

3次元ビデオ映像世界の開拓

情報学研究科知能情報学専攻知能メディア講座

教授 松山 隆司

tm@i.kyoto-u.ac.jp

助教授 和田 俊和

twada@i.kyoto-u.ac.jp

講師 杉本 晃宏

sugimoto@i.kyoto-u.ac.jp

1. 3次元ビデオ映像とは

1960年代から始められたデジタル画像処理研究は、この40年間に着実な進歩を遂げ、最近ではマルチメディア情報処理を支える基盤技術として更なる発展を続けている。

この間、処理対象となる画像データは、以下のように次第に多次元化されてきた。

2値画像 (2次元×1ビット)	: 文書、図面
濃淡画像 (2次元×8ビット)	: X線像
カラー画像 (2次元×24ビット)	: 衛星写真、カラーキャナ
3次元画像 (3次元×8ビット)	: 3次元CT像
動画像 (2次元×1次元×8 (24) ビット)	: ビデオ映像
距離画像 (2.5次元)	: ステレオ視、レンジファインダ
距離動画像 (2.5次元×1次元)	: 実時間ステレオ視、実時間レンジファインダ

最近では対象の周囲に多数の距離センサを配置し、得られた距離画像群から対象の完全な3次元形状を求める研究や、3次元形状とともに表面テクスチャ・色も同時に求める3次元像の復元、すなわち

3次元形状 (3次元)	: 全周レンジファインダ
3次元像 (3次元×24ビット)	: 全周ステレオ視

が実用に近づきつつある。

我々は、こうした多次元化をさらに一歩進めた3次元ビデオ映像 (運動対象の動的3次元像)、すなわち

3次元ビデオ映像 (3次元×1次元×24ビット) : 多視点ビデオ映像

の能動的実時間撮影・圧縮・編集・表示に関する研究を進めている。

3次元ビデオ映像は、コンピュータグラフィックスによる仮想的な3次元アニメーションではなく、ダンスやスポーツをする人間、自然界の動物などの生の姿・形・色の時間的変化を3次元的にそのま

ま記録した実写立体ビデオ映像で、実世界における対象の振る舞い・動作を余すところなく記録した究極の映像メディアである。

我々は、3次元ビデオ映像という新たな情報メディアを開拓し、DVDやデジタルテレビ放送、広帯域インターネットを活かした21世紀における映像メディア技術および文化を産み育てることを目的として研究を進めている。具体的には、3次元ビデオ映像を利用することによって、以下のような応用システムが実現できる。

- ・人間国宝やオリンピック選手の動作をそのまま記録再現できる身体技能・芸能デジタルアーカイブ
- ・動物のありのままの生態を多角的に観察できるDVD 3次元ビデオ映像図鑑
- ・デジタルテレビ放送や広帯域インターネットをインフラとして使った3次元テレビ放送

以下では、これまでに開発した3次元ビデオ実時間撮影システムを紹介し、今後の展望を述べる。

2. 3次元ビデオ実時間撮影システム

現在稼働中の3次元ビデオ実時間撮影システムは、人物の周囲に配置された9台のビデオカメラと10台のPCを持つPCクラスタから構成されている(図1)。各PCにはそれぞれ我々が開発した視点固定型パン・チルト・ズームカメラ【1】が備えられており、移動する対象を実時間で追跡してその映像を撮影することが可能である【2】(対象追跡機能は現在開発中)。また、PCは高速の(1.25Gbit/sec)ネットワークで結ばれており、PCクラスタを1つの並列計算機として利用することによって実時間で3次元形状復元計算を行うことができる。

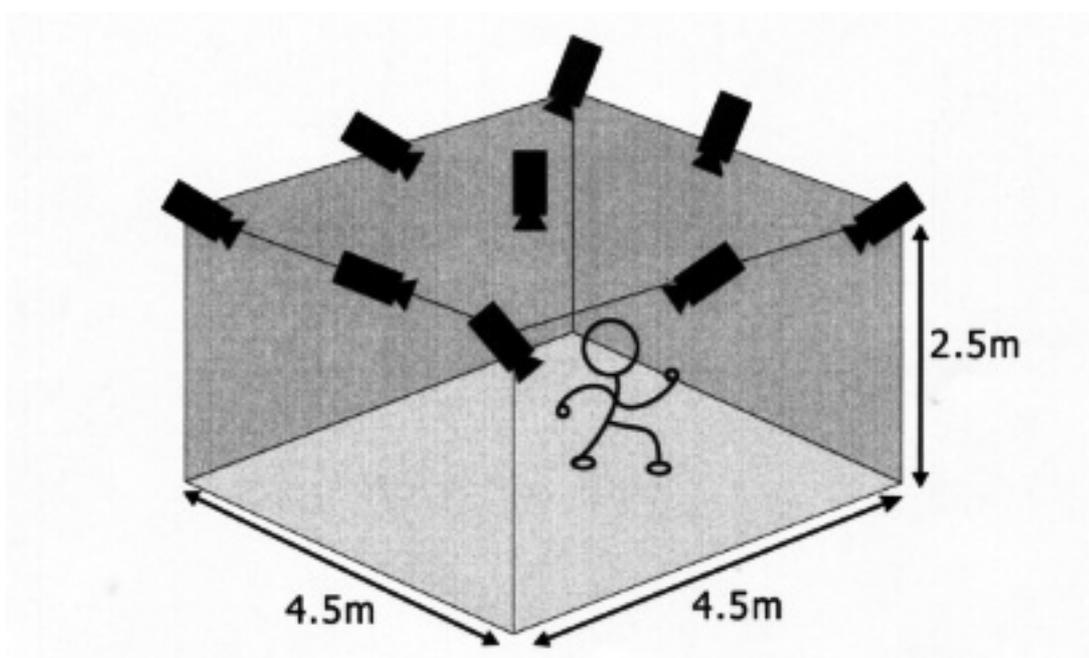


図1. 3次元ビデオ映像撮影システム

現在のシステムでは以下の方法で3次元ビデオ映像の撮影を行っている。

- (1)各PCにおいて、カメラで撮影されたビデオフレーム画像から背景差分によって対象のシルエットを抽出する。画像撮影は、ネットワークを通じた同期メッセージによって起動され、9台のカメラが同期して撮影を行う。
- (2)各カメラの投影中心(図2のPA、PB)を中心としてシルエットを3次元空間に逆投影し、得られた視体積の積集合によって対象形状の3次元ボクセル表現を求める(図2)。我々のシステムでは、この視体積交差法を高速に行うため、3次元空間を平行平面群によって表し、各平面上での対

象の断面形状を計算するというアルゴリズム（図3）を考案し、それをPCクラスタによって並列実行している【3】。

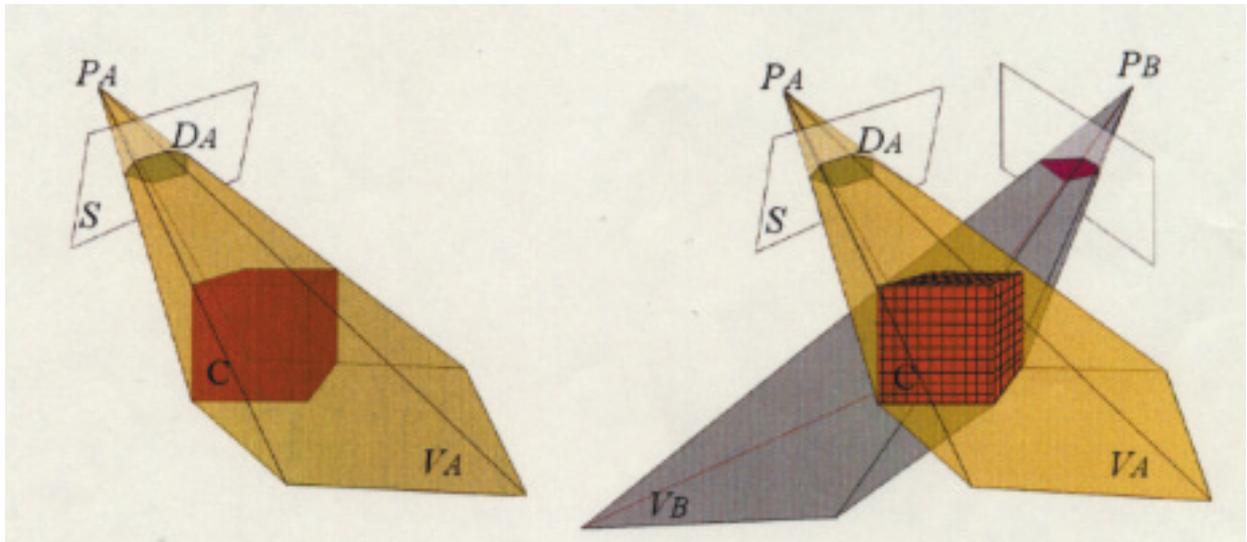


図2. 視体積交差法による3次元形状の復元の原理

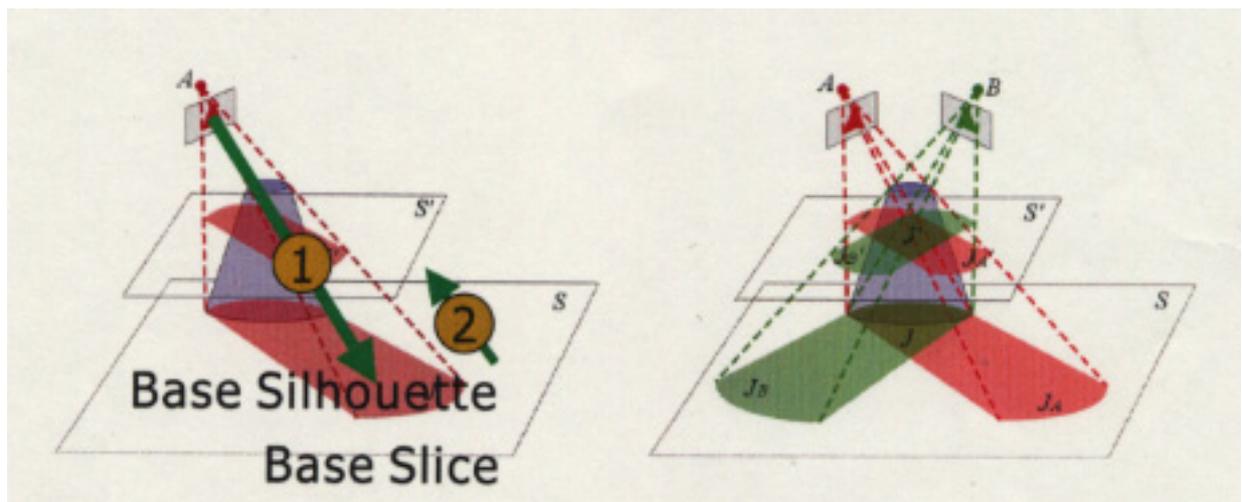


図3. 平行平面群に基づく視体積交差法

- (3) 視体積交差法で得られたボクセルデータ（図4左）の表面に3角形，4角形パッチを張り，対象の表面形状を求める（図4中央）。
- (4) パッチ上のテクスチャや色は多数のカメラによって写されているため，パッチが最もよく写っている画像を選び，その画像上のテクスチャ・色をパッチに張る（図4右）。
- (5) 以上の処理で，1フレームの3次元ビデオ映像が生成される。動画を作るには(1)～(4)の処理を繰り返えせばよい。

図5は踊っている女性の姿を撮影した3次元ビデオ映像を、適当に視点位置を移動させながら表示したものである。このように、3次元ビデオ映像の大きな特長は、対象を見る視点の位置・移動法，ズーム倍率を視聴者がその場で対話的に指定できることにあり、通常のビデオとは異なったより多角的な映像の楽しみ方が可能となる。また、対象の3次元的な運動が克明に記録されているため、スポーツや踊りにおける動作分析がより正確に行える。

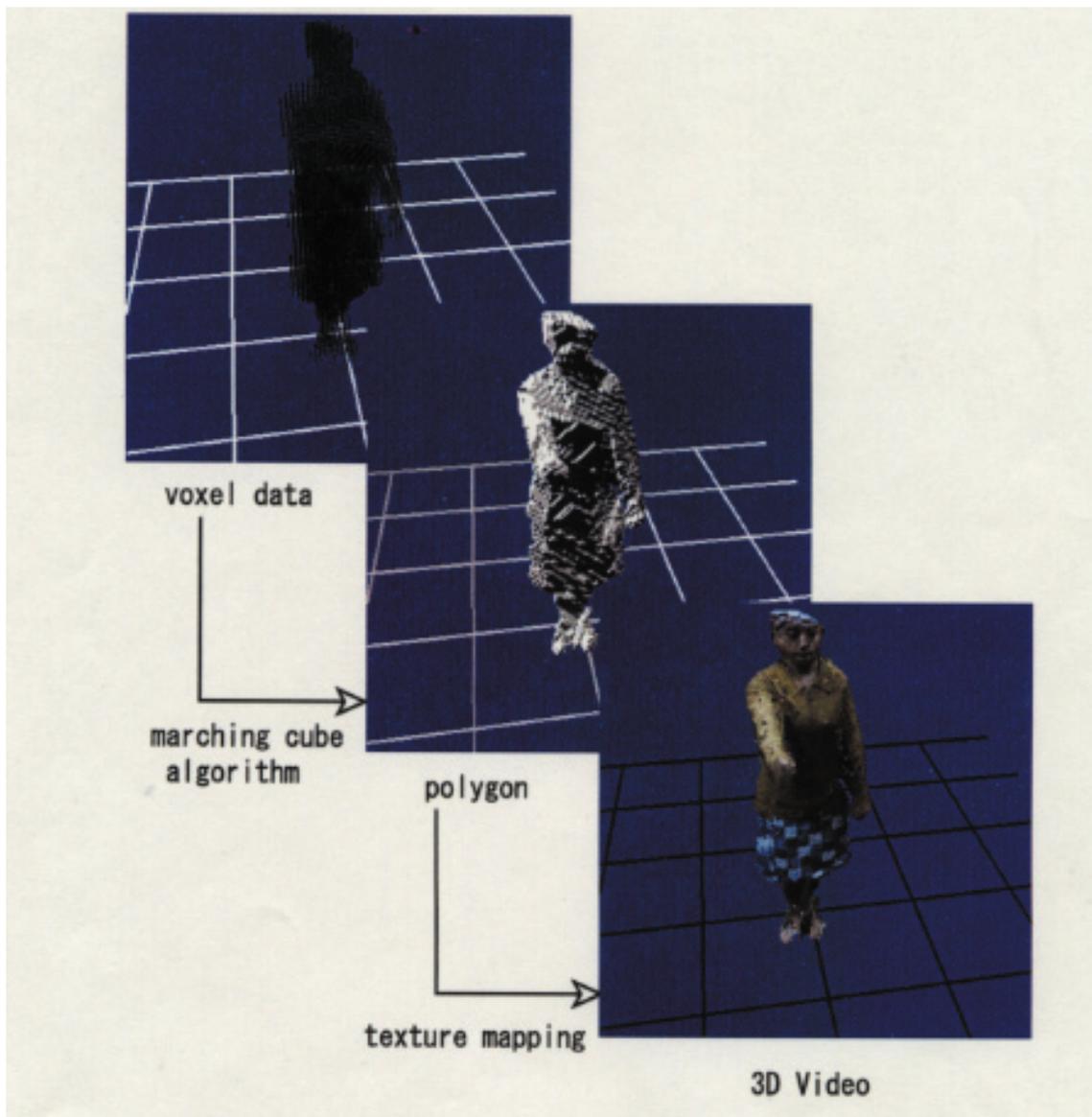


図4. 表面形状、テクスチャの復元

3. 今後の展望

現在のシステムで撮影できる3次元ビデオ映像は、空間解像度が $2\text{ cm} \times 2\text{ cm} \times 2\text{ cm}$ 、時間解像度が毎秒10コマと低く、かつカメラが小型のアナログビデオカメラであるため映像の画質も十分とは言えないのが現状である。さらに、カメラが固定されているため、対象の移動範囲も $2\text{ m} \times 2\text{ m} \times 2\text{ m}$ に限られている。

今後は、先に述べた運動対象の自動追跡機能を付加するとともに、デジタルビデオカメラを使った高画質映像の撮影、従来の2次元ビデオ映像と比べ桁違いに大量となるデータの圧縮・符号化法の考案、3次元ビデオ映像に映された対象の動作を解析し、その結果に基づいて複数の3次元ビデオ映像を編集し映像作品に仕上げるための映像編集法の考案、2次元あるいは3次元ディスプレイを使った3次元ビデオ映像の対話的表示、に関する研究を押し進め、3次元ビデオ映像を21世紀の映像メディアとして育てていきたいと考えている。



図5. 3次元ビデオ映像の表示

【参考文献】

- 【1】和田 俊和, 浮田 宗伯, 松山 隆司: 視点固定型パン・チルト・ズームカメラとその応用, 電子情報通信学会論文誌, Vol. J81-D-II, No.6, pp. 1182-1193, 1998
- 【2】松山 隆司, 和田 俊和, 物部 祐亮: 視点固定型パン・チルト・ズームカメラを用いた実時間対象検出・追跡, 情報処理学会論文誌, Vol. 40, No.8, pp.3169-3178, 1999
- 【3】T.Wada, X. Wu, S. Tokai, and T.Matsuyama: Homography Based Parallel Volume Intersection: Toward Real-Time Volume Reconstruction Using Active Cameras , Proc. of Computer Architectures for Machine Perception, pp.331-339, 2000