

バーチャルリアリティ世界構築のための 指の動きの検出法

中嶋 正之 柴 広有

東京工業大学 工学部

近年, 仮想現実世界を構築するための新しいマンマシンインタフェースとして, 手の動きをそのまま入力する研究が行なわれている. 中でも機器を装着しないで, 画像処理技術だけを使用する手法は関心が高く, また多くの手法が検討されている.

筆者らは, 指の動きに着目し, 実時間でかつインタラクティブな処理が可能な検出法を考察している. 本稿では, まず, 色テープを指先と指の付け根に貼り, 画像中から三次元位置座標の値を検出する処理を簡易にして検出処理時間を短縮する手法を提案する. そして, コンピュータ内に予め指の骨格線モデルと形状決定ルールを作製し, モデルに与える情報を指先と指の付け根の座標値の2つだけにして, 余分なデータのやりとりを極力抑える手法を考察する. さらに, 実世界での指の動きに合わせてコンピュータ内のモデルを逐一動かしていくモデルマッチング手法を検討する.

A STUDY ON DETECTING FINGER MOTION TO REALIZE A VIRTUAL WORLD

Masayuki Nakajima Hirokuni Shiba

Faculty of Engineering,
Tokyo Institute of Technology

2-12-1 Ookayama, Meguro-ku, Tokyo, 152 JAPAN

In recent years there are studies to input the data of a hand motion in itself as a new Man-Machine Interface to realize a virtual world. Especially, the motion detection method without wearing input sensors, using only an image processing engine, is recognized as effective one, so a large number of studies have been proposed on this method.

We are concerned with a finger motion. The study proposed in this paper to detect a finger motion in real time and interactively. The subjects of this paper are as follows. At first, we mark a base and tip of finger with a color tape to make easy a process to detect the coordinates of a finger in a space of 3-dimension in real time. Next, we make a finger flame model and a decision rule of a finger shape in a computer to make a finger model on display with minimum data. Lastly, we discuss the model matching method to move a model made in a computer in time to a finger motion in physical space.

1 はじめに

1.1 背景

近年、コンピュータの飛躍的な進歩に伴い、仮想現実世界を構築する研究が広く行なわれている。その構築作業を行なう際に、キーボードやマウスなどに代わる、新しくて使いやすいマンマシンインタフェースの開発が望まれている^[1]。

そのような要求に対して、自然な人間の動作それ自身を直接コンピュータに入力することが考えられており、データグローブなどの自由度の大きい入力デバイスの開発が行なわれている。しかし理想のインターフェースとは人にかかる負担を最小限に抑えたものである。

そこで筆者らは、手や身体の動きを捕らえる機器を装着せずに、画像処理技術だけを使用して動きのデータを計測することを検討し、またシステムとインタラクティブにコミュニケーションを行なうために、実時間処理が可能な手法を考察している。

新しいインタフェースの一つとして、現在、画像処理技術だけを使用して手振りを認識する研究が報告されている^{[2] [3] [4]}。しかし画像中の大量のデータから指の部分の輪郭を高速で検出するには、指の動きに多くの制限や条件が求められている。

1.2 目的

筆者らは、手や指の動きの検出と表現に関心があり、特に実時間で1本1本の指の動きをそれぞれ検出して表現することを目標としている^[5]。

そこで、指の動きの検出と表現を少ないデータをもとに実現する手法を考察している。

本稿では、そのための第一段階として、まずステレオ画像を用いて一本の指の空間位置、並びに屈伸状態の検出と表現を行なう手法を考察する。

次に、その応用として、同じ処理の方法で複数の指の検出と表現を行なう手法を考察する。

2 処理の概要

実時間でしかもできるだけ正確に指の形状を画像中から検出し表現するために、次のような処理の方法を考察する。

1. 指の骨格線モデルは予めコンピュータ内に作製しておき、同時に指の形状が一意に決まる形状決定ルールも作製しておく。
2. 指先と指の付け根の二箇所の色テープを貼り、画像中から指先と指の付け根にあたる部分の位置座標を検出する作業を容易にし、検出処理にかかる時間を短縮する。
3. また画像中から得るデータは指先と指の付け根にあたる部分の2点の座標値だけとし、データ量を最小限に抑える。
4. 得られた座標値にマッチさせるように骨格線モデルを形状決定ルールに基づいて変形していく。

手や足にマークをつけて身体の動きを撮影し、3次元空間内の動きのデータを計測する多くの手法が提案されている^[6]。しかしマークの数は何十個にも及ぶものであり、あくまで美しいコンピュータグラフィックスを生成するためのものであった。

筆者らは、画像中から指の特徴検出処理時間を短縮するために、指に色テープでマークをつけ、またコンピュータ内に予め骨格線モデルと形状決定ルールを作製することにより、画像から得るデータ量を少なくしても、指の形状を検出し、表現できるようにすることを考えている。

3 一本の指の動きの検出と表現

3.1 処理手順

処理手順を図1に示す。

処理の流れを大きく分けると、座標算出までの、データ抽出処理、得られたデータに基づいて指の形状を決定する処理、そして決定した形状をシリンドラモデルを用いて表現する処理の三つに分けられる。

3.2 対象とする指について

指骨は手のひらから近い順に、基節骨、中節骨、末節骨と並んでいる。

取り扱う指は、左手の人指し指とする。本来ならば指に何も負担をかけないことが理想ではあるが、

本稿では白色のテープを人指し指の先端と基節骨の骨底部分に貼る。この指を2台のビデオカメラを用いて撮影する。

また指はゆっくり動くようにする。

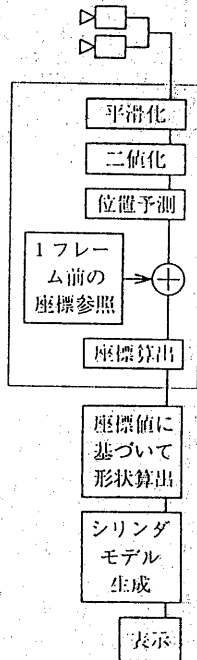


図1 処理手順

3.3 対象画像について

連続する動画像30フレームをそれぞれ横320×縦240画素、白黒256階調で計算機に入力し、これを対象画像とする。

対象画像に対しては、検出処理の一段階として、次に示す条件を設定する。

1. 第一フレームにおける指の向きおよび形状は、指をまっすぐ伸ばしてビデオカメラに手のひらを相対するように向けることにする。
2. 検出する左手人指し指の指先と基節骨の骨底部分は画像上にいつも存在していることにする。

対象画像の座標系は、左上を原点とし、x座標は水平右方向を、y座標は鉛直下方向を、z座標は紙面を裏から表へ突き抜ける方向をそれぞれ正とする。

3.4 指先と基節骨の骨底部分の検出

計算機に入力された対象画像をそれぞれ平滑化した後二値化して、面積が大きい2つの白い領域を検出する。そして、領域ごとに重心を算出する。

次に、左右のフレームを用いて、仮想空間上の座標を求める。そして求めた座標が、指先の座標か、それとも基節骨の骨底部分の座標かは、1つ前のフレームから求めた座標を使って判定する。その際、特にZ座標の値を重視する。

3.4.1 検出処理の方法

画像中から白い領域を検出する処理は、次に述べる処理をすることにより高速化をはかる。

1. 縦8×横8画素を1ブロックとして、それを縦に6個、横に6個並べた網を考える。
2. 次に、1フレーム前の画像から検出され、算出された指先と指の基節骨の骨底部分の座標を中心になるように、1の網をそれぞれ置く。
3. 各ブロックごとに、白色の画素の数を算出する。白色の画素数が設定した値よりも大きければ、そのブロックを選ぶ。
4. 網をx軸正方向へ2画素、y軸正方向へ2画素ずらし、再度3と同じ処理を行なう。選ばれたブロックが3で選ばれたブロックを縦に2画素、横に2画素ずらしたのならば、5の処理を行なう。もし異なる場合は、予め定めた方法で網をずらし、同じ処理を行なう。
5. 選ばれたブロックの8近傍を含めて、縦24画素×横24画素中から、白色領域の重心を求める。

3.5 指の骨格線の形状決定

形状決定は、予め定めておいたルールに従って行なわれる。

形状決定ルールは計算簡略化のため、次のように定める。

- 指の長さは基節骨の骨底のz座標の値ごとにあらかじめ定めていた値を割り当てる。
- 基節骨、中節骨、末節骨の長さの比を20:12:9とする。

- 形状決定は、予め計測しておいた画像上での指骨の長さ、および先に求めた指の先端と基節骨の骨底部分の位置座標を用いて行なう。

実際に用いるモデルを図2に示す。

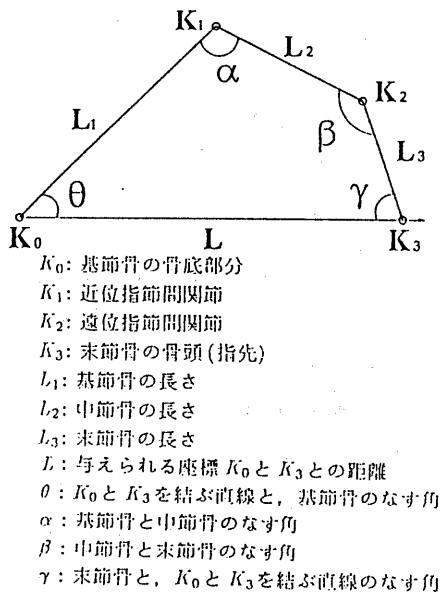


図2 指の骨格線モデル

得られた形状データは、シリンダモデルで表現する。

3.6 形状データの出力例

図3に第10フレームの入力画像を示す。そして、図4には図3を二値化した画像を、図5には、第10フレームの画像から得られた形状データを示す。



図3 第10フレームの入力画像



図4 図3を二値化した画像

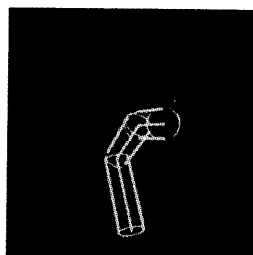


図5 第10フレームから得られた形状データ

同様に図6に第1フレームから得られた形状データを示す。また、図7に第30フレームから得られた形状データを示す。

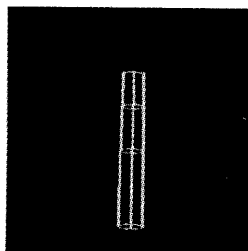


図6 第1フレームから得られた形状データ

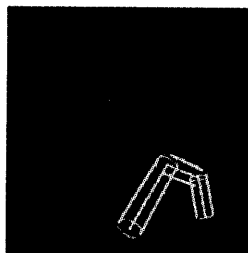


図7 第30フレームから得られた形状データ

3.7 結果

連続するステレオ画像を用いて、指の動きを検出し表示する研究の第一段階として、一本の指を対象にして、指の形状データを得ることができた。

二値化した時に、かならずしも処理しやすい画像が得られるわけではないが、例えば、図4の右画像を用いて処理を行なっても、結果として図5のような形状データを得られたのでこの手法は有効であると考えられる。

一本の指の検出と表現についての今後の課題を次に述べる。

- 画像中の検出点をよりきれいにするにより、形状検出の精度をあげる。
- 指の太さ、長さ、等の各個人の特徴をデータとして扱えるようにする。
- さらに処理時間を短縮するアルゴリズムを考察する。

4 複数の指の動きの検出と表現

次に、複数の指の検出と表現を行なう手法について考察する。

おおまかな処理手順は一本の指の動きの検出と表現の処理手順と同じである。

人指し指、中指、薬指、小指の先にそれぞれ青、緑、赤、黄色のテープを貼り、人指し指と小指の基節骨の骨底部分に白色のテープを貼る。それらを2台のビデオカメラで撮影し、各色領域を画像中から検出することで、各位置座標を検出する。

4.1 処理手順

処理手順を図8に示す。

指先と指の付け根の三次元座標の検出方法は一本の指を扱う場合と同様である。

4.2 対象とする画像

画像は、連続する動画像30フレームを、それぞれ横320×縦240画素のカラー画像で計算機に入力し、これらを対象画像とする。

検出条件として、第1フレームにおける各指の形状と向きは、指を伸ばして、指の腹の部分カメラに向けることとする。また、手の平が常にカメラに向くようにする。

4.3 指の骨格線の形状決定

まず、図9に示す骨格線モデルをコンピュータ内に作製する。次に、カメラを用いて検出した各座標をもとに、予め決めておいた指の形状決定ルールに従って、モデルの形状を変えていく。

指骨の長さ、各指の間の長さは第1フレームが与えられた時に計測する。得られたデータはシリンドラモデルで表現する。

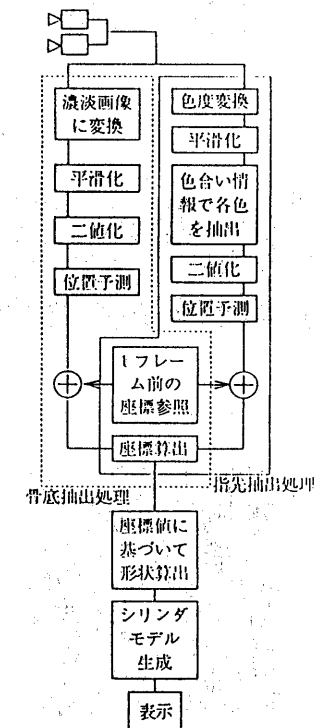


図8 処理手順

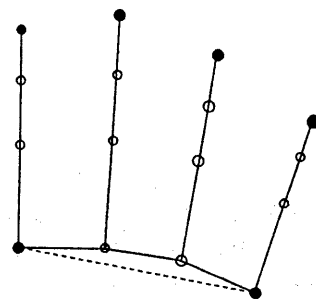
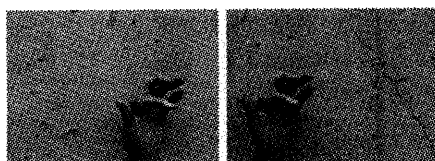


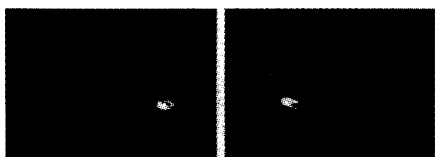
図9 指の骨格線モデル

4.4 形状データの出力例

図10に第10フレームの入力画像を示し、図10から青色の部分を検出した画像を図11に示す。そして図10を用いて得られた形状データの出力例を図12に示す。



(a) (b)
図 10 第 10 フレームの対象画像



(a) (b)
図 11 図 10 から青色領域を検出した画像

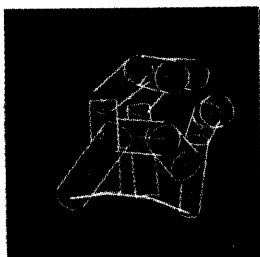


図 12 図 10 から得られた形状データ

4.5 結果

ステレオのカラー画像を用いて 4 本の指の形状データの検出と表現を行なう手法について考察を行なった。この手法は、処理時間はかかるものの各指の形状をとらえることができる。

実験では、指の先端が他の指に隠れて認識できない場合のために、第 3 のビデオカメラを上方に設置して、X-Z 平面上での座標を算出し、Y 座標は 1 つ前のフレームから求めた座標を割り当てる、というアルゴリズムを用意した。結果として、精度はよくないが、各指の形状をとらえることができた。但し、今のところ、この手法を適用するには、この状態があまり長く続かないことを条件にしなければならない。

5 おわりに

画像処理技術だけを用いて指の動きを検出し、表現する一手法を考察した。

指先にテープを貼るにより、求める部分の座標を検出しやすくして、処理時間を短縮する方法は有効であると思われる。しかし、指の形状を一意的に決定するルールを適用しているので、精度はあまりよくない。

そこで今後の課題は次のことである。

- 実際に指の動きを実時間で検出して表現するシステムの実現
- 精度をあげるための手法の考察。例えば、画像中からの検出点を増やす等。
- 新しいアルゴリズムの作成。例えば並列処理を用いる等。
- 手の負担をなくすため、テープを貼らずに検出する方法の考察。現在、動きベクトルを用いる方法を考察している。
- どこまで再現すればリアルに感じるのか、という「リアルの程度」の問題の考察。

参考文献

- [1] 岸野文郎:“4. 基礎技術 4-1 新しいマンマシンインターフェイス”, テレビ誌, Vol.47, No.3, pp293-296(1993)
- [2] 栄藤, 伴野, 岸野:“ステレオ輪郭像を入力とした物体の一般化円筒複合体による記述”, 信学論, J73-D-2, 9, pp1402-1412(1990)
- [3] 末永, 間瀬, 福本, 渡部:“Human Reader:人物像と音声による知的インタフェース”, 信学論, J75-D-2, 2, pp190-202(1992)
- [4] 末永康仁:“通信における画像処理-人物像の認識理解-”, テレビ誌, Vol.46, No.11, pp1443-1447(1992)
- [5] 中嶋, 柴, 鄭, 安居院:“ステレオ画像を用いた指の動きの検出の一検討”, 信学春季全大, 1993, D-600
- [6] 末永, 間瀬, 渡部:“CG 講座:人間の CG 表現と応用”, テレビ誌, Vol.46, No.8, pp1012-1020(1992)