

基礎論文

多視点動画画像解析に基づくパフォーマンスアニメーションシステム

川田 聡*¹ 近藤 拓也*² 菊池 亨*³ 山本 正信*⁴A Computer Animation System Based on Action Database Captured
from Multiple Image SequencesSatoshi Kawada,*¹ Takuya Kondo,*² Kou Kikuchi*³ and Masanobu Yamamoto*⁴

Abstract - This paper proposes a system to produce 3D computer animation based on the human performance from video image sequences. This system uses a 3D model of the human body for motion capturing as well as generating CG images. In fact, a human action is observed by video cameras allocated around the actor. Fitting the 3D human model to the human images, the human action are measured as series of 3D pose of human body. This method needs not to attach any equipment and marker to the human body. We have made a database of various kinds of human actions. An action required in a story-telling can be generated by warp and combination of actions from the database. We produced an animation video originated from a fairy tail in the selection of Grimm brothers by this animation system.

Keywords : performance animation, motion analysis, geometrical model, motion database

1. はじめに

本論文は、コンピュータを使ったアニメーション制作システムを提案している。現在、テレビや映画で多くのアニメーションが放映されているが、その制作はほとんど手作業で行われてきている。人件費が次第に高くなっていくにつれ、制作費用が高くなることから、コンピュータによるアニメーションの制作が進められてきている。

アニメーションの制作で最も重要なのは、キャラクターの動作の生成である。手書きのアニメーション制作においては、これまで技術やノウハウの膨大な蓄積がある。CAPS[13](Computer Animated Production System)やTicTacToon[4]などのシステムでは、アニメーション制作の手作業を可能なところからコンピュータに置き換えることが行われた。これらのシステムでは、アニメーターの描いた絵を基にしていることから、2次元コンピュータアニメーションと呼ばれる。

これに対し、コンピュータの中に3次元の仮想世界を構築し、その投影像としてアニメーション映像を作

成する方法が考えられている。この方法は3次元コンピュータアニメーションと呼ばれ、2次元コンピュータアニメーションとは異なった観点から動作の生成を行うものである。

その一つは、キャラクターをロボットとみなし、ロボットの作業計画として動作を生成させる方法である[14]。動作を自律的に行わせるためには、ロボットに意志や感情や動機などを与えることが試みられている[17]。簡単な動作の実行はできるが、複雑な動作を生成させるには至っていない。また、ロボットの作業計画は本来むだのない合理的な動作を目指しており、これはアニメーションで重要な個性的な動作を生み出すものではない。

個性を持った動作を簡単に実現するには、実際の人間の動作を模倣するのが近道である。人間の動作を測定する方法としては、既にモーションキャプチャーやデータグローブなどが市販されている。これらの装置は人体にマーカーや特殊な器具を取り付けるので、演技者が意識したり動きが制限されたりして自然な動作を測定することが難しい。また、得られた動作データとキャラクターの動作データの間には互換性が乏しいことが多くデータの変換が必要である。さらに、アニメーターはモーションキャプチャー及び映像制作システムという2つのシステムの使い方に習熟しなくてはならない。

演技者に一切の負担をかけない方法として、動画像

*1: 伊藤忠テクノサイエンス(株)

*2: 富士ソフトABC(株)

*3: 富士ソフトABC(株)

*4: 新潟大学 工学部 情報工学科

*1: Itochu Techno-Science Corporation

*2: Fuji Soft ABC Corporation

*3: Fuji Soft ABC Corporation

*4: Faculty of Engineering, Niigata University

の解析により動作を測定する手法が種々提案されている。これらの手法は、人間とコンピュータとのインターフェース (CHI) [19], ロボットの視覚や監視システム [6] などに用いられるものも多い。そこでは、人体の一部あるいは体全体を対象として追跡している。また、リアルタイム性を重視しているため、追跡動作の種類が限られている。

人間の動作をその手足など部分に至るまで正確に3次元的に復元するためには、人体の3次元モデルと多視点動画像を利用している [3], [5], [9], [16], [20]。

本論文では、多視点ビデオ映像からの動作測定法を用いたパフォーマンスアニメーションの制作システムを提案する。このシステムはアニメーション映像の作成のため、人体の3次元CGモデルを用いているが、このモデルを使って動画像の解析も行う。したがって、得られた動作データは、即キャラクターの動作データとして利用できる。また、汎用のビデオカメラ以外に特に特殊な機材を使用する必要が無いため、アニメーターは映像制作システムの使い方だけに習熟していれば良い。

2. システム構成

本論文で提案しているアニメーションシステムの構成要素をハードウェアとソフトウェアに分けて概略を述べる。

2.1 ハードウェア構成

本システムの使用機材は、ワークステーション (SUN, SparcStation 2), ビデオキャプチャボード (Parallax, Xvideo), ビデオディスク (松下, LQ-4100), ビデオカメラ, 同期信号発生器, 音響システムおよびVCRである。

演技者の動作は異なる位置に置かれた複数のカメラで観測する。それぞれのカメラは同一の同期信号で駆動されている。複数の映像信号は4画面分割ユニット (SONY, YS-Q400) に入力され、一本の映像信号として出力される。このユニットは4台までのカメラの映像を入力できる。出力映像は一旦ビデオディスクに記録される。個々の画像のサイズは半分になるが、複数の映像信号を同時に記録することが可能である。

ビデオディスクからの映像は、ビデオキャプチャボードを介してワークステーションのディスクに転送され、動作の解析が行なわれる。

一方、ワークステーションによって作成されたCG

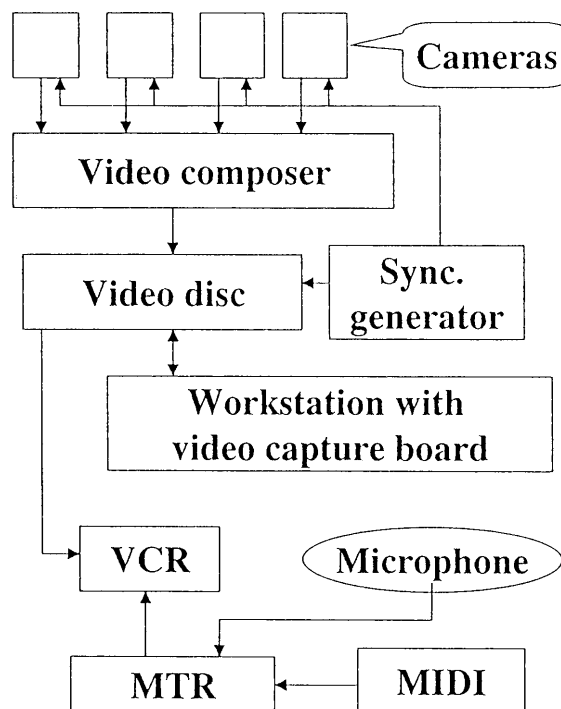


図1 コンピュータアニメーションのための器材構成

Fig.1 Organization of equipments for computer animation

映像は、キャプチャボードの出力機能を使ってビデオ映像として出力され、ビデオディスクに記録することができる。ビデオディスクに記録された映像にセリフや効果音・BGMを加えてVHSのビデオカセットに記録する。使用器材の構成図を図1に示す。

この機材構成では、アニメーターが映像を作成する場合には、ワークステーション、ビデオキャプチャボード及びビデオディスクを使用する。動作の測定には、さらにビデオカメラと同期信号発生器を必要とするが、これらの操作は接触型のモーションキャプチャに比べて容易である。

2.2 ソフトウェア構成

本システムのソフトウェア構成を図2に示す。ソフトウェアは大きく分けて3つにモジュール化することができる。すなわち、動画像解析によるモーションキャプチャ、幾何モデリングシステム、及びアニメーション作成ツールである。これらのモジュールは、全て同じコンピュータの画面上で操作可能である。

アニメーション製作者 (アニメーター) は幾何モデラによって人体や物体の幾何モデルをあらかじめ作成し、データベースに蓄積しておく。一方、動画像解析により人間の動きを運動パラメータの時系列として抽出し、これもデータベースに蓄積しておく。

次節からは、データベースの構築に必要な幾何モデ

1: 市販のモーションキャプチャで測定された動作データが、市販のキャラクターの動作表現法と一致しているとは限らない。また、測定された動作データには一般にノイズがかなり含まれている。キャラクターの動作データとして使用するためには、ノイズの除去に加えて、データのフォーマット変換を行わなくてはならない。これに対して、本論文では、モデラで作られたキャラクターを画像上の人体に一致させることによって動作を得ている。したがって、測定されたデータをそのままキャラクターに与えることによって動作の再現が可能である。また、動作は画像上の多数の特徴点から得られるので、ノイズに対して強いデータとなる。

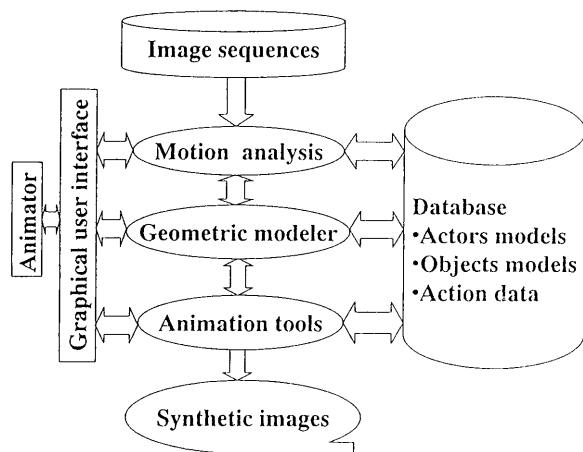


図2 ソフトウェア構成
Fig.2 Organization of software modules

ラと動画画像解析システムについて述べる。ついで、実際のアニメーションの製作過程を紹介しアニメ化に必要な各種ツールについて述べる。

3. 幾何モデラ

人体や物体のモデルの作成，モデルデータの管理及びCG映像の作成を行う幾何モデラについて述べる。

このモデラは，Xsolverと呼ばれ，ソリッドモデラのGEOMAP²とそのコンピュータビジョン用への拡張版であるSolver³を動画画像解析とコンピュータアニメーションのためにさらに拡張させたモデリングシステムである。このシステムは，現在UNIXマシン上のX-window環境で動作している。

簡単な立体は素立体(直方体，楕円球，回転体，凸多面体，一般錐台等)と呼ばれ，キーボードからの対話処理により作成し，立体同士の論理演算により複雑な立体を形成してゆく。

モデル作成のプロセスはUNIXのhistory機能を使って自動的にテキストファイルに記録される。このため，モデルはデータ構造としてだけでなく，作成の手続きとしても記録することができる。テキストを統合したり編集することにより，モデルの統合や改造も容易に行なうことができる。

モデルは画像に重ねて表示することもできる。基準物体をそのモデルと照合させることによるカメラキャリブレーションや，人体の追跡結果の表示と確認，などが容易になった。

この幾何モデラは，大学での計算機実習科目⁴で実際に使用されている。この実習科目は毎年100名程度

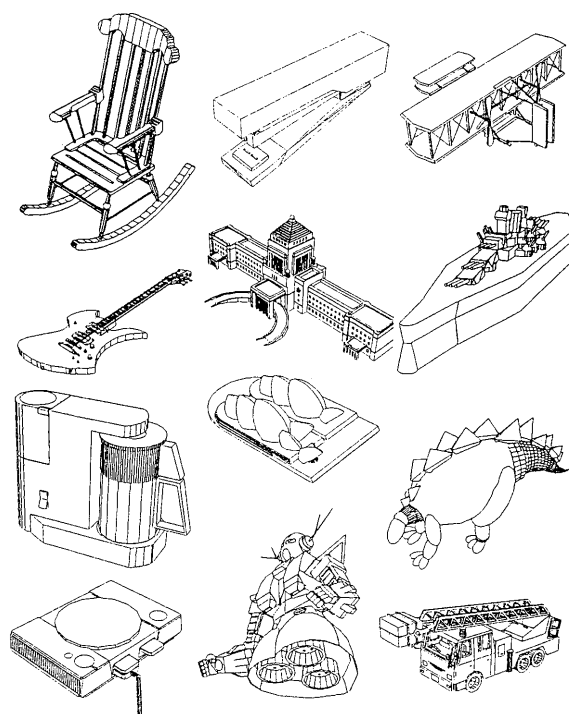


図3 モデルデータベースの表示例
Fig.3 Examples in a database of geometric models

の受講者があり，与えられた課題のモデルを3週間で作り上げている。作成されたモデルは，カテゴリ別に分類されデータベースに蓄積されている。著者らのグループで作成されたモデルも含めて，現在までに約500体のモデルが蓄積されている。図3に代表的なモデル例をカテゴリ別⁵に表示する。

このデータベース中の人物のモデルは，アニメーション映像の生成だけでなく動作追跡にも使用できる。モデルの共用は作業効率の改善に役立てることができる。

4. 多視点動画画像解析による動作測定

複数個の視点に配置されたカメラからの映像を使って人体の動きを3次的に測定する手法[16],[20]について述べる。

人体の3次元モデルを幾何モデラで作成しておく。人体は剛体ではないが，複数の剛体のリンク接続による多関節構造で表す。すなわち，頭部，胴体，上腕，下腕，上脚，下脚に相当する剛体の接続構造で表す⁶。接続順位は胴体を最上位とし，体の末端に向けて階層

2: 木村・穂坂がCAD/CAM用に開発したソリッドモデラ

3: 越川[12]がGEOMAPをコンピュータビジョンの研究のために機能を拡張したモデラ

4: 新潟大学情報工学科での実習科目“CADモデリング実習”。

5: 図3の左上から右方向に順番に，家具，大工道具・文具，飛行機，楽器・スポーツ用品，日本の建物，船舶，台所用品・食料品，外国の建物，人体・動物，家電・コンピュータ，ロボット・武器，自動車・鉄道車両

6: さらに詳しい構造として，下腕の先に手，下脚の先に足を付け加えたり，胴体を腹部，胸部，腰部のように分割することも可能である。

構造の下位部とする。このとき、人体の姿勢は各部の関節角で表すことができ、全体の位置・姿勢は胴体のシーン座標（ワールド座標）に対する位置と姿勢で表すことができる。

個々のカメラはシーン座標系に対してキャリブレーションしておく。人体モデルが人体の正しい位置と姿勢を表していれば、モデルを画像に投影したときモデルと人体像は重なるはずである。重ならなければモデルの姿勢と位置を変更し再び投影する。両者が一致するまで調整を行なう。このモデルマッチングは追跡開始画像上で行っておく。

一方、人体の運動は関節角速度と胴体の並進速度及び回転速度で表すことができる。これらを人体の運動パラメータと呼ぶ。人間の動作はこの運動パラメータを求めることにより定量化される。

カメラがシーンに対して静止しているとすれば、連続する画像間の差分画像には動いている部分の大きな値となって現れる。この差分値は人間の位置・姿勢の変化の結果であるから、逆運動学手法により、位置・姿勢の変化、すなわち運動パラメータを得ることができる。

得られた運動パラメータに基づきモデルを移動させる。移動したモデルは次のフレームでも人体と一致している。運動パラメータの推定と推定値に基づくモデルの移動を、追跡開始画像からフレーム順に繰り返すことにより人体の追跡を行う。結果として運動パラメータの時系列が得られる。これを動作のデータベースに蓄積しておく。図4には「歩いて、かがんで、うなずいて、立ち上がって、はっとする」といった249フレームの動作の追跡結果である。各列はそれぞれのカメラからの画像である。第1行は追跡開始フレームである。人体のモデルがワイヤフレーム表示で各画像に重ねられている。追跡結果を第2行以降3フレームほど選択して表示している。図5にはこの動作を3次元CGによって再生した結果を示す。

この測定法の特徴は、人間の運動パラメータ（あるいはその累積値としての姿勢パラメータ）が画像から直接得られることである。さらに、得られた動作データを使って任意のモデルに動作を再現することができることである。

5. アニメーションの作成

動作と物体のデータベースからアニメーションを作成する手順について述べる。これには、従来のメディアコンテンツの製作手順[10],[11]を参考にした。

アニメータは原作を選定し脚本（シナリオ）を作成する。シナリオには、場面（カット）毎に、舞台設定、登場人物の選定、動作、セリフなどを細かく記し

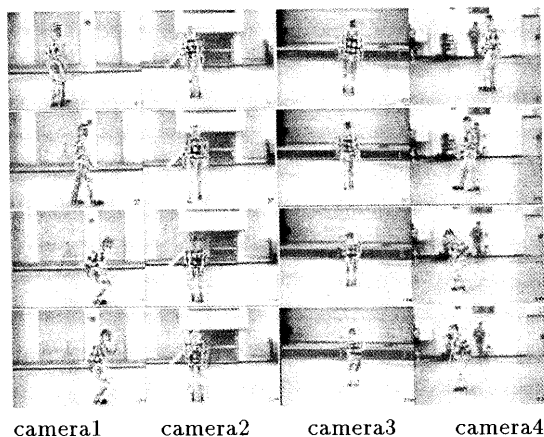


図4 4台のカメラからの動作追跡。全249フレームから4フレーム表示

Fig.4 Tracking results from four camera views. Four frames are depicted as samples from 249 frames.

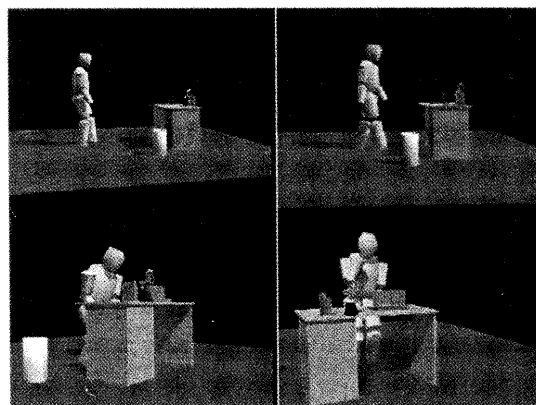


図5 追跡動作のCGによる復元結果

Fig.5 Reconstruction of the action captured from multiple camera views

ておく。舞台設定では、物体データベースの中から必要なモデルを検索し、それらを仮想空間内に配置する。

シナリオに記されている登場人物の動作をデータベースから検索し舞台の上で動作させる。動作データの長さからそのカットのフレーム数が決まる。このとき、人体モデルの初期配置、人体間の動作協調、人体と物体との衝突防止等に注意する必要がある。

カメラワークは、中心となる話題あるいは人物の演技にできるだけ注目させるように決定される。構成された演技は、ワイヤフレームによりプレビューを行なうことによってチェックされる。演技が決定されたならば、レンダリングを行ない、ビデオディスク上にCG映像を記録する。編集によりバランスのとれたアニメーションを作成する。最後に音声や音響効果を加える。

この処理の流れを(図6)に示す。次節から個々の手

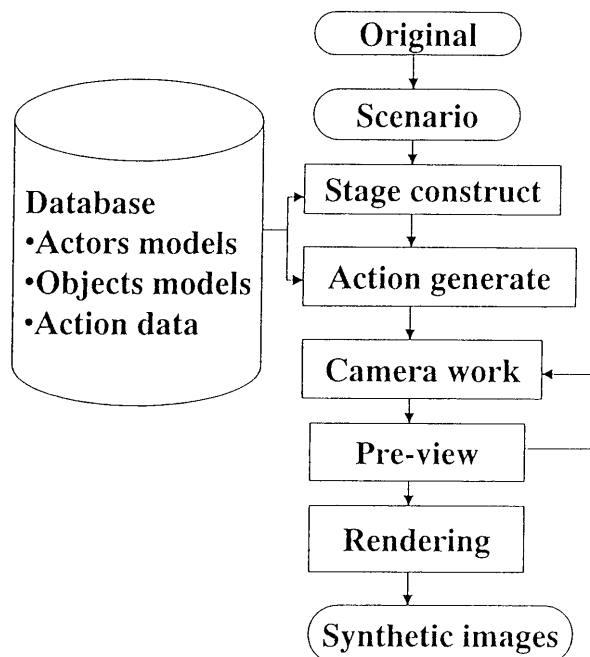


図6 パフォーマンスアニメーション作成のための製作手順
Fig. 6 Steps for performance-based computer animation

順について説明する。

5.1 作品の選定とシナリオの作成

今回、アニメーションの題材として、グリム童話の中から「きのおさら」⁷を選んだ。この作品は、起承転結が明瞭で行動が具体的である等の観点から比較的製作が容易と思われた。

シナリオの作成にはその作法[1]を参考にした。この作品は話の展開から大きく2つに分けることができる。それらを第1幕、第2幕として、それぞれをさらに細かく分けることにより全部で17のカットに分割した。付録の表A・1には、それらのカットの作成フレーム数、カメラで写す人物(場所)の順番、登場人物の動作、セリフ等によって構成されるシナリオを示す。

5.2 舞台構成

脚本に記された物体のモデルをデータベースから検索し、各カットの舞台を構成する。データベースに含まれない物体は新たに作成した。本作品の舞台は屋内になるので、まず、家屋(床、壁、屋根)のモデルを構成した。また、作品に見合った室内を装飾するために、椅子、机といったさまざまな家具を床面上に配置した。さらに、食卓上の食器や食物のモデルも配置した。構成した舞台を図7に示す。この舞台には、ベッド、人蔘、皿、暖炉、フォーク、台所、ナ

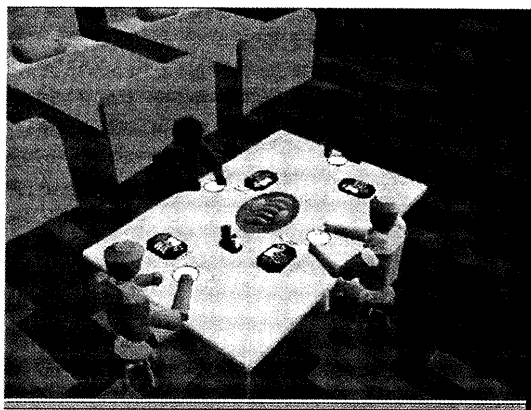


図7 カット1の舞台構成
Fig. 7 An arrangement of the 1st Cut in the fairy tale

表1 解析した動画像とフレーム数
Table 1 Image sequences analyzed to establish an action database

動画像	フレーム	動画像	フレーム	動画像	フレーム
1	249	8	99	15	199
2	528	9	294	16	139
3	369	10	180	17	115
4	733	11	156	18	201
5	599	12	120	19	308
6	378	13	141	20	82
7	99	14	141	21	150

イフ、お盆、ロールパン、ソーセージ、スライスポテト、スープ皿、スプーン、棚、椅子、洋服ダンス、木の椅子、家、調味料、テーブル、みかん箱等のモデルが含まれている。

5.3 動作の決定

この作品で必要と思われる動作を演技しその多視点動画像を解析した。解析した動画像の番号とフレーム数を表1に示す。#1の動画像の追跡結果と動作の復元結果は、すでにそれぞれ図4と図5に示した。その他に、#5「おじいさんがふるえた手でスープをこぼしながら飲む」の動画像の追跡結果を図8に示す。上半身のみの動作であるのでカメラは3台で十分であった。各列はカメラからの画像に対応する。各行は追跡結果のサンプル表示である。

本研究で使用した動作の測定手法は、初期姿勢に変位を積み重ねることにより追跡を行なっている。変化の推定に誤差が含まれていれば誤差も累積され追跡に失敗する。表1に示す追跡実験では、最長733フレーム(動画像#4)まで可能であった。動作が緩慢な場合には、これより長時間の追跡も可能であるが、スポーツ運動のような激しい動作の場合には100フレーム以下しか成功しないことがある。

映画やドラマでの演技の動作は一般に長くない。せいぜい数100フレーム程度である。したがって、激し

7: この原作のあらすじはつぎの通りである。子供と両親とおじいさんの4人家族が暮らしていた。おじいさんは老化のためへまばかりしている。とうとう、お皿を落して割ってしまった。そのため、おじいさん用には割れない木のお皿があてがわれてしまった。あるとき、子供が板切れから何やら一所懸命に作っている。それは、自分の両親のための木のお皿であった。両親は自分たちのしたことの非情さに気づき、おじいさんに謝った。



図8 食事動作の追跡結果. 3台のカメラを使用
Fig. 8 Tracking result of an action at the meal
from three camera views

い動作を含むアクション映画のような作品を目指さなければ、数百フレームの追跡が可能な本手法で十分である。

激しい動作の追跡にはキーフレーム追跡[15]が利用できる。この手法は、モデルが人体像と大きく離れた時点を新たな追跡開始時刻とし、モデルと人体像とのモデルフィッティングを行ない追跡を再開する。この手法はキーフレーム追跡と呼ばれるが、今回は使用しなかった。

動作のデータベースには、動画像から得られた一連の動作を基本的な動作に分割して登録しておく。例えば、動画像#1で得た動作は、#23「うなずく」、#24「歩く」、#25「かがむ」、#26「はっとする」の基本動作に分割して登録する。動作データベースを表2に示す。データベース中の動作データには、初期フレームでの姿勢データを先頭に、運動パラメータがフレーム順に並べられている。

データベース中の動作は、動作viewerによりデータベース作成者以外でも動作を直接確認することができる。操作パネル上で動作名とカメラを選択し実行させると、ワイヤフレームモデルで人体とその動作が表示される。

データベースに登録されている動作以外の動作を必要とする場合がある。このとき、次に述べる方法で既存の動作データから新たな動作を生成することができる。

1. 基本動作の組合せにより、動画像から得た動作データよりも多くの種類の動作を生成することができる。例えば、動作#1と動作#4を組み合わせる

表2 動作データベース
Table 2 An action database

動作番号	動作名	フレーム数	動画像
おじいさんの動作			
1	手を震わせる	40	5
2	スプーンを持つ	304	
3	スープを飲もうとしてこぼす	167	
4	大きな椅子に座ったまま左に回転する	139	16
5	小さな椅子に座ったまま左に回転する	116	17
6	椅子から引っ張ってもらい立ち上がる	120	12
7	人に手伝ってもらいながら椅子に座る	141	13
8	顔を上げる	99	7
9	パンを食べる	300	6
10	ため息をつく	76	
11	左に振り向く	99	8
12	皿を持って落とす	308	19
お嫁さんの動作			
13	テーブルに皿を置く	201	18
14	スープを飲む	146	9
15	左に振り向く	32	
16	椅子から立ち上がる	47	15
17	コーナーを右に曲がる	112	11
18	椅子から人を引っ張り上げる	156	
19	人を椅子に座らせるのを手伝う	141	14
20	皿を持って歩く	90	9
21	食卓に皿を置く	141	
22	首を左右に振る	82	20
23	うなずく	60	1
24	歩く	65	
25	かかんで子供に尋ねる	75	21
26	はっとする	41	
27	謝る	150	
息子さんの動作			
28	スプーンを持つ	31	3
29	スープを飲む	140	
30	座ったまま右手を上げて注意する	58	
子供の動作			
31	スープを飲む	209	2
32	物をつかんでのこぎりを引き顔を上げて話をする	733	4

ことにより、カット2で「お嫁さんと向き合うおじいさん」の動作を作ることができる。

2. 「歩く」「食べる」などの繰り返し動作は、その1周期のみをデータベースに登録しておく。このようにしておけば、繰り返し動作を連続して生成することができる。例えば、スープを飲むシーンは繰り返し動作で生成した。

3. 初期姿勢を入れ換えることにより他の人物の動作を作成する。例えば、歩く動作はお嫁さんの動作として登録されているが、上半身の初期姿勢を多少前屈にさせることにより、カット3のおじいさんの歩行動作(#24')を生成した。

4. 動作の一部を入れ換えることにより新たな動作を作成する。例えばカット1で使われた動作#30「椅子に座ったまま右手を挙げて注意する」の下半身の姿勢を立居にすれば、カット9での動作#30'

「右手を挙げて注意する」になる。

5. 対称的な動作はパラメータ値の操作で作成できる。例えば、椅子に座ったまま向きを変える場合には、左回転と右回転とでは左右のパラメータを入れ換えるなどして作ることができる。カット2のおじいさんの動作#4'は、そのようにして作成した。

6. 同じ動作でも環境を変えることによって異なる意味を持つ演技となる。例えば、#23～#26の一連の動作は、図5のように机の上にパソコンがある場合にはパソコンの画面を見て驚く演技となるが、シナリオのカット13中の子供がいる場合には、子供の意図を理解して自分達の非に気づく演技となる。

基本動作の組合せや繰り返しによって新たな動作を生成するとき、動作の繋ぎ目にギャップが生じる。ギャップを解消させるためには、動作パラメータをスプライン関数で補間する。また、動作データは数値列として記録されているのでデータを直接編集することも可能である。

動作データベースに載っている動作は互いに独立に得られたデータである。従って、それらを組み合わせる際には、人体同士や人体と物体との間で衝突が起きる可能性がある。作成されたカットのプレビューやモデラの干渉チェック機能の利用により衝突を検知し、モデルの再配置や動作の手直しを手動で行なっている。

全く新しい動作を必要とする場合には、新たに動作の解析を行なう必要がある。動作映像の解析と得られた動作データの再現が一つのシステム内で可能なため、この作業は楽に行うことができる。

5.4 カメラワーク

映画監督用の手引書[2],[11]が幾つか出版されている。その内容のほとんどがカメラワーク技法に割かれている。カメラワークは作品の主題を強調するために重要な技術である。本論文では、マウスとキーボードからカメラワークを容易に設定できるGUIを使用している。図9にGUIを示す。

GUIの左上のキャンパスには、舞台の俯瞰図が、また左下には側面から見た図がワイヤフレーム表現によって示されている。この二つのキャンパス内でマウスを使って、カメラの視点及び参照点の座標値を指定することができる。図には、視点と参照点を結ぶ直線が視野の範囲と共に描かれている。これらの座標値及び視野角はキーボードからも直接与えることが可能である。

これらの値を各フレーム毎に指定することによって、カメラの位置を時間的に変化させることが可能となる。しかし、全ての時刻において座標値を入力することは非常に多くの労力を必要とする。そこで、キー

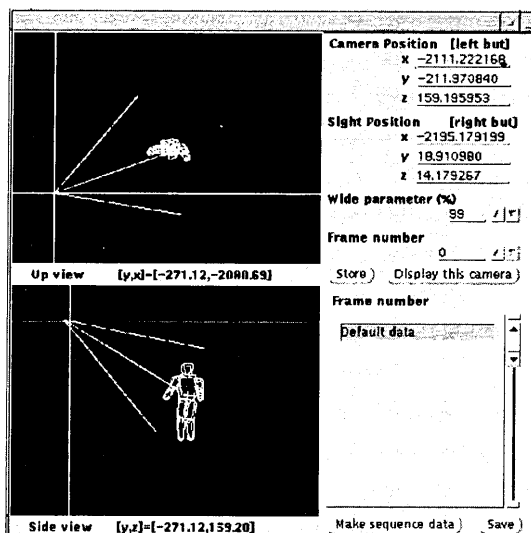


図9 カメラの設定用GUI

Fig.9 GUI for a setup of the camera work

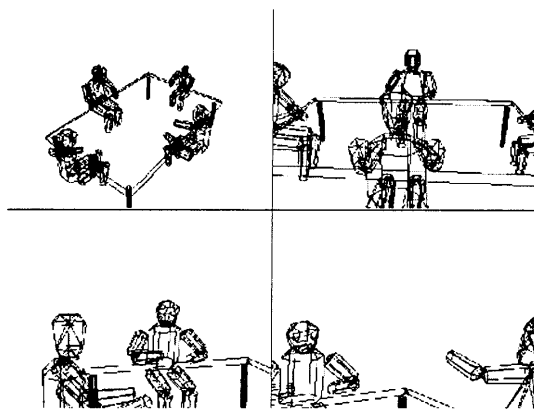


図10 アニメーションのプレビュー

Fig.10 Pre-view of the animation

となるフレームにおいて値を指定しその間をスプライン関数で補間することも可能である。

カメラワークのデータは、モデルと同様にテキストファイルとして保存しておくことができる。従って、過去に設定したカメラワークを試すことができ、最も良いカメラワークを見つけることが容易になる。

5.5 プレビュー

レンダリングは非常に重い処理となるので、最終映像の変更はできるだけ避けたい。そのためには、レンダリングの前に登場人物の動作とカメラワークを決定しておく必要がある。アニメータがCG映像を手早く確認するために、ワイヤフレーム表示によるプレビューを用意した。

図10は、カット1のシーンのプレビューから、その映像を400フレーム間隔で表示したものである。

5.6 レンダリンク

レンダリングには市販のソフトウェアの使用も考えられる。しかし、幾何モデラ Xsolver で生成されるモ

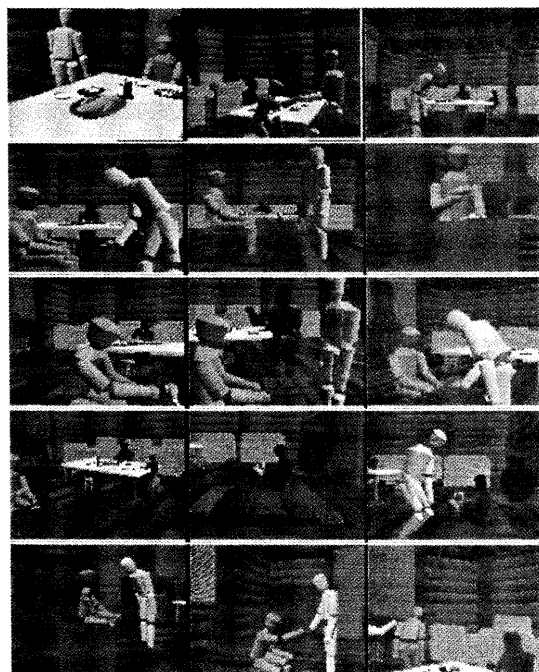


図11 作成したアニメーション映像。
Fig.11 Synthetic images in the animation.
Each image is sampled from the corresponding cut.

デルのデータ構造との親和性を考慮して、独自のレンダリングソフトを作成した。

このレンダリングシステムは、他のレンダリングソフトに見られるような基本的な機能は網羅している。すなわち、シェーディング処理 (Gouraud, Phong), ハイライトの処理, 影付け, 色指定, テクスチャマッピング, 光源選択 (点光源, 面光源, 複数光源), エリアシング除去等が可能である。

作成されたCG映像は一コマずつビデオディスクに記録される。ビデオディスクは映像をランダムに読み出すことができるので、全体の構成を再編集して一本のビデオ映像として出力することができる。今回の作品⁸は、全部で7,451フレーム作成した。これは約4分強の長さである。

図11に作成したCG映像を各カット毎に一枚ずつ表示する。それぞれの映像は、左上がカット2, その右はカット3, …, 右下がカット16といった順番で並べられている。これらの映像では影付けが省略されている。

5.7 音響効果

アニメーションにセリフを吹き込み、音楽や効果音を加えるためのシステムを図1の下部に示す。このシステムは、MIDI (Musical Instrument Digital Interface) と MTR (Multi Track Recorder) から構成さ

れ、ビデオディスクからの映像とマイクロホンからの音声を入力しサウンドトラック付きアニメーション映像をVCRに出力する。ここで、MTRとは複数のサウンドトラックを有し、音の編集や多重録音をデジタル的に行なう装置である。また、MIDIはパソコンとデジタル音源から構成され、キーボード等からの打ち込みによりBGMや効果音を作り出すことができる。MTR, MIDI及びビデオディスクは互いに同期されている。

音響効果は次のように加えられた。最初にアニメーション映像中のフレーム番号から、セリフや効果音のタイミングを決定しておく。声優はタイムキーパの指示に従ってセリフをMTRに吹き込む。皿が割れる音や鋸で木を切る音などの効果音は音源データから選択し、タイミング通りにMTRに録音した。BGMは物語の展開に応じて4曲ほど作曲した。セリフ, 効果音, BGMはすべてMTRにトラック別に録音しておく。最後に、MTRからの音響信号とビデオディスクからの映像をVCR上に合成する。このとき各トラック間の音量バランスの調整が重要である。作品は日本語版と英語版を作成した。

6. あとがき

人間や動物の動きを参考にして、キャラクターの動作を生成する試みは古くから行なわれてきた。しかし、身体の動きを測定する装置が、アニメーション製作システムに組み込まれてきたのはごく最近のことである。動作の測定装置は一般にモーションキャプチャと呼ばれているが、人体に特殊な器具やマーカーを取り付けて測定している。そのため演技者に余分な負担を強いることが多い。

本論文では、動画像解析に基づくアニメーション製作システムを提案した。本手法は現在のところ画像の処理に時間がかかるのが欠点である。しかしながら、人体に触れることなく動作が測定できる利点を生かして、実際に演劇などの舞台での演技を測定し、アニメーションの製作に役立てることも考えられる。

アニメーションをはじめとした映像コンテンツの製作には、これまでに膨大な技術が蓄積されている。それらの技術は製作者自身の技量によるものが多く、客観的な形式でマニュアル化 [2], [10], [11] されているものは少ない。しかし、最近ではこれらの技量をプログラム化 [8] することや、製作過程を自動化 [7] することが試みられている。なかでも TVML [7] は実際の製作過程に取り入れられつつある。TVMLと我々のシステムは競合するものではない。TVMLは動作やイベントのスケジューリングを記述する言語であって、動作そのものはモーションキャプチャ等で与える必要があ

⁸ <http://www.vision.ie.niigata-u.ac.jp/abst/abst1.html#four>
をアクセスすれば全映像を見ることができる。

る。一方、本システムは動作の生成（再生）を行っているが、様々な動作のスケジューリングはマニュアルで行っている。したがって、両者を組み合わせることができれば、より効率的な製作過程が実現できるものと思われる。

謝 辞

本研究は、一部文部省科研費重点領域研究「人工現実感」(#07244206, #08234207, #09220205)の支援を受けた。

参考文献

- [1] 新井一, “シナリオの技術”, ダヴィッド社, 1996
- [2] D.Arijon, “Grammar of the film language”, Focal Press, 1976. 日本語訳: 岩本憲児, 出口丈人, “映画の文法”, 紀伊國屋書店, 1980.
- [3] C.Bregler and J.Malik, “Tracking people with twists and exponential maps”, Proc. IEEE CVPR’98, pp.8-15, 1999
- [4] J-D.Fekete, E.Bizouarn, E.Cournarie, T.Galas and F.Taillefer, “TicTacToon: A papaerless system for professional 2D animation”, Proc. ACM SIGGRAPH’95, pp.79-90, 1995
- [5] D.M.Gavrila and L.S.Davis, “3-D model-based tracking of humans in action: A multi-view approach”, Proc. IEEE CVPR’96, pp.73-80, 1996
- [6] I.Haritaoglu, R.Cutler, D.Harwood and L.S.Davis, “Backpack: Detection of people carrying objects using silhoetts”, Proc. 7th ICCV, pp.102-108, 1999
- [7] 林正樹, “テキスト台本からの自動番組製作—TVMLの提案”, 1996 テレビジョン学会年次大会, S4-3, pp.589-592, 1996
- [8] L.He, M.F.Cohen and D.H.Salesin, “The virtual cinematographer: A paradigm for automatic real-time camera control and directing”, Proc. ACM SIGGRAPH’96, pp.217-224,1996
- [9] I.A.Kakadiaris and D.Metaxas, “Model-based estimation of 3D human motion with occlusion based on active multi-viewpoint selection”, Proc. IEEE CVPR’96, pp.81-87, 1996
- [10] 金子満, “メディアコンテンツの製作”, 画像情報教育振興協会, 1998.
- [11] S.D.Katz, “Film directing SHOT BY SHOT visualizing from concept to screen”, Michael Wiese Productions, 1991. 日本語訳: 津谷祐司 “映画監督術 SHOT BY SHOT”, フィルムアート社, 1996
- [12] K.Koshikawa and Y.Shirai, “A 3-D modeler for vision research”, Proc. ICAR’95, pp.185-190, 1985.
- [13] B.Robertson, “Desney lets CAPS out of the bag”, Computer Graphics World, pp.58-64, July, 1994
- [14] Y.Koga, K.Kondo, J.Kuffner and J-C.Latombe, “Planning motions with intentions”, Proc. ACM SIGGRAPH’94, pp.395-408, 1994
- [15] 大田佳人, 山際貴志, 山本正信: キーフレーム拘束を利用した単眼動画からの人間動作の追跡, 信学論, Vol.J81-D-II, vol.J81-D-II, no.9, pp.2008-2018, 1998
- [16] 佐藤明知, 川田聡, 大崎喜彦, 山本正信, “多視点動画画像からの人間動作の追跡と再構成”, 信学論(D), vol.J80-D-II, no.6, pp.1581-1589, 1997
- [17] D.Thalmann, H.Noser and Z.Huang, “Hou to create a virtual life?”, Interactive Computer Animation, pp.263-291, 1996

- [18] 土家由岐雄, “グリム童話”, pp.13-22, 偕成社, 1989
- [19] C.R.Wren, A.A.T.Darrell, A.P.Pentrand, “Pfinder: Real-time tracking of the human body, IEEE Trans. PAMI, vol.19, no.7, pp.780-785, 1997
- [20] M.Yamamoto, A.Sato, S.Kawada, T.Kondo and Y.Osaki: Incremental tracking of human actions from multiple views, Proc. IEEE CVPR’98, pp.2-7, 1998

付 録

1. シナリオ

表 A-1 「きのおさら」のシナリオ
Table A-1 A scenario for a fairy tail in Grimm brothers' collection

カット, フレーム	カメラ	演技 (# 動作番号), 台詞
～第1幕～ カット 1, 食事の場面, 1400Frame	全体(食卓) → お嫁さん → 息子さん → おじいさん	家族四人(#14, #28, #29, #31)で食事をしている。その中でもおじいさん(#1, #2, #3)は手足が震えていてスープを満腹に飲むことができない。お嫁さん(#15)は心配しておじいさんの方を振り向き、息子さん(#30)は手を指しておじいさんを注意する。
カット2, お嫁さんがおじいさんの所へ行く, 346Frame	おじいさん・お嫁さん	おじいさんの座る所をかえるためにお嫁さん(#4', #16, #17)がおじいさん(#1, #4')の所まで移動する。
カット3, お嫁さんがおじいさんを引っ張る, 241Frame	おじいさん・お嫁さん	お嫁さん(#18)がおじいさん(#6, #24')を椅子から引っ張り上げる。
カット4, お嫁さんがおじいさんを座らせる, 330Frame	お嫁さんとおじいさん → 食卓	お嫁さん(#19)がおじいさん(#7)をきたない椅子と机の所へ連れて行き座らせる。
カット5, お嫁さんが食事を運ぶ, 380Frame	食卓 → お嫁さん → おじいさん	お嫁さん(#20, #21)がパンとおかずをひとつのお皿にのせて、おじいさん(#1)の所へ運ぶ。
カット6, おじいさんが尋ねてお嫁さんが叱る, 241Frame	おじいさん → お嫁さん・おじいさん	おじいさん(#1, #8)「これで、おしまいかね。」と尋ねる。お嫁さん(#22)「とよよりが、たくさんたべてはどくです。」と首を振りながら叱りつける。

カット7, おじいさんがパンを食べる, 600Frame	おじいさん	おじいさん(#1, #9, #10)は悲しそうにパンを食べはじめる。おじいさん(#11)は息子さんたち(#14, #29, #31)のテーブルを見つめる。
カット8, おじいさんがお皿を落して割る, 367Frame	おじいさん	おじいさん(#1, #12)の震えた手がお皿を落す
カット9, 息子さんとお嫁さんがおじいさんを叱る, 380Frame	おじいさん → 息子さん・お嫁さん・おじいさん	息子さん(#30')「しょうがないおじいさん(#1)だ。」と手を指して叱る。お嫁さん(#23')「しっかりしなけりゃ、だめですよ。」と息子さんの後にうなづく。
カット10, お嫁さんが木のお皿をおじいさんの所へ運ぶ, 380Frame	おじいさん → お嫁さん → おじいさん	一番安い、みつともない木のお皿がお嫁さん(#20, #21)からおじいさん(#1)に渡される。
カット11, おじいさんがパンを食べる, 450Frame, 第1幕おわり	お嫁さん → おじいさん → おじいさん・食卓	ため息をつきながら、おじいさん(#1, #9, #10)はパンを食べる。全体の様子を写す。
～第2幕～		
カット12, ある日の様子, 250Frame	食卓 → おじいさん → 子供	お嫁さん(#14)と息子さん(#29)はテーブルで食事をし、おじいさん(#1)はきたない机で手足を震わせている。子供は床に座って遊んでいる。
カット13, 子供が何かを作りお嫁さんが尋ねる, 733Frame	子供 → 子供・お嫁さん	お嫁さん(#24, #25)が子供の所へ行き、に何を作っているのか尋ねる。子供(#32)は木のお皿を作っていると言い、それを聞いてお嫁さん(#26)は「はっ」とする。
カット14, お嫁さんがおじいさんに謝る, 310Frame	お嫁さん・おじいさん	お嫁さん(#27)がおじいさんの所へ行き、「ごめんなさい」と頭を下げ謝る。おじいさん(#5, #23')はうなづく。
カット15, お嫁さんがおじいさんを引っ張る, 241Frame	お嫁さん・おじいさん	お嫁さん(#18)がきたない椅子からおじいさん(#6, #24')を引っ張り上げる。
カット16, お嫁さんがおじいさんを座らせる, 250Frame	お嫁さん・おじいさん → 食卓	お嫁さん(#19)がおじいさん(#7)をもとのテーブルにつかせる。
カット17, みんなで仲良く食事をする, 536Frame 第2幕おわり	食卓 → (上から)全体	四人(#3, #14, #29, #31)で元通りに仲良く食事をする。全体の様子を写す。

(1999年7月30日受付)