることで、認識の堅牢性を高めている。つまり、1 回のハンドジェスチャの誤入力が発生しても、ハンドジェスチャ列とはならないため、誤認識が発生しにくいことから、通常のジェスチャ認識手法と比較して堅牢性が高い。通常のハンドジェスチャ認識を用いる場面を考えると、1 つの操作に対して複数のハンドジェスチャを用いることは効率が悪くユーザビリティの低下を招くと考えられる。しかし、ユーザが何かしらの動作を行っている状況下であれば、ハンドジェスチャ列を用いることで認識の堅牢性を高めることができる有用な手法であるといえる。そこで、本研究では、3 つのハンドジェスチャから構成されるハンドジェスチャ列を入力として用いることとした。

### 3.2 様々な把持条件への汎用性

センサベースのハンドジェスチャ認識は主に手や前腕部から発せられる電気信号をもとに内部構造、外部構造の変化を学習することで実現されている。そのため、モノを把持した状況においては、ユーザがどのような握り方でモノを把持するか、また重量など対象物の特性によって同じハンドジェスチャを行ってもユーザの入力が異なったものに変化する。Saponas ら [12] はモノを把持しながら入力可能なハンドジェスチャの学習手法を開発したが、学習していないモノに対して認識可能ではなかった。その原因の一つとして、人が把持する対象物は多種多様であり、全てのモノを学習することは難しい。また、対象物の形状や重量、重心などモノの特性を表すパラメータも多種多様であるため、「モノ」と「モノの特性」を組み合わせて学習を行うとなると現実的ではない学習データ量が必要となる。様々な把持条件に対して汎用的に入力を認識するためには、上述のモノやパラメータに依存しない学習方法が必要となる。そこで提案手法では、ハンドジェスチャ間の差分を学習すること手法を採用した。具体的には、2 つのハンドジェスチャを比較して、同じか異なるかを判断する学習を行う。差分を学習することで、同じハンドジェスチャは同様の入力が得られ、異なるハンドジェスチャは異なる入力が得られるため、把持するモノやその重量が変わっても入力を正しく認識できる。提案手法では、3 つのハンドジェスチャを入力として使用し、1 番目と 2 番目のハンドジェスチャ、2 番目と 3 番目のハンドジェスチャの連続したハンドジェ

表 1 3 つのハンドジェスチャからなる列の並び順 (表中の A, B は異なるハンドジェスチャを表す)。

1 番目	2 番目	3 番目	ハンドジェスチャ列
A	A	A	AAA
		B	AAB
	B	A	ABA
		B	ABB

スチャ間の差分からどのようなハンドジェスチャ列が入力されたかを特定する。本研究では、表 1 に示す AAB や ABA のようなモルルス信号に似たハンドジェスチャ列を定義して用いることとした (A と B はそれぞれ任意の異なるハンドジェスチャを表す)。

### 3.3 ハンドジェスチャの分割

ハンドジェスチャ列を入力として認識するためには、連続して入力されるハンドジェスチャを区切る必要がある。提案手法では、2 つの閾値を設定することで連続したハンドジェスチャから、3 つのハンドジェスチャに分割する。2 つの閾値を決定するために、2 名の協力者にハンズフリーの状況とモノを把持した状況 (傘, バッグ, 箱を把持) において、ハンドジェスチャを行ってもらい、その筋電位を計測した。その後、自身の最大限の力を前腕に入れてもらい、その筋電位も計測した。これらを用いて、随意収縮強度 (筋肉の収縮度合い、力を入れている度合いを百分率で示した指標。以後、%MVC とする) を計算し、ハンドジェスチャを行った際、またはモノを把持している際に %MVC がどの程度となるかを評価した。その結果、ジェスチャを行っている際は 70%MVC 以上、ジェスチャを入力していない (モノの把持などは含まれる) 時には 40%MVC 以下であることがわかった。結果より、この 2 つの値を閾値として 70%MVC 以上であればハンドジェスチャが入力され、40%MVC 以下となれば、ハンドジェスチャの入力がされていないとして、この 2 値を閾値として設定した。

### 3.4 ハンドジェスチャ列認識手法のインタラクションシーケンス

本研究では、提案手法であるハンドジェスチャ列認識手法の実現可能性を検証するために、プロトタイプを実装した。図 2 に提案手法のインタラクションシーケンスを示す。提案手法は以下の流れでハンドジェスチャ列の入力を認識し、識別する。1) ユーザは入力としてハンドジェスチャを 3 回行う。それぞれのハンドジェスチャは、筋電位が 40% 以下から 70% 以上の %MVC に遷移したい時点で入力として認識される。次のハンドジェスチャ入力は筋電位が 40% 以下になるまで認識されない。2) ハンドジェスチャ間の差分をユークリッド距離によって計算する。計算で得られた 2 つの距離 (1 番目と 2 番目のハンドジェスチャ、2 番目と 3 番目のハンドジェスチャからそれぞれ計算された距離)

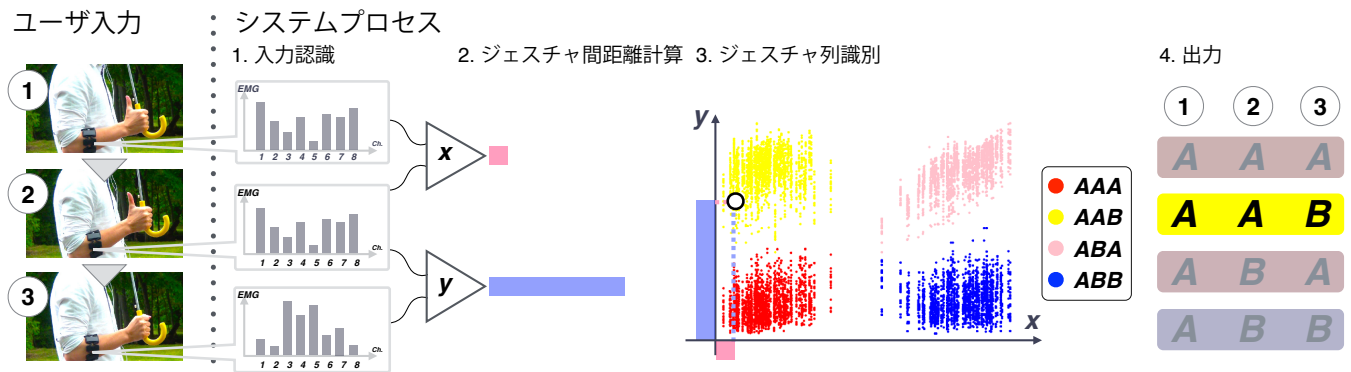


図 2 ハンドジェスチャ列認識手法

を機械学習の特徴量として用いる。3) k-NN アルゴリズムにより、ハンドジェスチャ列の識別を行う。4) 識別結果に応じて、特定の操作（音楽の再生など）を行う。

図 2 の例では、ユーザが「親指を立てる」ハンドジェスチャを 2 回行い、次に「こぶしを握る」ハンドジェスチャを行っている。この場合、識別されるべきハンドジェスチャ列は AAB となり、A と B はそれぞれ「親指を立てる」、「こぶしを握る」ハンドジェスチャを表す。ただし、提案手法は上記のハンドジェスチャのみを A, B とするわけではなく、例えば「手首を背屈\*1させる」、「手首を背屈させる」、「手首を掌屈\*2させる」のハンドジェスチャ列も AAB と識別される。これは、提案手法が特定のハンドジェスチャを学習しているのではなく、連続したハンドジェスチャ間の差分を学習しているからである。その結果、提案手法は未学習のモノを把持した状況においてハンドジェスチャ列を正しく識別可能である。

#### 4. 実験 1: ジェスチャ組の選定

提案手法では、ハンドジェスチャ列の認識のために 2 種のハンドジェスチャ（ここではジェスチャ A, B とおく）を要する。そこで、我々は予備実験として提案手法に適したハンドジェスチャの組み合わせ（以後、ジェスチャ組とする）の選定を実施した。まず、ハンドジェスチャを行った際の筋電位をハンズフリーの状況とモノを把持している状況から収集した。次に、12 種類のジェスチャ組からハンドジェスチャ列を作成し、それらの識別精度を比較することで、最も精度の高いジェスチャ組を決定した。

##### 4.1 実験手法

我々は、ハンドジェスチャ列の作成にシミュレーションによる手法を用いた。具体的には、ハンドジェスチャを収集し、それらを後で組み合わせることで擬似的なハンドジェスチャ列とした。通常、データ収集では実験参加者はジェスチャの種類 × 状況の種類 × 反復回数 の数だけ入

\*1 手の甲側に曲げること

\*2 手のひら側に曲げること

力を実行する必要があるため、多くの時間を要し、かつ実験参加者に負担をかける。さらにハンドジェスチャ列の収集では、一回の入力に三回のハンドジェスチャを実行する必要があるため、三倍の時間と負担がかかる。そこで、なるべく実験参加者に負担をかけない形式で少しでも多くの入力を得るために、ハンドジェスチャ列ではなく、ハンドジェスチャを収集し、それを後から組み合わせることでハンドジェスチャ列を作成した。結果として、320 回のハンドジェスチャを実験参加者ごとに収集した（4 種のジェスチャ × 4 種の状況 × 20 回）。

筋電位の測定のため、我々は 8ch の筋電位を測定可能な筋電センサである Myo を使用した。我々は筋電位およびハンドジェスチャ測定システムを Processing 3 を用いて開発した（筋電位の取得には、Myo for Processing ライブラリ\*3の“myoOnEmgData”メソッドを使用した）。結果として得られたデータを 200Hz でサンプリングされた 8bit の筋電値として扱った。

##### 4.2 実験参加者

12 名の実験参加者（男性 7 名、女性 5 名、平均年齢 23.5 歳、全員右利き）がデータ収集に参加した。全ての実験参加者は事前に実験内容、測定内容、公開内容について説明を受けたうえで、実験実施に同意した者が参加した。実験参加者には金属アレルギーの有無を確認した後（Myo の電極部が金属のため）に利き腕（右前腕部）に装着した。次に実験参加者に利き腕に可能な限り力を入れてもらい 100%MVC を計測した。その後、図 3 に示すハンズフリーとモノを把持した状況（傘、バッグ、箱を把持の 3 状況）で利き腕を使って 4 種のハンドジェスチャを行った。データ収集は休憩も含めて一人あたり約 1.5 時間の時間を要した。

##### 4.3 特徴量と機械学習

次に、作成したハンドジェスチャ列から機械学習のための特徴量を生成した方法について記す。まず、取得したハンドジェスチャの筋電位からそれぞれのチャンネルの RMS

\*3 <https://github.com/nok/myo-processing>





図 3 データ収集実験で用いた 4 つの状況におけるハンドジェスチャ。

値を次の計算式によって得た。

$$RMS_i = \sqrt{\frac{1}{N} \sum_{j=1}^N EMG_{ij}^2}, \quad (1)$$

$$RMS = [RMS_1, RMS_2, \dots, RMS_8], \quad (2)$$

ここで、 $EMG$  は筋電位の値、 $i$  はチャンネル数、 $N$  はサンプル数を表す。以上のようにして 8ch の筋電位から 8 次元の RMS 値のベクトルを得た。

次に、3 つのハンドジェスチャからなるハンドジェスチャ列から、2 つユークリッド距離を計算する。まず、ハンドジェスチャ同士の距離は以下の式によって計算した。

$$d(RMS_p, RMS_q) = \sqrt{\sum_{i=1}^8 (RMS_{pi} - RMS_{qi})^2}, \quad (3)$$

ここで、 $RMS$  は上で計算した 8 次元の RMS 値のベクトル、 $i$  はチャンネル数、 $p$  と  $q$  はハンドジェスチャの入力順を表す。3 章で述べたように提案手法の特徴量は 1 番目と 2 番目のジェスチャ間の距離、2 番目と 3 番目のジェスチャ間の距離からなるため、以下の式によって特徴量は以下ようになる。

$$feature = [d(RMS_1, RMS_2), d(RMS_2, RMS_3)]. \quad (4)$$

次に、上の特徴量を用いて機械学習を行う方法について述べる。提案手法では  $k$  近傍法 (以後、 $k$ -NN) によって学習することでハンドジェスチャ列の分類を行った。まず、学習に用いるデータは全てハンズフリーの状況のハンドジェスチャ列から作成した特徴量である。学習方法として我々は 10 分割交差検証によって、最も精度の高い  $k$ -NN のパラメータである  $k$  値と分類モデルを得た。具体的には、グリッドサーチと交差検証を組み合わせ、 $k$  値を 5~30 の範囲でそれぞれの分類モデルの精度を検証し、最も精度の高い  $k$  値と分類モデルを評価に用いた。この学習は実験参加者ごとに独立して行われたため、それぞれの  $k$  値は異

表 2 ジェスチャ組におけるハンドジェスチャ列の平均識別精度

A \ B	B					平均
	DE	PF	OH	MF		
DE	—	90.0%	73.4%	73.9%	79.1%	
PF	89.4%	—	78.2%	74.9%	80.8%	
OH	72.5%	82.0%	—	75.0%	76.5%	
MF	72.2%	78.4%	77.4%	—	76.0%	
平均	78.0%	83.4%	76.3%	74.6%	78.1%	

なり、 $k$  値の平均値は 12.9 であった (標準偏差は 10.7)。この分類モデルを用いて、12 種のジェスチャ組における、モノを把持した 3 状況 (傘, バッグ, 箱を把持) のハンドジェスチャ列における精度を評価した。

#### 4.4 結果

12 種のジェスチャ組における平均識別精度を表 2 に示す。ジェスチャ A, B は表 1 のハンドジェスチャの並び順を示す。平均識別精度はモノを把持した 3 状況における表 1 の 4 種のハンドジェスチャ列、つまり 12 種のハンドジェスチャ列の平均値を表す。表より最も精度の高いジェスチャ組は識別精度 90.0% の A=DE, B=PF であった。この結果から提案手法のジェスチャ組として A=DE, B=PF を用いることを決定した。

### 5. 実験 2: ハンドジェスチャ列認識手法の精度評価

前章では、未学習のモノを把持した状況において最も精度の高いジェスチャ組 (A=DE, B=PF) を選定し、90.0% の平均識別精度でハンドジェスチャ列が識別可能であることを示した。しかし、前章では擬似的にハンドジェスチャ列を作成しており、異なるハンドジェスチャ列であっても同じハンドジェスチャのデータが使用されているなど、データの均質化が生じているため過度に高い精度となっている懸念がある。さらにモノを把持した状況としては 3 種類と限られた条件で行われていた。そこで、提案手法であるハンドジェスチャ列認識をさらに検証するために、より多くの把持条件とモノを用いてより実用に近い形式で識別精度を評価する。

具体的には、ハンズフリーの状況におけるハンドジェスチャ列を用いて分類モデルの学習を行った後に、様々な把持条件において、実験参加者にハンドジェスチャ列を入力してもらい、その識別精度を評価するユーザ実験を実施した。加えて、ハンドジェスチャ列の入力時間を計測し、実用時における妥当性を評価した。実験条件として、我々は 6 つの把持条件におけるモノを用いた。実験参加者はそれぞれ順番にモノを把持してもらいその条件下でハンドジェスチャ列の入力を行った。なお、本実験は COVID-19 パンデミックの最中に実施されていたため、感染対策を行った上で実施された。

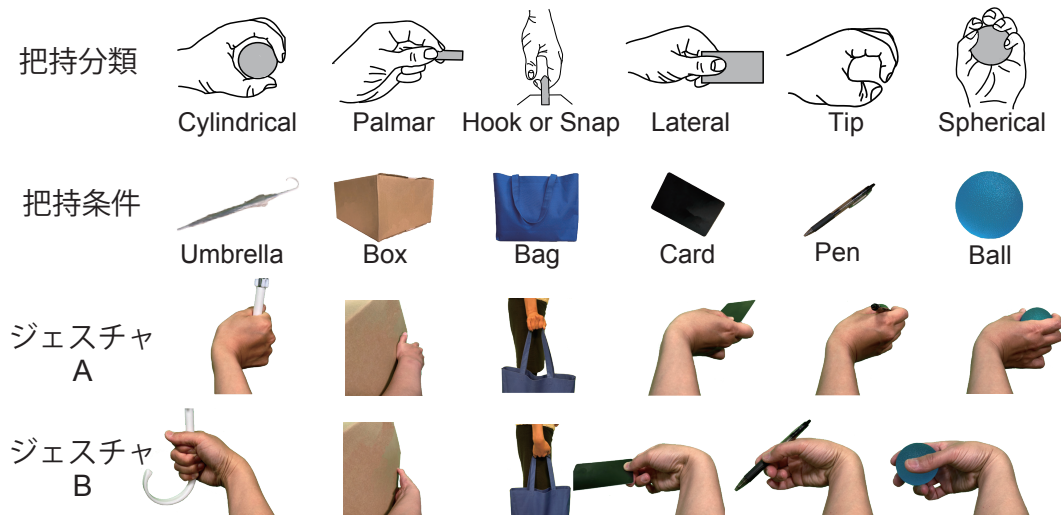


図4 ユーザ実験で用いた把持条件とハンドジェスチャ（ジェスチャ A, B はそれぞれ DE, PF ジェスチャを表す）。

### 5.1 実験参加者

12名（8名男性，4名女性，平均年齢23.2歳）の実験参加者がユーザ実験に参加した。全ての実験参加者は事前に実験内容，測定内容，公開内容について説明を受けたうえで，実験実施に同意した者が参加した。実験参加者には金属アレルギーの有無を確認した後）に利き腕（10名が右前腕部，2名が左前腕部）に装着した。ユーザ実験には実験参加者1人あたり，休憩含めて約1.5時間を要した。

### 5.2 把持条件

図4に本実験で実験参加者が把持するモノを示す。これらのモノはSchlesinger[13]の把持分類学を基に選定した。また図ではジェスチャ組A=DEとB=PFが各把持条件にてどのように行われたかを示している。我々は，実験参加者にこれらの条件において，表1にある4種のハンドジェスチャ列を入力させ，その識別精度を評価した。

### 5.3 実験手順

初めに，ハンズフリーの状況におけるハンドジェスチャ列の入力をトレーニングデータとして収集した。次に6つの把持条件におけるハンドジェスチャ列の入力をテストデータとして収集した。ハンドジェスチャ列の実行による疲労による識別精度への影響を考慮して，把持条件の順番は実験参加者ごとに無作為に決定した。実験参加者がハンドジェスチャ列を行う際は，3章に示した手法で，入力を分割することで記録した。収集したハンドジェスチャ列はジェスチャ組の選定時と同様の特徴量と機械学習の手法によって評価を行った。また，Myo デバイスは無線通信で筋電位の測定を行うが，実験時に無線通信や入力の識別処理による遅延は生じなかった。

### 5.4 結果

表3にそれぞれの把持条件における実験参加者ごとの平均識別精度を示す。まず，全体の平均識別精度は93.13%であった。これより，未学習の把持条件にて，高い精度でハンドジェスチャ列を識別可能であることが理解できる。次に，識別精度における一要因参加者内分散分析（要因：把持条件）の結果，Box条件が他の条件より有意に低い識別精度であることがわかった ( $F(5, 55) = 9.67, p < .05$ )。Box条件は他の把持条件において精度が劣るものの，精度低下の要因は4人の実験参加者に起因するものであった（表3下線部）。この原因についての詳細は，後の章で議論する。次に，表4にそれぞれの把持条件におけるハンドジェスチャ列の平均入力時間を示す。全体の平均入力時間は3.96秒であり，入力時間における一要因参加者内分散分析（要因：把持条件）の結果，条件間で入力時間に有意な差は認められなかった ( $F(5, 55) = 1.31, n.s.$ )。

ユーザ実験の結果より，実装したプロトタイプがハンドジェスチャ列をほぼ全ての把持条件において，ハンズフリーのハンドジェスチャ列で学習した分類器で正しく識別可能であることがわかった。また，入力時間について把持条件間で有意な差は認められなかったことから，ハンドジェスチャ列による入力は把持の形式に依らず，平均値である3.96秒から大きく乖離せず入力可能であることが理解できた。最後に，ハンドジェスチャ列入力による疲労について，実験参加者は実験の合間に平均0.86回（6人が休憩を取らなかった）の休憩を取り，ハンドジェスチャ列入力の継続が困難なほどの疲労感を訴えた実験参加者はいなかったことから，ハンドジェスチャ列入力には，実行困難なほどの高い負荷はかからないことがわかった。

表 3 実験参加者ごとの各把持条件におけるハンドジェスチャ列の識別精度の平均値 (列の P は実験参加者を示す)。

#	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	平均
Bag	96.30%	71.30%	93.80%	82.50%	97.50%	82.50%	97.50%	83.80%	98.80%	73.80%	91.30%	97.50%	88.88%
Ball	98.80%	98.80%	100.00%	100.00%	100.00%	98.80%	98.80%	98.80%	98.80%	97.50%	100.00%	100.00%	99.19%
Box	6.30%	8.80%	7.50%	97.50%	97.50%	87.50%	87.50%	91.30%	100.00%	78.80%	80.00%	1.30%	77.00%
Card	97.50%	98.80%	100.00%	100.00%	98.80%	95.00%	98.80%	97.50%	100.00%	97.50%	98.80%	100.00%	98.56%
Pen	100.00%	98.80%	98.80%	98.80%	97.50%	98.80%	100.00%	96.30%	100.00%	98.80%	100.00%	98.80%	98.88%
Umbrella	100.00%	93.80%	97.50%	100.00%	95.00%	95.00%	100.00%	97.50%	100.00%	80.00%	98.80%	97.50%	96.26%
平均	93.15%	86.72%	87.93%	96.47%	97.72%	92.93%	97.10%	94.20%	99.60%	87.73%	94.82%	89.18%	93.13%

表 4 各把持条件におけるハンドジェスチャ列の平均入力時間と標準偏差 (S.D.)。

把持条件	Umbrella		Box		Bag		Card		Pen		Ball	
	平均	S.D.	平均	S.D.	平均	S.D.	平均	S.D.	平均	S.D.	平均	S.D.
入力時間 (秒)	4.00	0.82	4.13	0.98	4.00	0.79	3.81	0.68	3.99	0.84	3.81	0.90

## 6. 議論

ユーザ実験では識別精度と入力時間の観点において提案手法であるハンドジェスチャ列認識手法の実現可能性を示した。一方で、実用化にはさらなる性能の向上が必要であるため、本研究で得られた結果に関する考察と実用化に向けて必要な事項について議論する。

### 6.1 ジェスチャと識別精度の関係

ユーザ実験では、提案手法が未学習のモノを把持した状況において 93.13%の精度で識別可能であることを示した。ここでは、Box 条件において低い精度となった 4 人の実験参加者について分析する。図 5 は 4 人の実験参加者の Box 条件におけるハンドジェスチャ列の識別精度の混合行列を示している。この混合行列からハンドジェスチャ列 AAB, ABA, ABB が AAA と誤認識されたことがわかる (AAA は 3 回同じハンドジェスチャを繰り返す)。この図より、箱を把持する条件では手や手首が固定されることから、十分に動かすことができず、その結果、2 種の異なるハンドジェスチャをうまく表現できなかった (異なる筋肉を収縮できなかった) ことが誤認識の原因ではないかと考えられる。この問題は、手首が固定された状態でも筋電位の変化がより明確なハンドジェスチャの組み合わせを用いることや、圧力センサのように微細な変化に敏感なセンサを組み合わせることで識別精度を向上できると考えられる [9]。

### 6.2 入力時間の妥当性

ハンドジェスチャ列の入力のために必要な時間はユーザビリティに大きな影響を与える。そのため、3 つのハンドジェスチャ列の妥当な入力時間について議論する必要がある。1 つのハンドジェスチャ入力と比較してハンドジェスチャ列入力はより多くの時間を要する。そのため、我々はユーザ実験によって計測したハンドジェスチャ列の入力時間の妥当性を分析する。

ユーザ実験より、ハンドジェスチャ列の入力にかかる時間は平均 3.96 秒であった。ハンドジェスチャ入力と比較するとこの入力時間は長い。しかし、モノを把持した状況で

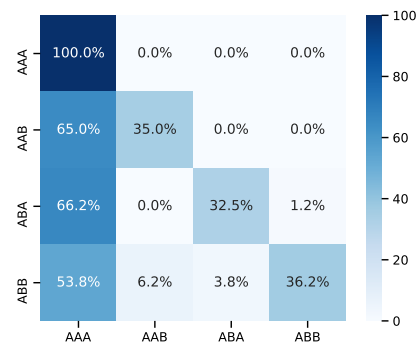


図 5 ユーザ実験の Box 条件時に識別精度が低かった 4 名のハンドジェスチャ列識別精度の混合行列

はハンドジェスチャ入力は利用できないことを考えると、音声入力のような同様の状況で入力可能な手法と比較することが適切であると考えられる。音声入力はウェイクワード (声による操作を開始するためにあらかじめ設定された特定の単語) を含めて約 3-5 秒程度の入力時間を要する。提案手法と音声入力による操作の種類やその多様性などが異なるため、提案手法が音声入力同等のユーザビリティを示すことはできないが、入力時間の観点ではハンドジェスチャ列入力は音声入力と同等であり、十分に妥当なものであると考えられる。

### 6.3 ハンドジェスチャ列入力による疲労の影響

ハンドジェスチャ入力と比較して、ハンドジェスチャ列入力は多くの動作を必要とするため、ユーザの疲労は重要な検討項目である。ユーザ実験では、実験参加者一人あたり 560 回のハンドジェスチャ列入力 (1 ハンズフリーの状況 × 80 反復回数 + 6 把持条件 × 80 反復回数) を 90 分間に行った (1680 回のハンドジェスチャ入力に相当)。実験中、実験参加者は平均して 0.86 回の休憩を取り、疲労によって入力の継続が困難となった実験参加者はいなかった。一方で、全ての実験参加者が実験中に疲れを感じたと述べた。また、全ての実験参加者が重量のあるモノ (Bag 条件) や不安定なモノ (Umbrella 条件) において他のモノより疲れを感じたと述べた。この結果より、ハンドジェス

チャ列入力は、タイピングのような頻繁な入力を要するタスクよりも、音楽プレイヤーの操作のような少ない入力で操作可能なタスクに適していると考えられる。また、把持する対象物について、モノの重量や安定性が疲労に影響を及ぼすことがわかっており、この点についてはさらなる調査が必要であることがわかった。

## 7. おわりに

本研究では、モノを把持している状況下で用いるためのハンドジェスチャ列認識手法を提案した。提案手法は、連続して入力された3つのハンドジェスチャからその差分を学習し、分類する。本研究では、ユーザの筋肉の非活動部位を入力モダリティとして用いたプロトタイプを開発し、モノを把持した状況下におけるプロトタイプの入力の識別精度、ユーザビリティを評価した。具体的には、ハンズフリーの状況における入力で作成された分類モデルを用いて、古典的な把持の分類学に基づいて、6つのモノを把持した状況において評価を行った。その結果、全体の識別精度として、93.13%の精度で分類可能であることがわかった。一方で、把持したモノの種類が識別精度に大きく影響することがわかった。把持条件における識別精度の低下を抑えるため、今後の課題として筋電位以外のセンサとの組み合わせによる改善を検討する。

**謝辞** 本稿の執筆にあたり、実験にご参加いただいた実験参加者の皆様に厚く御礼申し上げます。

## 参考文献

- [1] Christoph Amma, Thomas Krings, Jonas Böer, and Tanja Schultz. Advancing muscle-computer interfaces with high-density electromyography. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, page 929–938, New York, NY, USA, 2015. Association for Computing Machinery.
- [2] Koumei Fukahori, Daisuke Sakamoto, and Takeo Igarashi. Exploring subtle foot plantar-based gestures with sock-placed pressure sensors. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, page 3019–3028, New York, NY, USA, 2015. Association for Computing Machinery.
- [3] Donny Huang, Xiaoyi Zhang, T. Scott Saponas, James Fogarty, and Shyamnath Gollakota. Leveraging dual-observable input for fine-grained thumb interaction using forearm emg. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software and Technology*, UIST '15, page 523–528, New York, NY, USA, 2015. Association for Computing Machinery.
- [4] Scott E. Hudson, Chris Harrison, Beverly L. Harrison, and Anthony LaMarca. Whack gestures: Inexact and inattentive interaction with mobile devices. In *Proceedings of the Fourth International Conference on Tangible, Embedded, and Embodied Interaction*, TEI '10, page 109–112, New York, NY, USA, 2010. Association for Computing Machinery.
- [5] Hassaluk Jirathampradub, Chakarida Nukoolkit, Kaneungjit Suriyathumrongkul, and Bunthit Watanapa. A 3d-cnn siamese network for motion gesture sign language alphabets recognition. In *Proceedings of the 11th International Conference on Advances in Information Technology*, IAIT2020, New York, NY, USA, 2020. Association for Computing Machinery.
- [6] David Kim, Otmar Hilliges, Shahram Izadi, Alex D. Butler, Jiawen Chen, Iason Oikonomidis, and Patrick Olivier. Digits: Freehand 3d interactions anywhere using a wrist-worn gloveless sensor. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology*, UIST '12, page 167–176, New York, NY, USA, 2012. Association for Computing Machinery.
- [7] Zhiyuan Lu, Xiang Chen, Zhangyan Zhao, and Kongqiao Wang. A prototype of gesture-based interface. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*, MobileHCI '11, page 33–36, New York, NY, USA, 2011. Association for Computing Machinery.
- [8] Jess McIntosh, Asier Marzo, Mike Fraser, and Carol Phillips. Echoflex: Hand gesture recognition using ultrasound imaging. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, page 1923–1934, New York, NY, USA, 2017. Association for Computing Machinery.
- [9] Jess McIntosh, Charlie McNeill, Mike Fraser, Frederic Kerber, Markus Löchtefeld, and Antonio Krüger. Empress: Practical hand gesture classification with wrist-mounted emg and pressure sensing. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, page 2332–2342, New York, NY, USA, 2016. Association for Computing Machinery.
- [10] J. Rekimoto. Gestur wrist and gestur pad: unobtrusive wearable interaction devices. In *Proceedings Fifth International Symposium on Wearable Computers*, pages 21–27, 2001.
- [11] Jaime Ruiz and Yang Li. Doubleflip: A motion gesture delimiter for mobile interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, page 2717–2720, New York, NY, USA, 2011. Association for Computing Machinery.
- [12] T. Scott Saponas, Desney S. Tan, Dan Morris, Ravin Balakrishnan, Jim Turner, and James A. Landay. Enabling always-available input with muscle-computer interfaces. In *Proceedings of the 22nd Annual ACM Symposium on User Interface Software and Technology*, UIST '09, page 167–176, New York, NY, USA, 2009. Association for Computing Machinery.
- [13] G. Schlesinger. *Der mechanische Aufbau der künstlichen Glieder*, pages 321–661. Springer Berlin Heidelberg, Berlin, Heidelberg, 1919.
- [14] Adwait Sharma, Joan Sol Roo, and Jürgen Steimle. Grasping microgestures: Eliciting single-hand microgestures for handheld objects. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, page 1–13, New York, NY, USA, 2019. Association for Computing Machinery.
- [15] Katrin Wolf, Anja Naumann, Michael Rohs, and Jörg Müller. A taxonomy of microinteractions: Defining microgestures based on ergonomic and scenario-dependent requirements. In Pedro Campos, Nicholas Graham, Joaquim Jorge, Nuno Nunes, Philippe Palanque, and Marco Winckler, editors, *Human-Computer Interaction – INTERACT 2011*, pages 559–575, Berlin, Heidelberg, 2011. Springer Berlin Heidelberg.