

仮想スタジオのためのデジタル素材の オーサリングと検索

澤 扶美[†] 河内 美樹[†] 木俣 豊^{*} 田中 克己[‡]

[†]神戸大学大学院自然科学研究科情報知能工学専攻

^{*}通信・放送機構神戸リサーチセンター

[‡]神戸大学大学院自然科学研究科情報メディア科学専攻

{sawa,miki,kotaro,tanaka}@in4wolf.in.kobe-u.ac.jp

本稿では、コンピュータによる映像製作という広い意味での仮想スタジオのオーサリングのためのデジタル素材のデータベース化を行うための記述や検索の方法を提案する。デジタル素材とは仮想スタジオの大道具や小道具、仮想キャラクターのモーションデータといったものである。更にこれらをインターネット上で利用するために必要である著作権管理法を提案する。

Authoring and Retrieval of Digital Assets for Virtual Studio Systems

Fumi Sawa[†] Miki Kawauchi[†]
Yutaka Kidawara^{*} Katsumi Tanaka[‡]

[†]Division of Computer and Systems Engineering,
Graduate School of Science and Technology, Kobe University

^{*}Telecommunications Advancement Organization of Japan,
Kobe Research Center

[‡]Division of Media and Computer Sciences,
Graduate School of Science and Technology, Kobe University

In this paper, we propose a method of authoring and retrieving digital assets for creating digital video images by "virtual studio systems". By "digital assets", we mean the objects such as "chairs" and "screens" used for the stage, and motion data for "virtual characters" used in the "virtual studio". We then proceed further to propose a method for copyright protection for these digital assets considering their use on the internet.

1 はじめに

現在、TV局では放送セットを3DCGで作成する仮想スタジオが使われつつある。この仮想スタジオでは様々な大道具、小道具を3DCG化することによってコストの削減や実現不可能なセットの構築を可能としている[1]。しかし、現在のシステムではスタジオに配置する小道具の再利用性についての配慮がなされておらず、スタジオを構成する3Dオブジェクトのデータベース化や検索機能が不十分であり、仮想スタジオの作成には大変な労力が必要となっている。

また、仮想スタジオ内に様々な人物の3Dキャラクターを配置し、音楽に合わせてそのオブジェクトを動かすことで、実際のスタジオでは不可能な表現力豊かなコンテンツの作成が可能となるが、現在のシステムでは複雑な動きをする3Dオブジェクトを配置することは困難である。従って作成された3Dオブジェクトを管理し、必要な素材や動きを持つキャラクターをユーザの要求に基づいて、効率良く検索できる機能をもつオーサリングシステムが求められている。

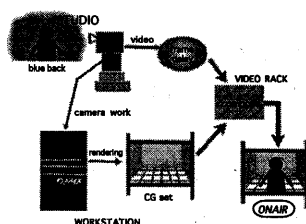


図 1: 現在一般的な仮想スタジオシステム

現在、製品化あるいは製作現場で独自に開発・運用されている仮想スタジオのシステムで一般的に指摘できる問題として、

- 出演者は仮想空間内でセットや小道具に対しイベントを起こす手段がない
- セットの構築は一般的なCG製作の技術で補えるが、システムの操作は容易とは言えない
- 仮想であることの本質的な利点を生かしきれていない
- ハードウェア的な制約
- 製作サイドのスタジオ利用経験の不足

などがあり、これを考慮し将来的な仮想スタジオの方向性を考えると以下ようになる[2]。

- フルポリゴンでの仮想空間の実現
映像のクオリティとのトレードオフ
- 出演者の位置情報をなんらかのセンサを使い仮想空間のオブジェクトに反映させる
- 従来のテレビとは異なるメディアでの利用
インターネットブロードキャスト、インタラクティブテレビでの利用

本稿で我々が提案するオーサリングシステムは、セット部品や仮想キャラクターのモーションデータ等を管理するデータベースを構築し、それらの中から映像を創造するために効果的な検索を行い必要なデータを効率良く取り出すことが可能にすることを目的とする。3次元CGオブジェクト・仮想キャラクターの生成に必要な人体などのモーションデータ・背景の生成や仮想空間と実空間を統合した映像の制作に用いる実写ビデオ映像などから成るデータベースを人間の印象度や必要とする動きを指定することで検索を行い使用することができ。

2 3次元素材オブジェクト

2.1 仮想スタジオの素材としての3次元オブジェクト

仮想スタジオは通常の静止画の3DCGの作成と同様に3次元のモデルを空間に配置していくことで作成される。作成時の任意の時点で個々のモデルの配置を変更することで異なるシーンが作成できるという点で他の映像製作手法にはない柔軟な製作ができるという利点がある。しかしその一方で、個々のモデルの作成は高度な専門の技術と作業時間を必要とする。したがって一度作成したモデルは再利用することが望ましい。再利用性を高めるためには仮想スタジオの大道具や小道具に相当するCGの部品に検索に必要な情報を付加しデータベース化を行う必要がある。

2.2 組み合わせ検索

以下に仮想スタジオを構成する場合に有効な検索手法を述べる。

- 外観の類似性、共通性による検索
- 過去の組み合わせ例を参考にした検索
- ふさわしい素材の組の検索

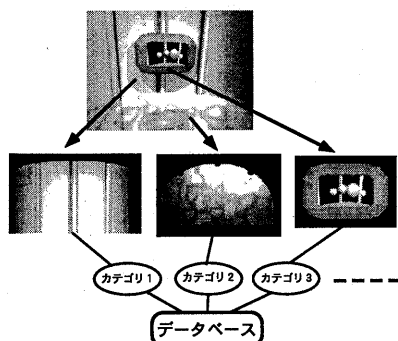


図 2: 3つの素材から構成されたセット

我々は、このような検索手法を用いた「不完全ベクトルによる組合せ検索」を提案する。この検索手法を仮想スタジオを構成する3次元素材オブジェクトに適用し、その有効性を検証した。

不完全ベクトルによるという問い合わせとは以下のようなものである。

$$q = (1, 4, *, 0, 2)$$

ここで質問ベクトル q の各要素はカテゴリ分けされた素材オブジェクトの集合を表し、その数値はカテゴリ内の要素を表す識別子である。また0は何も使用しないことを表しており、「*」は未決定の要素を表す。この問い合わせでは「*」として指定されたカテゴリ内でふさわしいと考えられる要素が検索され「ふさわしい部品の組」が得られる。

2.2.1 検索のための付加情報

- おおまかなカテゴリの定義
典型的な仮想スタジオの構成を参考にカテゴリライズする。(例: 背景・床・テーブル・仮想モニター・壁に配置する物・床に配置する物・その他小道具)

各カテゴリで部品は識別子を持つ。

- 印象語による特徴ベクトルの生成
あらかじめ定義した印象語への数値による入力の特徴ベクトルの要素とする。個々の部品は同次元の特徴ベクトルを持つ。(現代的な感じ、古典的な感じ、未来的な感じ、曲線的な感じ、装飾的な感じ、ポップな感じ、暖かい感じ等)
- オブジェクト相互の共出現頻度
二つの部品相互に「同一セット内で使われた回数」で定義される。
- セット構成 (オブジェクトの組)
あるセットを構成する全ての部品の組であり、カテゴリの数を要素数とし、セット識別子が要素になる。

2.3 検索手法の検討

2.3.1 特徴ベクトルのみの利用

質問ベクトル中の決定している要素に対応する部品の特徴ベクトルと類似した特徴ベクトルを持つ部品を選択する。特徴ベクトルの類似度はコサイン相関値を用いる。質問ベクトルは一つ以上の決定した部品を持つので、それらから基準にするべき特徴ベクトルを生成する。

- 特定の部品のみに着目し基準ベクトルを抽出する。
- 選択済部品の特徴ベクトルの平均をとり基準とする

の二つの方法が考えられるが、いずれの場合も単純な類似検索にとどまる。「ふさわしい組」を提供しているというには不十分であると考えられる。

また未確定の要素が複数ある場合は、それら同士の相関関係も考慮する必要がある。

2.3.2 共出現頻度の利用

これは過去の組合せ例を参考に組を完成させる方法である。共出現頻度とはある二つの部品が同一セット内で同時に採用された回数で定義される。質問ベクトル中の確定要素との共出現頻度の総和が最大になる部品を検索する。前出の質問ベクトル

ル $q = (1, 4, *, 0, 2)$ の場合、検索のアルゴリズムは以下ようになる。

ここで3番目のカテゴリーの識別番号 i である部品を $C_3(i)$ のように表すものとする。既に決定している要素は $C_1(1), C_2(4), C_5(2)$ であり、これらを決定済部品の集合 P の要素 P_1, P_2, P_3 とし、共出現頻度を $R(X, Y)$ (ただし X, Y は異なる部品) とする。

このとき検索の条件は、

$$\sum_{k=1}^3 R(C_3(i), P_k)$$

が最大となる i を選択することである。ただしこの例は未決定要素が一つである最も単純な例である。未決定要素が複数ある場合は、未決定要素相互の共出現関係も考慮しなくてはならない。

2.4 類似性と組合せ頻度の利用による検索手法

前出の二つの方法では類似生と共出現すなわち過去の組合せ履歴とを単独で参照していたが、それらを両方利用する方法を提案する。

この検索方法ではまず未決定要素の決定のために「参照すべき組」という過去の組合せデータの集合を抽出する。この「参照すべき組」というのは決定要素が完全に一致している組に加えて部品の特徴ベクトルの類似性から「ほぼ一致している組」を追加した組の集合である。

「参照すべき組」の選択方法を質問ベクトル $q = (1, 4, *, 0, 2)$ に適用する。

1. $(1, 4, [], \dots, 2)$ のような組。ただし“ $\dots$ ”はどのような部品でも一切関係が無いものとする。
 2. 決定要素が質問とは一部異なるが、その一致していない部品が質問中の同じカテゴリーの部品と類似性がある組。類似性の判定基準には特徴ベクトルのコサイン相関値が一定の値 θ を越えることとする。
- 決定要素のうち一つのカテゴリーの不一致を許すとすれば

- $(1, 4, [], \dots, X)$ ただし $R(X, C_5(2)) > \theta$
- $(1, X, [], \dots, 2)$ ただし $R(X, C_2(4)) > \theta$
- $(X, 4, [], \dots, 2)$ ただし $R(X, C_1(1)) > \theta$

コサイン相関値は、ある特徴ベクトルの i 番目の要素に与えられた数値を a_i 、他の特徴ベクトルの i 番目の印象語に対して与えられた数値を x_i とすると、 $\cos v = \frac{\sum a_i x_i}{\sqrt{\sum a_i^2} \sqrt{\sum x_i^2}}$ で与えられる。この値は0から1までの値となり、特徴ベクトルの類似性を反映している [3]。

未決定要素“*”はこれらの参照すべき組中の [] の最多出現部品が選ばれる。この時コサイン相関値は1以内の数値で類似度を反映しているので重みとして用いることで、過去の組合せ傾向をより繊細に反映できると考えられる。

3 モーションキャプチャーを用いた人物3Dオブジェクト

3.1 モーションデータの階層的な記述モデル

モーションデータの最大の特徴は、スケルトンモデルは part-of 関係であり、さらに part-of 関係で示された要素のそれぞれが、時間軸を持つことである [4][5]。

スケルトンモデルを構成するのはある関節から次の関節までのセグメント1つ1つであるが、通常はいくつかのセグメントからなる部位の動作としてとらえることによりはじめて意味のある動きとなる。ここで、我々は身体運動は以下に示す「機能部位動作」を行う5つの部位にわけ、「右手を大きく前に」といった基本的な動きを記述する。また、これらの組み合わせで、例えば、「手を(胸の前で)たたく」という動作のような、複数の部位による動作も記述できるように上位層を設けた。

これらのことを考慮し、スケルトンモデルに記述するための以下の4つの階層を設定した。

1. 全体動作階層: 体全体の動きに対する記述階層
「ターン」などの体全体の動きに対する記述はこのクラスを行う。
2. 半身動作階層: 機能部位動作階層で分けられた部位の組み合わせによる動作を表現する記述階層
 - 手(右手部、左手部)
 - 足(右足部、左足部)

「手をたたく」などの記述はこの階層に行う。

3. 機能部位動作階層:5つの部位に分けたある機能を実現する部分の動作に対する記述階層

- 中心 (Head, Neck, UpTorso, LowTorso)
- 右手部
(RCBone, RUpArm, RLowArm, RHand)
- 左手部
(LCBone, LUpArm, LLowArm, LHand)
- 右足部 (RThigh, RLowLeg, RFoot)
- 左足部 (LThigh, LLowLeg, LFoot)

「右手 (全体) を前に」などの記述はこの階層に行う。また、長い時間動作が行われていない場合は「待ち」、短い時間動作が行われていない場合には「ため」などのように、特に記述すべき動作が行われていない部分についても記述を行う。

4. セグメント階層:それ以上分割できない部位に対する記述

それぞれの部位に対して記述を行う。「ひじを曲げる」動作は、上腕と下腕の関係する動きである。よってこの動きに対する記述はセグメント”上腕”に行う。機能部位動作階層からの記述は、すべて継承するものとする。

また、これらとは別に「シーン層」には、そこにはダンスを大きな流れでとらえた層を用意する。ここには全体の流れを記述する。

下位の階層の記述は各部位のプリミティブな動作を記述しているが上位階層になるにしたがって抽象的、印象的が多くなる。よって、上位階層では特に明記する必要がないような動作で、記述が困難な場合には、参照の際、下位の階層が透過的に使用される。

3.2 動作記述・表示システムの開発

実空間からモーションキャプチャーを用いて得られた人物のダンスを、VRML空間では18本のセグメントと補助的なつなぎのセグメントを用いてスケルトンモデルを構成し、再現する。

モーションキャプチャーからの出力データは、3次元の位置情報やローカル座標系の変換行列の数値データである。これらのデータは最大1/120 secのサンプリング周期で計測されるために数値デー

タからその動作を推定することは困難である。したがって我々は、出力されたモーションキャプチャーからのデータを3Dモデルのアニメーションで確認し、各階層に記述を行うための記述支援ツールを開発した。このシステムではモーションキャプチャーからの出力データをVRML2.0データに変換し、3Dアニメーションとして人間の動きを再現する。このシステムの特長を以下に示す。

- 人間を21本のセグメントを持つスケルトンモデルで表現
- 音楽に同期してダンスを再現(時間軸の忠実な再現)
- セグメントをクリックすることにより、記述内容の参照が可能

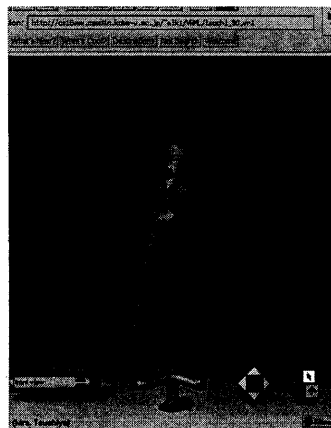


図3: VRMLによる動作記述・表示システム

4 3Dオブジェクトの著作権管理

4.1 JavaによるVRMLのカプセル化

我々が提案する仮想スタジオオーサリングシステムはインターネット、イントラネット上で使用し、仮想スタジオを生成する。そのため、3Dオブジェクトの記述や操作はインターネット上で操作可能な言語が望ましい。今回、我々は3Dオブジェクトの記述にVRML2.0を使用し、3Dオブジェクトの管理、操作にJava言語を用いることと

した。VRML2.0はSGIが提唱するEAI(External Authoring Interface)を用いることでJavaのアプレットへ組み込むことができる。本研究ではVRMLデータをEAIを用いてJavaのクラスファイルに組み込みデータベース化する予定である。

4.2 モーションコントロール

2次元のデータと違い、3次元のデータでは実空間を仮想的に構築することが望まれる。したがって、実空間の物体が持つ固有の動きを仮想空間上で実現することが必要となる。本研究では、使用するユーザのレベルを設定し、ユーザレベルによってオブジェクトの動作をコントロールする機能をJavaのメソッドとして組み込む予定である。

4.3 版權管理機能

EAIによってVRMLのシーングラフがJavaでカプセル化されることでユーザからシーングラフを隠蔽化し直接的な操作を防止することができ、従来のVRMLファイルの様な容易な編集や変更が防止できる。Javaのアプレットにデータをカプセル化し、ユーザ認証機能をメソッドとして組み込めば、認証されていないユーザに対してオブジェクトの振る舞いを制限することが可能となる[6]。この手法を用いることで使用を許可したユーザだけにオブジェクトの動作や複数のオブジェクトから構成される仮想空間上での動作を許可する機能が実現でき、不正なユーザからの使用を制限できる。本システムではネットワーク上にユーザ認証を行うための認証サーバを配置し3Dオブジェクトの版權を管理する機能を実現する予定である。

5 今後の展開

人物3Dオブジェクトの検索においては、記述を用いた検索を行っているがモーションキャプチャーによって得られたデータは正確な時刻と全身の3次元空間内の位置等が記録されている。従って指定された動作の類似検索を数値データを解析する事によって実現できる可能性がある。また、今回は動作のスピードを表現するような形容詞について

は、記述者の主観により判断したが、将来的には数値計算によりセグメントの動作の速度等をもとめて、適当な閾値を設けることによりこれらの形容詞との対応づけをはかることを考えている。最終的には「回転」「曲げ」などの動きも各セグメントのなす角などを計算することにより、比較的プリミティブな動作記述であるセグメント階層や機能部位動作階層への記述が自動的に行えるシステムへと発展させる予定である。

謝辞

本研究は、TAO 神戸リサーチセンター次世代デジタル映像通信の研究開発プロジェクトの一貫として行われた。また本研究の一部は、日本学術振興会未来開拓学術研究推進事業における研究プロジェクト「マルチメディア・コンテンツの高度処理の研究」および文部省科学研究費重点領域研究(課題番号 08244103)による。ここに記して謝意を表す。

参考文献

- [1] <http://www.studio.sgi.com/Features/VirtualSets/>
Silicon Studio Feature : Virtual Sets
- [2] 田村 秀行, 池田 克夫編, 「知能情報メディア」総研出版
- [3] Peter Ingwersen, ph.D., 「情報検索研究 認知的アプローチ」 監訳 藤原 鎮男, 訳 細野公男, 後藤智範, 岸田和明
トッパン 1995 年
- [4] 柴田 正哲, 「映像の内容記述モデルとその映像構造化への応用」 信学論 (D-2), Vol.J78-D2, No.5, pp754-764, May 1995
- [5] 小林 みな子, 今井 さやか, 有澤 博, 「リアルワールドデータベースにおける3次元情報の取得と表現」, ADBS 1996
- [6] Yutaka Kidawara, Katsumi Tanaka, Kuniaki Uehara, 「Encapsulating Multimedia Contents and A Copyright Protection Mechanism into Distributed Objects」, DEXA 1997