

ジェスチャ認識システムを用いた観測に基づく 共通ジェスチャの抽出と検討

松澤 荘太^{†1} 木村 浩章^{†1} 中島 達夫^{†1}

ジェスチャ認識を用いたインタラクションは、ユーザにとって直感的で学習が容易な手法の1つであると言える。現在ジェスチャ認識システムについては数々の研究がなされているが、その中で「各アプリケーションの機能に対して定義されるジェスチャをどのようなものにするのが最適か」という問題をうまく解決した事例は少ない。本稿ではこの問題を解決することを目的に、ユーザがジェスチャ認識システムを使用する過程を観測することによって、特定の機能に対して定義されるのに適したジェスチャやその性質を抽出する。さらに実験後にアンケートを行うことで定性的な評価も含めた検討を行い、デザイナーや開発者にとっての判断材料となる要素を提案する。

Observation-based study of common gestures for desktop applications

SOTA MATSUZAWA,^{†1} HIROAKI KIMURA^{†1}
and TATSUO NAKAJIMA^{†1}

In a ubiquitous computing environment, people are surrounded by hundreds of mobile or embedded computers which may be used to support one or more user applications, and the situation of using the services become diverse. In this environment, the methods of human-computer interaction will change. And gesture recognition is one of the intuitive way of interaction. When a application of gesture interaction is designed, the designer must invent a set of possible meanings for each of the gesture actions that users perform. Currently, however, there is no reference point to get such informations. In this paper, We propose the universal gesture set which can be some help for designers.

^{†1} 早稲田大学基幹理工学研究科情報理工学専攻
Department of Computer Science, Waseda University, Japan

1. はじめに

近年、コンピュータの高性能化やネットワークインフラの発達によってコンピュータの利用形態やその目的は変化しつつある。今後はこれまでとは異なった種類のサービスが創出されることが予想され、質の高いユーザ体験を提供するために、情報の入出力、すなわちコンピュータとのインタラクションの形態も変化していくと考えられる。コンピュータとのインタラクションとしては一般的に、ユーザにとって直感的で学習容易なものが理想的である²⁾とされ、本稿で取り上げるジェスチャ認識を用いたインタラクションはこれらの要件を満たす有効な手法の1つであると言える。このことはジェスチャ認識によってゲームを操作する任天堂 Wii^{*1}がユーザに広く受け入れられ、高い評価を得ていることから伺える。

ジェスチャ認識技術については数々の研究がなされており、東芝による Qosmio G50^{*2}などのように、アプリケーションを操作するためにジェスチャ認識を採用したものが商用化されている例もある。だが、このようなジェスチャ認識によるインタラクションを用いたシステムの開発において、「各アプリケーションの機能に対して定義されるジェスチャをどのようなものにするのが最適か」という問題をうまく解決した事例は少ない。この問題はアプリケーションのユーザビリティに大きく影響し、ひいてはシステム全体の評価にまで波及する可能性がある。それにもかかわらず、現状ではデザイナーや開発者の意図のみによってデザインされている場合が多い。

この問題はジェスチャ認識における体系化されたデザインパターンや、広く普及したガイドラインといったものが存在しないことに起因する。すなわちデザイン時の判断材料が存在しないため、開発者は自分達の意図をユーザに強制するか、もしくは煩雑でコストのかかるユーザビリティテストを繰り返すかという選択肢を迫られることになる。ジェスチャ認識は未だ発展途上の技術であることから、開発対象のシステムがクリティカルなタスクを遂行するためのものでないことがほとんどである。多くの場合コストのかからない前者が選択され、知識・経験の蓄積やコミュニティの形成が行われないという悪循環が起きている。本稿では上記の課題に取り組むために、実際にユーザがジェスチャ認識システムを使用する過程を観測することで、特定の機能に対して適したジェスチャやその性質を抽出し、デザイナーや開発者にとっての判断材料となる指針を提案する。

*1 任天堂 Wii <http://www.nintendo.co.jp/wii>

*2 東芝 QosmioG50 <http://dynabook.com/pc/catalog/qosmio/080820g5/index.j.htm>

コンピュータとのインタラクションとして使用するジェスチャを定義するとき、そのジェスチャがあらゆるユーザにとって使いやすいものであるためには、ユーザの属性に依存しない共通したジェスチャであることが必要だと考えられる。多くのユーザが特定のコマンドから共通して想起することのできるジェスチャには、そのコマンドからジェスチャへのなんらかの共通したきっかけや特徴が存在していると言える。そのような特徴が反映されたジェスチャを用いることで、コマンドによって行われるタスクを達成するためのメンタルモデルは容易に形成できる。すなわち複数ユーザ間において共通性が存在するジェスチャは直感的で学習負荷が少なく、誰もが使用することのできるジェスチャであると言える。

本稿では共通ジェスチャを抽出するために、実際にユーザが自由なジェスチャを各コマンドに対して定義し、そのジェスチャを用いて各タスクを実行することのできるアプリケーションを開発した。ユーザはアプリケーションを使用しながら、より直感的であり、認識率が高いジェスチャを模索することができる。このアプリケーションを用いてユーザが特定のコマンドから想起するジェスチャのデータを収集し、それらの統計や変遷の傾向を検討・考察する。実際にユーザがアプリケーションを使用する過程のデータを収集することによって、アンケートによる調査では得られないシステムの制約の中における実用的なジェスチャが抽出できるものとする。実験ではアプリケーション実装の容易性とアプリケーションを実行するコンピュータに要求されるスペックという理由から加速度センサを用いたジェスチャ認識を対象とする。携帯電話等の身近なデバイスに加速度センサが埋め込まれることを考えると、加速度センサを身体の一部に取り付けるよりも、手でデバイスを掌握した状態でジェスチャを行うほうがユーザに受け入れられやすいと言える。また手は微細な運動機能を有するため、多様で正確なジェスチャを行うことが可能である。したがって、本稿においては携帯端末を持った状態でジェスチャを行うことを想定し、特に手および腕の動作（ハンドジェスチャ）に焦点をあてる。

2. 実験用アプリケーションの設計と実装

2.1 設 計

まず我々はアンケート調査によって、ユーザが想起できるジェスチャのおおよその種類と操作対象とすべきアプリケーションの知見を得た。その結果からジェスチャによる操作の対象として動画再生と画像閲覧を実装することとした。これらのアプリケーションに対する機能要求を分析すると表 1 のタスクが存在し、実験用アプリケーションでは表中の各タスクを実行するコマンドに対してジェスチャを定義して操作できるようにした。また表 2 のタ

表 1 タスク内容

動画再生	再生, 停止, 一時停止, 早送り, 巻き戻し, 音量上げ, 音量下げ, フル画面表示, ダウンロード
画像閲覧	拡大, 縮小, 回転, 印刷, メール送信

表 2 プリミティブなタスク

プリミティブ	確定, 上, 下, 左, 右
--------	----------------

スクはアプリケーションに依存せず、共通のものとして用いられる。本研究ではこれらをプリミティブなタスクと呼ぶことにし、他のタスク同様にユーザが定義したジェスチャを用いた操作を可能とした。

ジェスチャ認識のためのアルゴリズムには隠れマルコフモデル (HMM) を用いる。HMM は確率モデルを用いたパターン認識手法である。HMM によるパターン認識を行う場合、一般に多数のトレーニングデータが必要とされる。実装に用いたライブラリでは 15~20 のトレーニングデータを学習させることを推奨している。しかし特定のコマンドに対してユーザがジェスチャを定義する際に、20 回もジェスチャを行わなければならないというのは非現実的である。もしこのような仕様にするとジェスチャを定義すること自体へのモチベーションが下がり、正しい結果が得られない。

そこでトレーニングデータにガウスノイズを付加することで疑似的に学習データを増やすこととした。ガウスノイズとは正規分布に従うノイズのことをいう。意図的にノイズを付加することで基準となるオリジナルデータの傾向を残しつつ、ゆらぎを持った別のデータを生成することができる。このとき負荷するノイズの大きさがオリジナルデータに対して大きすぎるとオリジナルデータの特徴が失われてしまうし、小さすぎるとフィルタリングによって同じデータとして扱われてしまう。したがって、まず実験的に最適なパラメータを見出し、そのパラメータを用いて生成したガウスノイズをトレーニングデータに付加した。こうすることでユーザへの負担を軽減した。

2.2 実 装

実験用アプリケーションは Java 言語と Adobe AIR (Adobe Integrated Runtime, 以下 AIR) *1 によって実装した。MVC モデルに則って設計を行い、データマネージャ部 (Model), プレゼンテーション部 (View), センサ部 (Controller) から構成される。これらコンポーネント間はすべてソケット通信によってデータの送受信を行っているため、将来的にコン

*1 AdobeAIR <http://www.adobe.com/products/air/>

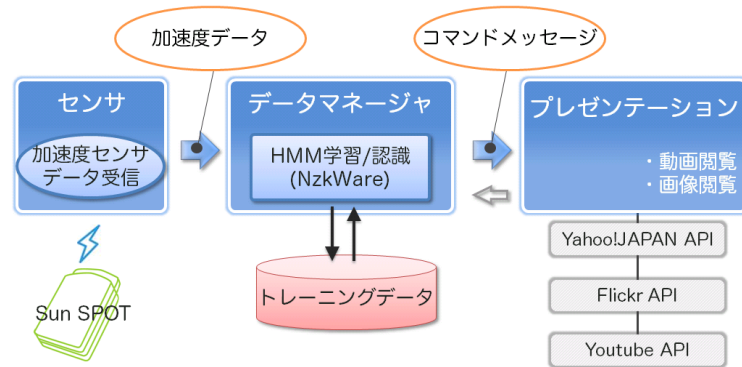


図 1 アプリケーション構成図

ポーネントを差し替える必要が生じた場合に対応できるようになっている。例えば用いるセンサを変更したい場合や、プレゼンテーションを別のものに変更したい場合に容易に変更可能である。具体的な構成を図 1 に示す。

データマネージャ データマネージャはセンサからのデータを解釈し、結果をプレゼンテーションに反映する役割を担う。この部分は Java 言語で実装した。各コンポーネントの中心に位置し、それぞれのコンポーネントと繋がっている。

データマネージャはジェスチャ学習・認識部を包含している。HMM による学習・認識には筆者らの所属する研究室で開発した NzkWare^{*1}を利用した。NzkWare は元々、複数センサを用いた際のコンテキスト取得における認識率の変化を測定するために開発されたものである。NzkWare はシンプルなインタフェースで API が提供されているため、非常に扱いやすいのが特徴といえる。HMM 学習・認識のコア部分にはオープンソースの wii リモコン用ジェスチャ認識ライブラリである wiigee⁵⁾ が組み込まれている。

プレゼンテーション プレゼンテーションは主にジェスチャによって実行されたタスクの結果を表示する役割を担う。この部分は Adobe AIR を ActionScript3 によって実装した。AIR の開発には Adobe Flash の技術をそのまま利用できるため、非常に優れた表現力とインタラクティブ性を実現できる。プレゼンテーションの各表示を図 2~図 5 に示す。各々の画面状態でジェスチャを行うことで学習/認識が行われてタスクが実行される。

*1 Nzkware <http://www.dcl.info.waseda.ac.jp/~nzk.msk/NzkWare/>



図 2 動画再生画面のスクリーンショット



図 3 画像閲覧画面のスクリーンショット

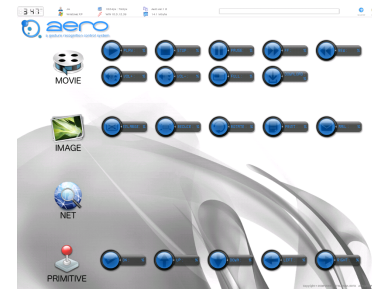


図 4 ジェスチャ登録画面のスクリーンショット

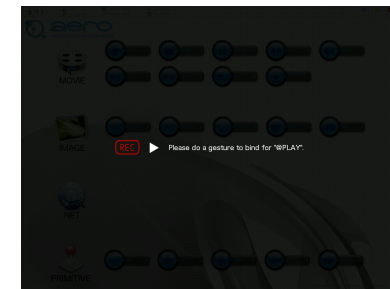


図 5 登録中のスクリーンショット

センサ センサはユーザが行ったジェスチャをデータとして取り込み、データマネージャへ渡す役割を担う。センサには Sun Microsystems, Inc. が開発した無線センサネットワークデバイスである Sun SPOT^{*2}を使用した。実験用アプリケーションでは Sun SPOT を用いて加速度センサのデータを無線でパーソナルコンピュータへ送信する。無線での通信によって、ユーザがジェスチャを行う際の制約が軽減されるというメリットがある。

*2 SunSPOT <http://jp.sun.com/products/software/sunspot/>

3. 実験

3.1 実験における狙い

はじめに述べた目的を果たすために、前節で述べた実験用アプリケーションを用いて検証実験を行った。その使用に際してユーザは図6の過程を繰り返す。この過程を繰り返すことで、ユーザは自分にとってより直感的であり、かつ認識率が高いジェスチャを模索する。一定期間の模索を経た後、最終的に満足のいくジェスチャが見つかる、ユーザの中にメンタルモデルが形成され、ユーザはそれ以上定義したジェスチャを更新しなくなると考えられる。このような過程を経て収束したジェスチャのデータを見たときに、各ユーザの間で共通したジェスチャが用いられていれば、そのジェスチャはコマンドに定義するのに適当な性質を持っていると言える。

3.2 内容

被験者は全部で8名で、5名は情報工学を専攻する20代の学生、残り3名は情報工学を専攻していない20代の学生である。実験は被験者が所有するパーソナルコンピュータ上に実験用アプリケーションをインストールし、各被験者の馴染んだ環境下において実施した。被験者には最初に本研究の目的を説明した上で、アプリケーションを実際に操作しながら口頭で操作方法を教示した。また説明内容をまとめたマニュアルも配布した。被験者は各自が実験を行う環境にアプリケーションをインストールし、以下のことを条件にアプリケーションを使用する。

- 実験期間は4日間。



図6 想定されるユーザアクションの過程

- 最低でも1日15分以上はジェスチャを用いた操作でアプリケーションを使用する。
：一定量以上のログデータが回収できるよう、被験者には意識的にアプリケーションを使用する時間を取ってもらう。
- 1日1回は各コマンドに対して定義されているジェスチャを見直す。
：使いづらいつと感じるジェスチャを中心に定義しなおしてもらう。
- 定義したジェスチャはその都度手書きで記録を取る。
：定義したジェスチャの加速度データはログとして収集するが、そのデータからジェスチャの正確な概形を復元することが困難であったため、同時に被験者による手書きのデータを収集した。

4日間の実験終了後、被験者に対して表3のアンケートを実施した。アンケートはWeb上でを行い、Q1～Q5は7段階評価及び自由回答、Q6～Q10は選択式（複数選択可）、Q11～Q12は自由回答の形式で回答を得た。

表3 実験後アンケート項目

質問	回答軸
Q1: 各タスク（機能）に割り当てたジェスチャを覚えることはできましたか	覚えられなかった: 1 すべて覚えられた: 7
Q2: ジェスチャ認識は操作方法として直感的だと思いますか	直感的でない: 1 直感的である: 7
Q3: ジェスチャの個人による登録機能は必要だと思いますか	不必要: 1 必要: 7
Q4: アプリケーション全体としての使いやすさはどうでしたか	非常に使いづらい: 1 非常に使いやすい: 7
Q5: 今回の実験中に認識率の高いジェスチャを見つけることができましたか	全く発見していない: 1 多く発見した: 7
Q6: あなたがジェスチャを用いた操作を許容できるシーンをすべて選択してください	・パソコンのアプリケーションの操作 ・屋内のさまざまな家電や機器の操作 ・公共空間におけるパブリックディスプレイやショーウィンドウのようなものの操作
Q7: 自分にとって直感的だと思うジェスチャを見つけられたタスク（複数選択可）	
Q8: ジェスチャで操作することが不適切だと感じるタスク（複数選択可）	
Q9: 自分が最後に割り当てたジェスチャを思い出せないタスク（複数選択可）	
Q10: タスクからジェスチャをイメージすることが容易だったタスク（複数選択可）	
Q11: ジェスチャ認識によって操作するアプリケーションを使用しての感想（自由回答）	
Q12: ジェスチャの応用分野に関するアイデア（自由回答）	

3.3 結果

被験者には実験時に、各コマンドに対して定義したジェスチャを手書きで記録するよう依頼した。図7～図10は定義されたジェスチャの概形を示しており、図中のアルファベットは各被験者を表す。これらは4日間のうち最後に定義したジェスチャに関して記載しており、変更の過程を示したほうが良いと判断したものについてはそれも併せて記載した。なお定義されたジェスチャの概形は共通した特徴が見出された4つのタスクの結果のみを示すこととし、残りのタスクに関しては省略する。

表4～表6に実験後行ったアンケートの結果を示す。表4は質問に対して7段階で回答してもらった結果であり、表中の数字は回答数を示している。表6は各質問においてそのタスクが選択された数を示している。

この他にアプリケーションの使用状況に関するログを収集したが、ここでは省略する。

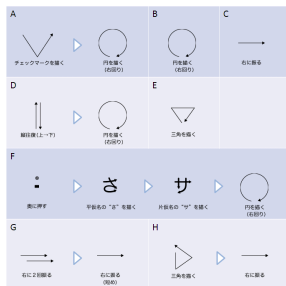


図7 再生に定義されたジェスチャ(動画再生)

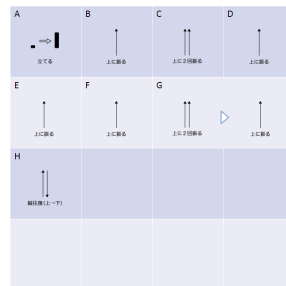


図8 音量上げに定義されたジェスチャ(動画再生)

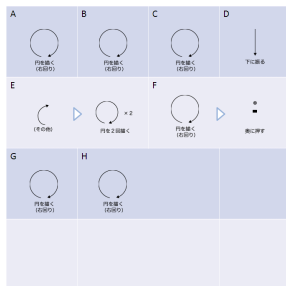


図9 回転に定義されたジェスチャ(画像閲覧)

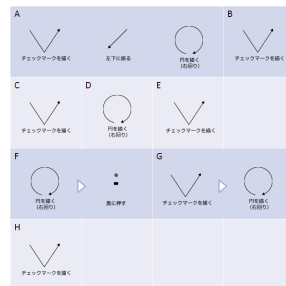


図10 確定に定義されたジェスチャ(プリミティブ)

表4 実験後アンケート結果 Q1～Q5

質問事項	回答	回答数 (人)							
Q1: ジェスチャの記憶	覚えられない	1	0	2	0	2	2	1	覚えられた
Q2: 直感性	直感的でない	1	0	1	3	0	3	0	直感的である
Q3: ジェスチャ登録の要/不要	不必要	2	0	0	3	2	0	1	必要
Q4: 使いやすさ	使いづらい	0	2	3	1	0	2	0	使いやすい
Q5: 認識率の高いジェスチャ	発見していない	0	0	0	3	3	2	0	多く発見した

表5 実験後アンケート結果 Q6

質問事項	回答	許容できると答えた人数 (人)
Q6: 許容できるシーン	パソコンのアプリケーションの操作	2
	屋内のさまざまな家電や機器の操作	6
	公共空間における操作	7

表6 実験後アンケート結果 Q7～Q10 (単位: 人)

	PLAY	STOP	PAUSE	FF	REW	VOLUP	VOLDOWN	FULL	DOWNLOAD	ENLARGE	REDUCE	ROTATE	PRINT	MAIL	OK	UP	DOWN	LEFT	RIGHT
Q7	4	1	2	4	4	4	4	1	1	3	3	3	0	0	6	8	8	8	8
Q8	1	1	2	2	2	3	3	1	3	1	1	2	5	4	0	0	0	0	0
Q9	0	1	0	2	2	1	1	2	4	2	2	0	4	4	0	0	0	0	0
Q10	2	2	0	7	7	5	5	3	1	3	3	7	0	0	6	8	8	8	8

4. 考察及び将来課題

4.1 各タスクにおける共通ジェスチャ

まず本稿の中心的な関心事である共通ジェスチャの抽出についての考察を述べる。実験の結果から共通性を見出せるタスクと見出せないタスクが存在することがわかった。それぞれ次のような結果となった。

- 共通性を見出せるタスク
 PLAY, FF, REW, VOLUP, VOLDOWN, ENLARGE, REDUCE, ROTATE, OK, UP, DOWN, LEFT, RIGHT
- 共通性を見出せないタスク
 STOP, PAUSE, FULL, DOWNLOAD, PRINT, MAIL

以下各タスクについてアンケート結果も加味しながら分析する。

PLAY 動画の再生については2種類の共通ジェスチャを見出すことができる。半数の被験者が円を描くジェスチャ、もう半数が右方向へ振るジェスチャで共通している。アンケートを見てみると、再生というタスクからジェスチャをイメージすることが容易であると感じた被験者は25%と少数である。それにもかかわらず、直感的だと思えるジェスチャを発見できたという回答が比較的高ポイントであり、また定義したジェスチャを思い出せないという回答にいたっては1つもなかったという点から、メンタルモデルの形成は行われたと考えられる。この結果に関して、アプリケーション使用状況のログデータから再生というタスクは動画再生に不可欠なタスクであるために使用頻度が高いことが分かった。したがって何度もジェスチャを繰り返すことによってメンタルモデルが形成されたと言える。さらに使用頻度の高いコマンドに対しては、なるべく単純で短いジェスチャを定義したいという心理が働くために、共通のジェスチャが定義されたと考えられる。以上より、タスクに対してジェスチャを定義する際はメタファなどを用いた対応付けによる学習性の他に、タスクの使用頻度による学習性も考慮する必要がある。

STOP/PAUSE 動画の停止、一時停止について、今回の実験では全く共通性を見出せなかった。しかしジェスチャでの操作が不適切だと感じた被験者は少なく、定義したジェスチャを思い出せないという回答も少ない。ログデータを見ても使用頻度の高いコマンドではないため、開発者がジェスチャのプリセットを決定することが許容されるタスクである。

FF/REW 動画の早送り、巻き戻しについては、半数以上がそれぞれ横方向に2回振るというジェスチャで共通した。すなわち、早送りと巻き戻しのタスクに関しては共通の対応付けが行われており、恐らくそれはAV機器におけるアイコンが三角形を2つ重ねたものであることのメタファによると考えられる。また実験後のアンケート結果を見ると、動画再生のタスク中、被験者が最もジェスチャをイメージしやすいタスクであった。

VOLUP/VOLDOWN 音量の操作についても、共通ジェスチャを見出すことができた。音量を上げるのは上方向に1回振るジェスチャ、音量を下げるのは下方向に1回振るジェスチャである。しかしアンケートではジェスチャを用いた音量の操作が適切でないという回答が他のタスクよりも多い。実験用アプリケーションでは、音量を上げる/下げるのコマンドを行った場合、1回の実行につき25%ずつ音量を変化させたのだが、この方式に不満を持つという被験者が複数おり、このことがアンケートの結果に影響している。音量のように連続的な量を持つものは、ジェスチャの移動量や傾き加減によって変化量を調整できるべきであり、そのほうがより直感的であると考えられる。

FULL 動画のフル画面表示については、被験者のうち2人が四角を描くジェスチャ、3人

が円を描くジェスチャ、2人が右斜め上方に振るというジェスチャであったが、どれも共通ジェスチャとして提案するには確証に乏しい。またアンケート結果からもこれらのジェスチャを直感的だと感じている被験者は少ないことが分かる。

DOWNLOAD 動画のダウンロードについては被験者によって全く異なるジェスチャを登録するという結果が出た。アンケート結果を見ると、直感的か、ジェスチャでの操作が適切か、定義したジェスチャを思い出せるか、タスクからジェスチャをイメージすることが容易かという4つの質問すべてにおいて、ネガティブな回答が多い。これらを踏まえると、このタスクに関して今回の実験ではジェスチャを用いた操作のメリットを享受することができなかったと言える。しかし動画をダウンロードするというタスクが性質的にジェスチャ操作に適していないと結論付けるのは尚早である。これについては後述する画像の印刷機能およびメール機能に同じであり、4.2節で詳細に述べる。

ENLARGE/REDUCE 画像の拡大および縮小について、共通ジェスチャとして断言できるものは見つからなかったが、概ね拡大は上方向を指し、縮小は下方向を指すという共通要素の傾向を見出すことができた。また画像の拡大と縮小に関しては、先述の動画の音量操作における問題と同じことが言える。すなわち、拡大率や縮小率はジェスチャの変化量に基づいた量で調整できるべきだと考えられる。

ROTATE 画像の回転について、半数以上の被験者が最終的に円を描くジェスチャを登録した。捻るジェスチャや傾けるジェスチャも用意に想起することができると推測されるが、今回の実験ではそのようなジェスチャを定義した被験者はいなかった。人間工学的な観点から分析すると、捻るジェスチャや傾けるジェスチャが手首の関節を用いた運動であるのに対して、円を描くジェスチャは肘の関節を用いる。人間の身体構造は、体幹から指先の方に向かうにつれて、筋肉の太さ(横断面積)が小さくなるという性質を持ち、筋肉の疲労のしにくさは筋肉の横断面積の大きさに比例する。したがって後者のジェスチャのほうが前者よりも疲れにくいということが言え、被験者にとってより自然だと感じられたのだと考えられる。

PRINT/MAIL 画像の印刷およびメールでの送信機能については、共通のジェスチャが見られた。印刷に関しては“Print”の頭文字である“P”を描くジェスチャ、メールに関しては“Mail”の頭文字である“M”を描くジェスチャである。しかしながら、アンケートの結果を見ると、タスクからジェスチャをイメージすることが容易だったと回答した被験者、および定義したジェスチャを直感的であると感じた被験者は皆無である。またこれら2つのタスクに関してはジェスチャを用いた操作が不適切であるという回答が圧倒的に

多い。これらを踏まえると、ジェスチャが共通したのは、定義すべきジェスチャが想起されず、結果的にジェスチャに変換できる要素が単語のスペルしかなかったというネガティブな要因によるものだと考えられる。したがって先述の動画のダウンロードと同様に、画像の印刷およびメールでの送信というタスクについても今回の実装方法ではジェスチャ操作のメリットを表現することができなかったと言える。これについては4.2節で詳細に述べる。

OK アプリケーションに依存しない確定を行うタスクについては、2種類の共通ジェスチャが見出せた。1つは円を描くジェスチャで、もう1つはチェックマーク（V字）を描くジェスチャである。円を描くジェスチャは日本において正解を意味する丸のメタファ、チェックマークを描いたジェスチャは国際的に“ok”を意味するチェックのメタファであると考えられる。円からチェックマークに定義するジェスチャを変更した被験者や、逆にチェックマークから円に変更した被験者もいたため、両ジェスチャのどちらを選択するかは本人の好みや認識率に依存する。

UP/DOWN/LEFT/RIGHT アプリケーションに依存しない上下左右への移動や選択を行うタスクについては、被験者全員が各々の向きに振るジェスチャを定義した。アンケート結果を見ると被験者全員が直感的だと感じていて、イメージも容易だったと答えている。

4.2 アンケートから得られた知見

次にアンケートの結果から得られた事項に関して考察を行う。

認識率の高いジェスチャ Q5の回答の中から被験者が発見した認識率の高いジェスチャの特徴を抜粋すると以下ようになる。

- (1) 単純である
- (2) ジェスチャを行う時間が短い
- (3) 排他的である（他のコマンドに定義したジェスチャと競合しない）

またアプリケーションの実装に関しては、認識率のフィードバックを返し、ユーザがジェスチャを見直す基準を与えることによって認識失敗時のフラストレーションを回避できると考えられる。アンケート結果の中に、ジェスチャを定義した直後は認識率がいいが、1日経過後に再度同じジェスチャをやっても認識率が悪いとの回答があった。これは時間が経過するほどユーザが定義したジェスチャを正確に再現できなくなることに起因すると考えられる。このようなヒューマンファクタによるエラーをどのように扱うべきかという事項も今後の課題となる。

ジェスチャの学習 ジェスチャ認識システムの作成において、デザイナーや開発者はコマンドに対してあらかじめプリセットを定義しておくか、もしくは実験に用いたアプリケーションのようにジェスチャを自由に定義できる機能を持たせるかを選択することになる。Q3において自由に定義できる機能の必要性を尋ねたところ、最初から基本となるプリセットがシステムによって提示され、そのうえで補助的な機能として自由に定義できると良いとの回答が複数あった。この原因として、すべてのコマンドに異なるジェスチャを定義することの手間が挙げられる。したがって最初に提示するプリセットのジェスチャを複数用意して選択できるようにしておき、それでも違和感のある場合には自分で定義できるようなシステムにすることでさらにユーザビリティの向上が見込める。

また今回の実験ではジェスチャを定義するコマンドの数が多く、ジェスチャを覚えることに対する負荷は高いものであると予想していた。しかしQ1とQ9における回答から、ユーザはそれほど負担を感じずにジェスチャを記憶できることが分かった。ただし使用頻度の低いジェスチャや再登録したばかりのジェスチャは覚えられないことが多いという意見もある。したがってジェスチャを用いたタスク実行のメンタルモデル形成のためには、ある程度反復してそのジェスチャを使用することが有効であり、このことは動画を再生するタスク（PLAY）の結果からも言える。また定義したジェスチャを確認できるようなフィードバックを返すことで、たとえジェスチャを覚えていなくてもタスクを達成できる仕組みも必要である。

ジェスチャを用いた操作対象 本稿ではパーソナルコンピュータ上のデスクトップアプリケーションをジェスチャによる操作の対象として検証を行った。しかしQ2やQ6の回答を見ると、本来マウスとキーボードによって使用されることを前提に作られたWIMPを採用したアプリケーションをジェスチャによって操作しても、被験者はそのメリットをあまり感じていなかったと言える。今回の実験では特に動画のダウンロード、画像の印刷、画像のメール送信の各タスクについて、ユーザがジェスチャを用いた操作は直感的ではないと感じている。これには2つ原因が考えられる。1つはダウンロード先ディレクトリの指定やメールアドレスの指定などの操作が必要であり、タスクをジェスチャのみで完結できなかったこと、そしてもう1つはコマンドによって実行するというデザインに問題があったということが挙げられる。この2点を考慮したインタラクション方法/インタフェースデザインであれば、直感的だと感じられなかったタスクに関してもジェスチャインタフェースが有効となりうる。例えば動画のダウンロードの場合、手で掴むというジェスチャによって動画ファイルを選択し、それをダウンロード用のスペースまで持っていつ

てから手を離すことによってダウンロードが実行されるといったデザインにすると、コマンドに対して定義したジェスチャによってタスクを行うよりも格段に直感的だと感じられる。また Q11 の回答から実験用アプリケーションについてはあまり直感的だと感じてもらえなかったものの、Q12 における被験者の回答にはジェスチャインタフェースに対して肯定的なものが多かった。例えば以下のような意見があった。

- 3次元が認識できるという特徴を生かして、Google Street View^{*1}のような3Dインタフェースの操作に応用するのが良い
- パブリックスペースでの一時的なインタラクションには適しているのではないか
- セキュリティ認証への応用や神経障害を持っていて細かい指先の動きができない人のためのインタフェースには使えそう

これらを踏まえると、キーボードやマウスで扱いつらい領域、もしくはジェスチャを用いた操作のほうが面白いと感じられるサービスであれば、十分にジェスチャインタフェースが普及する余地はあると言える。

4.3 将来課題

本稿における成果をより良いものにするためには、以下の課題を検討する必要がある。

(1) 他の認識方式/他のタスクによる検証

本稿では加速度センサによるジェスチャ認識を対象とし、パーソナルコンピュータ上のデスクトップアプリケーションを操作する際のジェスチャについて検討した。もしカメラによるジェスチャ認識などを対象とした場合、またデスクトップアプリケーションではないものを操作する場合、どのような結果となるのかを比較及び検証する必要がある。

(2) ジェスチャ認識を用いたインタラクションが生きるサービスおよびインタフェースデザインの創出

ジェスチャインタフェースのメリットを最大化できるサービスや適用領域を考案する必要がある。例えばアンケートによればパソコンのアプリケーションの操作よりも公共空間におけるパブリックディスプレイやショーウィンドウのようなものの操作のほうがジェスチャインタフェースが受け入れられるという結果が出ている。また今回の実験で直感的でないと感じられたタスクに関しても、適切なインタフェースデザインを構築することでよりジェスチャのメリットを享受できるものになると考えられるため、それらについても検討を行う必要がある。

5. おわりに

本稿では、ユーザがジェスチャ認識システムを使用する過程を観測することで、デスクトップアプリケーションにおける動画再生および画像閲覧のタスクについて、共通ジェスチャを抽出し、デザイナーや開発者にとっての判断材料となる要素を提案した。

まず、アンケートによる事前調査で、ユーザが想起できるジェスチャのおおよその種類と操作対象とすべきアプリケーションの知見を得た。この知見をもとに実験用アプリケーションを実装し、被験者に使用してもらうことでデータを収集した。データを分析した結果、各タスクにおいて、共通して定義されたジェスチャや、その特徴および性質を見出すことができた。また、実験終了後に行ったアンケートから、ジェスチャ認識システムを構築する際の手がかりとなる要素やデザインの方針を決めるのに有用な事項に関して考察を行った。

将来課題としては、他の認識方式による検証やジェスチャ認識を用いたインタラクションの効果を最大化できるようなサービスやインタフェースデザインの創出などが挙げられる。

参考文献

- 1) Dix, A., Finlay, J., Abowd, G.D. and Beale, R.: *HUMAN-COMPUTER INTERACTION*, Prentice Hall (2003).
- 2) Norman, D.A.: *誰のためのデザイン?*, 新曜社 (1990).
- 3) Prekopcsák, Z., Halácsy, P. and Gáspár-Papanek, C.: Design and development of an everyday hand gesture interface, *MobileHCI '08: Proceedings of the 10th international conference on Human computer interaction with mobile devices and services*, New York, NY, USA, ACM, pp.479-480 (2008).
- 4) Saffer, D.: *Designing Gestural Interfaces*, O'REILLY (2008).
- 5) Schlömer, T., Poppinga, B., Henze, N. and Boll, S.: Gesture recognition with a Wii controller, *TEI '08: Proceedings of the 2nd international conference on Tangible and embedded interaction*, New York, NY, USA, ACM, pp.11-14 (2008).
- 6) Ward, J.A., Lukowicz, P. and Tröster, G.: Gesture spotting using wrist worn microphone and 3-axis accelerometer, *sOc-EUSAI '05: Proceedings of the 2005 joint conference on Smart objects and ambient intelligence*, New York, NY, USA, ACM, pp.99-104 (2005).
- 7) 平川康史: Wizard-of-Oz 法に基づいたジェスチャ認識システムの研究 (2004).

*1 Google street view <http://maps.google.com/help/maps/streetview/>