

〔非公開〕

TR-C-0008

手の動き認識とこれを利用した
マンマシンインタフェース
の研究について

栄藤 稔 伴野 明 小林 幸雄

MINORU ETOH AKIRA TOMONO YUKIO KOBAYASHI

1988. 1. 25

A T R 通信システム研究所

手の動きの認識とこれを利用したマンマシンインタフェースの研究について

栄藤 稔 伴野 明 小林 幸雄
(ATR通信システム研究所)

あらまし

言語以外の情報伝達手段として、手の役割に注目して手をポインティングデバイスとしてどのように位置づけるか、また手の位置・形状をどのように認識していくかについて検討し研究の方向を示す。

研究のアプローチを手による指示特性の評価、画像処理による手の3次元認識に大きく分け、各課題について研究の現状を述べた上で、問題点を整理し、具体的な研究課題を抽出する。当面の研究課題として手の指示精度評価と画像からの手の3次元モデル変形の2つを挙げ、この研究計画を述べる。

1. まえがき

今日の情報化社会の高度化に伴い、我々をとりまく情報はマスメディアによる情報に代表される身近なものからデータベース検索にみられるような機械操作を必要とする個別的情報にまで広がってきている。ここにマンマシンインタフェースに関する課題として、人間の意志や要求をシステムにどのように伝達するか、またそれがどれだけ自然に行えるかという問題が生じる。従来、人とシステム間の情報伝達手段はキーボードやマウス、タッチパネル等が主流であった。従来の伝達手段を否定せずも、

- 1) 3次元画像データ操作等にみられる計算機が扱う対象の高度化
- 2) システム利用者の大衆化

の要求に対応する一つの方策としてマンマシンインタフェースに手を直接用いることを提案する。ここで、手による指示の概念をこれが有効と思われる3次元画像のレイアウト操作を例に(図1)述べる。

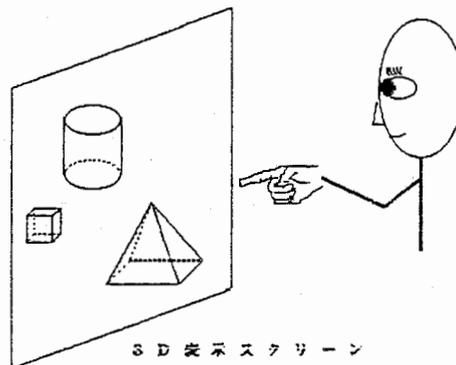


図1 3次元画像レイアウト

3次元表示技術の進歩は、その空間に居るが如くの視覚を我々に与えてくれる。このような空間では、表示されるオブジェクトをより高い現実感で操作する、すなわち実物のように見、触り、変更を加えることが必要となる。これを”直接操作”と仮に呼ぶとする(従来の2次元表示の操作環境でこれを可能としている代表例はタッチパネルである。)。本報告では、このような”直接操作”を実現する第1歩として、手による空間位置指示に

着目する。これは手による空間位置指示が人間の指示行為の中で基本的なものであり、以下の特徴を持つからである。

1) 指示の高い自由度

3次元空間中の位置、方向を指先、指の向き等で表現可能である。

2) 場知覚性の向上

手による直接指示によって表示に対する一体感が向上する。従って高臨場感のある操作環境が得られる。

ここで研究課題として2つを挙げる事ができる。第1は上記の特徴を持つ手をインタフェースに用いることにより操作性がどの程度向上するかということ、第2は以上の特徴を損なうことなく指示情報を抽出するには何が必要かということである。前者は手による直接指示の評価技術、後者は手の形状、動きの認識技術と要約できよう。第2節で現状を述べた後、これらの研究課題を段階別に整理し具体的課題を抽出する。また第3節で研究計画を第4節で当面の研究内容を紹介する。

2. 手の空間指示利用の現状と課題

2.1 手の空間指示利用の現状

手の位置、形状などを画像処理により認識し、これをマンマシンインタフェースに用いようとする試みは数例を除いて見つからない。表1に手をインタフェースに利用する試みおよび上肢の動作を認識しようとした試みを示す。

表1

	目的	認識対象
文献 [1]	手の指示による コンピュータへの情報入力	指先の検出 指先と目を結ぶ仮想指示棒
文献 [2]	手話の認識	手の形、手の軌跡、 手の顔に対す相対位置
文献 [3]	人間の動作解析 (正確な人体運動の再構成)	頭、胴体、上肢、下肢の動き 人体のスティックモデル、円筒モデル 特徴点の追跡
文献 [4]	手の身振り理解 (身振り認識)	手の形状 手の3次元モデル

文献[1]は唯一手による空間指示入力を画像により試みた例であるが認識手法は発見的であり、本文献では人差指の検出に必要な画像サイズを議論するにとどまり、マンマシンインタフェースへの応用展開には非力である。これは文献[2][3]の例においても同様で、顔写真の解析に関する過去の文献[14]に見られるように人体部位の特徴点を予見される特徴とのマッチングによって手を認識しようとしているが空間位置と形状は正確に得られていない。文献[4]では手を18個のボディーセグメントと17個の関節から構成される三次元モデルを用意し、観測された画像とこの3次元モデルとの照合により形状認識を行うと

している（詳しくは第4節で述べる）。形状を正確に得ようとする試みは上肢に関して他に例がない。ただし、顔画像に対していくつか為されており[18][19]、目的の多くはテレビ電話、会議等に顔画像を超低ビットレートで伝送しようとするもので、まだ基礎研究の段階である。

結果として言えることは手の空間位置、形状を把握するための試みは未だ不完全であり、観測の複眼化や3次元モデルの変形手法等多くの研究の余地が残されている。

2.2 具体的課題の抽出

第1節に述べたように研究課題として大きく評価と認識の2つに大別したが以下に述べる。

・評価（手を用いたマンマシンインタフェースの操作性評価）

手を用いたマンマシンインタフェースが我々の情報化社会に貢献するかどうかの問いかけに答える1つの方法は具体的な利用イメージを提示しその応用モデルを評価することである。MITのメディア・ラボの実験[20]に見られるように人間の視線や手の指示を磁気センサにより検出して画面制御に用いる試みは一つの応用モデルであろう。

近年になって表示技術の進歩により両眼立体視による画像表示が可能となり、建築物やインテリアのレイアウトなどの意匠設計や遠隔操作に利用しようとする試みが生まれてきている[20][21][22]。そこで、第1節に述べたようなインテリアレイアウトを3次元表示する応用モデルを設定し、手による物体の指示特性を評価することによって手による指示操作の有用性を明らかにする。ここで以上の評価に必要な基礎的課題として以下の2項が挙げられる。

- 1) 手の空間指示特性（空間知覚）の分析
- 2) 手の指示動作（意味行動）の分析

・認識（手の認識のための画像処理技術）

マンマシンインタフェースに手を用いるための画像の認識は以下のように性格づけられる。

- 1) 指示に用いる指の正確な3次元空間位置を検出
- 2) リアルタイム処理
- 3) 運動軌跡の追跡

以上の3つが課題となる。1) について、必要な精度は手の指示精度の評価によって決まるものであり、前記の評価における分析結果に依存している。

3. 研究計画

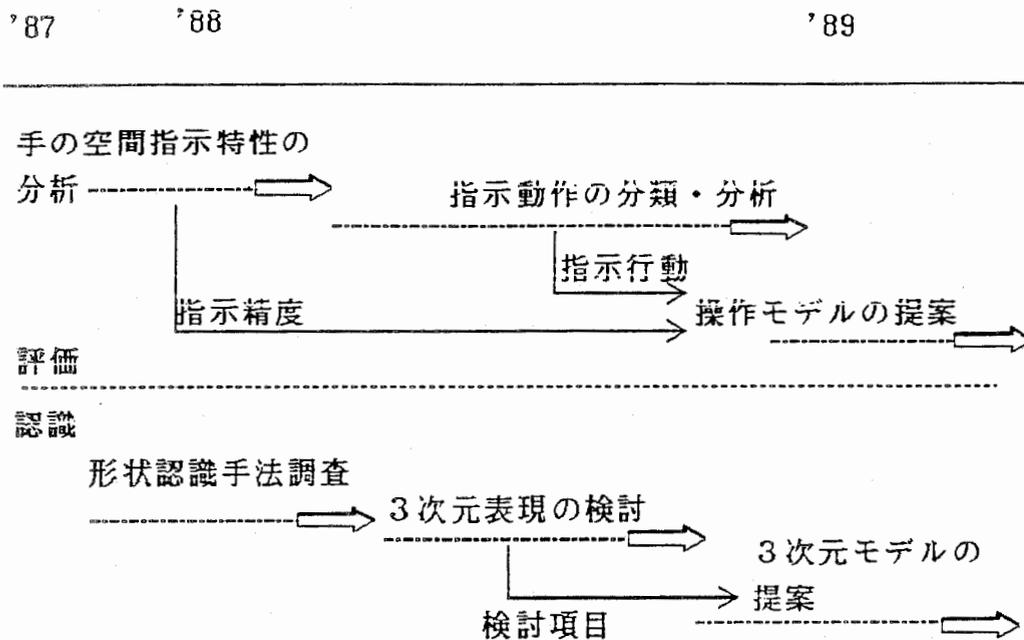
前節で明らかにした2つの研究課題についてそれぞれの研究計画を表2に示す。評価に関する研究のステップは先ず最初に人間の手のもつ指示特性を明らかにし次に人間の指示行動のパターンを類別する。以上の基礎データによって人間の手の操作モデルの提案を行うことができる。この操作モデルに対して操作性の評価を行い手によるマンマシンインタフェースの有用性を示す。

手の撮像による位置、形状、動きの測定・認識についての研究ステップはまず認識に必要な3次元モデルの提案を行い、手の形状認識手法の検討を考える。次のステップとして

- ・手の領域画像の切り出し、特徴点抽出アルゴリズム
- ・特徴点位置の計測アルゴリズム
- ・手の領域追跡、リアルタイム化

の各検討を行い、以上のアルゴリズムの実際のシステムへの応用化を目指す。

表2 研究計画線表



4. 当面の研究内容

4. 1 手の空間指示特性（空間知覚）の分析

手の指示入力の前準備として以下の予備実験を行なう。実験の主なねらいは、人間の手の何の特徴として観察すればその指示方向を得ることができるか、またどのような条件下でどの程度の精度が得られるかを知ることにある。

4. 1. 1 指示精度のパラメータ

ここで指示精度に影響を与えるであろう項目を以下に列挙する。

1)指示方向のフィードバック

カーソル表示等により指示方向を表示するかしないか

2)指示の表現方法

指示方向を (a)人差指の方向で表現するか あるいは
(b)眼と指先の延長線で表現するか

3)指示方向

眼あるいは手を基準とした位置と目標の相対位置によって精度に変化があるか

4)顔と手の相対位置

指示の表現方法が(a)の場合、顔と手の相対位置によって精度変化があるかどうか

5)指示方向の履歴

例えば右方向の対象を指示するときその前の指示方向が中央方向である場合と、左方向である場合では指示誤差が異なることが予想される。直前の指示方向により指示誤差に偏位が存在するかどうか

4. 1. 2 人間系のみによる指示精度実験

(指示方向をカーソル表示しない)

指先によるインタフェースを実現するためにはカーソル表示等のなんらかのフィードバックが必要と考えられるが、この方法を検討するためには、その前提として人間が本来持っている指示精度を明らかにする必要がある。そこでまず人間系のみによる指示精度実験を行う。

我々は日常生活において注視の度合に応じて指示の形態を変化させている。本実験ではこれを図2に示すような”体側”，”眼前”，”視線”の3形態に分類した。この指示の3形態とは指示の表現方法と顔と手の相対位置のパラメータを約するものであり、以下にこれを説明する。

1)指示方向を視覚により確認しない指先方向指示（体側指示と略す）

2)指示方向を視覚により確認する指先方向指示（眼前指示と略す）

3)視線による方向指示（視線指示と略す）

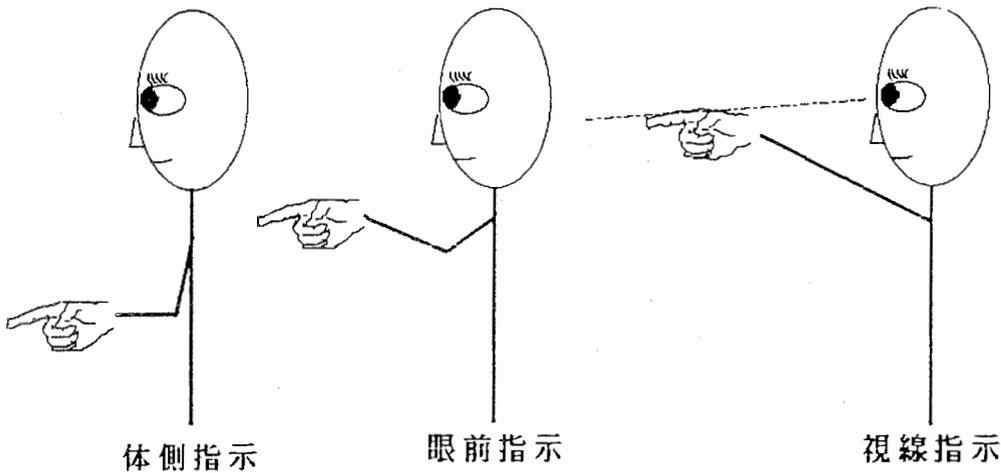


図2 指示3形態

以上の指示形態の特徴を以下の表3にまとめる。

表3 指示形態の分類

指示形態	指先の確認	顔との相対位置	精度	指示速度	方向偏位
体側指示	無	遠 視野周辺	低?	高?	大?
眼前指先	有	近 視野内	中?	中?	中?
視線指示	有	視線上	高?	低?	小?

体側指示は目で捉えた対象を即座に指し示す動作に対応し視覚により指示方向が妥当か否かの判断により指示方向の修正を行わない。指先の位置は肩より下の位置にあり視野外にあるかまたは視野内であっても方向修正が行えるほど十分に指先が確認できない周辺部となる。これに対して眼前指示は目で捉えた対象位置に指先が向いているか否かの確認を行う。従って指先の位置は対象の視野内に有り、指先の方向と対象の位置の二つが同時に認知できる。最後の視線指示は目の位置と指先により対象を指示するもので、指先の位置は視線上となる。

経験的に理解できるように、指示形態は後者ほど注視の度合いが強くなり指示精度が向上するものと思われる。

実験1

内容

IRISワークステーション上の複数位置に目標を順次表示しこれを被検者に指さして対象を指示してもらう。測定は6次元磁気センサを用い、誤差としてカーソル中心と指示位置を結ぶ線と実際の指示方向との角度を指標に用いる。このときセンサの取り付け位置を各指示形態に合わせて、誤差角度を測定する。

また十字カーソル表示から指示（マウス、キーボード等のスイッチをクリックする）までの時間により指示時間を測定する。実験環境を図3に示す。

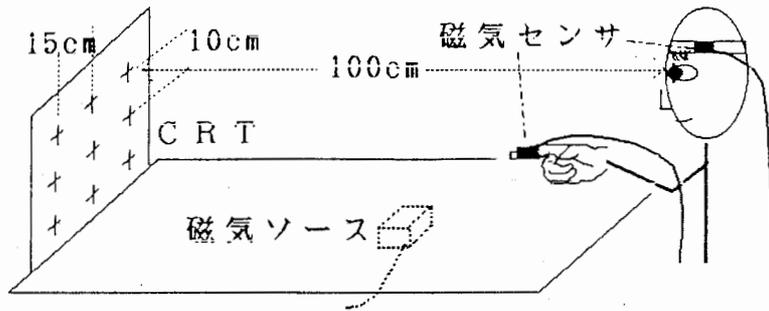


図3 実験環境

4. 1. 3人間・機械系による精度実験

(指示方向をカーソル表示する)

カーソル等で指示方向を表示することによって人間・機械系にフィードバックを与える場合、人間系のみによる指示動作に比べて、前述の体側、眼前、視線の指示形態による指示精度の差はなくなり、高精度となることが予想される。またその反面カーソルによって可能になった補正操作のため指示速度が低下すると思われる。

ここで問題となるのは、

1) 限界精度

2) 実際の画像の計測による測定誤差(カーソル表示方向のずれ)の許容範囲の2点である。図3に示すように指示対象を円で表現し、これに対する指示を試みる。カーソルを円内に誘導するに用する時間は円の直径が小さくなるにつれて大きくなり、遂には人間の手の変動要因によりカーソルが対象円内に安定して位置しなくなると予想される(図3グラフ)。この状態における指示精度を限界精度と呼ぶ。

実験2

本実験では限界精度と時間に関する指示特性データを得ることを狙いとし、目標となる円の直径と配置、および指示方向とに画像の測定誤差を意味するオフセットをパラメータとし、時間特性を観測する。使用ツールは実験1と同じである。

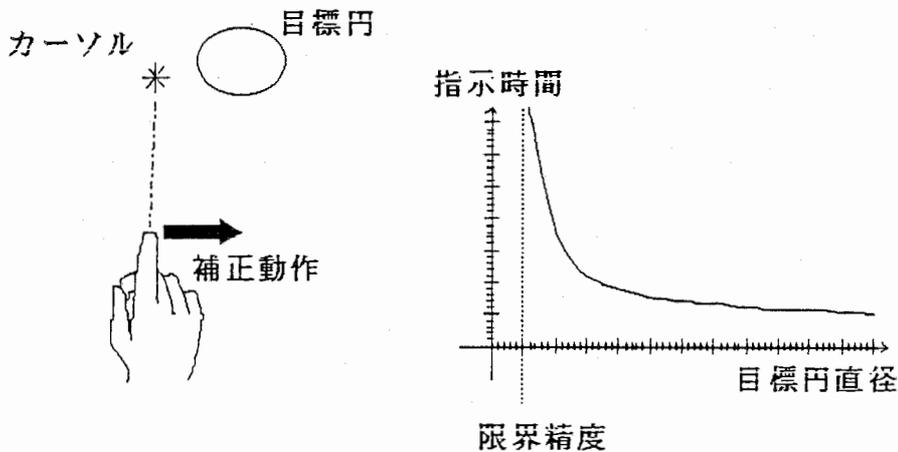


図4 カーソル表示による補正行動と限界精度

4. 2 手の3次元モデル構築のための基礎検討

人体の3次元モデルを記述しこれに操作を加えることによって望む姿態の人体をグラフィック表示する試みが活発におこなわれている[26][27]. この3次元モデルに対する操作は通常, 言語による. 例えば”笑っている”とか”手を伸ばしている”という表現に対して操作者がそれを下位の詳細な記述に分解するかあるいは高度なインタプリタの解釈により3次元モデルを変形させる. この変形させられた3次元モデルに対して画像を生成することによって一連の作業が完了する. これに対して手をマンマシンインタフェースに用いる過程はこれとは逆で以下のステップとなる.

- 1) 撮影画像の解析
- 2) 解析結果と3次元モデルとの対応づけによるモデル変形
- 3) 変形されたモデルの意味解釈

上記の3次元モデル変形手法による典型例として, 文献[4] に述べられている3次元ポーズ再構成システムがある. この例では認識を画像解析部とモデル変形部に分けて行っている(画像解析は2次元単一画像による). このような画像解析部で得られた結果とモデル変形部における3次元モデルの拘束条件を照合する手法は, モデルが既知である認識の一般的な手法であるが, 画像解析, モデル変形には様々な手法が考えられる. 一方, インタフェース利用のためには認識と併せて位置と指示方向の検出が必要となる. このための手法を以下の表4にまとめる.

表4 3次元位置計測のための手法

手法	内容	特徴
光投影等による能動的距離計測法 [28]	スポット光, スリット光, バターン光投影による三角測量の原理による. 距離画像が直接得られる.	精度, 分解能に関して適用範囲が広い. 使用環境の制約とリアルタイム化に難点がある.
投影幾何に基づく位置推定法 [29]	投影像から計測対象の幾何学的特徴に関する情報を拘束条件として3次元空間位置を推定する.	単眼視で可能. 対象の幾何情報が既知である必要がある. 適用できる対象に制約(剛体)がある.
ステレオ画像法 [30]	立体視による受動的な三角測量	対象物体の適用範囲が広い. 対応点探索が課題となる. 複眼視が必要.
錐体相関法 [15][16][17]	複数の視点からの投影像輪郭より, 3次元物体の近似領域を得る.	近似精度向上のためには多眼視が必要.

現在, 画像解析に錐体相関法を応用した手法を候補として検討している.

錐体相関法は図5に示すように複数の異なる位置にある視点から得られた複数の画像より3次元の共通部分を演算することによって3次元形状を得ようとするものである. 元来は

3次元計測のための手法であるが、対象が既知であることから2眼視あるいは3眼視でインタフェース実現に必要な形状認識と位置計測が可能であると考えられる。

他の計測手法に比べて以下の特徴がある。

- 1)形状認識と位置計測が同時に可能である。
- 2)直接3次元構造を得ることができる。
- 3)シルエット（輪郭像）による画像解析であることから画質の影響が少ない。

以上錐体相関法による案を述べたが、当面の研究課題として錐体相関法の画像解析への応用、拘束条件を与える3次元モデルの検討、このモデルと画像との照合手法の検討が挙げられる。

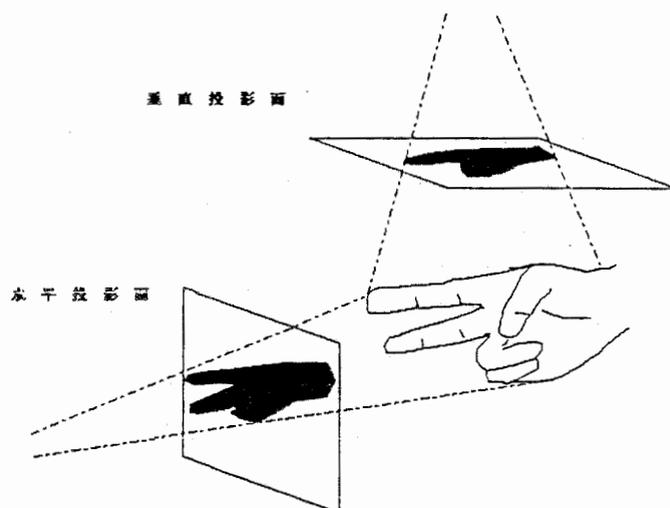


図5 錐体相関法の応用

5. むすび

本報告では手をマンマシンインタフェースに応用するための技術的背景を述べた。これらは手をインタフェースとして評価することと、手の位置、形状を非接触で測定することの2つに大別されるが、どちらの分野においても解決すべき課題は山積している。手を直接マンマシンインタフェースに用いようとする試みはまだ端緒についたばかりである。従って、研究の最初のステップは指示精度等の基礎データの収集および認識のための3次元モデルの構築となるがその重要性は本報告で指摘できたと思う。今後はポインティングというプリミティブな動作理解によるインタフェース実現を目標とし、非接触・画像認識の手によるインタフェースの有用性を明らかにしていく予定である。

本文では

引用せず参考文献として

マンマシンインタフェースの評価方法として[5][6][7][8][9][10]

手の空間指示特性(空間知覚)として[11][12]

3次元表示に対する視覚特性として[23][24]

を挙げた。

参考文献

- [1]末永, 間瀬; "手の動きの映像処理による情報入力法の検討", 情処学会30全大, pp1249-1250, 1985 (横須賀通研)
- [2]田村, 川崎; "手話動画認識システム", 情処学会研資, Vol.86, No59(CV-44), pp44.1.1-44.1.8, 1986.8 (阪大)
- [3]K.AKITA; "Image Sequence Analysis of Real World Human Motion", Pattern Recognition Vol17, NO.1, pp73-83, 1984 (三菱電気Computer Systems Works)
- [4]大橋, 黒川, 芝田; "身振り理解を目的とした単一画像からの3次元人体ポーズの再構成", proc. 2nd sympo. on Human Interface Tokyo, pp345-350, Oct., 1986
- [5]S.K.Card, T.P.Moran and A.Newell; "The Psychology of Human-Computer Interaction", Lawrence Erlbaum Associates, New Jersey, 1983.
- [6]S.K.Card, W.K.English and B.J.Burr; "Evaluation of Mouse Rate-Controlled Isometric Joystick, Stepkeys, and Text keys for Text Sekection on a CRT", Ergonomics, Vol.21, No.4, pp.601-613, 1978.
- [7]D.Whitfield, R.G.Ball and J.M.Bird; "Some Comparisons of On-Display Off-Display Touch input Devices for Interaction with Computer Generated Displays, Ergonomics, Vol.26, pp.1033-1053, 1983.
- [8]P.Reisner; "Formal Grammer as a Tool for Analysing Ease of Use", Human Factors in Computer Systems, pp.53-78, ABLEX, New Jersey, 1984.
- [9]P.Reisner; "Formal Grammer and Human Factors Design of an Interactive Graphics System", IEEE Trans. Soft. Eng., Vol.SE-7, No.2, pp.229-240, 1981
- [10]中山 剛; "ヒューマンファクタの課題", 情報処理会誌, Vol.27, No.10, pp.1121-1128, Oct., 1987. (日立中研)
- [11]内田 ひろみ他; "人の身体空間知覚能力 指関節における空間2等分割, 最小分割ならびに2点識別", 体力科学, Vol.35, No.1, pp.22-30, 1986. (慈恵医大)
- [12]旭敏之, 辻三郎, 笠井健; "眼-手協調動作時における空間知覚の役割", 信学技報, Vol.83, No.46, pp.9-16, 1983. (阪大)
- [13]佐藤宏介, 井口征士; "ビデオタブレット-CCDカメラを用いた座標入力装置", 信学論D, Vol. J67D, No.6, 1984. (阪大)
- [14]坂井利之, 長尾真, 金出武雄; "計算機による顔写真の解析", 信学論, Vol.56D, No.4, 1973. (京大)
- [15]宇佐美芳明, 仁尾都; "3次元デジタル画像入力方法の基礎検討", 情処学会31全大, pp1319-1320, 1985. (日立)
- [16]三宅哲夫, 土井順多; "立体形状の多面体近似システム", 情処学会論Vol.25, No.5, pp745-754, Sep. 1984. (東大)
- [17]登尾啓史, 福田尚三, 有本卓; "錐体相貫法を用いた2次元画像からのオクトツリーの作成", 信学技報PRU86-86, pp49-58, 1/16 1987. (阪大)

- [18]相沢清晴 他；”分析合成符号化におけるモデル構成と表情合成”，信学技報IE87-2, pp9-pp15, 1987.2 (東大)
- [19]村上伸一, 市原英也；”知能画像通信方式の一構成法”，信学技報IE87-14, pp7-12, 1987. (N T T 通研)
- [20]C.Schmandt;”Spatial Input/Display Correspondence in a Stereoscopic Computer Graphic Work Station”,Computer graphics,vol.17,No.3,July 1983.
- [21]Proceedings of SPIE;”True Three-Dimensional Imaging Techniques and Display Technologies”,15-16 Jan. 1987 L.A. Calif.
- [22]稲葉雅幸, 井上博ちか；”ロボットによるひものハンドリング”，日本ロボット学会誌, Vol.3, No.6, pp538-547, 1985 (東大)
- [23]広瀬通孝 他；”3次元立体視における奥行情報の認知精度に関する研究”，情処学会 32 全大, pp1431-1432,
- [24]畑田豊彦 他；”画面サイズによる方向感覚誘導効果”，テレビジョン学会誌, vol.33, No.5, pp407-413, 1979. (NHK)
- [25]樋渡涓二；”感覚と工学”，共立出版, 1976.
- [26]N.M.Thalman, D.Thalman;”The Direction of Synthetic Actors in the film Rendez-vous a Montreal”, IEEE Computer graphics and Applications, 1987.12.
- [27]根本啓次 他；”顔の3次元アニメーション作成システム”，NIKKEI Computer Graphics, 6/23 1987.
- [28]井口征士；”三次元画像入力システム”，第18回画像工学コンファレンス, 13-1, pp243-248, 1987 (阪大)
- [29]北橋忠宏 他；”固定カメラ映像による物体の平面上運動の認識”，信学論, Vol.J68D, No.4, pp584-591, 1985. (豊橋技科大)
- [30]丸山稔 阿部茂；”立体視に基づく面の抽出と面内ベクトルの認定による多面体構造の抽出”，信学論D, Vol.J70D, No.12, pp.2581-2589, 1987. (三菱)