

ジェスチャインタフェースを用いた 3次元モデリング手法の提案

○横川 健 石井 雅博 唐 政 山下 和也

富山大学大学院 理工学研究科 知能情報工学専攻

抄録 - ユーザが頭の中にイメージしている 3次元形状をロボットやエージェントに伝達させること目的として、3次元モデリングシステムを構築する。人同士のコミュニケーションに近づけるために我々が日常使用しているジェスチャを用いる。システム構築における 3つの課題①手領域抽出、②リサンプリング、③ポリゴン生成、の解決について述べる。また作成したプロトタイプシステムの動作実験について述べる。

Proposal of 3D modeling technique that uses gesture interface

○Takeshi Yokogawa, Masahiro Ishii, Zheng Tang and Kazuya Yamashita

Graduate school of Intellectual Information Systems Engineering, University of Toyama

Summary - 3D modeling system is built as previous paragraph, in order that robots is made to transmit 3D form. In case we input 3D form into computer, we use 3DCG software, such as Light wave 3D. Much 3DCG software uses a mouse and a keyboard as input interface, however in order to bring close to men's communication, it uses the gesture which we perform every day. In this research, since gesture is acquired using stereo camera, it becomes an un-contacting type interface. It conducted the prototype operation experiment of the solution of three themes in a system configuration, and a modeling system.

1. はじめに

今後我々の生活の中にロボットやエージェントとのインタラクションが必要となってくる。本研究では、ロボットやエージェントとのインタラクションを人同士のコミュニケーションに近づけることを目標としている。ロボットやエージェントに言葉だけでは、うまく伝えられない形状を伝達させるため 3次元モデリングシステムを構築する。

現在一般的に 3次元の形状を作る場合、3DCG ソフトウェアを用いる。モデリングソフトウェアとしては Light wave 3D などがある。これらのソフトウェアの多くはマウスやキーボードなどの接触型入力インタフェースが必要であり、熟練した操作も必要である。簡単に 3次元形状を作るモデリングソフトウェアとしては、teddy^[1]があるが、マウスやタブレットが必要である。また、外形を入力すると自動的にふくらみを持たせてくれるので操作は簡単であるが必ずしもユーザの意図した形状が得られるわけではない。本研究では非接触型入力インタフェースとしてジェスチャを用いる。バードウィステル^[5]によるとジェスチャなどのノンバーバルコミュニケーションはメッセージ伝達全体の 65%を占めており、ロボットやエージェントとのインタラクションに取り入れられている。本研究で

はジェスチャを用いた 3次元モデリングシステムを構築する。

2. 概要

ジェスチャをインタフェースとするためにステレオカメラを用いる。ユーザがイメージした物体をステレオカメラの前で手振り動作してもらい、ディスプレイにイメージしたポリゴンを表示させるシステムを構築する。図 1 にシステムのイメージを示す。

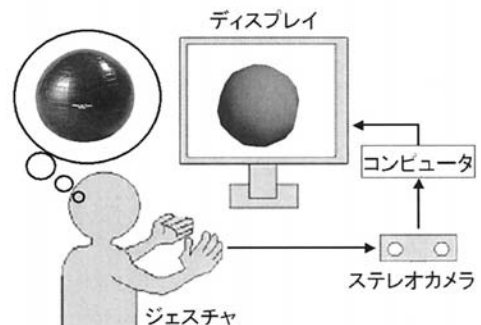


図 1 システムのイメージ
Fig.1 General Idea of proposed system.

3. 提案するシステム

3.1 ジェスチャの利点

ジェスチャは空中でモノの形を描く動作で、エックマン^{[9][10]}によると例示的動作に分類される。ジェスチャには以下の3点の利点がある。

- ① 手指による形状描写が可能である。
- ② 人が日常使用しているコミュニケーション手段なので、直感的に操作がわかる。
- ③ 非接触型のインタフェースとして使える。

3.2 入力装置

本研究では入力装置としてステレオカメラ（図2）を用いる。ステレオカメラは色情報（RGB値）と奥行き情報（レンジデータ）が取得できる。利点は非接触型であり、動作に影響を与えずに3次元空間位置座標のデータ（点群データ）が取得できることである。また時系列データの取得により、動作による形状取得が可能である。一方、欠点としては、重なりによる隠蔽部分は観測できないということがある。



図2 ステレオカメラ
Fig.2 Stereo camera.

3.3 出力装置

本研究では、ジェスチャによって入力された3次元形状を最終的に2Dディスプレイ上にポリゴン表示させる。ポリゴンはパーソナルコンピュータにおける3次元形状表現の標準方式であるし、ユーザによる3次元形状の認知にも適しているからである。

前述したように、ジェスチャ入力時に取得されるのはレンジデータであるため、ポリゴンとして表示するためには適当な3つの点を組としてまとめ、ポリゴンを生成しなければならない。

3.4 システム構築における課題

本研究で提案するシステムを構築するにあたっては、解決すべき課題が3つある。

- ① 手振り動作から形状を取得するため手指部分の情報のみが必要であり、背景などは必要ではない。そのため、手領域抽出を行う必要がある。
- ② ステレオカメラで取得される点群データの数はポリゴン表示には多すぎるため、リサンプリング（点の数を減らすこと）が必要である。
- ③ 点群データから適切な点を選び出し、それらを組み合わせることにより、三角ポリゴンを生成

しなければならない。

4. プロトタイプシステムの構築

4.1 システム環境

ジェスチャインタフェースとしてユーザが操作する空間は直径1mの球の内部であると仮定した。そこでステレオカメラの焦点距離と視野角度より、ユーザから約2mの位置にカメラを置くことにした。またカメラは手前のみしか取得できないので隠蔽部分を少なくするため、ジェスチャを見下ろすようにカメラを15°傾けた。ユーザとカメラの位置関係を図3に示す。

カメラの解像度や撮影のフレームレートは手の形や動作が取得可能かつ画像処理のコストを考慮し、表1のようなものとした。本研究では表示ディスプレイは液晶ディスプレイを用い、ポリゴンの描画はDirectXで行う。システム環境を表2に示す。

表1 ステレオカメラの仕様
Table 1 Specification of stereo camera.

CCDサイズ	1/3インチ
解像度	320×240ピクセル
フレームレート	30フレーム/秒
焦点距離	6mm
レンズ間距離	12cm
視野角度	40°

表2 システム環境
Table 2 System environments.

開発環境	Visual C++ .net
開発言語	C++
OS	Windows XP
CPU	ADM Athlon64processor 2.0GHz
メモリ	1.0GBRAM
ポリゴン描画	DirectX SDK

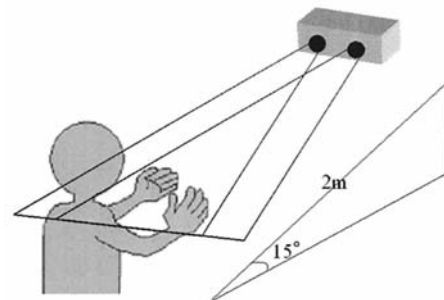


図3 ユーザとカメラの位置
Fig.3 Position of user and camera.

4.2 手領域抽出

4.2.1 課題

ステレオマッチングを適用することによって、3次元空間位置座標の点群データを取得することができる。しかしこのまま全領域を点群データに変換してしまうとジェスチャに必要でない部分（体や背景など）も取り込んでしまう。つまり、必要のない部分の点群データを取り除く必要がある。

4.2.2 手領域抽出方法

手領域抽出には、ステレオカメラから取得される色情報（RGB 値）と奥行き情報（レンジデータ）を用いる。色情報では RGB 値を HSV^[7] [変換式:(1)(2)(3)式] と YcrCb [変換式:(4)(5)式] に変換して H（色相）と CrCb（色差）の値から肌色抽出する。さらに奥行き情報では、手振り動作とカメラの有効範囲から必要となる点群データのみを取得する。手領域を抽出するための閾値内を手領域とする。それぞれの閾値を以下の(6)～(9)に示す。閾値は文献^[7]とカメラによる色情報の調整のため修正したものである。色情報と奥行き情報を用いた手領域抽出の結果を図 4,5 に示す。

[HSV 変換式]

$$\text{if: } \max = R \quad (1)$$

$$H = 60.0(G - B) / (\max - \min)$$

$$\text{if: } \max = G \quad (2)$$

$$H = 60.0(B - R) / (\max - \min) + 120.0$$

$$\text{if: } \max = B \quad (3)$$

$$H = 60.0(R - G) / (\max - \min) + 240.0$$

[YCrCb 変換式]

$$Cr = 0.500R - 0.418G - 0.081B \quad (4)$$

$$Cb = -0.168R + 0.331G + 0.500B \quad (5)$$

[H(色相)の閾値]

$$0.0 < H < 30.0 \quad (6)$$

[CrCb（色差）の閾値]

$$Cr < 60.0 \quad (7)$$

$$Cb < -10.0 \quad (8)$$

[Z（奥行き距離 [m]）の閾値]

$$0.5 < Z < 2.0 \quad (9)$$



入力

出力

図 4 手領域抽出（色情報）
Fig.4 Hand area extraction.(color)



入力

出力

図 5 手領域抽出（レンジデータ）
Fig.5 Hand area extraction (rang data).

4.3 リサンプリング

4.3.1 課題

ステレオカメラは数秒のジェスチャだけでも膨大な数の点群データになってしまうため、点の数を減らす必要がある。図 6 は手領域抽出後の取得されたレンジデータである



図 6 取得されたレンジデータ
Fig.6 example of acquired range data.

4.3.2 リサンプリング方法

リサンプリングの方法を以下に述べる。点群全体を直方体で包み、一番長い辺の中点で 2 分割する。分割した直方体中の点の数が規定値以上なら分割を繰り返す。規定値内であれば直方体中の点群を重心に代表させる。図 7 にリサンプリング方法のフローチャートを示す。また図 6 の点群データをリサンプリングして点の数を 1577 にした結果を図 8 に示す。

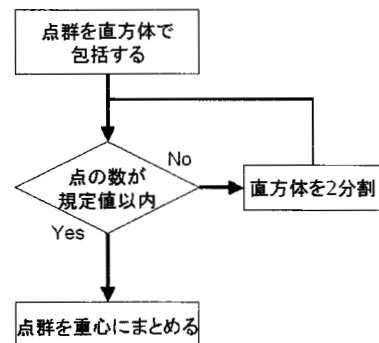


図 7 リサンプリングフローチャート
Fig.7 Flow chart of resampling.



図8 リサンプリングした結果
Fig.8 Result of resampling.

4.4 ポリゴン生成

4.4.1 点群データからのポリゴン生成

現在のビデオカードは、3次元表示がポリゴン表示に特化している。ジェスチャから得られた点群データをそのまま3次元で表示することは可能だが、モデリングツールとして考えると3次元上に不規則に並んでいて、かつ膨大な数がある点群データは編集がしづらくデータとしての汎用性がないといえる。一般的な3次元モデリングシステムでは計算量を減らすためにサーフェスマデルが用いられている。点群データの表面部分のみをポリゴンに変換し、データとしての汎用性を高める必要がある。また、点群の表示は3次元物体としての視認性が悪く、ポリゴンによる表示が有効である。

4.4.2 課題

サーフェスマデルとして点群からポリゴンを生成する場合以下の点に注意する必要がある。

- ポリゴン同士が重ならない。
- すべての点を使ってポリゴンを生成する。
- ポリゴンが鋭角・鈍角にならない。

図9にポリゴン生成の良い例・悪い例を示す。

図9は2次元上の6点が与えられている。良い例では、ポリゴン同士が重なることなく全ての点を包括するようにポリゴンを生成されている。悪い例では、ポリゴンが重なったり、ポリゴンが鋭角になったりしてしまっている。

4.4.3 ポリゴン生成法

本研究では、ポリゴン生成をロシアの数学者ボリス・ドロネーが考案したドロネー三角分割⁴⁾を3次元に拡張することで実現する。ドロネー三角分割は平面上で与えられた点を頂点とする三角形のうち、三角形の外接円の頂点以外に他の点を含まないように三角形に分割する方法である。ドロネー三角分割は2次元で行う手法である。本研究では点群データが3次元のデータとなるので、ドロネー三角分割を3次元に拡張

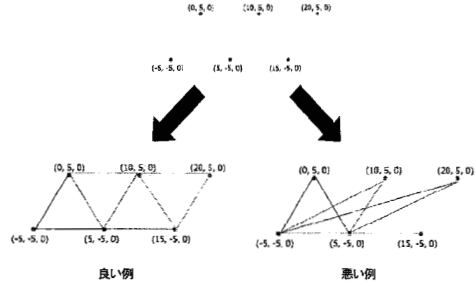


図9 ポリゴン生成の良い例・悪い例
Fig.9 good and bad example of generated triangle

した。3次元ドロネー三角分割のフローチャートを図10に示す。3次元のドロネー三角分割を行った結果を図11に示す。図11の図形はドロネー三角分割を2回に分けて行うため20枚のポリゴンで生成されている。

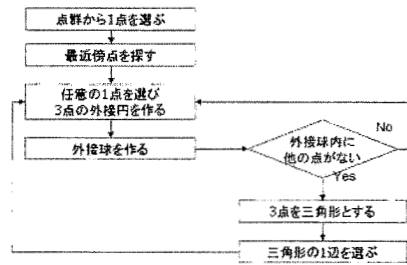


図10 3次元ドロネー三角分割のフローチャート
Fig.10 Flow chart of 3D Delaunay Triangulation.

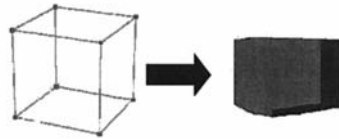


図11 3次元ドロネー三角分割
Fig.11 3D Delaunay Triangulation.

5. 動作実験

検討したシステムのプロトタイプの実験を行った。ユーザにステレオカメラの前で球と直方体をイメージして手振り動作をしてもらいシステム全体を通してどのように動作しているのかを検証した。また全体の平均動作時間は約3分で入力に約1分、出力に約2分となる。動作実験でディスプレイに表示されたポリゴンをそれぞれ図12,13に示す。

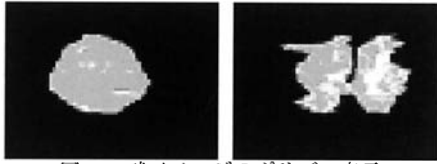


図 12 球イメージのポリゴン表示
Fig.12 Polygon display of ball image.

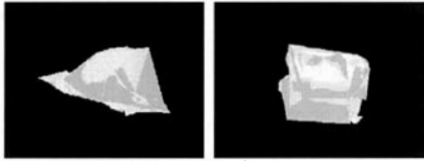


図 13 直方体イメージのポリゴン表示
Fig.13 Polygon display of hexahedron image.

6. 考察

本実験では、手領域抽出による手振り動作の取得は十分な結果が得ることができた。しかし、図 9 右にあるように球状にはなるものの、ポリゴンが欠けてしまった。これは点群にノイズが残っていたり、必要な点がなかったりしたためと考えられる。

7. おわりに

本研究では、ジェスチャを用いた 3 次元モデリングシステムを提案し、実際にシステムを検討しその有効性を確かめた。今後の課題として、リサンプリングの性能向上がある。ユーザのジェスチャの軌跡はユーザのイメージする 3 次元形状のすべてを表現するものではない。つまり、ユーザのジェスチャには欠如部分が必ず生じる。そこで、欠如部分を効果的に補完するリサンプリング手法の構築が必要である。

本研究で用いられた手振り動作はエックマン^[5] [6] によると例示的動作に分類され、発話と共に用いられている。今後の研究で発話による形状の判別を加え、精度を高められると考えられる。

参考文献

- [1] T.Igrashi, et al: Teddy-a Sketching Interface for 3D Freeform Design; ACM SIGGRAPH, (1999).
- [2] 黒川隆夫: ノンバーバルインターフェース; オーム社, (1994).
- [3] Corlett.N: The ergonomics of working postures; Taylor & Francis, (1986).
- [4] 谷口健男: FEMのための要素自動分割 - デローニー三

角分割の利用 - ; 森北出版, (1992).

- [5] 岡田謙一 他: ヒューマンコンピュータインタラクション; オーム社, 第 5 章 (2002).
- [6] W.ラフリー・エンジェル: ノンバーバルコミュニケーション; 大修館書店, (1981).
- [7] 森田哲也 他: カラー濃淡画像を対象とした手領域抽出のための表色系の検討; 社団法人電子情報学会, Vol.98 No.528, (1999).