

実写映像への3次元人体モデルによる身体運動の合成表示

滝澤朋美[†] 佐々木大[‡] 合谷国宏[‡] 河口孝義 八村広三郎
立命館大学 理工学部

人文科学の分野でも、3次元CGのシステムと技術を利用して、民族舞踊、古典芸能などのさまざまな身体動作の記録したり、解析したりする試みが行われるようになってきている。このようにして入力した身体運動を、動きのある3次元CG像として再現する際には、背景などの情景画像は、別の何らかの方法で与える必要がある。これらもCGで作成することは、演算コストや表示能力の点で問題がある。これらを考慮し、ここでは、背景として撮影した実写のビデオ映像の中に、CGで作成した人体モデルによる身体運動を合成することを行った。

実写映像には大きさと形状が既知のレファレンス長方形を写し込み、この4頂点が画像中でとる位置を画像処理によって求め、これからレファレンス長方形のカメラ座標系での座標値を導く。さらにこの情報を使って、人体モデルのCG像をおなじカメラ座標系での実写映像中に合成する。身体運動データはキーフレームにおける姿勢を与え、中間フレームの運動は線形補間により作成する。実験により、実写映像中にCGによる人体像の運動を、違和感なく自然に合成できることが確かめられた。

Composition of Moving CG Human Body Model and Real Video Image

Tomomi Takizawa[†] Hiroshi Sasaki[‡] Kunihiro Goya[‡] Takayoshi Kawaguchi and
Kozaburo Hachimura
Department of Computer Science, Ritsumeikan University

Human body motion information processing is being used even in the Humanities field like record and analysis of ethnic dance or traditional performance. Human body motion can be recorded and re-displayed by using 3D graphics technology. In these circumstances a background scene has to be displayed together with a 3D body model. Rendering these scenes by computer graphics is ideal but expensive in terms of processing cost. Our approach is to use real pictures taken by video camera, a rendered 3D human body model being fused into the video picture.

Background scene including a reference rectangle, whose actual size is known, is taken by a video camera. The four corner points of the rectangle are extracted by image processing, and 3D coordinates of these corners can be obtained by solving simultaneous equations. 3D human body model can be placed in the same 3D coordinate system, and the rendered image of the model is fused into the background video picture keeping geometrical consistency. Movement of the model is specified by a key frame based description.

[†]現在 (株) 松下ソフトリサーチ

[‡]現在 オムロンソフトウェア (株)

[†]Currently with Matsushita Soft-Research, Inc.

[‡]Currently with Omron Software Co., Ltd.

1 はじめに

計算機の処理能力の向上や3次元グラフィックス技術の性能向上により、3次元人体モデルによる人間の身体動作の表示ができるようになった。そのため、人文科学の分野でも、このようなシステムと技術を、民族舞踊、古典芸能などのさまざまな身体動作の記録に利用しようという試みが行われるようになってきている。

われわれは、舞踊の記述に使用されている舞踊譜 Labanotation を計算機で取り扱えるようにして、これを動作入力に利用したり [1]、入力した動作データを編集して再利用するためのシステムの開発 [2] など、総合的な人間の身体運動処理システムの構築を目指してさまざまな観点から研究を行っている。

何らかの方法により、人間の身体運動を入力・記述したデータは、3次元のデータとして記録されるので、表示の際には、任意の人体モデルに対してこのデータを適用し、動きのある3次元CG像を作ることができる。身体運動の入力・記述の際には、人体動きのデータだけが対象となるので、この時の周囲の風景や環境については特に考慮が払われずに入力が行われる。したがって、入力した身体運動をグラフィックスとして再現する際には、これらの情景画像は、別の何らかの方法で与える必要がある。一つは、これらもCGにより作成することであるが、これでは、無機質な映像になる傾向があり、また、実写に近い忠実な映像を作るのは、演算コストや表示能力の点で問題がある。このような点を考慮し、ここでは、背景として撮影した実写のビデオ映像の中に、CGで作成した人体モデルによる身体運動を合成することを検討した。

2 実写映像とCG像の合成

近年、拡張現実感 (Augmented Reality)、あるいは、複合現実感 (Mixed Reality) と呼ばれる技術が注目されている。Augmented Reality (AR) とは、実世界の環境に、たとえばCGで作った仮想的な世界をスムーズに融合させ、増強された新しい空間や環境を実現するための技術である。

ARにおける最も基本的な手法は、実世界の映像とCGによる映像を合成して表示するものである。本研究はこの技術を身体運動の3次元表示に応用す

る。すなわち、1台のビデオカメラにより撮影される実写映像に、人体モデルの3次元CG像による身体動作を合成する。

ビデオカメラによる実写映像は2次元の画像であり、CGの人体モデルは3次元データである。これらを、違和感なく自然に合成するためには、実写映像を撮影する際のカメラ位置や視線方向などのカメラのパラメータが分かっていることが必要である。

このため、ここでは、実写映像中に形状が既知のレファレンス物体を写し込み、この物体上の複数の特徴点が画像中でとる座標を求めて、これらの関係からカメラパラメータを推定する。実写のカメラパラメータが推定できれば、このパラメータを使って3次元人体モデルを同じ画面内に自然に合成することができる [3][4]。

本研究では、レファレンス物体には長方形を用い、長方形の4頂点を特徴点として利用する。

システムの全体構成を、図1に示す。システムは、実写映像に対する画像処理の部分と、CGモデルによる身体運動の表示の2つの部分から成り立っている。

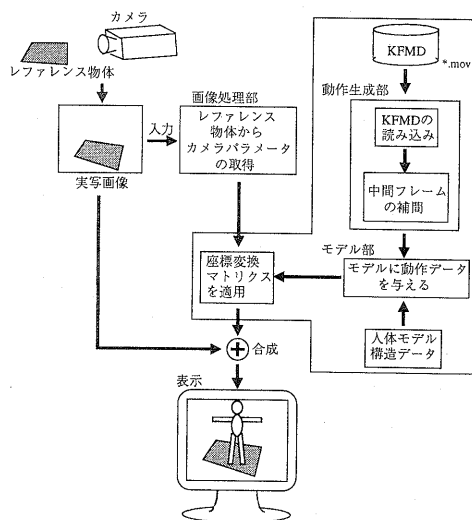


図1: システムの概要

画像処理の部分では、形状が既知のレファレンス物体の特徴点の2次元の投影面座標から、カメラパラメータを推定する。本研究では長方形のレファレンス物体 (以下では、これをレファレンス長方形

と呼ぶ)を用い、この全体が実写画像中に必ず写し込まれているものとする。

一方、モデル表示部では、人体モデルを OpenGL によりレンダリングして、この CG 画像を作成する。前述の画像処理により得られたカメラパラメータを用いて、人体モデルに座標変換を施し、実写画像と CG 画像による人体モデル像の幾何学的整合性をとった合成を行う。身体の動作データには、キーフレームにおける姿勢を記述した、図 1 中の KFMD (Keyframe Movement Data) が参照される。

3 実写映像からの 3 次元カメラ座標の推定

3.1 レファレンス長方形の検出

図 2 に示すような、実写映像中のレファレンス長方形を検出するために、以下のような画像処理を行う。

1. カラー画像の白黒濃淡化
2. 平滑化とエッジ検出
3. 2 値化、孤立点除去と細線化
4. Hough 変換による線要素の抽出
5. 線分の端点の抽出
6. レファレンス長方形の頂点の検出

すなわち、微分処理によってエッジ画像を求め、これからさらに、2 値化と細線化によりレファレンス長方形の輪郭を示す線図形を抽出する。図 2 に対してエッジ検出と細線化処理を行った例を図 3 に示す。

さらに、この輪郭画像に対して、Hough 変換によりレファレンス長方形の辺をなす 4 直線を導いて、長方形の頂点の位置を求める。

以下では、Hough 変換処理による線抽出処理を中心に述べる。

3.1.1 Hough 変換

Hough 変換とは、画像中の線要素(直線や曲線)を抽出するための基本的な手法である。

いま直線を検出することを考える、 x - y 空間にある直線の式を、パラメータ θ と ρ を用いて、式 (1) で表

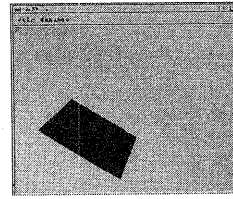


図 2: レファレンス長方形を含む実写画像

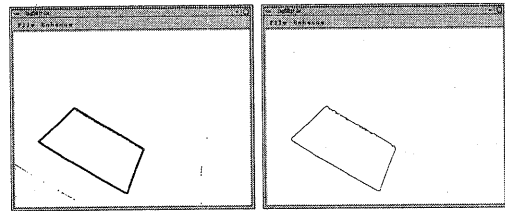


図 3: エッジ検出と細線化

現する。ここで ρ は原点から直線までの距離、 θ は原点から直線への垂線が x 軸となす角度である。

$$\rho = x \cdot \cos \theta + y \cdot \sin \theta \quad (1)$$

このとき、直線上の任意の点 (x_0, y_0) は次式を満足する。

$$\rho = x_0 \cdot \cos \theta + y_0 \cdot \sin \theta \quad (2)$$

この関係を用いると、 x - y 空間内の直線上の 1 点は θ - ρ 空間では 1 本の正弦曲線として表され、同一の直線上に存在する点に対応する正弦曲線は 1 点で交わる。この点の座標 (θ, ρ) が求める直線のパラメータの値となる。

このような直線のパラメータ (θ, ρ) を求めるために、 θ - ρ 空間に対応する 2 次元配列を用意し、正弦曲線が通る配列要素に 1 を加える(投票する)操作を行う。これを画像中のすべての黒画素について行なって、投票数の大きな要素を抽出すれば、画像中に存在する直線のパラメータ θ と ρ を求めることができる。

3.1.2 θ - ρ 空間から極大点の抽出

実際の実写画像では、ノイズが存在するため、Hough 変換によるパラメータ空間 (θ - ρ 空間)でも、傾きや

切片のわずかに異なる直線群が存在することになる。このようにパラメータのわずかに異なる直線群は、ひとつの直線であるとしてまとめる必要がある。ここでは、 θ - ρ 空間において、投票数が極大値をとり閾値以上である点を抽出することによってこれを行う。すなわち、 θ - ρ 空間の各点において 9×13 近傍を調べ、この範囲における極大点を抽出していく。したがって、この範囲内に収まる直線は、極大値の位置の直線で代表され、まとめて1つの直線として抽出されることになる。得られた θ, ρ の値を利用して直線の解析式(1)を求める。

以上のようにして行なった直線の抽出結果の例を図4(a)に示す。

3.1.3 線分の抽出

次に、得られた直線のうちレファレンス長方形の辺をなす部分(線分)を求め、その両端の座標を得る。これは、Hough変換により求められた直線を、エッジ画像によりマスクして取り出すことにより行う。線分を抽出した結果を図4(b)に示す。

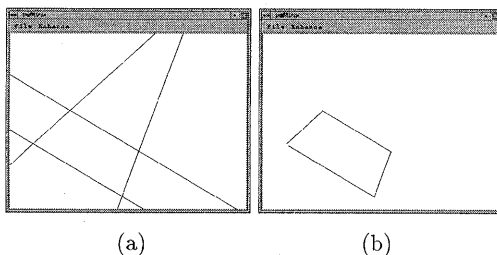


図4: Hough変換による直線の検出とレファレンス長方形の辺の抽出

3.1.4 投影面における頂点の抽出

Hough変換によって求めたすべての直線の組合せについて、まずその交点を算出する。このようにして求めた交点には、画像の範囲外に存在するものや、レファレンス長方形の頂点に対応しないものも含まれている。このため、それらを取り除く処理を行う必要がある。

まず最初に画像の範囲外の交点は除外する。また、レファレンス長方形の、本来一つの辺であるべきと

ころに、2本以上の直線が検出されるときがあるが、このようなときには、レファレンス長方形の頂点付近で複数の交点が得られる。ここでは、マスク処理により、これらの交点を1つにまとめる。

さらに、すでに求められている長方形の辺の近傍内に存在する交点だけを選び出す。

このようにして残った交点を、レファレンス長方形の頂点とする。

3.2 レファレンス長方形の3次元カメラ座標の推定

図5に示すような、世界座標系と、カメラのレンズ中心を原点としたカメラ座標系との関係は、モデルビュー行列 A を用いて、以下のように表される。すなわち、齊次座標系を用いて、ある点の世界座標系での座標を $w = (w_x, w_y, w_z, 1)^t$ 、カメラ座標系での座標を $c = (c_x, c_y, c_z, 1)^t$ とすると、

$$c = A \cdot w \quad (3)$$

となる。ここで、

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (4)$$

である。

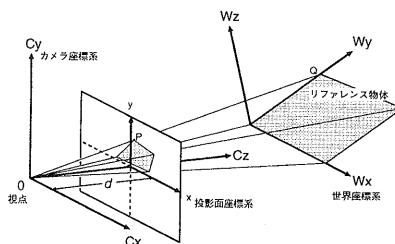


図5: 座標系の関係

一般的には、既知の点の投影面座標の値からこのモデルビュー行列を求めることができれば、世界座標系で表現された任意のCGモデルに対して、撮影カメラと同じ条件でCG像をレンダリングすることができる。

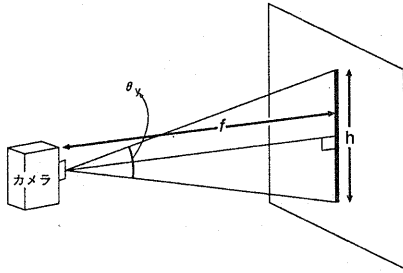


図 6: カメラの垂直画角の測定方法

しかしながら、ここでは、カメラによって撮影されている実写映像に CG 像を合成することが目的であるから、CG 像の座標を、カメラの中心を原点とするカメラ座標系で表現できればよい。このためには、レファレンス物体(長方形)のカメラ座標系での座標値を求めて、これを基準にして CG 像のカメラ座標値を与えればよいことになる。

すなわち、レファレンス長方形の、3次元カメラ座標系での座標値を求めることがここでの処理の中心となる。以下ではこれについて述べる。

ここでは、次の条件が成り立っているものとする。

1. レファレンス物体には長方形を用い、この形状と大きさが既知である。
2. 長方形の4頂点は同一平面上にある。
3. カメラの垂直画角は事前に測定しておき、変更はしない。

カメラの垂直画角 θ_y は以下の手順で測定する。

- (a) 垂直面内に長さ h の物体を置く。
- (b) 図 6 のように、物体と垂直になるようにカメラを向け、長さ h の物体がちょうど画面の上端から下端に収まるように距離をとり、その際の距離 f を測定する。
- (c) $\theta_y = 2 \times \tan^{-1}(\frac{h}{2f})$ により θ_y を求める。

さて、長方形の4頂点が同一平面上にあることから、この平面を物体座標系の $z=0$ 平面として考えることができ、このことより、 $w_{zi} = 0$ となり、式

(3)は式(5)とおける。

$$\begin{pmatrix} c_{x_i} \\ c_{y_i} \\ c_{z_i} \\ 1 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} w_{x_i} \\ w_{y_i} \\ 0 \\ 1 \end{pmatrix} \quad (5)$$

ここで、 a_{13}, a_{23}, a_{33} はどんな値を持ってよい。そこでこれを整理すると、

$$\begin{pmatrix} c_{x_i} \\ c_{y_i} \\ c_{z_i} \\ 1 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{14} \\ a_{21} & a_{22} & a_{24} \\ a_{31} & a_{32} & a_{34} \end{pmatrix} \begin{pmatrix} w_{x_i} \\ w_{y_i} \\ 1 \end{pmatrix} \quad (6)$$

となる。

また、カメラ座標系での座標値 $Q(c_{x_i}, c_{y_i}, c_{z_i})$ と、Z軸に垂直で $z=d$ の位置においた投影面上での2次元座標 $P(X_i, Y_i, Z_i)$ には、透視投影の関係から以下のような式が成り立つ。

$$X_i = \frac{c_{x_i}}{c_{z_i}} \times d \quad (7)$$

$$Y_i = \frac{c_{y_i}}{c_{z_i}} \times d \quad (8)$$

$$Z_i = d \quad (9)$$

式(6)(7)(8)より、

$$X_i = \frac{a_{11}w_{x_i} + a_{12}w_{y_i} + a_{14}}{a_{31}w_{x_i} + a_{32}w_{y_i} + a_{34}} \times d \quad (10)$$

$$Y_i = \frac{a_{21}w_{x_i} + a_{22}w_{y_i} + a_{24}}{a_{31}w_{x_i} + a_{32}w_{y_i} + a_{34}} \times d \quad (11)$$

これを整理すると、

$$\begin{aligned} a_{34}X_i &= da_{11}w_{x_i} + da_{12}w_{y_i} + da_{14} \\ &\quad - a_{31}w_{x_i}X_i - a_{32}w_{y_i}X_i \\ a_{34}Y_i &= da_{21}w_{x_i} + da_{22}w_{y_i} + da_{24} \\ &\quad - a_{31}w_{x_i}Y_i - a_{32}w_{y_i}Y_i \end{aligned} \quad (12)$$

となり、 $i=1, \dots, 4$ の場合をまとめて以下の式が得られる。

$$a_{34} \begin{pmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \\ Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{pmatrix} = B \begin{pmatrix} a_{11} \\ a_{12} \\ a_{14} \\ a_{21} \\ a_{22} \\ a_{24} \\ a_{31} \\ a_{32} \end{pmatrix} \quad (13)$$

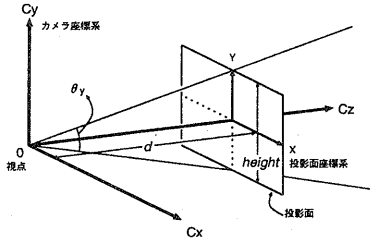


図 7: カメラの投影面サイズと画角

ここで、

$$B = \begin{pmatrix} dw_{x1} & dw_{y1} & d & 0 & 0 & 0 & -w_{x1}X_1 & -w_{y1}X_1 \\ dw_{x2} & dw_{y2} & d & 0 & 0 & 0 & -w_{x2}X_2 & -w_{y2}X_2 \\ dw_{x3} & dw_{y3} & d & 0 & 0 & 0 & -w_{x3}X_3 & -w_{y3}X_3 \\ dw_{x4} & dw_{y4} & d & 0 & 0 & 0 & -w_{x4}X_4 & -w_{y4}X_4 \\ 0 & 0 & 0 & dw_{x1} & dw_{y1} & d & -w_{x1}Y_1 & -w_{y1}Y_1 \\ 0 & 0 & 0 & dw_{x2} & dw_{y2} & d & -w_{x2}Y_2 & -w_{y2}Y_2 \\ 0 & 0 & 0 & dw_{x3} & dw_{y3} & d & -w_{x3}Y_3 & -w_{y3}Y_3 \\ 0 & 0 & 0 & dw_{x4} & dw_{y4} & d & -w_{x4}Y_4 & -w_{y4}Y_4 \end{pmatrix} \quad (14)$$

である。 d は視点から投影面までの距離であるが、図 7 より、以下の式を使って与えることができる。

$$d = \frac{height}{2 \tan(\frac{\theta_y}{2})} \quad (15)$$

ここで、 θ_y はカメラの垂直画角、 $height$ は投影面の高さ (ピクセルサイズ) である。

また、式 (13) において、パラメータ a_{34} はスケールファクターとなるので、これは任意の値にとることができる。そこで $a_{34} = 1$ とすると、式 (13) は、

$$\begin{pmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \\ Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{pmatrix} = B \begin{pmatrix} a_{11} \\ a_{12} \\ a_{14} \\ a_{21} \\ a_{22} \\ a_{24} \\ a_{31} \\ a_{32} \end{pmatrix} \quad (16)$$

となる。

式 (16) は、 a_{ij} を未知数とする線形連立方程式であるので、これをガウス法により解くことができる。

このようにして求めた a_{11} 、 a_{12} 、 a_{14} 、 a_{21} 、 a_{22} 、 a_{24} 、 a_{31} 、 a_{32} と式 (6) を用いて、レファレンス長方形の 4 頂点の 3次元カメラ座標系での座標値 (c_{x_i} 、 c_{y_i} 、 c_{z_i}) ($i = 1, \dots, 4$) を求めることができる。

4 実写映像への身体運動の合成

4.1 人体モデルの動作表示

4.1.1 人体モデル

本研究で用いた人体モデルは、図 8 に示すように、それぞれが 3 自由度を持つ 21 個のジョイントからなる多重リンク構造で作成している。各ジョイントには図 8 に示す名前が付けられている。人体モデルの各セグメントの形状は、OpenGL により八角柱でモデリングしている。

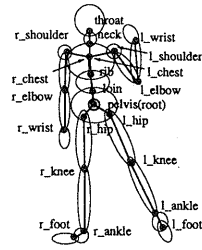


図 8: 人体モデル

このモデルは、Pelvis(腰) をルートとした階層構造となっており、Pelvis を移動させることによってモデル全体を前後、左右、上下に移動させることができる。

4.1.2 キーフレーム動作記述データ

人体モデルの身体運動の記述には、キーフレーム動作記述データ (KFMD) を用いた。KFMD とは、キーとなるフレームの姿勢のみを記述するもので、キーフレーム間のフレーム (中間フレーム) の姿勢は、補間して生成する。

図 9 にキーフレーム動作記述データの例を示す。#Frame 0 に続く 2 行は動作開始時点での姿勢を表している。ジョイント名の行の 3 つの数値は、3 軸まわりの回転角度を示している。#From 0 To 25 に続く行の記述は、フレーム 0 から動作を開始し、フレーム 25 において、記述された姿勢をとることを表している。また、Move の行の 3 つの数値パラメータが、人体モデル全体の移動量を表している。

4.1.3 動作生成

KFMD による人体モデルの姿勢を補間して人体モデルの動作を生成する。ここでは、2つのキーフレーム間の対応するジョイントの位置を1次式で補間する線形補間を用いて、中間フレームの動作を生成する。

```
#Frame 0
R_Shoulder 0,0,20
L_Shoulder 0,0,-20
#From 0 To 25
R_Shoulder -180,0,90
L_Shoulder -180,0,-90
#From 25 To 50
R_Shoulder -180,-90,180
L_Shoulder -180,90,-180
#From 50 To 55
R_Shoulder 0,-90,0
L_Shoulder 0,90,0
#From 55 To 75
R_Shoulder 0,0,20
L_Shoulder 0,0,-20
#From 75 To 100
Move 0.0,0.0,0.0
R_Shoulder -53,44,0
R_Elbow 0,63,0
L_Shoulder -31,-39,0
L_Elbow 0,-63,0
R_Hip 0,0,0
R_Knee 0,0,0
R_Ankle 0,0,0
L_Hip 0,0,0
L_Knee 0,0,0
L_Ankle 0,0,0
```

図 9: キーフレーム動作記述データの例 (一部)

4.2 実写映像と人体モデル CG 像の合成

4.2.1 環境設定

実写映像と人体モデルを合成するためには、実写映像と CG 空間の環境を描える必要がある。本研究で用いたビデオカメラと、得られる画像の特性を表 1 に示す。一方、人体モデルの像をレンダリングする際の OpenGL の視界環境を図 10 に示す。パラメータ θ 、 h 、 w を、表 1 で示したビデオカメラの特性に合わせる。

表 1: 使用したカメラの特性

垂直画角	37 度
水平方向サイズ	640
垂直方向サイズ	480

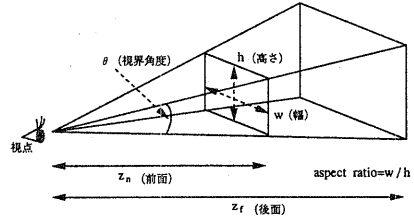


図 10: OpenGL における視界環境

4.2.2 座標変換

実写映像への CG 像の合成の際には、図 11 に示すように、レファレンス長方形のカメラ座標系での位置を基準にして、座標変換により人体モデルの位置と方向を決める。

たとえば、レファレンス長方形の重心位置に、長方形に垂直に立つようにするには、以下の手順により、人体モデルの座標系 (モデル座標系と呼ぶ) に座標変換を適用する。

1. 長方形の重心位置に平行移動
2. 長方形の法線方向に回転
3. 長方形のサイズに合わせてスケーリング

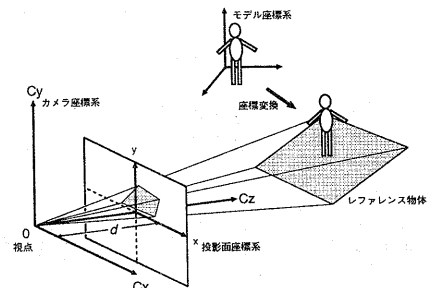


図 11: レファレンス長方形と人体モデル

4.2.3 隠面消去処理

人体モデルがレファレンス長方形に隠れる場合、隠される部分を消去した人体モデルを実写映像と合成する必要がある。そのために、検出したレファレンス長方形の位置に、同じ大きさの長方形を定義し、これをマスクとして用いる。この長方形は OpenGL の混合処理を用いて実写映像の長方形部分と同色で表示する。これにより、人体モデルがレファレンス長方形によって隠される場合でも、前後関係を正しく保ったまま、身体運動を表示することができる。

5 実験結果

実写映像中に写し込まれたレファレンス長方形を基準として、人体モデルの身体運動を合成表示した結果を図 12 に示す。人体モデルの歩行運動のデータは KFMD で与える。図 12 は、両腕を上に向けてポーズを作りながら、レファレンス長方形上を歩行している動画像のスナップショットであるが、幾何学的整合性を保ったまま、CG による人体像の運動が表示できていることが分かる。

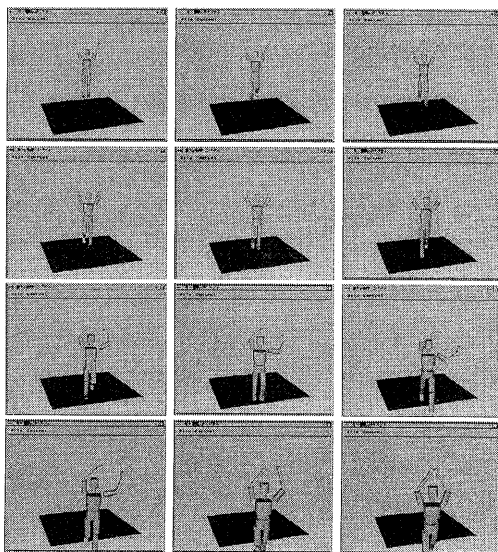


図 12: 実行結果

6 おわりに

実写映像中に写し込まれた、形状が既知のレファレンス長方形の頂点を画像処理により抽出し、これから、これらの頂点の、カメラ座標系での 3 次元座標を導いた。この情報を用いて、実写の画像と 3 次元人体モデルを合成し、身体運動アニメーションを表示させることができた。

本研究では、レファレンス長方形は固定しておき、カメラの方が動くことを前提にした。レファレンス長方形の頂点の座標値が計算されるたびに、人体モデルに適用する座標変換マトリクスを更新し、これを人体モデルに適用して、合成画像中で人体モデルが取るべき姿勢を決めている。当初は、動画像の実写映像に人体モデルの動作を合成することを目標にしたが、処理速度や、画像中のノイズの問題などのために動画像に対して満足できる結果を得られなかった。このため、今回の実験では静止した実写映像を用いた。特に、Hough 変換による長方形の頂点検出に大部分の時間を費やしているの、今後は、アルゴリズムの見直しによって処理速度の向上をはかり、実写動画像に対しても、ビデオレートでの処理が行えるようにする予定である。

また、レファレンス物体を写し込む実写映像についても、今回のような、白背景に黒っぽい長方形という実験的なものではなく、たとえば、実際の屋内の廊下の映像や、屋外の道路や建物形状をレファレンス物体として利用して、このような画像中に CG による人体モデル像の身体運動を合成表示できるようにすることを考えている。

参考文献

- [1] 吉田康行, 松岡洋介, 八村広三郎: 舞踊譜 Labanotation に基づく身体運動の処理 - 譜面読み取り LabanReader と 譜面エディタ LabanEditor -, 情報処理学会研究報告, 98-CH-38, pp.61-68, 1998.
- [2] 平松尚子, 八村広三郎: 身体動作の記述と表示のための実行制御環境の開発, 情報処理学会論文誌, Vol.40, No.3, pp.939-948, 1998.
- [3] 大隈隆史, 清川清, 竹村治雄, 横矢直和: ビデオシンセ型拡張現実のための実画像からのカメラパラメータの実時間推定, 電子情報通信学会技術報告 PRMU97-113, pp.13-20, 1997.
- [4] 中沢裕二, 中野誠巳, 小松隆, 斎藤隆弘: 画像中の特徴点に基づく実写画像と CG 画像との動画合成システム, 映像情報メディア学会誌, Vol.51, No.7, pp.1086-1095, 1997.