

表情追跡システムと人形型入力デバイスを用いた 人物上半身のリアルタイムアニメーション

新井 清志* 坂本 浩**

* (株)日立製作所 中央研究所

** (株)フジテレビジョン 美術制作局CGセンター

本報告では、モーションキャプチャの新技术を用いて日立製作所とフジテレビジョンが共同開発した人物のリアルタイムアニメーションシステムについて述べる。モーションキャプチャの技術は、人物アニメーションの動作データの作成に幅広く利用されている。しかし、従来の手法では、データに対する後処理の発生が原因となり、この技術によって人物のリアルタイムアニメーションを生成することは困難だった。本開発では、この問題を解決するため、表情追跡システムからの入力を高速表情生成手法で処理するとともに、人形型入力デバイスからの入りに予約姿勢を混ぜ合わせる処理を行なった。

REAL-TIME ANIMATION OF THE UPPER HALF OF THE BODY USING A FACIAL EXPRESSION TRACKER AND AN ARTICULATED INPUT DEVICE

Kiyoshi Arai* Hiroshi Sakamoto**

* Central Research Laboratory, Hitachi, Ltd.
1-280 Higashi-Koigakubo, Kokubunji, Tokyo 185, Japan

** CG Center, Art Production Dept., Fuji Television Network, Inc.
3-1 Kawada-cho, Shinjuku, Tokyo 162, Japan

This paper describes a real-time human animation system using a new method of motion capture developed by Hitachi, Ltd. and Fuji Television Network, Inc. Motion capture technology is widely used for making motion data of human animation. In existing methods, however, this technology was difficult to apply to real-time human animation because of data post-processing that was necessary. In our work, we have solved this problem by introducing a rapid facial deformation method for processing the input from a facial expression tracker, and reserved poses blended with real-time input from an articulated input device.

1. はじめに

インターネットの普及に伴い、個人から公共機関、企業によるホームページが多数開設されている。これらの目的はそれぞれ多岐にわたるが、共通の課題は「アクセス数を確保できる魅力あるコンテンツ制作」である。コンピュータグラフィクス(CG)を用いたアニメーションは、このような映像表現に有効である[1]。しかし、CGアニメーションの制作には一般的に時間と費用がかかる。より手軽にCG映像を制作するには、背景を含めたすべての映像を3次元CGで表現するのではなく、写真や絵を背景としてCGを合成する方が効果的である。実写とCGの合成は、映画や商業フィルム等で多く用いられているが、煩雑な操作が必要である。本報告では、2次元映像と3次元CGを簡単に合成するのに有効な、任意の視点移動を表現する対話的画像合成手法を提案する。

2. CGと実写合成の課題

CGと実写の自然な合成映像を得るためには、実写撮影のカメラの位置、方向、画角といったカメラアングルを、CG映像制作におけるカメラアングルと一致させる必要がある。このために用いられる一般的な手法は大別して次の3通りある[2]。

- (1) 実写を撮影した後CG映像を制作：
撮影時にカメラワークデータを記録、または後処理でカメラパラメータを抽出。
- (2) CG映像を制作した後実写を撮影：
CG映像制作で用いられたカメラパラメータを使って、自動的に駆動する「モーションコントロールカメラ」により実写を撮影。
- (3) CGの3次元空間内に実写映像を合成：
舞台装置をCGで表現して登場人物のクロマキー実写映像を合成。

上記のいずれの手法もCGと実写の合成映像制作には有効な手段であるが、時間と費用を要する。そこで、CGと実写のカメラアングルを一致させるのではなく、静止画(またはカメラ固定の動画)をCGのカメラアングルに合わせて変形できれば、実写とCGの合成は飛躍的に容易になると考えた。本研究では、簡単なユーザインタフェースにより、画像の3次元形状モデルを推定し、別のカメラアングルから見たような画像に再合成する手法を試みた。

3. 形状推定と画像再合成

本手法では、写真や絵などの2次元画像から、カメラアングルを移動した画像を生成するために、オリジナル画像の3次元形状モデルを推定する。入力に用いる画像は、撮影時の特別な処理は不要であり、後処理も無い。ユーザーは、消失点と架空の正面壁を簡単な2次元のユーザインタフェースで設定する。これら2つの情報から、多くの背景写真や絵に適用可能な暗黙の仮定を取り入れ、3次元形状を推定する。図1に本報告で述べる「一枚の画像を別のカメラアングルから見たように再合成する」手順を示す。

- (a) 写真や絵などの2次元画像をスキャナーなどで計算機に取り込む。
- (b) 市販のペイントツール等で、人物や街灯など背景と区別されるべきオブジェクトを消去した画像と、オブジェクトを白く塗りつぶした濃淡のマスク画像を作成する。
- (c) 放射状ガイドラインの2次元ユーザインタフェースを用いて、画像の消失点などを設定する。
- (d) 計算機は、設定された消失点などの情報から、3次元形状モデルを自動推定する。
- (e) ユーザにより、背景と区別されるべきオブジェクトの領域を2次元の四角形で設定する。

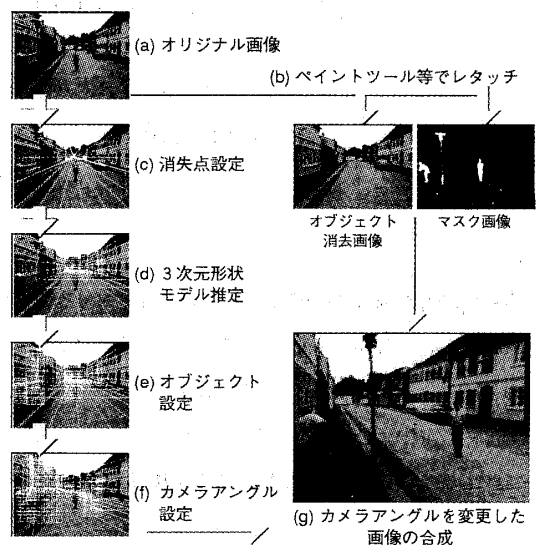


図1 形状推定と画像再合成の手順

以上が前処理である。以下の(f), (g) を対話的に繰り返し行う。

(f) カメラアングル（視点および視線方向）を設定する。

(g) 与えられたカメラアングルから見たような画像を合成する。

ここで視線方向とは、視点から注目点（参照点）の方向へのベクトルを示す。(f), (g) を対話的に行い2次元画像の中をウォークスルーする他に、(f) のカメラアングルを複数指定し、これら之間を補間したカメラアングルを用いて (g) の処理を行い、動画画像を生成することもできる。次節以降に各ステップの詳細について述べる。

3. 1 消失点の設定

本節では、オリジナル画像の消失点と架空の正面壁を設定する方法について述べる。これらの設定は次節で述べる3次元形状モデル推定の手がかりとして用いられる。

図2 (a) は画像を読み込んだ直後のウィンドウを示す。画像に上書きされた放射状の線は、消失点VPと画像の端を結んだガイドラインである。ユーザは、このガイドラインを道や建物の稜線に沿うようにして、消失点を設定する。また架空の正面壁の領域を長方形Rにより設定する。図2 (b) に設定終了時のウィンドウを示す。通常、初心者でも1分以内でこれらの操作を完了することができる。



図2 消失点の設定

また、消失点が画面の外にある場合、長方形Rは消失点の位置により自動設定される。

3. 2 背景の3次元形状モデル推定

本節では、3次元形状モデルを推定する手法について述べる。前節で設定された消失点と架空の正面

壁の2つの情報と暗黙の仮定より、3次元形状モデルを推定する。ここで用いる仮定を以下に示す。

- (1) 背景となる領域は、左/右/正面の壁、天井、床の5つのいずれかに投影可能。
- (2) 左と右の壁、天井と床はそれぞれ互いに平行。
- (3) 左と正面の壁、正面と右の壁はそれぞれ互いに垂直。
- (4) 左/右/正面の壁は天井および床に対して垂直。
- (5) 正面の壁は画面に対して平行。
- (6) 床は3次元空間上の $y=0$ 平面上に存在。
- (7) 初期カメラの視線方向は左右の壁、天井、床に平行。

一般には、視線方向と画面（スクリーン）が交わる点は画面の中心であるといった仮定を用いることが多いが、本手法では、消失点を視線方向と画面の交点とする。これらの仮定のもと、画像を左/右/正面の壁、天井、床の5つの領域に分割し、それぞれ領域の3次元座標を求めることにより、3次元形状モデルを推定する。まず、図3 (a) に示すように、5つの領域に分割する。

- 床 : 長方形Rの下辺を上底とし、消失点VPと上底の両端を結ぶ直線が画像の下辺と交わる点を両端とする台形。
- 左壁 : 長方形Rの左辺を上底とし、消失点VPと上底の両端を結ぶ直線が画像の左辺と交わる点を両端とする台形
- 右壁 : 長方形Rの右辺を上底とし、消失点VPと上底の両端を結ぶ直線が画像の右辺と交わる点を両端とする台形
- 天井 : 長方形Rの上辺を上底とし、消失点VPと上底の両端を結ぶ直線が画像の上辺と交わる点を両端とする台形
- 正面壁 : 長方形R。

次に、各領域は3次元空間では長方形とし、それぞれの頂点の3次元座標を求める。一般にカメラアングルが既知の場合、2次元座標(h, v)に投影された点の3次元座標(x, y, z)は、この点を含む3次元空間上の任意の平面($ax + by + cz + d = 0$)を規定すれば求められる。

まず、図3 (b) に示すように、 $y=0$ 平面に含まれている1から8までの頂点の3次元座標を求める。次に、 $y=0$ 平面に垂直かつ頂点5および6を含む

平面を規定する。この平面に含まれる頂点9および10の3次元座標を求める。これらの頂点のy座標の平均をHとする。最後に $y=H$ 平面に含まれる頂点11から16までの頂点の3次元座標が求められる。

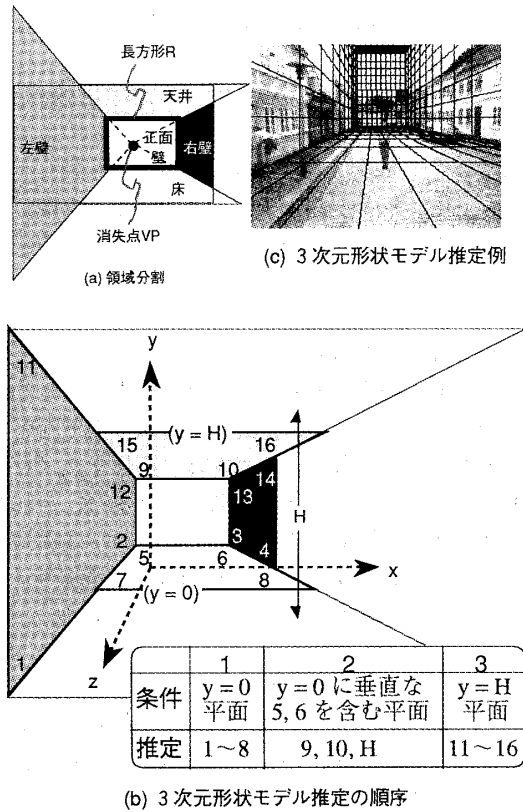


図3 3次元形状モデル推定

以上により、2次元画像を手がかりにして与えた消失点と正面壁の長方形領域の2つの情報から、背景の3次元形状モデルを自動生成できる。図3(c)に推定後の3次元形状モデルをワイヤーフレーム表示した例を示す。

3.3 オブジェクトの3次元形状モデル推定

前節までで、左/右/正面の壁、天井、床の5つの領域に投影できる背景画像の3次元形状モデル推定を行った。本節では、これらの領域に投影できないオブジェクトの3次元形状推定について述べる。ここでオブジェクトとは、例えば人物や街灯など、床や壁から突起している物体を示す。本手法では、オブジェクトの画像領域を2次元の四角形で大まかに

に設定し、3.2節と同様の手法で3次元推定する。あらかじめオブジェクトを白く塗りつぶした濃淡のマスク画像が用意されているため、大まかに設定された領域内でマスク情報を用いることによりオブジェクトの切り出しが可能となる。以下操作と処理の手順について述べる。

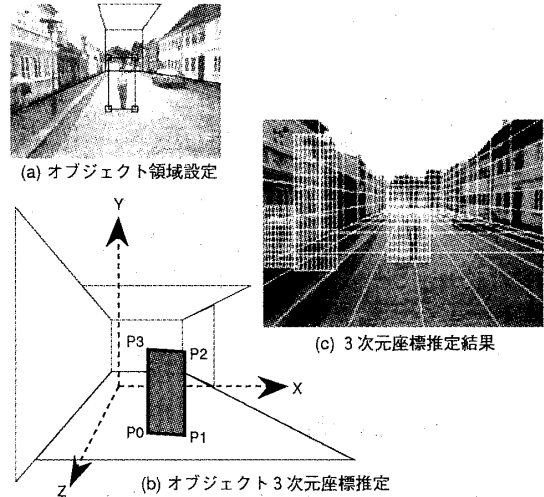


図4 オブジェクト3次元形状モデル推定

まず、図4(a)に示すようにオブジェクトを内接する四角形を設定する。この四角形の各頂点は自由に座標を変更することができ、長方形である必要はない。四角形の設定後、このオブジェクトが、左/右/正面の壁、天井、床の5つのうちのどの面に接するかを設定する。次に設定された四角形の4頂点と接する面から3次元形状を推定する。図4(b)に床に接する場合を示す。まず、床($y=0$ 平面)上にある頂点 P_0 および P_1 の3次元座標を求める。次に頂点 P_0 および P_1 を含む、床に垂直な平面を規定し、この面に含まれる頂点 P_2 および P_3 の3次元座標を求める。

図4(c)に中央と左右の人物、街灯の3次元座標を推定した結果を示す。以上の前処理により、背景(壁・床・天井)およびオブジェクトの3次元形状モデルが得られる。

3.4 対話的画像再合成

本節では、前節までに得られた背景とオブジェクトの3次元形状モデルに対するカメラアングル(視点の3次元座標、視線方向ベクトル、画角)の設定

と、与えられたカメラアングルから見たような画像に再合成する方法について述べる。

3次元形状モデル推定において、カメラアングルの初期値は与えられており、ここでは、この値に差分を加える操作を施すことになる。マウスなどにより、以下の方法でカメラアングルを設定する。

- (1) 上下左右回転
- (2) 上下左右移動
- (3) 前後移動
- (4) 画角
- (5) 回り込み

このユーザインタフェースの他、直接カメラアングルのパラメータを入力したり、3次元CGオーサリングソフトのカメラファイルから読み込むことも可能である。

次に、推定された3次元形状モデルを、別のカメラアングルから見た画像に再合成する手法について述べる。図5(a)に、カメラアングル変更後の3次元形状モデルを示す。表示領域の全てのピクセル(h_0, v_0)について、以下の処理を行う。

- (1) ピクセル(h_0, v_0)が背景(左/右/正面の壁、天井、床のいずれか)の場合、このピクセルの3次元空間における座標(x, y, z)を逆透視変換により求める。次に図5(b)に示すように、この3次元座標(x, y, z)のカメラアングル変更前の2次元座標(h_1, v_1)を透視変換により求める。さらに、図5(c)に示すように2次元画像のうちオブジェクト除去画像から座標(h_1, v_1)に対応する色情報を取得し、この色で元のピクセル(h_0, v_0)を描画する。

- (2) ピクセル(h_0, v_0)がオブジェクトの場合、このピクセルの3次元空間における座標(x, y, z)を逆透視変換により求める。次に、この3次元座標(x, y, z)のカメラアングル変更前の2次元座標(h_1, v_1)を透視変換により求める。さらに、2次元画像のうちマスク画像から輝度を、オリジナル画像から色情報Aを、オブジェクト除去画像から色情報Bを取得する。色情報AおよびBにそれぞれ輝度の重み付けをおこない、色情報Cを得、この色で元のピクセル(h_0, v_0)を描画する。

以上の処理により、カメラアングル変更後の合成画像が得られる。上記手法の原理は一般のテクスチャマッピング法と同様である。しかし、正面から見たテクスチャ画像を用意しない。このため、手前と奥での解像度は保持される。

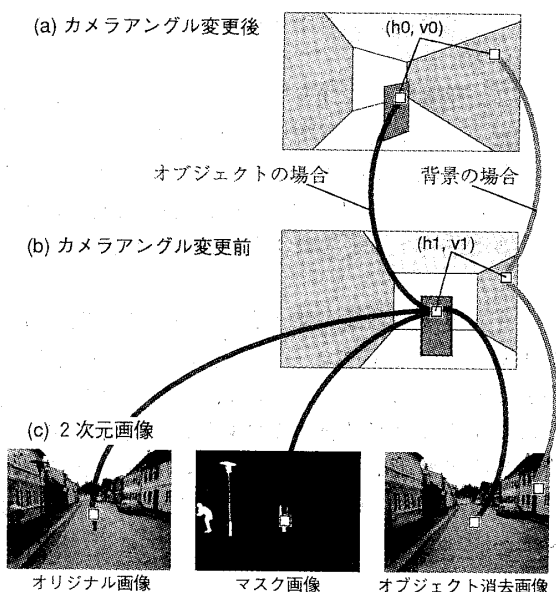


図5 画像再合成

上記の処理をカメラアングルの設定と交互に行うことにより、インタラクティブに画像の中をウォークスルーすることが可能となる。しかしながら、計算機の計算速度の制約から、全てのピクセルを描画すると、快適にウォークスルーできないことがある。このため、場合によってはファットピクセル(2x2や4x4ピクセル)を用いて解像度を低くする。また、入力画像と出力画像のサイズが異なる場合、図5(b)の(h_1, v_1)を拡大(または縮小)して対応する色情報を得る。

4. 画像再合成結果

本節では、本報告で述べた視点移動画像再合成と3次元CGオーサリングソフトと組み合わせて用いた例について述べる。図6(a)は3次元CGによる動画である。歩いている人物に向かって、カメラが遠くから近くまで移動した例である。図6(b)は本手法による動画で、カメラアングルの始点と終点は、(a)で用いたカメラアングルのデータを用いて、間を補間した映像を生成した。図6(c)は、(a)を(b)にオーバーラップして合成した動画である。本手法を用いることにより、煩雑な操作なしに実写

や手描き画像と3次元コンピュータグラフィクスを自然に合成することができる。

本手法のユーザからみた特徴は、以下の3点である。

- (1) 普通のカメラやビデオ、手描きの画像を用いる。
- (2) 平易なユーザインタフェースで消失点を指定し、3次元形状モデルを自動推定。
- (3) 3次元形状モデルに対するカメラアングルを設定し、別の視点から見た画像に変換。また、複数設定することにより動画像を生成。

(1) 普通の画像が使えるメリットには以下の4つがある。(i) 特殊なハードが不要で、素材収集が容易である。(ii) 事件、古写真、記念写真など再度と撮影するのが困難な画像にも対応できる。(iii) 手描きの絵が使える。これはセルアニメーション制作に適する。(iv) すでに保有している素材を有効利用できる。

(2) 平易なユーザインタフェースはスピーディーなコンテンツ制作に必須である。本報告で述べる手法は、消失点と架空の正面壁を2次元の平易なユーザインタフェースで与えるだけで、3次元形状モデルを自動推定する。

(3) カメラアングルの設定は、あたかもその場所に居て、あたりを見回す感覚で行うのが好都合である。また、このカメラアングルの情報は3次元CGオーサリングソフトと互換性があり、実写とCGの合成を容易にする。

5. おわりに

本報告では、「一枚の画像を別のカメラアングルから見たように再合成する」手法を開発した。本手法により一枚の画像の中をウォークスルーできる他、動画像を生成できる。本手法を用いて3次元CGと合成した結果、少ない操作で合成映像が得られることが確認できた。

参考文献

[1] 新井、堀井：二次元ドローデータによる三次元多関節構造体のモデリング；情報処理学会グラフィクスとCAD 79-3, pp 15-22.

[2] 曾利：「乱歩～妖しき女たち～」におけるデジタル映像技術；テレビジョン学会技術報告 pp.13-18(1995).

※本稿は、情報処理学会 グラフィクスとCAD 研究報告 No.81-11 (1996.8.23) を再掲載したものです。

(a) 3次元CGによる動画像



(b) 本手法による動画像
(カメラは(a)を使用)



(c) (a)と(b)の合成動画像



図6 3次元CGとの合成