

3次元人体動作の認識・生成とインタラクション

1B-03

中野 裕史 米元 聡 谷口 倫一郎
九州大学工学部 九州産業大学 九州大学大学院システム情報科学研究所

1. はじめに

実空間と仮想空間とをシームレスに融合する高度なインタフェースを実現するために、実世界におけるユーザから仮想空間への情報入力を効率的に行うための装置や技法の開発が進められている。また、近年の高度情報化社会の振興、発展に伴い、個人レベルの生活環境においてもインターネットを介した情報収集など、高度な情報操作を要求されることが多くなってきている。人間中心の情報操作インタフェース機器としては、デスクトップ型、モバイル型の計算機などが普及しているが、操作性に関して見てみると、人間すなわちユーザがシステムに合わせる形で行われている。しかしながら、人間の活動は主として3次元操作であり、人間がシステムに合わせるのではなく、システム側が人間の直接的な操作に合わせるスマートなインタフェースの実現が望まれている。特に、人間の3次元動作の直接的な入力を可能にするものには、モーションキャプチャに代表される3次元動作計測機器があり、装着するセンサ数を増やすことで詳細な動作を計測できるものまで開発され、実用レベルに達しているといえる。しかし、上述のスマートなインタフェースの実現という観点からも、なるべく装置の存在を意識させず[1]、装着、脱着の手間をとらない効率さを有し、しかもユーザの意図した動作をシステムが自動的に理解し仮想空間上で適切に再現できることが望ましいといえる。本研究では、そのような直接的な3次元動作入力を可能にする技術として、

コンピュータビジョンによる非接触な入力方式を用いる。

非接触での人体動作の計測は、効率やそのスマートさから仮想空間と実空間とのシームレス化に重要である[2][3]。ビジョンによる動作入力をを用いる場合、人体動作を精度よく推定することが一般に困難であるとされているが[4][5][6]、本研究ではこの問題に対処するために、比較的安定に推定可能な手や顔の動作情報から、リアルな人体モデルの動作を仮想空間上に生成することができる方法を提案する。また、3次元動作入力をもとに仮想空間の静的・動的なオブジェクトと円滑にインタラクションを行うための枠組みについても述べる。

2. 3次元人体動作の解析と生成

2.1 3次元位置の追跡

本研究では、人間の動作を画像より推定するための画像特徴として肌色領域を用いる。肌色領域としては、上半身では顔、両手部分が相当し、肌色という均一色の領域を人体部位とみなして観測する。2視点、あるいはそれ以上の複数視点の画像について同一部位の肌色領域が観測できる場合、ステレオ視の原理によってその領域重心の3次元位置を容易に計算することができる。領域重心位置を追跡するという問題においては、画像特徴の性質上、正確な3次元位置を推定することはあまり重要でないため、本研究では簡易な多視点ステレオ計算法[2]を用いる。対応のとれたそれぞれの視点における重心位置、及びカメラキャリブレーション情報より得られるカメラ座標原点を通過する直線の交点として重心の3次元位置を計算する。したがって上半身の動作においては、顔、両手の重心の3次元位置を求めることになる。(図1(上))。

Recognition and generation of human action in 3D and its application to human-machine interaction
Hiroshi Nakano: Faculty of Engineering, Kyushu University
Satoshi Yonemoto: Kyushu Sangyo University
Rin-ichiro Taniguchi: Faculty of Information Science and Electrical Engineering, Kyushu University

2.2 3次元人体動作の生成

仮想空間において、アバターの動作をユーザの動作に合わせてリアルに再現することは、インタラクションを遂行する際に正しく情報伝達を行うために必要不可欠である。そこで本研究では、画像解析により得られた少数の位置情報よりリアルな動作を生成するための姿勢推定法を導入する。リアルな動作を再現するには、身体構造系の厳密な物理特性に基づいて姿勢を計算することが望まれるが、すべてのユーザに共通する物理特性を予め獲得しておくことも困難であり、一般に画像解析により求まる位置、速度などの物理特性の精度も悪い上、少数部位の位置情報しか用いることができないことから、厳密なダイナミクス系として3次元人体モデルの動作生成を行うことは難しいと思われる。そこで本研究では、画像解析により得られる位置入力だけをもとに、近似的ではあるが物理法則に基づいた動作生成を行う方法を導入する。

まず人体の骨格モデルに対し、物理的な制約として重力及びリンクの長さは一定で互いに接続するという多リンク系の制約を条件に、画像解析により得られる位置変位をもとに全リンクの端点、すなわち関節点位置の均衡する位置を推定する。その後、各関節点のペアについて各々の形状を表す部位モデルの当てはめ計算を行い、並進、回転パラメータを求める。この手法により、画像解析では求めることが困難な肩や肘の位置などを物理法則に従うという範囲でリアルに再現することが可能となる。

人体の骨格モデルという多リンク系について、隣接する関節点を結ぶリンクを強力なバネモデルであると仮定し、入力される位置変位が生じてもリンク長を一定に保つよう均衡することを考える。また、計算に必要な速度成分は位置変位から近似的に求め、加速度については陽に考慮しないものとする。

具体的には、各リンクの端点である関節点 i, j のペアについて、以下の計算を収束するよう繰り返す。

$$(1) \mathbf{d} = \mathbf{p}_i - \mathbf{p}_j \text{ 及び } \mathbf{f} = \frac{\mathbf{d}}{\|\mathbf{d}\|} (\|\mathbf{d}\| - r_{ij})K \text{ を計算す}$$

る。ここで、 \mathbf{p}_i と \mathbf{p}_j は関節点の3次元位置であり、 r_{ij} は予め定義する $\mathbf{p}_i, \mathbf{p}_j$ 間のリンク長を表す。 K はバネモデルの弾性定数を表す。

- (2) $\mathbf{p}_i, \mathbf{p}_j$ を(1)の \mathbf{f} に基づき、刻み幅 Δ_f だけ更新する。

$$\mathbf{p}_i = \mathbf{p}_i + \mathbf{f}\Delta_f \quad (1)$$

$$\mathbf{p}_j = \mathbf{p}_j + \mathbf{f}\Delta_f$$

さらに、すべての関節点位置 $\{\mathbf{p}_k\}$ ($k=1, \dots, N$) について、重力に従って刻み幅 Δ_g だけ更新する \mathbf{g} は重力方向のベクトル。

$$\mathbf{p}_k = \mathbf{p}_k + \mathbf{f}\Delta_g \quad (2)$$

この繰り返し計算により、全関節点位置 $\{\mathbf{p}_k\}$ ($k=1, \dots, N$) が入力される位置変位に応じて均衡を保つ。ここではリンク長の制約、重力のみを考慮したが、衝突判定のための制約などの他の制約を式(1), (2)と同様に考慮することが可能である。

図1(下)に本手法により生成した上半身姿勢の例を示す。図中の3つの丸印が動作入力として与える位置であり、この位置とモデルの状態によりすべての上半身の姿勢が決定される。

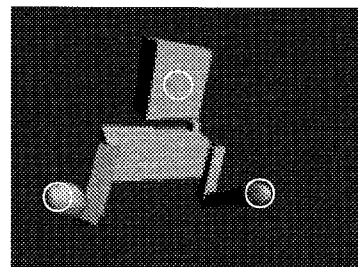
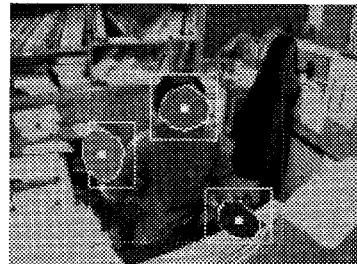


図1 上半身動作：(上)解析，(下)生成

3. インタラクションへの応用

我々は、ビジョン入力による 3D インタラクションを円滑に行うための枠組みとして、アフォーダンスの概念を利用する。実世界に対して提案されたアフォーダンスの概念を応用し、仮想空間中のオブジェクトが人体モデルに対し意味・価値のある情報、ここでは動作情報をアフォードするという捉え方である。アフォーダンスの概念を利用することは、実空間からの人体動作入力をもとにリアルな動作を生成するために以下の点で有用であると考えられる。

- 計測できない詳細動作をアフォードされる付加的な動作として追加できる
- アフォードされる動作情報として仮想空間上にシーン制約を課すことで、実空間ではありえない動作を抑制し、生成する動作をリアルにすることができる

前者の例としては、仮想オブジェクトを把持する、すなわち把持されるとアフォードした際に手指の詳細な動作を自動的に生成することが挙げられる。後者の例としては、ドアの開閉のように物体の運動が一定の規則に拘束される場合、アフォードされるドアの開閉動作もその運動規則に従う、などが挙げられる。本研究におけるインタラクションの様式として以下を実現している。

- 仮想空間上に置かれた静的な物体の把持、移動などの基本操作
- 動的オブジェクトとのインタラクション
- シーン制約を考慮したインタラクション
- 実物体を利用したインタラクション

4. 実験

プロトタイプシステムとして、ノート PC に IEEE1394 規格のデジタルカメラ 2 台を接続した、デスクトップ型の実時間インタラクションシステムを構築した。図 2 にシステムの概要、図 3 にカメラからの入力画像の例を示す。このシステムはデスクトップ型であるため、上半身動作のみを計測、生成の対象とする。

4.1 静的・動的オブジェクトの操作

図 4 に示すように仮想空間上の棚に置いたオブジェクトを、ユーザの動作に応じてアバターが把持、移動を行うことが可能である。また、動的オブジェクトとして人体フィギュアである仮想ロボットを考え、簡単なインタラクションの例としてアバターとの握手動作を実現した。これは、仮想ロボット側に握手をアフォードするよう実装することで実現している。

4.2 シーン制約を考慮したインタラクション

シーン制約利用の例として、仮想空間上のドアの開閉動作を考え、ドアの開閉動作はその運動軌道に正しく従うという制約を課した。図 5 に示すようにユーザの動作入力に従って、ドアの開閉に関する動作が手とドアノブの距離に応じてアフォードされ、アバターがドアの開閉を行う。ドアの動作は円運動に拘束されることより、アバターの手の位置もその軌道に再現されることになる。すなわちビジョン入力をきっかけにより正確な動作が再生されたことになる。

4.3 実物体を利用したインタラクション

実物体を利用することで、人間が直接入力することが難しい仮想オブジェクトの回転操作を実現した。実物体として、2つの識別用マーカを付与した平面パネルを用いた(図 6)。ユーザが実物体を回転させることで、対応する仮想空間上のディスク及びその上に置いた仮想オブジェクト(dino)を回転することができ、同時に自らの動作も再生することが可能であった。

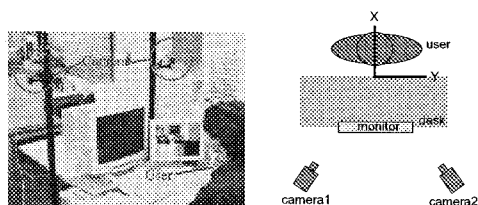


図 2 システムの概要



図3 2台のカメラからの入力

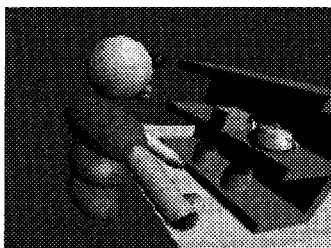


図4 仮想オブジェクトの操作

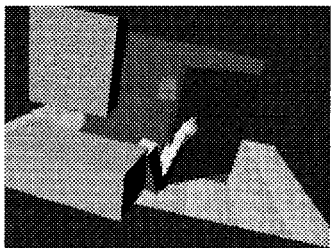


図5 ドアの開閉

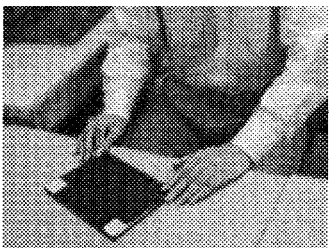
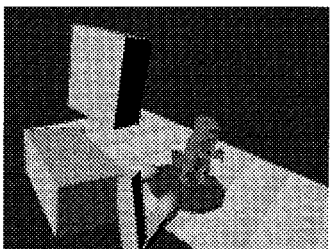


図6 仮想オブジェクトの回転操作

5. おわりに

円滑なインタラクションを実現することを目指し、人間の3次元動作を直接に入力すること

が可能で、しかも少数の入力情報からリアルな動作の生成が可能な3次元人体モデルの動作制御技術を開発した。また本手法では、アフォーダンスの概念を利用することによりシーン制約を考慮し、よりリアルなアバターの動作表現を行うことができる。今後、この枠組みを拡張し、より高次の動作の認識・生成を実現する予定である。

謝辞

本研究の一部は科学研究費補助金特定領域研究(C)13224074「個人適応能力を有する高次マルチモーダルユーザインタフェースに関する研究」の補助を受けた。

参考文献

- [1] M.Weiser, "Some Computer Science Issues in Ubiquitous Computing", *Communications of the ACM*, Vol. 36, No.7, pp75-84, 1993.
- [2] S.Yonemoto, D.Arita and R.Taniguchi, "Real-Time Human Motion Analysis and IK-based Human Figure Control," *Proc. Workshop on Human Motion (HUMO2000)*, pp.149-154, 2000.
- [3] S.Yonemoto, R.Taniguchi, "High-level Human Figure Action Control for Vision-based Real-time Interaction", in *Asian Conference on Computer Vision*, 2001.
- [4] C.Wren, A.Azarbayejani, T.Darrell, A.Pentland, "Pfinder: Real-Time Tracking of the Human Body," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol.19, No.7, pp.780-785, 1997.
- [5] C.Bregler, "Learning and Recognizing Human Dynamics in Video Sequences," *Proc. Computer Vision and Pattern Recognition*, pp.568-574, 1997.
- [6] D.Metaxas, *Physics-based Deformable Models Applications to Computer Vision, Graphics and Medical Imaging*, Kluwer Academic Publishers, 1997.