

運動解析を用いたヒューマンインタフェースシステム

Human Interface System Using Visual Motion Interpretation

岡本 恭一
Y. Okamoto風間 久
H. Kazama

テレビカメラで撮影した画像から手などの未知形状の物体の運動を解析して、三次元空間の操作を行うことができるヒューマンインタフェースシステムを開発した。

このシステムの特長は、従来の運動解析手法に比べて、運動の大きさに関する情報は完全ではないが方向については安定して求めることのできる定性的な運動解析手法、並列画像処理プロセッサを用いたリアルタイムでの特徴の追跡と、両者の特徴の結合によるリアルタイムでの認識結果の利用者への視覚的フィードバックである。利用者は、ディスプレイを見ながら所望の運動量になるまで手の動作を続けることによって、三次元の運動を指示することができる。

This paper presents a new man-machine interface system that enables 3D operations on graphic objects by automatically interpreting the 3D motions of an unknown object such as a human hand. Image features are tracked in real time (video rate) using parallel image processors. The system uses a qualitative algorithm to interpret visual motion of the image features. The solution of this algorithm is reliable in terms of the direction of movement, although incomplete in terms of the degree of motion. The results of visual motion interpretation are communicated to a graphic workstation, which responds by changing the position and orientation of a computer graphics model. When watching the displayed results, the user can rotate or translate an object to the desired degree by continuing the motion in the desired direction.

1 まえがき

近年、仮想現実という分野でコンピュータグラフィックスを使って計算機のヒューマンインタフェースを作成しようという研究が盛んに行われている⁽¹⁾。このようなシステムではオペレータが三次元空間を操作することが必要となる。

これに対して、オペレータの手や頭など体にセンサをつけて、センサの位置や姿勢情報を計測して、三次元空間の操作を行わせる研究が進められている。センサを使うことによって、精度よく手や頭の位置を計測できるが、一方、オペレータになんらかの装具の装着を強いることになる、センサの測定範囲が限られるなどの問題がある。

画像処理を用いることによって、オペレータにはセンサをつけず、カメラで撮影することによってユーザの手の位置を計測し、ヒューマンインタフェースに用いようとする研究も進められている。

これらの研究では、ステレオ計測によって手の三次元位置と指の方向を計測しどこを指で指し示したかを求めている。ところで、通常、人間が三次元的な物体や空間に対して行う操作には、三次元位置を指示する以上に、三次元方向に物を並進運動させたり回転させたりする三次元空間中での運動が、

より頻繁に含まれている。

これに対して、当社は、手の運動を画像処理によって解析し、その結果を使って三次元空間の操作を行うヒューマンインタフェースシステムを試作した(図1)。

試作したシステムで検出するのは、各時点の間の短い時間



図1. 三次元空間の操作を行うヒューマンインタフェースシステム 左にあるカメラで撮影したオペレータの手の動きを解析した結果で、ディスプレイ中の飛行機を操作することができる。

System overview

間隔における相対的な運動であり、しかもそれを定性的に解析した結果である。したがって、ステレオ視覚のように三次元の絶対位置を検出することはできず、また運動の大きさについても絶対値を求めることはできない。

しかし、二次元空間の操作装置であるマウスを見てもわかるように、十分な反応速度で視覚的フィードバックを与えることができれば、絶対的な位置や姿勢は必要なく、相対的な運動情報を解析しただけでも十分ヒューマンインタフェースに利用できる。

ここでは、この目的に沿って試作したシステムについて述べる。

2 インタフェースシステム

今回試作したシステムは、次のような特徴をもっている。

- (1) 定性的な運動解析アルゴリズム
- (2) リアルタイム画像処理
- (3) 視覚的フィードバック

以下、この特徴について順に説明する。

2.1 定性的な運動解析アルゴリズム

運動の解析アルゴリズムについては、文献(2)、(3)ですでに詳しく述べたのでここでは概略を説明する。

図2に示した指先と手の甲に四つのボール(特徴点)をつけた手袋を着用して、手が図2の(a)の状態から(b)の状態に動いたとする。4点の特徴点を三次元空間で平面を構成する3点A、B、Cと、この平面上にない1点P(参照点)に分け、これらの点のカメラで撮影した画像上への投影をa、b、c、pとa'、b'、c'、p'とする。図2(a)の画像ではカメラと点Pを結んだ視線の上にある点はすべて点pに投影されるので、1枚の画像だ

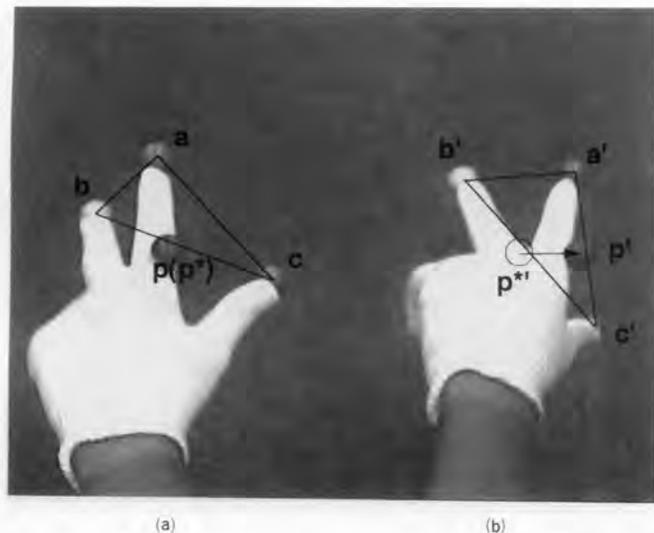


図2. 仮想運動視差 図の(a)から(b)への動きで仮想運動視差が発生する。この視差の方向は運動の回転軸の方向と垂直である。

Pseudo motion parallax

けでは点Pが平面ABCに対して前後にあるかどうかかわからない。ここで、図2(a)で点Pを通る視線上にあつて平面ABC上にある点P*を考える。点P*の投影p*の座標はpと一致する。

この図2の物体が三次元運動を行ったとする。“弱”透視投影という簡略化した投影系では、平面上の点の運動による位置関係の変化は、アフィン変形(一次変換と平行移動)で表されることが知られている(4)。ABCと点P*は同一平面上にあるので、これが投影された点a、b、c、p*からa'、b'、c'、p*'への移動もアフィン変形で表される。したがって、点p*はベクトル \vec{ab} と \vec{ac} を使って、 $p^* = \alpha \vec{ab} + \beta \vec{ac}$ と表すことができ、p*'も同様にベクトル $\vec{a'b'}$ と $\vec{a'c'}$ を用いて $p'^* = \alpha \vec{a'b'} + \beta \vec{a'c'}$ と表すことができる。

物体の運動を図3に示したカメラの視線(光軸)をz軸とした座標系で、各軸方向への並進運動と回転運動については、z軸の回りの回転、xy平面上の回転軸の向きとその回転軸の回りの回転によるオイラー角で表すことにする。並進運動とz軸の回りの回転では点PとP*の位置関係は変化しないので、p'の位置もp*'の位置と一致する。

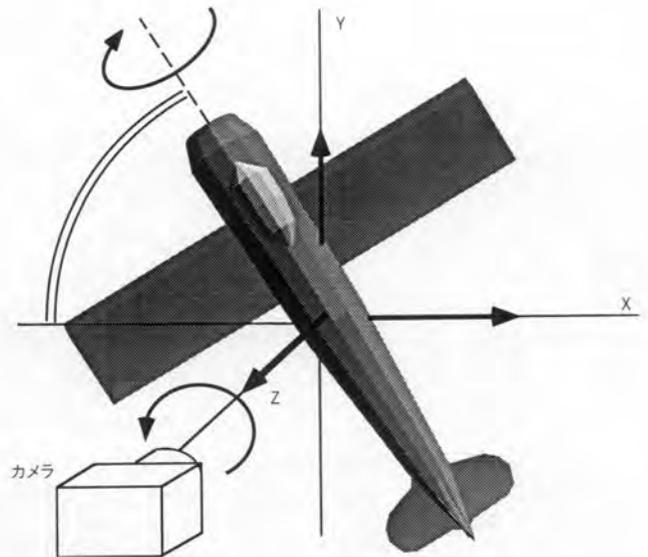


図3. 物体の運動パラメータ 物体の運動をカメラの視線をz軸とした座標系上で表現する。回転運動はオイラー角で表す。

Object motion parameter

しかし、図2のように縦軸の回りに回転する運動では、PとP*の位置関係が変化するため、運動の結果p'とp*'の位置に差が生ずる。この差を仮想運動視差と呼ぶ。

仮想運動視差p*'p'の方向は、オイラー角で表現した物体の回転軸の方向と垂直であることから、仮想視差から回転軸の方向を導くことができる。これによって、回転軸の回りの回転角以外の運動を求めることができる。回転軸の回りの回転については、奥行き大きい物体が小さく回転した場合と奥

行きの小さい物体が大きく回転した場合とで同じ結果が得られるため、物体の形状と回転量の間にあいまい性があるため、2時点の間の運動から検出することはできない。

以上述べた方式は、弱透視投影という実際のカメラの投影を簡略化した投影の下で運動を求める。実際のカメラは透視投影(中心投影)であって、カメラと物体の間の距離が非常に短ければ回転軸の周りの回転量も求めることができるが、実用的には難しい。以上述べた方式では弱透視投影でも正確に求められるパラメータと正確には求められないパラメータとを分けた表現をすることによって確実にわかる動きを安定してとらえることができる。

2.2 リアルタイム画像処理

このヒューマンインタフェースシステムは、画像処理専用プロセッサ、ワークステーション AS4075、グラフィックス画像を表示するグラフィックスワークステーションの三つで構成される。

このうち、画像処理プロセッサは、M68030 マイクロプロセッサを搭載したボード 16 台で構成される並列計算機である⁽⁵⁾。各ボードは図 4 のようなハードウェア構成をとり、VME (Versa Module Eurocard) バスを介して AS4075 と接続されている。図のイメージバスには、つねにカメラから入力された動画像がデジタル化されて流れている。各ボードには、マイクロプロセッサがウィンドウコントローラに必要とする領域を指示することによって、イメージバスを流れる画像から任意の部分領域をボードのもつウィンドウメモリに取り込むことができる。これによって、部分領域ごとに各マイクロプロセッサに動画像を振り分けて、並列に処理できる構成になっている。

画像処理を単純化してリアルタイムに動作させるため、オ

ペレータは、4 色の色を塗ったボールが先端についた手袋を着用し、画像中の各ボールの中心位置を特徴点の位置座標とした。このシステムでは、画像処理プロセッサの 16 台のボードのうち 4 台だけを使用し、それぞれのボードに自分の担当色のボールを独自に追跡させて特徴点座標を計算する。

なお、各ボードにはあらかじめ追跡対象のボールの色を教示して、RGB (赤, 緑, 青) の画素値から、追跡対象の色を識別するルックアップテーブルを作成しておく。

実行時には、各マイクロプロセッサは、まず、画像の一部分の長方形領域を取り込み、この領域中の各画素について、ルックアップテーブルで色の識別を行う。教示した色と一致する画素が一定数以上あればその重心と半径を求め、重心の位置座標を特徴点の座標とする。マイクロプロセッサは、計算した特徴点の位置、半径とそれまでに検出した特徴点の位置から、次の画像での特徴点の位置を予測し、次の画像では予測位置を中心にした領域で画像を取り込み、運動しているボールのトラッキングを行う。

教示した色と一致する画素がない場合は、領域の面積を広げて時間的に連続した 2 枚の画像を取り込む。2 枚の画像の差をとって、差が大きく、かつ教示した色と一致する画素を探す。

ボールが見えている間は、その近傍領域だけに注目して追跡し、見失うと探索範囲を広げて画像全体から運動しているボールを探すことによって、背景の中に類似の色をもつ物体があっても影響されずに特徴点の追跡を行うことができる。

2.3 視覚的フィードバック

AS4075 では画像処理プロセッサから特徴点の追跡情報を受け取り、オペレータが手を動かして行った運動を特徴点の運動から解析する。解析結果は図 3 に示した並進運動と回転運動の六つのパラメータで表されるが、このパラメータは物体の運動を完全に表すものではなく、例えば回転軸の周りの回転は前述した深さのあいまい性のために正確な回転量はわからない。しかし、回転の有無と回転軸に対して右回り、左回りの運動の判定は適当なしきい値との比較で判定できる。このように運動が起きたと判定すれば、回転並進の種類と方向を付加したイベント情報をグラフィックスワークステーションに送る。

グラフィックスワークステーションは AS4075 から受信したイベント情報に従ってグラフィックスの表示内容を変更し、その画像をオペレータに表示して視覚的フィードバック情報を返す。

イベント情報には、運動の種類と方向しか指示されない。したがって、グラフィックスワークステーションでは運動の大きさについては、あらかじめ決めた量を与えて三次元の運動を計算する。計算結果に合わせてモデルの姿勢や視点の位置などを変更したグラフィックス画像をオペレータに表示する。

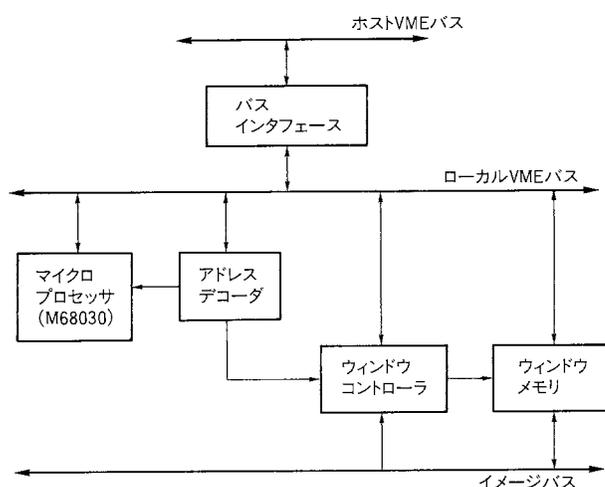


図 4. 画像処理プロセッサのハードウェア構成 イメージバスを流れる動画像から任意の部分領域をウィンドウメモリに取り込むことができる。

Hardware configuration of image processor

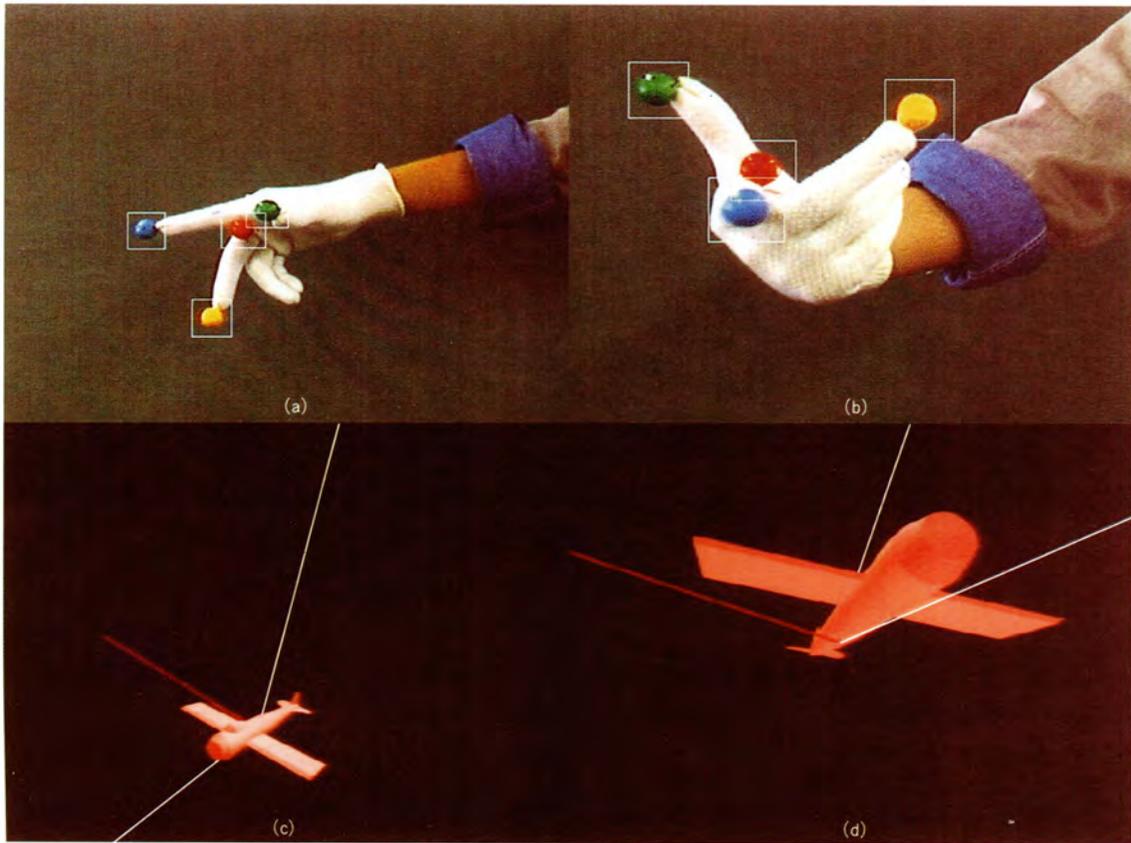


図5. システムでの三次元モデルの操作例 手を図の(a)の位置から(b)の位置に動かすことによって飛行機の三次元モデルを(c)から(d)の姿勢に操作できる。

Result of hand motion interpretation

3 三次元 CAD モデルの操作例

試作したシステムで、飛行機の三次元モデルを操作する実験を行った。図5は、手の動きを解析して三次元モデルを操作した結果である。

オペレータの手の動きによって三次元モデルに回転、並進などの操作を行い、モデルを希望する姿勢に置くことができる。特徴点は、親指、人差し指、中指の指先と人差し指の付け根の4か所につけている。実際には手は剛体ではないが、意識的に指を動かさない限り1/30sの短い時間間隔では剛体と仮定して運動を検出することができる。

4 あとがき

カメラの画像から手の動きを解析し、その結果を使って三次元モデルを操作することのできるヒューマンインタフェースシステムを試作した。今後は、手袋などを使わず、素手で操作することのできる簡便なインタフェースシステムの開発を行っていく。

文 献

- (1) J. Nomura: Visual System and Its Application to Consumer Showrooms, Visual Computing, pp.183-198 (1992)
- (2) 岡本恭一, 他: 定性的運動認識を用いたヒューマンインタフェースシステム, 電子情報通信学会論文誌 D-II, J76-D-II, 8, pp.1813-1821 (1993)
- (3) R. Cipolla, et. al: Robust Structure from Motion using Motion Parallax, Proceedings on ICCV '93, pp.374-382 (1993)
- (4) J. Koenderink, et. al: Optic flow, Vision Research, 9-2, pp.161-179 (1986)
- (5) H. Kubota, et. al: Vision Processor for Moving Object Analysis, Computer Architecture for Machine Perception, pp.461-470 (1992)



岡本 恭一 Yasukazu Okamoto

1986年入社。画像理解技術の開発に従事。現在、研究開発センター関西研究所研究主務。
Kansai Research Lab.



風間 久 Hisashi Kazama

1992年入社。画像理解技術の開発に従事。現在、研究開発センター関西研究所。
Kansai Research Lab.