

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

人物モデルのCG映像は、映像制作の大切な素材の一つである。放送[1][2][3]や家庭用ゲーム機[4]等のコンテンツ制作においては、低コスト、短期間で映像の量産が要求されるので、人物モデルのCGアニメーションを少ない工数で作る技術が必要である。人物モデルのアニメーションで要求されるリアルな、複雑な、あるいは長時間の動作データを効率良く作成するため、実際の演技者の顔[5][6]や体[3]や人形[7]の動きを測定して動作データとする、モーションキャプチャと呼ばれる手法が用いられている。

本研究では、モーションキャプチャにおけるデータの処理方法を改良し、最も効率良く映像を制作する手段であるリアルタイムアニメーションシステムの開発に適用した。

2. モーションキャプチャを用いたリアルタイムアニメーション

モーションキャプチャを用いた人物モデルのリアルタイムアニメーションシステムの構築にあたり、以下の課題に取り組んだ。

(1) 表情を含む人物映像のリアルタイム化

人物モデルの複雑な表情計算は、モーションキャプチャの後処理で行なうシステムが多い。このようなシステムでは、映像制作者の意図に合わせた表情の演技の調節が難しい。

本研究では、制御点を用いた高速表情生成手法[8]を用いて、顔のモーションキャプチャと同時に複雑な表情の映像を生成できるようにした。

(2) 腕の動き等に対する後処理の防止

一般に、演技者や人形と、人物モデルの腕の長さは互いに異なるため、モーションキャプチャにおいて手を合わせたり、手が顔に触れたりする動作の入力は失敗しやすく、得られた動作データを修正する後処理が発生する。リアルタイム映像を生成するためには、後処理を回避する工夫[3]が必要である。

本研究では、難しい姿勢を予約姿勢データとしてあらかじめ登録し、リアルタイムで入力されるデータと自然に混合することにより、後処理の発生を防止した。体の動きを計測するよりもコストの面で有利な人形型の入力デバイスを用い、人物モデルの上半身をモーションキャプチャの対象とした。

以下、開発したリアルタイムアニメーションシステムと、本システムを用いた人物CGの長編映像制作について述べる。

3. 表情追跡システムを用いた顔の制御

顔のモーションキャプチャは、被測定者の顔に貼ったマーカの動きを追跡する市販のシステムを用いて実現し、高速表情生成手法によって顔を制御した。

3.1 表情追跡システムの概要

ドイツ VIERTE ART GmbH の表情追跡システム X-IST は、図 1 に示すように顔の被測定者が装着するヘルメットと、ヘルメットの先に取り付けられた赤外線カメラ、および PC から成る。被測定者は顔の表面にマーカを貼る。赤外線カメラからの入力に対して PC が簡単な画像処理を行ない、RS232C によってマーカの位置情報を毎秒30回程度シリコングラフィクス社製の GWS ONYX (4 CPU) に送信する。マーカの位置情報を受信した表情用通信ソフトは、受信内容を表情生成用共有メモリに書き込む。この内容を表情データ等を生成するソフトが読み、高速表情生成手法の制御点の移動量に変換して、毎秒30回程度人物モデル描画用共有メモリに書き込む。人物モデル描画ソフトウェアは、人物モデル描画用共有メモリの内容を読みながら、毎秒15フレーム程度人物モデルを描く。

マーカは最大32個まで貼ることができる。ただし、画像処理を単純かつ高速にするため、赤外線カメラの入力画像上で、各々のマーカの移動範囲の外接矩型が互いに重なってはならない、という制約がある。検討を重ねた結果、眉毛に4個、上まぶたに2個、鼻の脇に2個、頬に2個、口のまわりに4個、計14個のマーカを顔の被測定者に貼った。顔に対する赤外線カメラの位置が不安定だと、画像処理の精度が落ちるので、被測定者の頭部を圧迫せずに赤外線カメラの位置を長時間安定させるため、ヘルメットには大幅な改造を施した。なお、X-IST では視線の追跡はできないので、顔モデルの視線制御は人形型入力デバイスの補助入力機器で実現した。

3.2 顔モデルの変形処理

14個のマーカの位置情報は、高速表情生成手法[8]の制御点の移動量に変換される。この表情生成手法は、顔モデルの表面を展開して得られる二次元パラメータ空間上に制御点を配置し、制御点に与えた移動量を線形内挿して顔モデルの表面を変形し、表情を生成する手法である。本システムでは、表情生成に適した約4500個の三角形ポリゴンから成る顔の「基本モデル」を顔モデルとして用い、基本モデルのパラメータ空間上に99個の制御点を配置して表情を生成した。基本モデルとその表面上にマッピング

グした表情生成用の制御点を図2に示す。

マーカの縦横それぞれの移動量がしきい値を超えた分に対して、移動の方向の正負によって別のスケールリングを行ない、マーカの「補正移動量」を求めた。そして、各々の制御点の縦横それぞれの移動量は、マーカの縦横それぞれの補正移動量の重み付きの和に基づいてリアルタイムで決定した。

4. 人形型入力デバイスを用いた上半身の制御

体のモーションキャプチャは人物上半身モデルを対象とし、新たに開発した人形型の入力デバイスを用いて実現し、予約姿勢を利用して後処理を防いだ。

4.1 人形型入力デバイスの概要

人形型入力デバイスは、図1に示すように人形師が手で操作する人形本体、足で操作する補助入力機器、およびこれらからの入力を変換するコントローラから成る。人形本体は、人形師が操作しやすいように、実際の操り人形と同じ形状をしている。人形本体からは、肩、ひじ、手首からそれぞれ左右2自由度ずつ、腰から3自由度、首の複数の位置から4自由度、合計19自由度の関節角度を検出する。補助入力装置は2自由度、1自由度、1自由度の3つのペダルから成り、それぞれを用いて人物モデルの視線、肩と胴体の上下運動、予約された姿勢の「重み」を制御する。以上合計23次元の数値情報をコントローラがデジタル信号に変換し、RS232Cによって毎秒30回程度 ONYX に送信する。コントローラから数値情報を受信した人形用通信ソフトは、受信内容を

毎秒30回程度人物モデル描画用共有メモリに書き込み、このメモリの内容に基づいて人物モデル描画ソフトウェアが毎秒15フレーム程度描画を実行する。

4.2 予約姿勢データの登録

人形型入力デバイスを使用するには、まず人形本体をゆっくり動かして難しい姿勢の各々を作り、これを「予約姿勢データ群」に登録しておく。そして、動作を収録する際には、予約姿勢データ群から選ばれた1つの予約姿勢データと、人形本体からリアルタイムで入力される姿勢データとを、補助入力装置のペダルから入力された重みを用いて混合する。

難しい動きを予約姿勢によって実現することの利点は、予約姿勢を作る作業はリアルタイムではなくゆっくりの動きでできることと、同じ姿勢が確実に再現されることである。

4.3 予約姿勢データの利用例

姿勢の予約および動作の収録の際の、人形本体の状態とGWS上の人物モデル映像との関係の一例を図3に示す。状態1は、人物モデルの左手が顔に触れる姿勢を予約姿勢データとして登録する瞬間の人形本体の姿勢を表わしている。映像1は、このときの人物モデル映像を表わしている。この姿勢を作るのに必要な関節は、首、左肩、左ひじ、左手首の関節である。これらの関節の指定はユーザが行なう。姿勢予約の際には、人形本体から測定される関節角度と、各々の関節が必要か否かを1と0で表わす判別値とから成るデータを、予約姿勢データ群の中の1つの予約姿勢データとして登録する。判別値が1

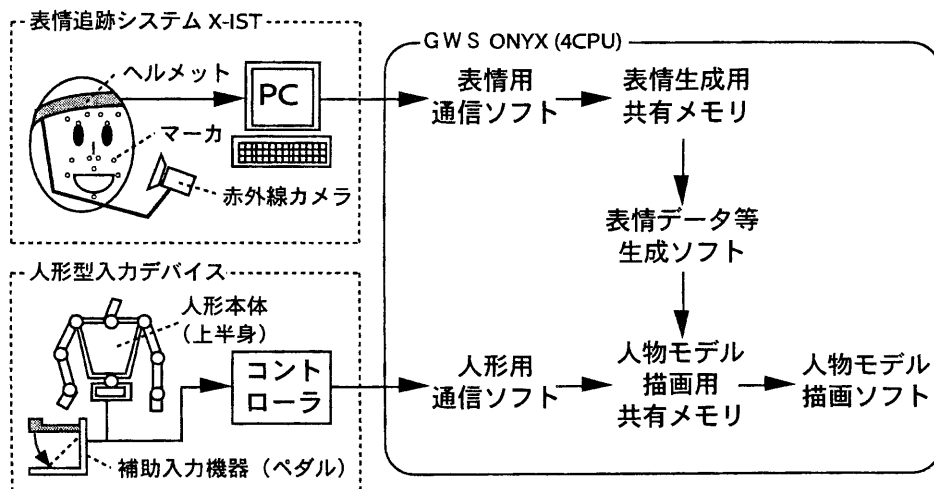


図1 人物上半身のリアルタイムアニメーションシステム

である関節の関節角度だけが、動作収録処理の際に混合処理の対象となる。判別値を変更すれば、混合処理の対象となる関節を変更することもできる。

状態2は、動作を収録する処理のある瞬間における人形本体の姿勢を表わしている。このとき、予約姿勢データ群の中から、状態1によって登録した予約姿勢データが選ばれているものとする。この予約姿勢データの判別値が、人物モデルの関節の現在の判別値となる。補助入力装置のペダルから入力される1次元の実数値を0以上1以下に正規化し、これを選ばれている予約姿勢データの重み W とする。人物モデルの関節の中で、現在の判別値が1である関節の角度は、選ばれている予約姿勢データの関節角度の W 倍と、状態2によってリアルタイムで入力されている姿勢データの関節角度の $(1-W)$ 倍との和によって決定される。これにより、ペダルから入力される値の変化に応じて人物モデル映像が変化する。映像2、映像3、映像4はそれぞれ $W=0$ 、 $W=0.5$ 、 $W=1$ のときの映像である。状態1によって登録した予約姿勢データにおいては、首、左肩、左ひじ、左手首の判別値が1であるから、状態2における人形本体の首、左肩、左ひじ、左手首の関節角度が、予約姿勢データの関節角度と混合され、人物モデルの映像に反映されている。このようにして、あらかじめ登録された姿勢が確実に再現される。

5. 映像制作

本システムを用いたCG映像制作について述べる。この映像は、中部電力川越電力館（三重県）の常設展示「サミットスタジオ」の中で用いるものであり、古今東西の偉人達が会場の大画面に登場し、テレビ会議形式でエネルギー問題について語り合うストーリーになっている。12体の人物モデルの動作データを本システムを用いて1体分ずつ収録し、合計約100分の映像をレーザーディスクに収めた。

5.1 顔と上半身のモデリング

物体表面の形状と色を同時に計測できる米国Cyberware社の三次元デジタイザ[9]は、顔のモデリング作業[10]等の効率化に役立つ。この三次元デジタイザは、約30cm立方の物体表面を、円柱座標の角度方向、高さ方向それぞれ512ステップの解像度で計測し、半径で表わされた形状データと、半径計測点におけるフルカラーの色データを出力する。これらの計測データのうち、色データは512画素×512画素のテクスチャの素材としてそのままの解像度で利用できる。一方、形状データの半径計測点を全て

結んで多面体を作ると、そのポリゴン数は10万～20万になり、リアルタイムでの表示は不可能である。このため、少ないポリゴン数で形状データを表現する必要がある。しかし、この多面体に一般的なポリゴン削減[11]を施すと、変形を伴うアニメーションの際に多面体の表面形状の滑らかさが損なわれるため、ポリゴン削減は顔モデルの作成には適さない。

本映像制作においては、3.2で述べた約4500ポリゴンの顔の「基本モデル」を、デジタイズした顔の形状データに整合させる「フィッティング」と呼ばれる処理によって、リアルタイムで表示でき、かつ表情生成にも適した顔のモデルを作成した。フィッティングには、3.2で述べた高速表情生成手法[8]を利用した。基本モデルのパラメータ空間上に配置した86個の制御点を動かして、基本モデル上の点と、デジタイズした顔の形状データ上の点の対応付けを行なった。制御点に与えた移動量を線形内挿して基本モデルを変形し、デジタイズした顔の形状データに整合させた。基本モデルとその表面上にマッピングしたフィッティング用の制御点を図4に示す。

人物上半身モデルの中で、胴体のモデリングにも三次元デジタイザの計測データを利用した。胴体は顔に比べると形状が単純であり、複雑な形状変形も必要としなかった。そこで、形状データの半径計測点を等間隔で間引いたものを結んで約500ポリゴンの胴体の多面体を作った。胴体部分の衣服の詳細は、色データに市販のペイントソフトで加筆して表現した。計測データを利用して作成した顔と胴体に、首、腕、手、帽子やアクセサリ等の付属品を市販のモデラで作成して追加し、12体の人物上半身モデルを作成した。1体のモデル全体のポリゴン数は7000～9000程度とし、全ポリゴンにテクスチャを貼った状態で毎秒15フレーム程度の描画ができるようにした。

5.2 アニメーション映像の生成

表情追跡システムX-IST、人形型入力デバイスの写真をそれぞれ図5、図6に示す。図5の写真は、顔の被測定者が14個のマーカを貼り、頭に確実に固定されるように改造を施したX-ISTのヘルメットを装着しているところである。ヘルメットから顔の前に伸びたアームの先に赤外線カメラが固定されており、被測定者が若干首を動かしても、赤外線カメラが同じアングルでマーカの映像を入力できるようになっている。図6の写真は人形本体を人形師が両手で操っているところである。これらの装置をONYXに接続して人物上半身を1体表示できるリアルタイムアニメーションシステムを構築し、図7に示すような12体の人物モデルの映像を制作した。

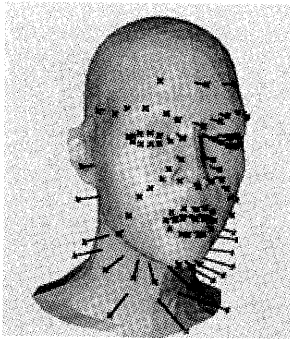


図2 顔の基本モデルと表情生成用制御点

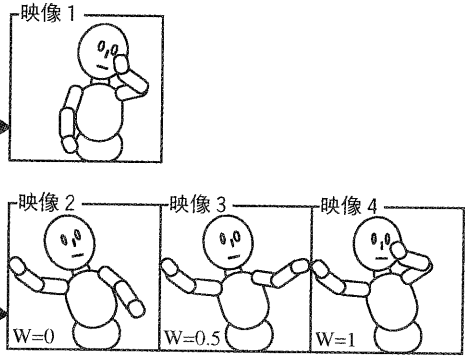
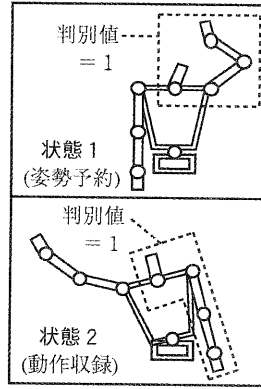


図3 姿勢予約と動作収録

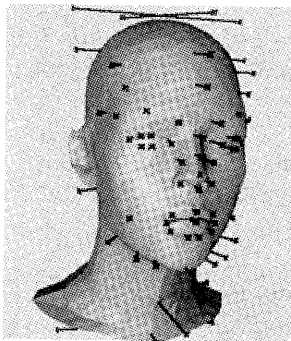


図4 顔の基本モデルとフィッティング用制御点



図5 マーカを貼り表情追跡システムのヘルメットを装着した被測定者

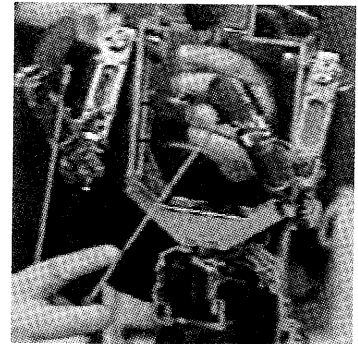


図6 人形型入力デバイスの人形本体を両手で操っている様子



図7 「サミットスタジオ」に登場する12体の人物モデル
 (上段左端から順に、キューリー夫人、レオナルド・ダ・ビンチ、エジソン、平賀源内、クレオパトラ、
 マリー・アントワネット、モーツァルト、ナポレオン、ニュートン、ノーベル、シェイクスピア、聖徳太子)

各々の人物モデルの映像をレーザディスクに収める作業は、まず台本にしたがって声優がセリフを話し、これを録音することから始めた。次に、録音したセリフを聞きながら、顔の被測定者と人形師が同時に人物モデルを制御した。このとき、人物モデル描画用共有メモリの内容の1/30秒毎の時間変化を、動作データファイルとして ONYX のハードディスク上に格納しておいた。顔の被測定者、人形師、および演出家は、ONYX がリアルタイム表示する人物モデルの映像をモニタしながら、演技のリハーサルやブラッシュアップをその場で行なった。顔の被測定者は、セリフを聞いてから顔の表情を作り、これが画像処理されるまでのディレイを考慮し、録音内容を人形師よりも400ミリ秒先に聞くようにした。人形師にとって難しい動作が必要な部分では、リハーサルの際に4. で述べた姿勢予約を行ない、後処理の発生を防いだ。12体の人物モデルの合計約100分の動作データファイルの作成に要した作業日数は8日間であった。ONYX がリアルタイム表示した映像は毎秒15フレーム程度であり、そのままでも最終映像として使えるものであった。しかし、映像の時間軸方向の品質をさらに上げるため、動作データファイルの内容をスローモーションで共有メモリに書き込み、このときの映像をテープに録画したものを早送りして、レーザディスクに収録した。

6. おわりに

本報告では、モーションキャプチャを用いた人物上半身モデルのリアルタイムアニメーションシステムについて述べた。表情追跡システムからの入力データに、制御点を用いた高速表情生成手法を適用し、映像制作者の意図に合わせて表情の演技を容易に調節できるようにした。人形型入力デバイスからの入力データに、予約姿勢データを自然に混合し、動作データの後処理の発生を防いだ。本システムを用いて人物CGの長編映像制作を行なった結果、本システムが低コスト、短期間での人物CG映像の量産に適したシステムであることを確認できた。

参考文献

- [1]上瀬, 坂本, 小滝, 長島: リアルタイム3次元CGシステムの開発とキャラクターの生放送出演; 第5回ニコグラフ論文コンテスト pp. 317-324 (1989) .
- [2]町田, 田中, 山田: パピュレーションシステムについて; 第8回ニコグラフ論文コンテスト pp. 194-202 (1992) .
- [3]久野, 坂本, 佐藤, 坂井, 福井: リアルタイム・モーションキャプチャの放送利用; 情報処理学会 グラフィクスとCAD 79-5, pp. 31-37 (1996) .
- [4]薄井, 中村, 鈴木, 青木, 大場: ポリゴンアイドル・薄井幸; 情報処理学会 情報メディア 25-5, pp. 33-39 (1996) .
- [5]L. Williams : Performance-Driven Facial Animation ; ACM Computer Graphics, Vol. 24, No. 4 (SIGGRAPH 90), pp. 235-242 (1990) .
- [6]S. Glenn : VActor Animation Creation System ; SIGGRAPH 93 Visual Proceedings (1993) .
- [7]C. Esposito, W. B. Paley, and J. Ong : Of Mice and Monkeys: A Specialized Input Device for Virtual Body Animation ; ACM 1995 Symposium on Interactive 3D Graphics, pp. 109-114 (1995) .
- [8]K. Arai, T. Kurihara, and K. Anjyo : Bilinear interpolation for facial expression and metamorphosis in real-time animation ; The Visual Computer, Vol. 12, No. 3, pp. 105-116 (1996) .
- [9]Y. Suenaga, Y. Watanabe : A Method for the Synchronized Acquisition of Cylindrical Range and Color Data ; IEICE Transactions, Vol. E74, No. 10, pp. 3407-3416 (1991) .
- [10]K. Waters and D. Terzopoulos : Modelling and Animating Faces using Scanned Data ; The Journal of Visualization and Computer Animation, Vol. 2, No. 4, pp. 123-128 (1991) .
- [11]P. Hinker and C. Hansen : Geometric Optimization ; Proceedings of Visualization '93, pp. 189-195 (1993) .