

〔非公開〕

TR-C-0049

仮想空間における手振り理解に関する研究

大西 剛
TAKESHI OONISHI

竹村 治雄
HARUO TAKEMURA

1990. 5. 21

A T R 通信システム研究所

仮想空間における手振り理解に関する研究

大西 剛 竹村 治雄
ATR 通信システム研究所

4月27日, 1990

概要

これまでのユーザインタフェースよりも、よりユーザフレンドリなインタフェースを構築する一つの手段として、ジェスチャを用いる研究が始まっている。これらの研究についての現状を概説した上で、ジェスチャをマンマシンインタフェースとして利用した場合の利点を挙げ、システムとして実現するに於ける具体的な問題点をあげる。最後にこれらをふまえて今後の研究の方向を述べる。

1 まえがき

我々は日常生活の中で様々なことを表現するためにジェスチャを使っている。指さしてものを指示することから、手話で用いられるような高度に抽象化されたものまでその幅は広い。しかし概してジェスチャは、直観的なものであると言えよう。

ジェスチャをマンマシンインタフェースとして利用することを目的とした研究は、これまであまり行われてきていない。その理由として、ジェスチャを入力できるデバイスがなかったこと、ジェスチャの利用の有用性が不明であること、等が挙げられる。ジェスチャを利用する利点をして、直観的・直接的であること、これに関してうまく利用すれば、ユーザフレンドリなインタフェースとできること、パラメータが多いことなどが考えられる。逆に短所として、コンテキストに依存する度合いが強いこと、認識方法が確立していないこと、などがあると思われる。

コンピュータグラフィックスをはじめとするコンピュータ利用技術の発展、これに並行したコンピュータ利用の一般化は、使い易いマンマシンインタフェースを実現するための手段であり、動機である。臨場感を伴った通信や遠隔操作などの仮想空間操作と言った新しい環境中では、ジェスチャの利用価値も高いものと思われる。仮想的な空間を構築し、その中でジェスチャに何らかの意味を持たせ、利用するのである。このときジェスチャの意味の持たせ方は、様々な方法があり得るが、重要なことはユーザにとって自然な定義かどうかである。

本研究では、注目する点をジェスチャのうち右手に関するものに限定し、マンマシンインタフェースとしての有用性を実験により示す。この実験と相互に関連を持ちながら、手振りを利用する環境を仮想空間上に構築するとともに、この環境中での手振り認識手法を確立することを目標とする。本報告では以下、第2節でこれまでになされたジェスチャを利用したマンマシンインタフェースの研究とその問題点を述べる。第3節では、これらの現状を踏まえて本研究の今後の進め方の概要を述べ、第4節では当面の研究を具体的に示す。

2 手振り理解の研究の現状

2.1 研究動向

手の位置・形状およびそれらの変化を何らかの方法によって認識し、マンマシンインタフェースとして利用する試みは、未だ緒についたばかりと言える。以下にマンマシンインタフェースとして手振りを実際に用いた試みを以下の表に示す。

	使用デバイス	特徴
文献 [1]	VPL DataGlove	手をマウスのように使用 (ボタン、メニュー選択等)
文献 [2]	Sensor Flame	ノブを回す、スイッチを動かす動作 — 方向を考慮
文献 [3]	VPL DataGlove	対応物体の移動・回転・拡大縮小と 3D モデリングにおける直接操作
文献 [9]	タッチパッド	パッド表面での指の操作。ハードウェアから認識手法まで 総合的に議論

表 1. 実際に手振りを用いたシステムの例

[1]では、仮想的画像環境に対するインタフェースとしての手によるジェスチャを、手と指におけるそれぞれの動きの有無によって分類している。その分類にしたがって、実際のインプリメント例を示している。[2]はオブジェクト指向によるインプリメント手法を中心に述べている。手と他のオブジェクトとの位置に加えて、動きの方向を考慮している点が注目される。[3]では、3Dモデリングにおける直接操作を行い、手振りによる入力の有用性を示している。これら[1][2][3]に共通して、非常に限られた手振りのみが対象とされていること、および動的な変化に充分対応されていないことが挙げられる。[9]では、タッチパッド上での指の動きを認識する入力装置を作成し、総合的にそのシステムを検討している。ただし、入力される装置として2次元のタッチパネルが使用され、ジェスチャの特徴を十分に生かしているとは言い難い。この方式ではパネルへの接触によって、ジェスチャのセグメンテーションが容易にできるが、3次元で行なうときには難しい問題となる。ジェスチャをマンマシンインタフェースとして用いることが何らかの点で有効であると、直観的には理解できるが、いざ用いるとなるとどのように行うのか模索の状態にあると言うのが現状であろう。

2.2 問題点

問題となる点は、大別すると2つある。第一は、ジェスチャをマンマシンインタフェースとして用いるのに有効な環境の構築である。第二に、定められた環境の中でのジェスチャの認識に関するものである。

2.2.1 ジェスチャ使用環境の構築

我々は日常生活において様々なジェスチャを使って、意志の疎通を行っている。同一の体の動きであっても複数の意味を持っている場合があるが、その中から我々が正しく意味を理解できるのは、そのジェスチャを使用した状況から判断できるからである。マンマシンインタフェースとしてジェスチャを考えると、我々はこの状況を設定する必要がある。たとえば、物理的3次元空間を仮定するならば、仮想空間に物理的性質を採り入れることが必要となる。この方向で特に手への力学的フィードバックを重視した研究に文献[5]がある。この設定には複数の実現法が考えられるが、重要なことはユーザの観点からの検討である。文献[4]は、その意味で示唆的である。

2.2.2 ジェスチャの認識

ある環境で意味のあるジェスチャが決定されれば、ジェスチャの認識はパターン認識の一つと考えられる。すなわち目的は、手の動きからの定められたパターンの抽出である。手からの情報を一般的に扱うために、

- 手の位置・形状の表現法
- 上記の時間的変化の表現法

を定める必要がある。文献[6]では、「仮想骨格モデル」を定め、手の形状の表現を行っている。また文献[7]では、指関節の曲げ角度と手の向きからコード化を行っている。これらの研究では、ともに静的なモデルのみを扱っている。動的なモデルは今後の課題となっている。

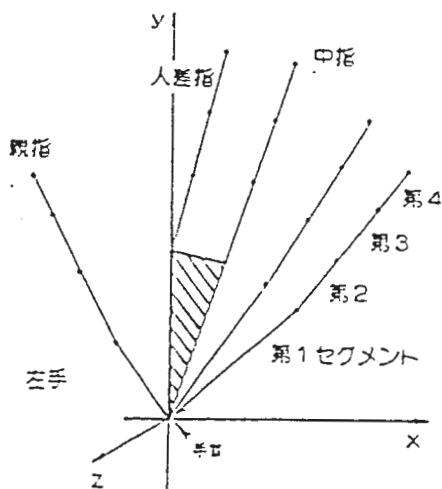


図1. 手の仮想骨格モデル (文献[6]より)

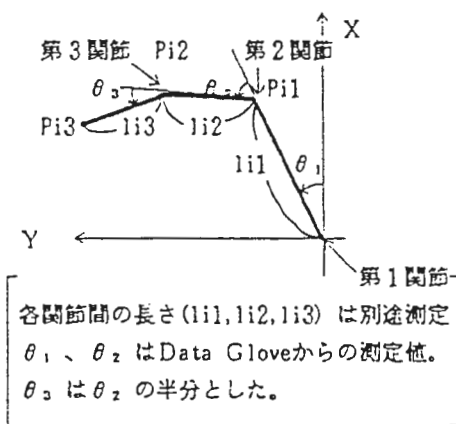


図2. 指のモデル (文献[7]より)

手の動きの表現法が確立されたとして次の課題は、認識方式の確立である。この段階は現在のところ全く新しい分野と言える。具体的に問題点を挙げれば、少なくとも次の3点が考えられる。

- セグメンテーションの問題 — 一つの時間的広がりを持ったジェスチャがどこから始まりどこで終わるかをどのようにして決定するか。
- 特徴抽出の問題 — ノイズ、歪み、同一なジェスチャを意図した様々な形態などにどう対応するか。
- 識別の問題 — 感覚的・視覚的に近いと思われる手指の形態が近く、遠いと思われる形態が遠いと識別できる方法をどう構築するか。また時間的な同時性 — 例えば、回転と移動を意図するジェスチャがある場合、回転しながら移動するジェスチャを考えるのは自然であろう — をどのように考慮するか。

これらは相互に関連し、それぞれ多くの障壁を含んでいる。その上、ユーザインタフェースとして用いることを目標とするためには、リアルタイム処理を考慮することが必要である。

3 研究計画

前節で述べたように、ジェスチャを用いたマンマシンインタフェースを実現するにあたって、環境の構築と認識という2つの研究課題がある。それぞれについての研究計画を以下に示す。

環境の構築に関する研究では、マンマシンインタフェースとしての使用に適したジェスチャをあげることから始める。その点を考慮して、この研究で扱うジェスチャは、仮想空間内にあるオブジェクトの操作に限定する。この環境においてジェスチャとして取り扱うべきモデルを構築するために実験を行う。

認識に関する研究は、構築した環境に依存する部分があると考えられるが、当面は環境としては初歩的なモデルを採用し進めて行くものとする。この研究の基本的順序として

- 手の位置・形状の表現法の確立
- 上記に時間的要素を加えた表現法の確立
- 以上2つの項目に対して、使用ハードウェアとの整合
- 手の動きからの手振りの認識法（特徴抽出法など）の確立
- 構築したモデルとの融合

を目安として行く予定である。

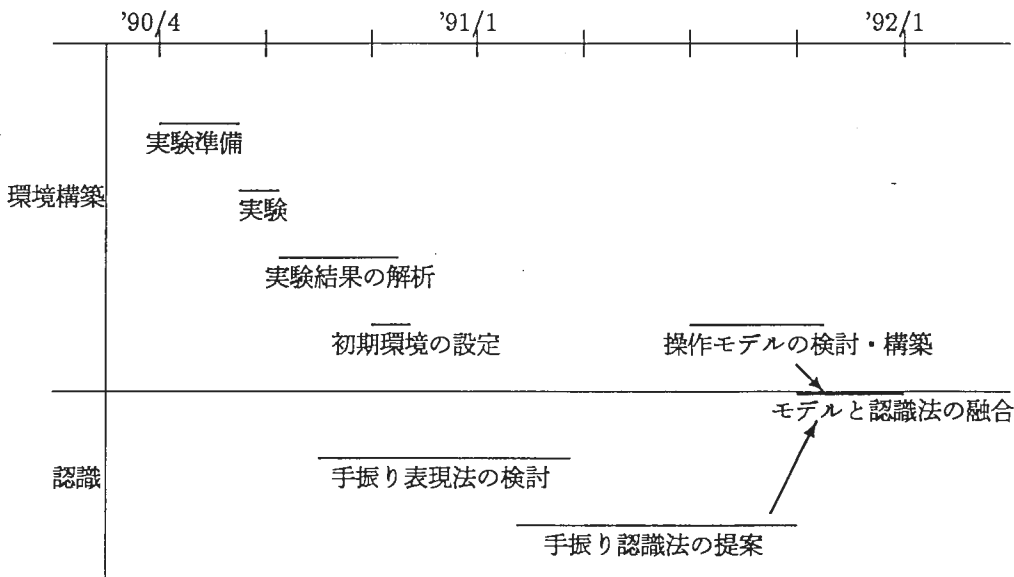


表 2. 研究計画スケジュール

4 当面の研究課題

4.1 手振り利用環境の構築の研究

手振りをマンマシンのインプットに利用する準備として以下の実験を行なう。詳細は付録として最後に記すが、その目的と基本的な手法をここに示す。

実験の目的は、自然なインタフェースとしてシステムにとり入れるべき手振りを検討することと、手振りを定量的に分類するのにどんなデータが必要であるかを調べることである。実験は大きく分けて二つの段階に分かれる。その環境は、ほぼ共通している。被験者は、IRIS ワークステーションのディスプレイ上に描かれた画像オブジェクトを指示によって操作しようと試みる。指示には、オブジェクトの回転・移動・拡大縮小がある。この時、被験者は右手にデータグローブと3次元磁気センサを装着し、また身振りは複数のビデオカメラによって記録される。実験者はビデオカメラからの画像を見て被験者の身振り判断し、被験者の見ているディスプレイ上のオブジェクトを操作し、被験者に視覚的フィードバックを与える。データグローブ・3次元磁気センサからのデータは毎秒60回または30回サンプリングされる。

第1の実験では、手振りを限定せずに被験者の思いのままに動作を行う。この実験によって、ある指示に関して被験者に共通した動作があるかどうか分かる。また、手振りを認識するのに必要な手の形状・運動のデータを知るのに役立つと期待できる。

第2の実験では、被験者に意味のある手振りをあらかじめ知らせることによって、手振りを限定する。被験者が同じ手振りであると考えたものであっても、ばらつきが出ると予想される。

4.2 手振り認識の研究

4.2.1 手振りの記述法

身振りは空間的・時間的広がりを持ったパターンであると考えられる。よってその記述は、これら両方の要素を含んだものでなければならない。

静的な手指の形状を記述するには、その時の手の位置・掌の向き・指の形状が必要である。その中で特に問題となるのは、指の形状を記述する方法であろう。考えられる方法として、

1. 指の各関節の角度
2. 各指の先端の手首の一点からの距離
3. 各指の先端の座標
4. 手の形態のコード化

等がある。1の方法は、各人の手の大きさに関係ないという一般性を持っている。ただし、指の曲げ角が異なっても全体としての形状に意味がある場合—例えば、親指と人指し指で円を作る場合—の判別が難しい。また、指相互の接触や交差といった関係の判別も難しいと思われる。指の第2、第3関節の自由度は1つであるが、第1関節は2自由度あり、完全に記述するには複雑になる。2、3の方法では手の大きさに依存することから、一般化が難しい。4の意味での形状記述は、1から3のデータを何らかの手法によって処理した一つの結果と言える。

時間的要素を導入する方法として、認識装置へのデータの与え方の観点から図1のように、(1) 時間とともに変化する入力を与える方法と、静的な手指の形状を時間軸に沿って記述したデータを与える方法、すなわち時間を空間的に表現する方法がある。後者には (2-1) 処理すべきデータをすべて同時に提示する方式と、(2-2) ある時間的な範囲のみを提示する方式がある。(1) では、認識装置内に一時的にデータを記憶する機能を組み込み時間的関係を処理する必要がある。(2-1)の方法で「データをすべて同時に提示する」とは、ある手振りがすべて終わった後でその手振りの認識を行なうことを意味する。最終的な目標はインタラクティブなインタフェースを構築することであるから、(2-1)の方法は問題がある。

手振りを記述する方法としては、文献 [8] で示されるように時間軸を明示せず速度を含む動作の系列として記述する方法がある。この場合、記述法としては有効であるが、認識のための手段とするには適当でないようである。

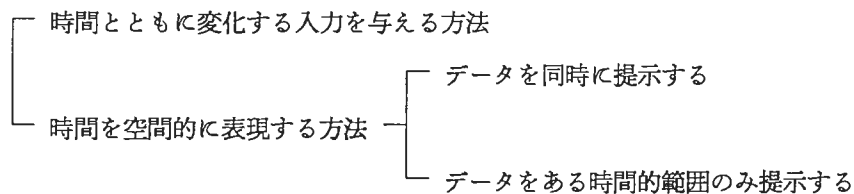


図3. 時間的広がり表現する2つの方法

身振りを記述する方法としては、文献 [8] で示されるように時間軸を明示せず速度を含む動作の系列として記述する方法がある。この場合、記述法としては有効であるが、認識のための手段とするには適当でないようである。

4.3 手振りの認識方式

認識方式に関する問題点として、少なくともセグメンテーション・特徴抽出・識別の3つの問題があるとすでに述べた。しかし現時点ではそれぞれについて具体的に論じる段階にない。

5 むすび

本報告では、ジェスチャをマンマシンインタフェースとして使用する上での問題点を中心に述べた。問題点は大別して、環境構築の問題と認識手法の問題の二つである。前者の問題については、ある程度環境を限定しなければならないこと、その環境がユーザにとって自然であることを重要と考えていることを示し、実験により右手の手振りのみを使った環境を決定することを述べた。後者については、手振りをなるべく普遍的なデータによって記述するために、手の動きを指の関節の曲げ角・手の位置・掌の向きを時間的広がりを持ったデータとして扱うことを述べた。しかし、その具体的な認識手法については未だ流動的であり、研究を進めるに当たっての重要な課題となっている。

参考文献

- [1] D.J.Sturman, D.Zeltzer, and S.Pieper: "Hands-on Interaction With Virtual Environment", Proceedigns of the ACM SIGGRAOH/SIGCHI Symposium on User Interface Software and Technology, 1989, pp.19-24
- [2] R.B.Dannenberg, D.Amon: "A Gesture Based User Interface Prototyping System", Proceedings of the ACM SIGGRAPH/SIGCHI Symposium on User Interface Software and Technology, 1989, pp.127-132
- [3] D.Weiner, S.K.Ganapathy: "A Synthetic Visual Environment With Hand Gesturing And Voice Input", CHI'89 Proceedings
- [4] D.Weimer, S.K.Ganapathy: "Speech and Gestures for Graphic Image Manipulation", CHI '89 Proceedings
- [5] 黒川、紙谷: "コンピュータ身振り辞書の試作", Human Interface N&R, 1989 Vol4 pp.87-92
- [6] 岩田洋夫: "力感覚に対応した人工現実感 — 仮想空間の構築", Proceedings of the 5th Symposium on Human Interface, 1989, pp.1-4
- [7] 高橋、小林: "指文字を例にした手振り認識実験", Human Interface N&R, 1989 Vol4 pp.151-156
- [8] G.Marino, P.Morasso, E.Troianom and R.Zaccaria: "NEM: A language for the representation of motor knowledge", Human Movement Understanding, North Holland, 1986, pp.127-155
- [9] R.L.Shackelford: "Gesture-Based Communication with Computer: Feasibility Study", PB Rep NO. PB-88-235908 pp.53, 1985

付録 手振りによるディスプレイ上オブジェクトの操作実験計画

目的

自然なユーザインタフェースとしてシステムに採り入れるべき見振りを検討する。ある操作を意図した身振りを定量的に分類するのにどんなデータが必要かを調べる。

実験環境

1. この実験の環境には、指示者・実験者・被験者の3者が含まれる。
2. 被験者は、データグローブを右手に装着する。
3. 右手の位置計測用に磁気センサを装着する。
4. データグローブ・磁気センサはIRISに接続し、一定時間ごとの記録をとる。
5. 被験者観察用のカメラを3箇所用意する。
6. カメラの画像はVTRで記録する。
7. 被験者はIRISディスプレイの前に座り、指示者の指示によって、ディスプレイ上のオブジェクトの身振りでの操作を試みる。オブジェクトに以下の制限を加える。
 - オブジェクトは常に1つとする。
 - 異なる大きさのオブジェクトを用意する。
 - オブジェクトの形は数種類とする。ただし、一操作中は変化しない。
 - オブジェクト・背景の色はともに一色とする。
 - オブジェクトは2次元表示または3次元表示で行なう。ただし立体視は用いない。
8. 実験者は被験者の身振りをカメラの画像から判断し、ディスプレイ上のオブジェクトを操作する。このとき実験者は、指示者の行なった指示を知らない状態で操作する。これは、実験者が指示を知ることによって被験者の身振りを予想することを、排除するためである。

実験手順

被験者にディスプレイ上のオブジェクトの操作を音声で指示する。または、ディスプレイ上に目標となる画像を表示する。指示する基本操作は、

1. オブジェクトの回転（回転の中心はオブジェクトの中心とし、これを表示する。）
2. オブジェクトの移動
3. オブジェクトの拡大・縮小（相似の中心は、オブジェクトの中心）

とする。指示する操作は、基本操作とそれらの組合せとする。これらの操作を以下の条件で実験を行なう。ただし、(a)は必ず(b)の前に行なう。

- (a) 身振りを指定しない場合（ただし、右手のみで身振りを行なうことは伝えておく）— 実験者が、被験者が行なった身振りをカメラからの画像を見て意図を理解できれば、オブジェクトを操作する。
- (b) 身振りを指定する場合 — 実験を行なう前に、被験者にそれぞれの指定された身振りを絵あるいはVTRで見せておく。実験者は、指示した身振りを被験者が行なったと認めた時のみ、オブジェクトを操作する。

また、指示者は操作の終了を被験者・実験者に伝える。(a)、(b)それぞれについて同一被験者に対し、オブジェクトの条件・操作をかえて複数回繰り返す。

予備実験として2、3人の被験者を対象として実験し、その後10人程度の被験者で本実験を行なう。

分析方法

(a) 身振りを指示しなかった場合

- どのような身振りが実際に行なわれたかを分類する。
- オブジェクトの大きさや形状に依存するかどうかを調べる。
- 異なる指示に対して、個人の身振りにおける好み・傾向があるかを調べる。
- 個々の指示に対して、複数の被験者に共通な身振りがあるかどうかを調べる。

(b) 身振りを指示しておいた場合

- 被験者にアンケートを行ない、決められていた身振りの印象を検討する。
- 被験者の身振りを行なり速度・指の動きの関係等を解析する。
- パラメータの変化に、上記結果が依存するかどうかを調べる。

上記の点に注目しながら、使われる指の頻度・指の各関節の曲げ角と相互の関係・手指の動作速度などの計測値について、統計的手法を用いて解析を行なり。またビデオ画像の指示に対して意味のあると考えられる身振りを分類・解析する。

予想結果

身振りを指示しなかった場合、それぞれの操作においていくつかの身振りに分類されると思われる。オブジェクトの大きさによって、同じ操作であっても違う身振りが行なわれることが考えられる。身振りをあらかじめ指示しておいた場合、同じ身振りであっても指の動かし方・動作の速さなどに個人差が現れると予想される。

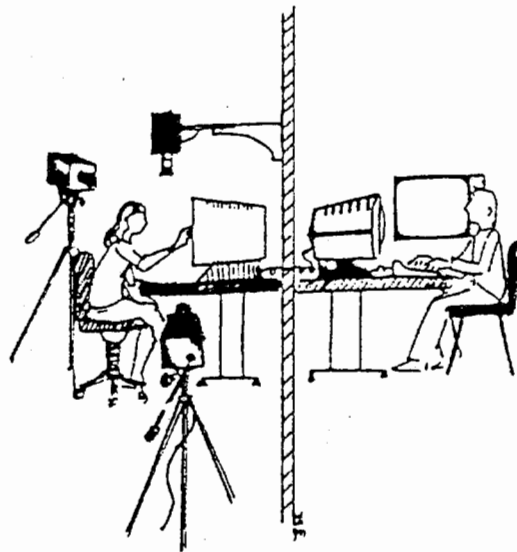


図 4. 実験風景 (文献 [4] より)

我々の実験も、基本的にはこの図のような環境を予定している。ただし、被験者は右手にデータグローブと磁気センサをつけること、実験者と操作を指示する人が別なことがこの図と異なる。