

温度画像のパターン認識を用いたジェスチャーの3DCG作成

(知能情報システム学) 浅田 太郎

1. 序論

近年、コンピュータグラフィック(以下「CG」と表記)によって作られたアニメーション(以下「CGアニメーション」と表記)をテレビや映画で多く目にすることができる。それらのCGアニメーションの中で、特に映画に登場する物は人間と違和感なく共演し、生きてるように動いている。そのようなCGアニメーションでは、人間の動作解析を行い、そのデータを用いてアニメーションを作成している。

人物動作解析の手法として、2つ以上のカメラを用いたモーションキャプチャー方式やデータグローブ等が利用されている。そして、これまで静止画像から人体姿勢の記述を得ようとする研究^[1-2]が多くなされてきた。しかし、それらの方法では動作の対象に環境的制約(複数台のカメラの準備、明るさの制御)や、物理的制約(マーカーやデータグローブの装着)が課せられることが前提になるため、人間に負担をかけない自然な状態での動作認識とは言い難い。これらの制約を解消するため、入力情報として赤外線温度画像を用いる方法が検討された^[3]。しかしながら、この方法では、2次元情報から3次元情報を作成することに伴う問題が残されている。

本研究では、人間に特殊な負担を課すことなく人物全身像の動きをCGに表現する方法を提案した。手段として赤外線温度画像を認識し、予め作成した3次元モデル(以下「モデル」と表記)を変形することにより3DCGアニメーションを作成した。本手法により、従来の環境的制約や物理的制約の問題が解決できる可能性を示した。

2. 本法の概要

アニメーション作成には次の2通りの手順を検討した。

2.1 特徴ベクトル空間におけるユークリッド距離を用いたアニメーション作成

初めに予め用意した人物姿勢パターンの学習用温度画像を読み込み、特徴ベクトルを求めることによりそれぞれの温度画像を学習する。そして、各学習用画像に対応する人物モデルを作成する。次にアニメーションにしたいジェスチャーの開始と終了の温度画像(以下、「開始画像」、「終了画像」と表記)を読み込む。読み込まれた開始画像に特徴ベクトル空間で最も近い人物姿勢パターン学習用画像(以下、「最近接学習画像」と表記)を計算で求め、また2番目に近い人物姿勢パターン学習用画像(以下、「第2近接学習画像」と表記)も求める。特徴ベクトル空間での開始画像、最近接及び第2近接学習画像の位置情報を用いて、開始画像に対応させる人物モデルにおける各関節の角度を求める。終了画像についても同様にして、人物モデルにおける各関節の角度を求める。

最初と最後のポーズの間をアニメーションにするため、2つの画像に対応させた人物モデルの違いを所定の軸回りの回転角として計算する。計算結果をもとに、VRML(Virtual Reality Modeling Language)アニメーションを作成する。

2.2 ニューラルネットワークを用いたアニメーション作成

初めに予め用意した人物姿勢パターンの学習用温度画像に対応する人物モデルを作成する。各学習用温度画像の人物領域をモザイク化し、そのモザイクのブロック数を入力層のユニット数、ブロックの平均濃度を入力値、人物モデルの関節の回転軸数+1(位置情報)を出力層のユニット数、出力値を関節の回転角度とする学習データを作成し、BP(Back Propagation)ニューラルネットワーク(以下、「NN」と表記)を用いて学習する。次にアニメーションにしたいジェスチャー

の開始画像と終了画像を読み込む。学習画像と同様にモザイク化を行い、各ブロックの平均濃度を入力値とし、NN を用いて各画像に対応させる人物モデルにおける各関節の角度を求める。終了画像についても同様にして、人物モデルにおける各関節の角度を求める。

最初と最後のポーズの間をアニメーションにするため、2つの画像に対応させた人物モデルの違いを所定の軸回りの回転角として計算する。計算結果をもとに、VRML アニメーションを作成する。

3. 特徴ベクトル空間におけるユークリッド距離を用いた認識

本研究では、濃淡画像を元に特徴ベクトルを作成する。まず、複数個の基本動作パターンを設定しそれを赤外線温度画像として撮影し、静止画としてコンピューターに入力後、人物なしの背景画像とで差分をとる。この時、2値化前の中間値処理や2値化後の膨張、収縮、孤立点除去を行うことで、ノイズを効率良く除去し、濃度“0”となった画素において、原画像の濃度を“0”とした濃淡画像を作成、アフィン変換による位置とサイズの規格化を行う。人物のみの領域を取り出し、画像をモザイク化して各ブロックの平均濃淡値を要素とする特徴ベクトルを作成する。そして、認識用画像に対して予め学習した各クラス（人物姿勢）との特徴ベクトル空間におけるユークリッド距離を求め、その距離値が最も小さなクラスを認識結果とする。本法では距離値が2番目に小さいクラスも人物モデル作成に用いる。

4. NNを用いた認識

3節記載の方法を用いて学習画像をモザイク化し、ブロック数を入力層のユニット数、人物モデルの関節軸数+1(位置情報)を出力層のユニット数としたNNに、各ブロックの平均濃淡値を入力値、学習画像に対応する人物モデルの関節での回転角度と位置情報を出力値として学習する。認識用画像において、同様の方法で作成した入力値を、NNに入力して得られる出力値を認識用画像に対応する人物モデルの関節での回転角度と位置情報とする。

5. 人物モデルの作成

人物モデルはVRMLを用いて作成した(図1)。関節を12箇所(回転軸数24)設定し、各関節に用意された角度ファイルの角度情報をJavaによって読み込むことにより関節が曲がるようにした(図2)。同様に人物モデルの階層構造で最も上位である腰の位置情報ファイルも用意しjavaで読み込み、屈伸のような上下運動も可能にした。

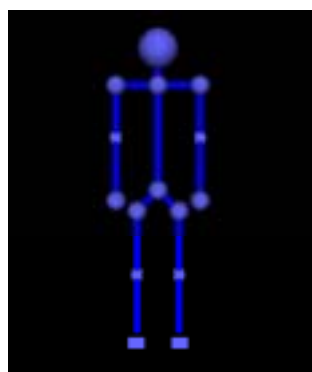


図1 人物モデル

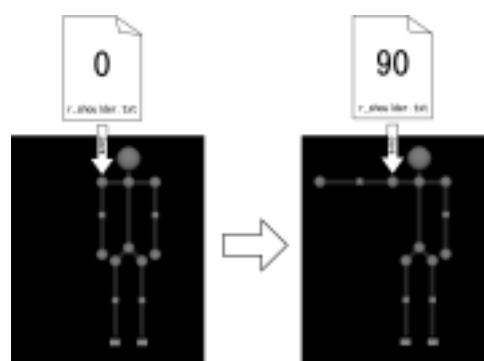


図2 関節を動かした例

6. 特徴ベクトルを用いた認識画像対応人物モデル作成方法

3節で記載の方法を用いて最近接学習画像と、第2近接学習画像を求める。特徴ベクトル空間で最近接学習画像の点と第2近接学習画像の点を結ぶ直線に認識画像の点から垂線を下し、垂線と上記直線との交点と、最近接学習画像の点及び第2学習画像の点との距離の比を計算し(図3)、この内分比を用いて人物モデルの各関節の角度を2つの学習画像に対応付けた2つの人物モデル

の関節の角度の内分で求める。後述する実験では、最近接学習画像だけを用いたアニメーションも比較のため作成する。

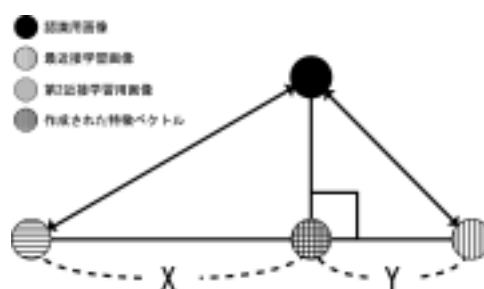


図3 特徴ベクトル作成

7. 実験

7.1 温度画像の作成条件

赤外線画像装置(ニコン製サーマルビジョンシステム LARD-3ASH)を用いて、検出温度領域を人物全身が目視で認識でき、かつ、人物の周囲ができるだけ現れない温度範囲(27.6~36.4)にし、71個の基本姿勢パターンの温度画像を3回分8mmビデオに録画した。しかる後、mpeg2ファイルを作成し、所定のタイミングで256階調380×460画素のビットマップ画像を取得した。この71個の基本姿勢パターンは体操やリハビリ等の文献により決定された姿勢である。

7.2 特徴ベクトルを用いたパターン認識の精度検証実験

特徴ベクトルを用いたパターン認識の精度を検証する実験を行った。実験に用いる温度画像は7.1節の実験条件で作成したものをを用いる。25×25画素を1つのブロックとし、そのブロックでの濃淡値の平均を求める。実験で使用したブロック数は14×14である。対応する特徴ベクトル作成用モザイク画像の作成例を図4に示す。学習用基本姿勢パターン各2回分の特徴ベクトルを平均することにより学習用の特徴ベクトルを作成した。71個の基本姿勢パターンを認識した結果、正答率は91.1%(65/71)であった。誤認識は、上半身のポーズは同じだが下半身のポーズが違うような、よく似ている画像の間で生じていた。

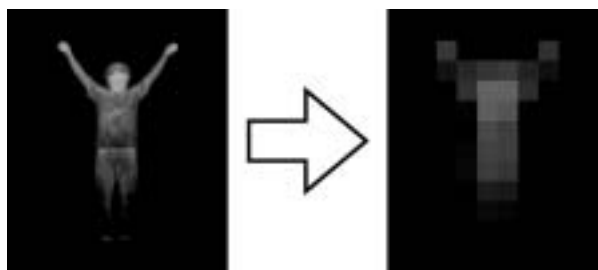


図4 特徴ベクトル作成に用いるモザイク画像の作成例

7.3 特徴ベクトル空間のユークリッド距離を用いた認識によるアニメーション作成実験

7.2節と同様の実験条件で数秒間の人物動作を赤外線動画画像として撮影し、その動作を2通りの作成条件で3DCGアニメーションにする実験をした。条件は以下の通りである。

- (1) 開始画像と終了画像及び、中間時点の画像1枚をパターン認識する。最近接学習画像のクラスを認識結果とし、対応する人物モデルを作成して、開始から中間時点を経て終了までの人物モデルをつなぐアニメーションを作成する。
- (2) (1)と同様に開始画像と終了画像及び、中間時点の画像1枚をパターン認識する。最近接学習画像のクラスと第2近接画像のクラスを用いて6章記載の方法で対応する人物モデルを作成して、開始から中間時点を経て終了までの人物モデルをつなぐアニメーションを作成する。

条件(2)の方が(1)より入力画像に近いアニメーションとなった(図5)。

7.4 ニューラルネットワークを用いた認識の精度実験

7.1節の実験条件で作成した71個の基本姿勢2パターンをクラスタリングした後、NNで学習し、3パターン目を認識し認識結果の精度を4段階の評価で表した(表1)。クラスタリングを行

ったものについては、認識用画像が正しいクラスタに分類されているかを検討した(表2)。

クラスタリングを行うことにより良好な姿勢が作成される割合が増す傾向がみられた(表1)。この傾向は学習画像の数の増加に従い、より顕著にあらわれると考えられるので、学習画像の増加するにつれクラスタリングの効果がさらに顕在することが期待される。

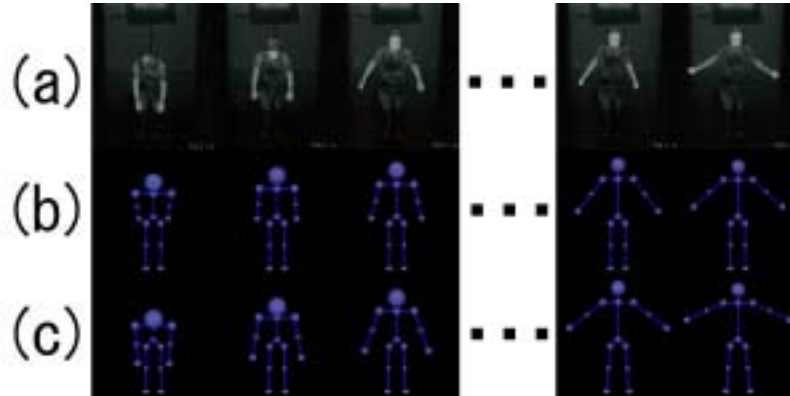


図5(a) 赤外線温度画像, (b)7.3(1)の方法でのアニメーション, (c)7.3(2)の方法でのアニメーション

表1 評価結果

		クラスタ数			
		1 (クラスタリングなし)	3	5	10
認識結果	(良好な姿勢)	41	45	43	51
	(正しくない箇所がある)	26	20	21	17
	(姿勢があまり良くない)	2	3	4	1
	× (異なった姿勢をとっている)	2	3	3	2
	合計	71	71	71	71

表2 認識用画像のクラスタリング結果

	クラスタ数		
	3	5	10
クラスタ正しく認識	67	62	66
クラスタ誤認識	4	9	5
合計	71	71	71

8. 結論

赤外線温度画像をパターン認識し、その情報を用いてジェスチャーを3Dコンピューターグラフィックスに表現する手法を提案した。今後、よりリアルなアニメーションを作成するために人物姿勢の合成法の改善などを行っていく。

参考文献

- [1] A. Pentland: Automatic Extraction of Deformable Part Models; International Journal of Computer Vision, 4, pp.107-126 (1990).
- [2] 亀田, 美濃, 池田, : シルエット画像からの関節物体の姿勢推定法; 信学論 (D-11), J79-D-11, 1, pp.26-35 (Jan. 1996).
- [3] 岩澤, 海老原, 大谷, 中津, 森島, : 熱画像を用いた人物全身像の実時間姿勢推定; 映像情報メディア学会誌, Vol51, No. 8, pp. 1270-1274 (1997).