１．研究活動の概要

１．１　研究の概要

本報告書の「はじめに」において述べたように、本プロジェクトでは、
（１）能や日本舞踊などの無形文化財のデジタル化
（２）大仏殿や遺跡などの大型有形文化財のデジタル化
のためのソフトウェアを開発することを目的としており、
（１）については京都大学大学院情報学研究科松山研究室で開発してきた３次元ビデオに関する技術
（２）については東京大学生産技術研究所池内研究室で開発してきた大仏像の３次元デジタル化技術
を基礎として、より高度な技術を開発し、実用に耐えうるレベルの高精度デジタル化ソフトウェアを開発することを目指している。

ここでは、無形、大型有形文化財のデジタル化を目指した研究開発の全体的枠組みおよび平成１６年度、１７年度に得られた成果の概要について述べる。なお、成果の技術的詳細については、本報告書後半に関連論文をまとめてあるので、それらを参照して頂きたい。

１．１．１　無形文化財の高精度デジタル化ソフトウェアの開発

（ａ）３次元ビデオ：多視点映像からの３次元映像の生成

ここ数年、分散配置された多数のカメラによって、動的に変化するシーンや運動対象を異なった視点から同期撮影し、得られた多視点ビデオから、シーンや対象の３次元映像を復元したり、任意の視点から見た映像を合成する研究が盛んに行われている。

多視点ビデオからの３次元・自由視点映像生成のための手法を大別すると、画像ベース法とモデルベース法に分けられる。表１は両者の特徴を比較したもので、前者はシーン全体を写した映像の変換、後者は撮影対象の切り出し・３次元形状復元を基本としている。

京都大学松山研究室では、能や日本舞踊などの無形文化財のデジタル・アーカイブ化を目指した研究を進めており、人物の３次元的な動作そのものを記録・保存するという目的のためにはモデルベース法が適していると言える。また、モデルベース法では、対象の３次元的コピーや幾何学的・光学的編集が自由に行え、多様な映像コンテンツが創作できるほか、対象の３次元運動解析など映像化以外の応用も可能で利用範囲が広い。

このような観点から、我々は無形文化財の高精度デジタル化の基本的枠組みとして、モデルベース法による3次元・自由視点映像生成を行う３次元ビデオを取り上げ、その高精度化を目指して研究を行っている。

	処理の基本的方式	カメラ台数 カメラ間距離	画質	撮影対象	視点移動	照明変化	コンテンツ編集	**対象の運動解析**
画像ベース法	各視点から撮影された画像を基に、視点間の任意の位置から見た画像を合成する。	100台規模以上 密な配置が必要	高	シーン全体、複雑な自然物も可	カメラ近くに限定	不可	不可	**不可**
モデルベース法	多視点画像から、対象の3次元形状を復元し、3次元像を合成する	数台～数十台 疎な配置でも可能	中	人物、動物などの3次元対象	完全に自由	自由	自由	**可能**

表１　画像ベース法とモデルベース法の比較

3次元ビデオは、CGによる仮想的・人工的な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実写3次元映像メディアであり、次世代映像メディアとして以下のような応用が期待されている。

・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・芸能ディジタルアーカイブ
・ 自分の身体動作と先生による手本とを３次元的に比較しながらトレーニングできるリハビリ、スポーツ練習システム
・動物のありのままの生態を多角的に観察できる３次元ビデオ図鑑
・ディ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
・ 遠隔地の雰囲気をそのまま伝えることができる高臨場感遠隔会議・講義・実験システム

以下では、マルチカメラのキャリブレーション、多視点ビデオの同期撮影、３次元ビデオの生成、全方位パノラマ画像を背景とした３次元映像編集、インタラクティブ立体表示までの一連の処理を行うシステムの概要について述べ、このシステムを用いた舞妓による日本舞踊の３次元ビデオ生成・編集・表示結果を示す。

なお、平成１６年度には、本プロジェクト開始までに開発した種々のソフトウェアを統合し、研究開発の基盤となるベースライン・システムを開発し、平成１７年度には、ベースライン・システムに改良を加えたレベル１・システムを開発した。以下では、ベースライン・システムの処理過程を順次説明し、各過程において改良された機能・性能についてそのポイントを示すことによってレベル１・システムの性能を示すこととする。

（b）３次元ビデオ撮影システムの機能と性能評価

【１】処理の概要

現在のシステム（ベースライン、レベル１・システムとも）では図１に示す処理過程に従って処理が行われ、3 次元ビデオの撮影・生成・表示・配信が実現される。ここでは、処理過程の概要を述べる。

(1) １０m四方の部屋で、撮影対象を取り囲むように配置した２５台のカメラ（天井：１２台、床：１３台）を備えた PC クラスタ（PC：３０台、ネットワーク：Myrinet(1.2Gbps)、図２）を用いて、対象の多視点同期ビデオを撮影する（図３、図１最上段)。カメラ群は、我々が独自に開発した視点固定型パン・チルトカメラモデル[3]に基づく高精度キャリブレーション法で内部・外部キャリブレーションが行われており、約１／１０００（５m先で５mmの誤差）の精度で校正されている。また、このシステムでは、12 フレーム/秒（カメラのハードウェアの制限による）で２５の異なった視点から同期の取れたビデオ（ＶＧＡ画質）が撮影される。舞妓さんの踊りを撮影した２５視点映像の例を図３に示す。この撮影システムは、平成１６年度に構築した宮崎スタジオのもので、平成１７年度には、高解像度３次元ビデオ撮影用として、別途西陣スタジオを構築した。西陣スタジオは、直径６mの円筒状の壁面、天井に合計１５台の高解像度（ＸＧＡ画質）の固定カメラが設置されている。今後は、宮崎スタジオ：首振りカメラによる自動追尾撮影、西陣スタジオ：高解像度撮影、という機能別の実験のために両スタジオを使っていくことになる。

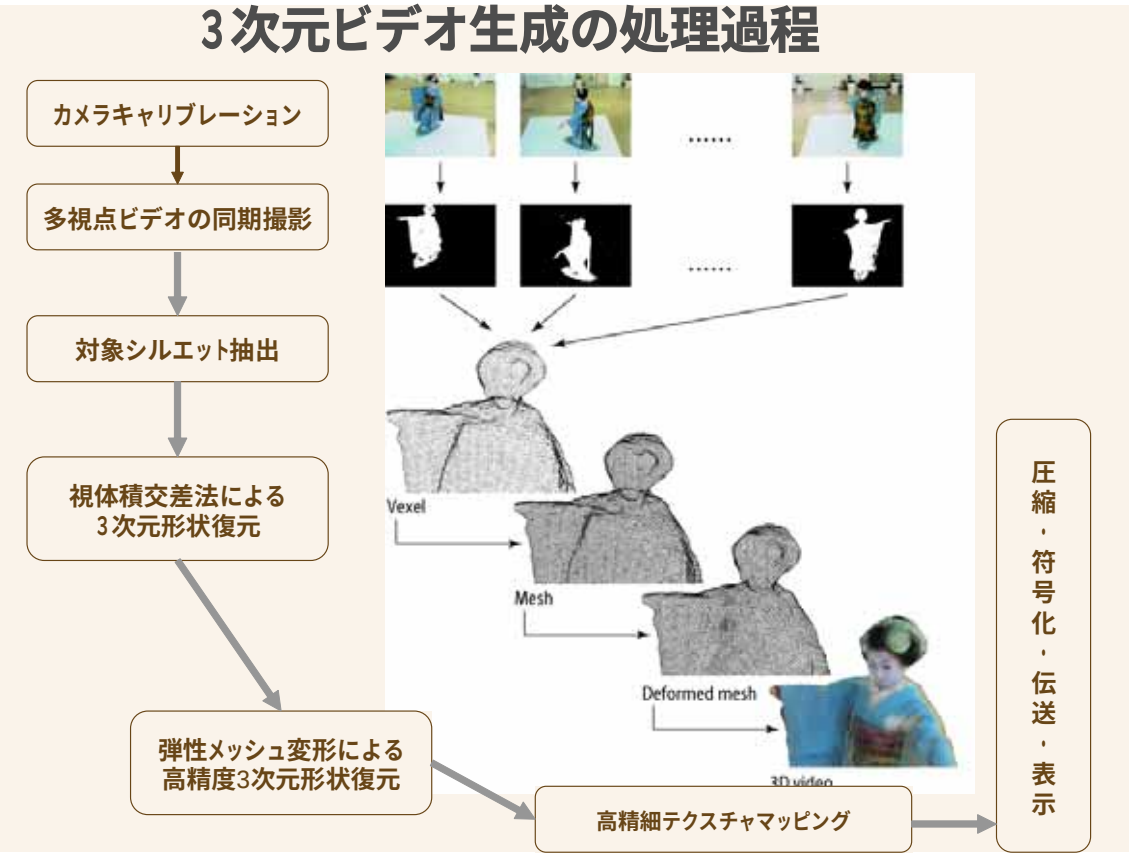


図１　３次元ビデオ生成の処理プロセス

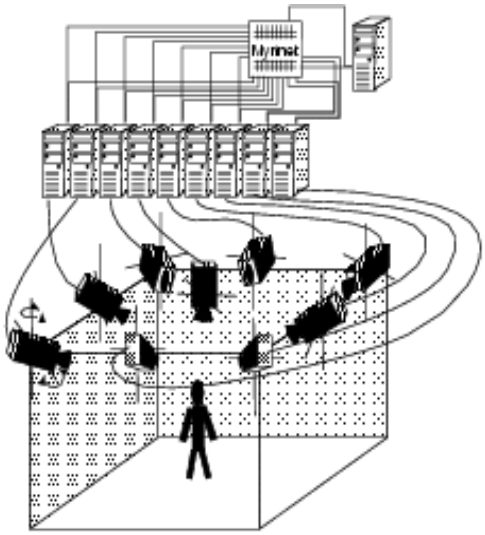


図２　３次元ビデオ撮影用ＰＣクラスタ

図３　撮影された２５視点のビデオ映像

(2) 撮影されたビデオフレーム画像から対象のシルエットを抽出する（図１上から2段目)。
(3) 視体積交差法（図４）によって、対象の visual hull をボクセル形式で求める（図１上から3段目)。
(4) 視体積交差法で得られたボクセルデータ（図１上から3段目）を3角形メッシュモデルに変換する（図１下から3段目および図１０左)。
(5) (4)で得られた3次元形状は、visual hull の計算原理上、カメラの視線方向に対して凹な部分は復元されない（図１下から3段目および図１０左)。そこで3次元メッシュがゴムのような弾性を持っていると考え、各3角形の頂点の位置が真の対象表面に張り付くようにメッシュの変形を行うことで、より高精細な

3 次元形状を得る（図１下から２段目および図１０右）。

(6) 各 3 角形面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、その面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・色を面に張る（図１最下段）。

(7) 以上の処理で、1 フレームの 3 次元ビデオが生成される。動画を作るには(1)～(6)の処理を繰り返えせばよい。また、得られた３次元ビデオは膨大なデータ量となるため、配信のためにはデータ圧縮のアルゴリズムを新たに開発する必要がある。

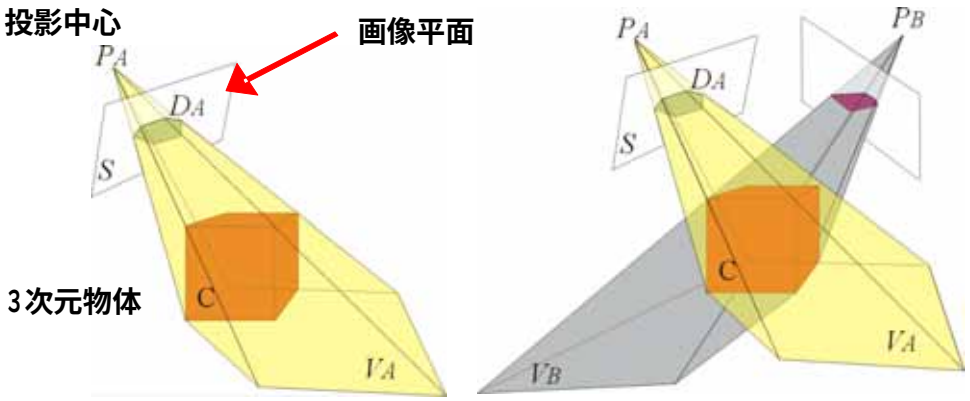


図４　視体積交差法による３次元形状復元の原理（左図：画像中のシルエットをカメラの投影中心を基点として３次元空間に逆投影することによって、３次元対象をその中に包含する視体積が計算できる。右図：異なった視点から得られた複数の視体積の共通部分を求めることによって、対象の粗い３次元形状が計算できる。）

以下では、各処理過程の技術的概要と処理結果を示す。

【２】ＰＣクラスタを用いた実時間並列３次元形状復元

図４に示す通常の視体積交差法では、計算を簡単にするため、計算アルゴリズムとして以下の方法が用いられる。

① 対象が存在すると仮定される３次元空間中の全ての点に対して、各点をそれぞれの撮影画像面上に投影し、対象のシルエット領域の中に含まれているかどうかを計算する。

② そして、すべての撮影画像面上への投影像がシルエット領域内に含まれる点を、対象を構成するボクセルとすることによって、視体積の交差計算を実現する。

こうした計算には多数の点に対する３次元座標変換が必要で、高い空間解像度での実時間 3 次元形状復元を実現することは困難であった。

これに対し、我々の考案した方法では、以下のようにして高速処理を実現している。

(1) まず、3 次元空間を平行な平面群として表し（図 5）、

(2) 各撮影画像から抽出されたシルエットを平行平面群中の基準平面に投影する（図 6 左図の①）。

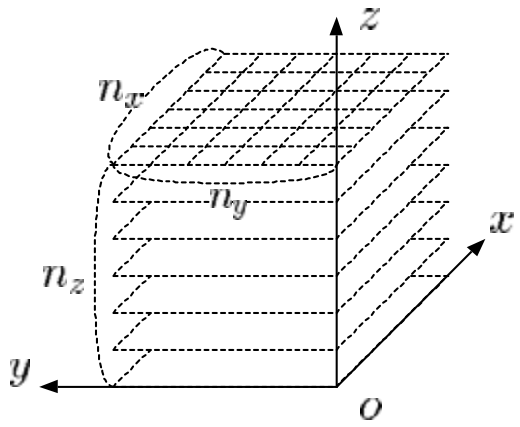


図５　平行平面群による３次元空間の表現

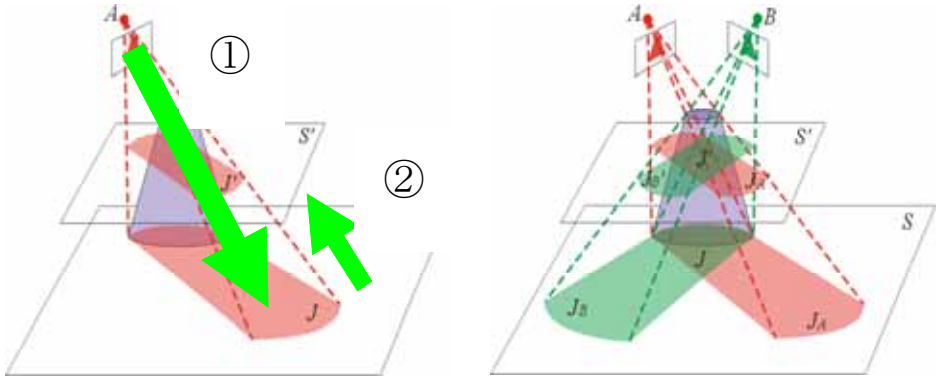


図６　平面間透視投影に基づく視体積交差法

(3) 次に基準平面上のシルエットを他の平行平面群に投影する（図 6 左図の②)。
(4) それぞれの平面上で各撮影画像から得られた 2 次元シルエットの積領域を求め、これを対象の断面形状とする（図 6 右図)。
(5) 得られた断面形状が平行平面間の間隔分の厚みを持つと考え、それらを 3 次元的に積み重ねて対象の 3 次元ボクセル形状を求める。

以上述べた平面間透視投影に基づく視体積交差法は、以下の利点を持つ。

(A) 投影計算が平面から平面へのもののみであるため、シルエットの投影計算が高速に行える。
(B) 特に、基準平面から他の平面へのシルエット投影のように、平面が平行である場合は、画像の 2 次元的なスケーリングと平行移動で投影計算が実現でき、より高速な計算ができる。
(C) さらに、基準平面から他の平面へのシルエット投影および各平面上での断面形状の計算は他とは独立に行えるため、並列処理の効果が高い。

以上述べた方法に基づき、図２に示した PC クラスタを用いて平面間透視投影に基づく並列視体積交差法を実装した。図 7 にその並列処理方式を示す（図 7 の各列は 1 台の PC での処理過程を示している)。

（１）カメラを備えた PC に撮影コマンドを出し、多視点画像の同期撮影を行う。
（２）各 PC が背景差分により対象のシルエットを抽出し、それを基準平面に投影し

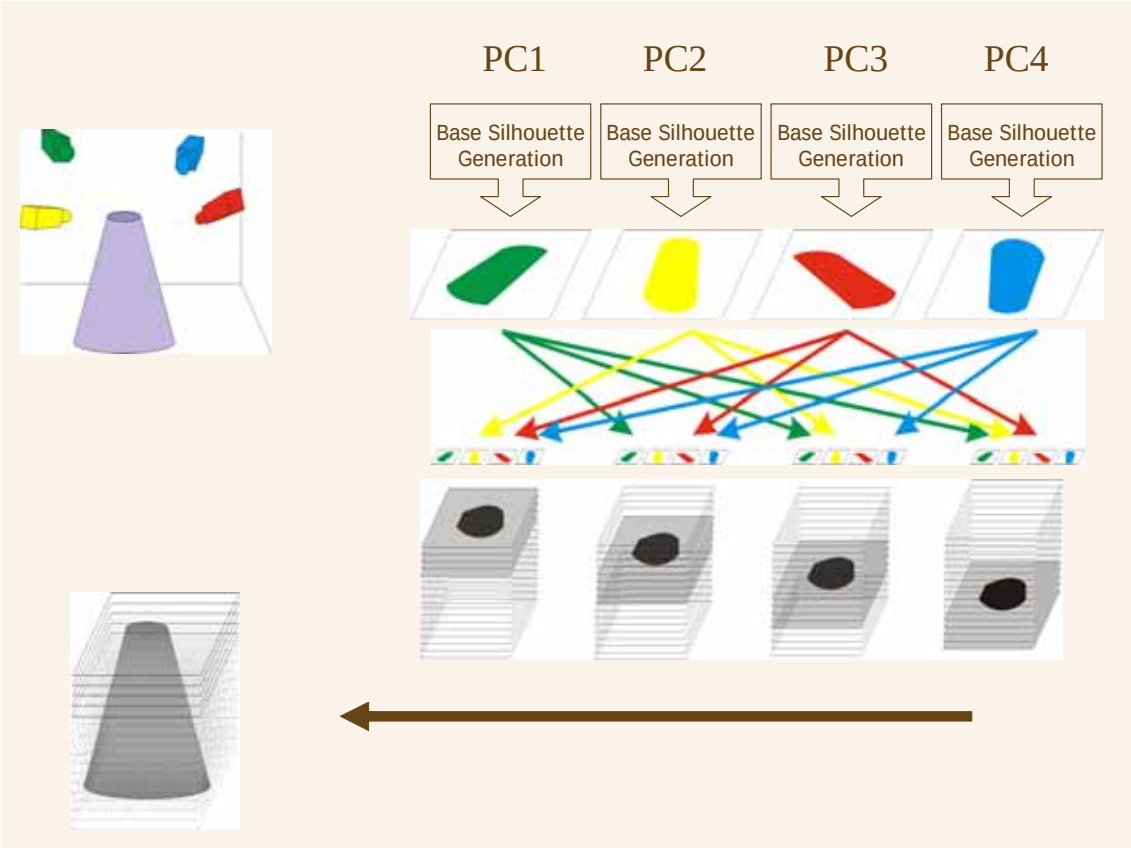


図７　並列視体積交差法の処理アルゴリズム

て基準シルエット求める。

（３）カメラを持たない PC も含め、全ての PC が全ての基準シルエットのコピーを持つように相互通信を行う。

（４）平行平面群をグループに分け、各グループをそれぞれ１台の PC に割り当てる。各 PC は割り当てられたグループ内の平行平面群を対象として基準シルエット群の平行平面への投影とその上でのシルエット交差計算を行い、対象の部分断面形状を求める。

（５）最終的に、各ＰＣで計算された部分的な３次元形状データを統合することによって、対象の全体像のデータが得られる。

平成１７年度上期に開発した最新のシステムでは、さらなる高速化を図るとともに、カメラの配置がどのような場合でも安定かつ効率的に３次元形状復元ができるようにするため、基準となる平面として互いに直交する３つの平面を利用した３平面並列視体積交差法（図８）を開発した。具体的な拡張点としては、以下の点がある。

(1) 各 PC にはそれぞれ２つのＣＰＵがあり、ＰＣで実行される処理を並列パイプライン化する。その際、パイプラインの編成を柔軟に制御できる機能を付加することによって、パイプラインの各処理ステップの負荷に応じたパイプライン処理の最適化を図る。

(2) 基準平面として、ｘ、ｙ、ｚの各座標軸に平行で互いに直交する３つの面を設け、基準平面と撮影面が平行に近いときに基準平面上のシルエットが極端

に大きくなることを防ぎ、計算量を削減する（図８）。

(3) 撮影対象を取り囲む直方体を設け、その中においてのみ視体積交差法を適用する。

このシステムでは、9台のカメラからのビデオ入力の場合、解像度１cmボクセルで毎秒６０フレーム程度、８mmで３０フレーム程度、５mmで１５フレーム程度の処理速度によって、人物の3次元動作をリアルタイム復元できることが実証され、ビデオレートでの実時間３次元形状復元という研究目標が達成された。なお、オンラインビデオ撮影・形状復元では、カメラの撮影速度のハードウェア的性能限界から、毎秒１２フレームの３次元形状復元となる。

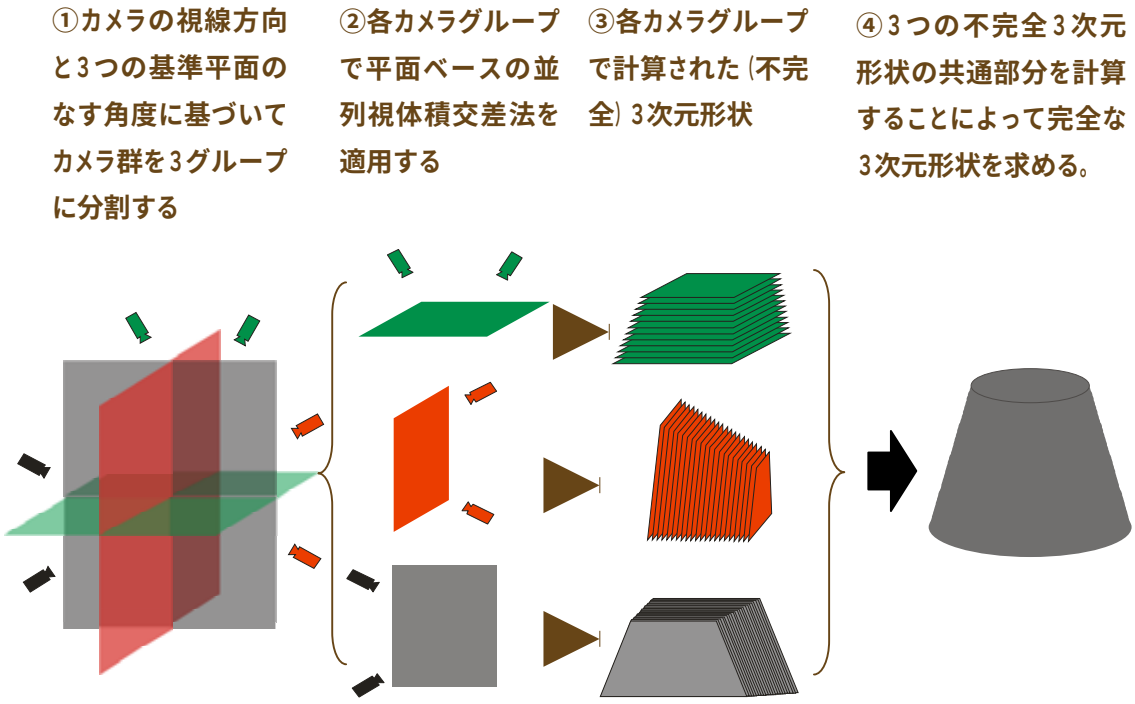


図８　３平面並列視体積交差法の原理

【３】弾性メッシュモデルを用いた高精度３次元形状復元

視体積交差法では、対象の粗い形状しか得られず、カメラ視線に対して凹となる部分の３次元形状が復元できないという本質的問題がある。この問題を解決し、対象の正確な3次元形状を求めるため、本研究では、まず視体積交差法で得られたボクセルデータを離散マーチングキューブ法[2]によって3角形メッシュデータに変換し、このメッシュデータを初期形状としてメッシュの動的変形を行う弾性メッシュモデルを考案した。

メッシュ変形は以下の反復計算によって行う。

Step 1 各頂点に働く力（後述）を計算する。

Step 2 各頂点をそれぞれ力に沿って微少量移動させる。

Step 3 全ての頂点の移動が許容量以下なら終了。そうでなければ Step1 へ。

弾性メッシュモデルでは、ステレオ法や視体積交差法では利用されていなかった対

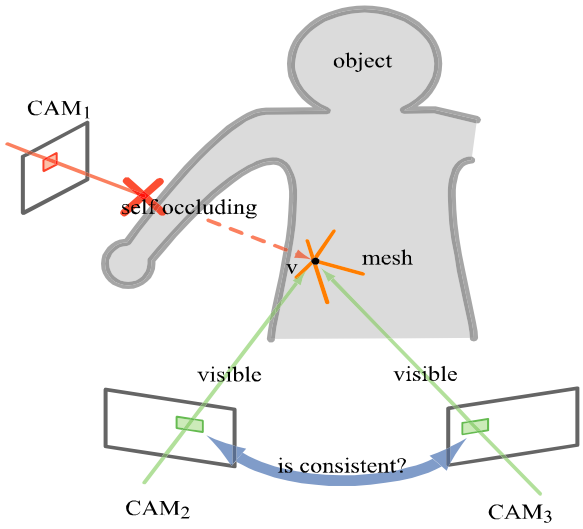


図９　メッシュ変形のためのphotometric consistency制約に基づいて生成される力の計算法

象表面形状の連続性や滑らかさといった幾何学的情報を利用することができ、正確な形状復元が可能である。

我々の弾性メッシュモデルでは、

- photometric consistency制約：各カメラから対象表面の3角メッシュを見た場合、各撮影画像中のテクスチャが一致する。
- シルエット保持制約：3次元メッシュを各カメラの撮像面に投影した場合、そのシルエットが観測された画像中のシルエットと一致する。
- 単純閉曲面制約：メッシュ表面はねじれのない滑らかな曲面となる。

という3つの制約条件を満たしながら変形を行うように設計されている。すなわち、各制約を満たすような位置へと３次元メッシュの頂点を移動させる力を定義し、３つの力が各頂点で平衡するようにメッシュの変形を行う。図９は、photometric consistency制約に基づいて生成される力を計算するアルゴリズムを示したものである。

(1)現在の3次元メッシュ中の各頂点に対して、その3次元座標から、頂点が写されているカメラを求める。図９では、CAM_2とCAM_3が求まる。

(2)(1)で求めた各カメラに対して、頂点が写されている画像中の位置を計算し、画像中における頂点像の付近の画像パターンを求める。

(3)各カメラで撮られた画像パターン間の類似度を計算し、類似度が大きくなるように頂点の3次元座標を移動させる力を生成する。

他の制約に関しても同様に、現在のメッシュ頂点をどちらに移動させれば、制約が満たされるのかを計算し、その移動を実現する力を生成する。最後に、各制約に基づいて生成された力の重み付き平均によって最終的に頂点に働く力とする。

図１０右は、上記の方法で求めた舞妓の高精度3次元形状で、同図左の視体積交差法だけの場合と比べ、精度が向上していることが分かる。また、図１１は、弾性メッシュ変形前後の３次元形状を用いて生成された３次元ビデオを表しており、メッシュ

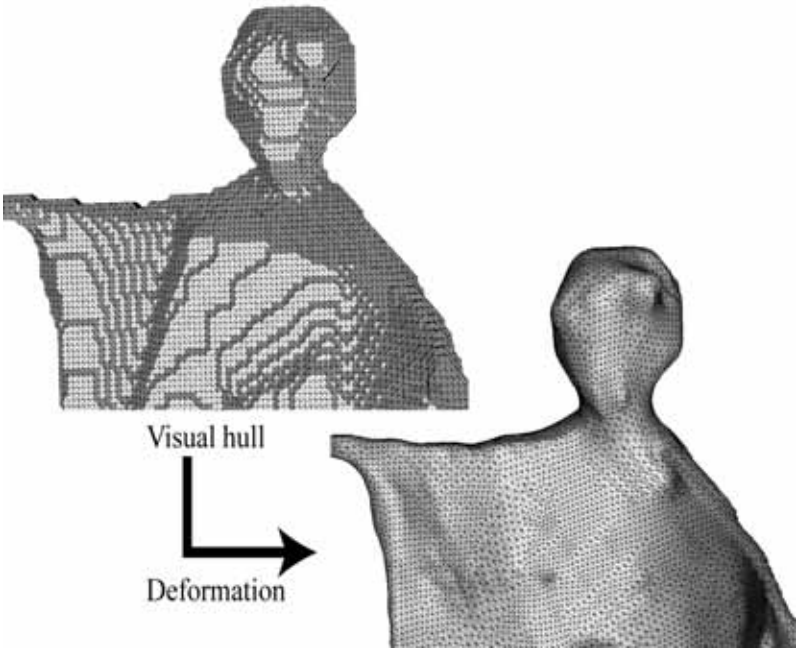


図１０　弾性メッシュによる高精度３次元形状復元

図１１　弾性メッシュ変形の映像生成上の効果

変形によって３次元形状の精度が向上することによって、映像としても見栄えが大幅に向上していることが分かる。

平成１７年度後期から開発を進めているより高度な弾性メッシュモデルでは、対象の動作に合わせてメッシュを時間的に変形させることを目指して、以下の拡張を行っている。

(A)メッシュの各部分の硬さ・柔らかさを過去の運動状態から求め、それに基づいて弾性変形か剛体運動かを判別してメッシュ変形を制御する。
(B)入力される映像は、対象の動きを撮影したものであり、入力映像から視体積交差法によって計算されたボクセル時系列に応じてメッシュを時間的に変形させ、対象の動作に従ってメッシュが変形するようにする。
(C)人間が手を胴体に添えたり、離したりするような複雑な動作（3次元形状の大局的なトポロジーの変化）にも対応できるようなメッシュの動的変形を実現する。

こうした機能拡張によって、

- すべての時刻における対象形状を、同一のメッシュトポロジー（3角形パッチの接続関係）をもったメッシュで表現することが可能となり、対象の詳細な３次元運動解析や後に述べる３次元ビデオのデータ圧縮が容易に実現できるようになる。
- 舞妓による舞踊のように長い振袖や帯がしなやかに揺れるような動作でもかなりの精度で3次元形状が復元できるようになる。
- 複数人物による舞踊などを同時に３次元ビデオとして撮影する場合、人物同士がふれ合うことによって、視体積交差法では一人ひとりの３次元形状を区別することができなくなるが、時間的な弾性メッシュ変形を用いれば個々の人物のメッシュが絶えず区別され、一人ひとりの動作が復元できるようになる。

平成１８年２月現在、これらの機能を備えたソフトウェアがほぼ開発でき、現在最終調整、性能評価実験を行っており、中間評価においてその結果を報告する予定である。

【４】高精度３次元ビデオレンダリング

以上の処理で対象の動きを表す3次元メッシュの時系列データが得られ、メッシュ中の各3角形パッチに、撮影されたビデオから抽出したテクスチャを貼り付けることによって、3次元ビデオが完成する。

最も単純なテクスチャマッピング法としては以下のものが考えられる。

■視点独立面ベース法

(1) 各3角形パッチに対して、その面の法線と最も近い視線を持つカメラを選び、
(2) そのカメラによって撮影された画像上に3角形パッチを投影し、3角形パッチ上のテクスチャを求め、それを対象表面パッチに張付ける。

しかし、計算された3次元メッシュは対象の形状を完全に正確に表しているものではないため、この方法では、以下の問題が生じる。

- カメラ間の撮影映像の差異によるちらつき：異なったカメラで撮られた画像が対象表面の隣接した3角形パッチに張られ、その境界線上でテクスチャが不整合となり、ちらつきが生じる。
- 隣接パッチ間でのテクスチャの不整合によるちらつき：同じカメラで撮影された映像を基にしてテクスチャマッピングを行った場合でも、テクスチャはパッチ毎に独立して生成されるため、各パッチの継ぎ目において不整合が生じることがある。

これらの問題を解決し、滑らかなテクスチャマッピングを実現するために、映像を表示する際の視点位置を利用し、対象表面パッチの頂点色の補間によってパッチ上のテクスチャを生成する方法を開発した。

■視点依存頂点ベース法

(1) 対象を見る、すなわち３次元ビデオの映像化のための視点位置・視線方向を決める。

(2) 定められた視線方向と撮影に用いた各カメラの視線方向との一致度を計算する。

(3) メッシュの各頂点について、頂点を各カメラで撮影された画像上に投影し、各画像中での RGB 値を求め、(2) で求めた一致度に基づいた重みづけを用いて荷重平均を計算し、その値を頂点の RGB 値とする。

(4) パッチ上のテクスチャをその 3 頂点の RBG 値の線形補間によって生成する。

図１２は、視点独立面ベース法、視点依存頂点ベース法によるテクスチャマッピングの結果を表したもので、後者の優位性がよく分かる。図１３は、それぞれのテクスチャマッピング法で生成された画像と撮影画像とのRGB値の平均誤差（縦軸）を各フレーム毎（横軸）に計算したもので、このグラフからも提案手法の優位性が実証された。

視点独立面ベース 法　　視点依存頂点ベース 法

図１２　テクスチャマッピング法の定性的性能比較

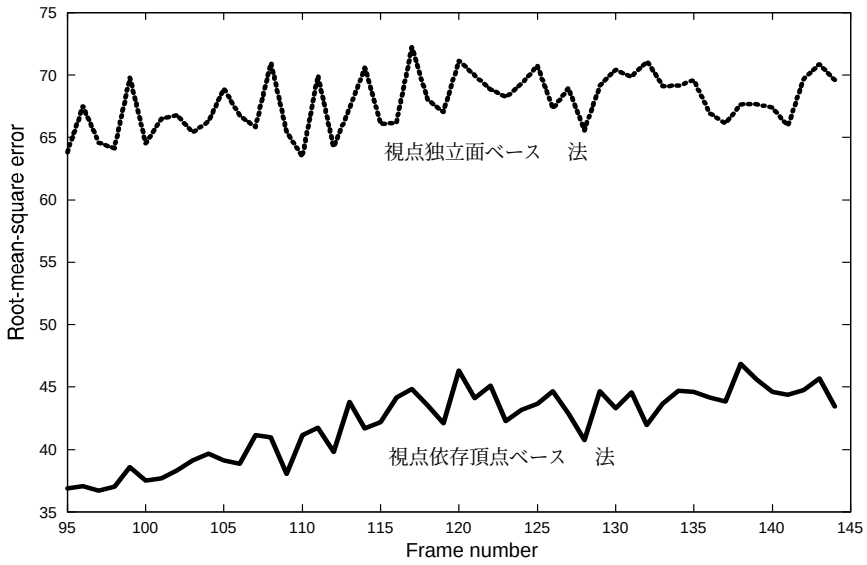


図１３　テクスチャマッピング法の定量的性能比較

上記のように、視点依存頂点ベース法は、滑らかな映像を生成することができるが、その反面各パッチ内では内挿計算によって細かなテクスチャが失われ、着物の織り目や毛皮の風合いといった質感がなくなってしまう。

この問題点を解決するため、現在新たな高精度テクスチャマッピング法として視点依存面ベース法を開発しており、中間評価においてはその結果を示すことができると考えている。

【５】３次元ビデオの編集・表示

従来のビデオと異なり、3次元ビデオは3次元の形や運動情報を持っているため、その編集は4次元空間（3次元座標軸+1次元時間軸）において行われる。すなわち、図１４のような対象、背景、仮想カメラからなる3次元シーンを設け、その中で対象の移動・拡大縮小・コピー、背景の拡大縮小・移動、カメラの移動・ズームなどを行い、仮想カメラで撮られた映像を編集結果として出力・表示する。仮想カメラとしてステレオカメラを使い3次元ディスプレイに表示すれば立体映像として見ることもできる。

本研究では、背景として全方位パノラマ画像（静止画）・ビデオ（動画）を用いて３次元ビデオをインタラクティブに表示する３次元ビデオビューア・ソフトを開発した。全方位パノラマ画像は非常に高解像度であるため、映像化においてズーム倍率を大きくしても十分鑑賞に耐えることができる。

全方位パノラマ画像と、撮影した人物をコピーし３次元的に並べて作った３次元ビデオ・コンテンツを映像化した結果を図１５に示す。また、図３の舞妓による踊りを撮った３次元ビデオと、水中の熱帯魚の生態を写した全方位パノラマビデオとを編集した映像例を図１６に示す。

これらの３次元ビデオ・コンテンツは、立体ディスプレイを備えたノートPCによって、インタラクティブに視点やズームを変えながら鑑賞できるようになっている。

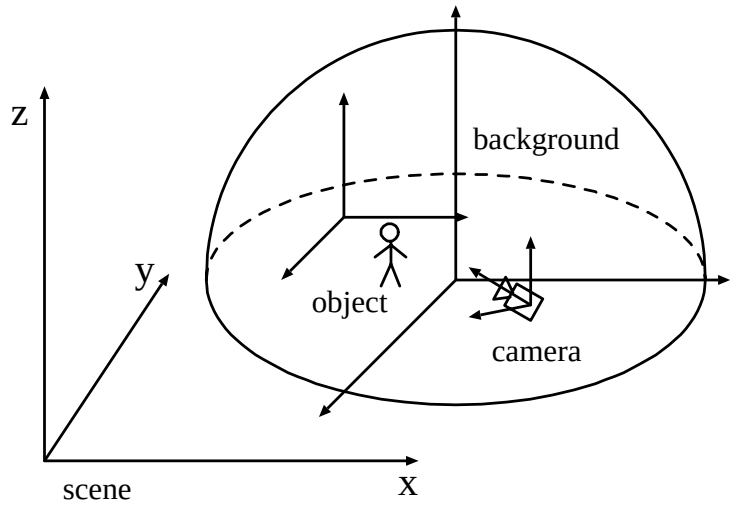


図１４　３次元ビデオ編集のための３次元シーンの構成

frame #68

frame #103

frame #157

frame #138

図１５　３次元ビデオのコピーと全方位パノラマ画像を編集した映像例

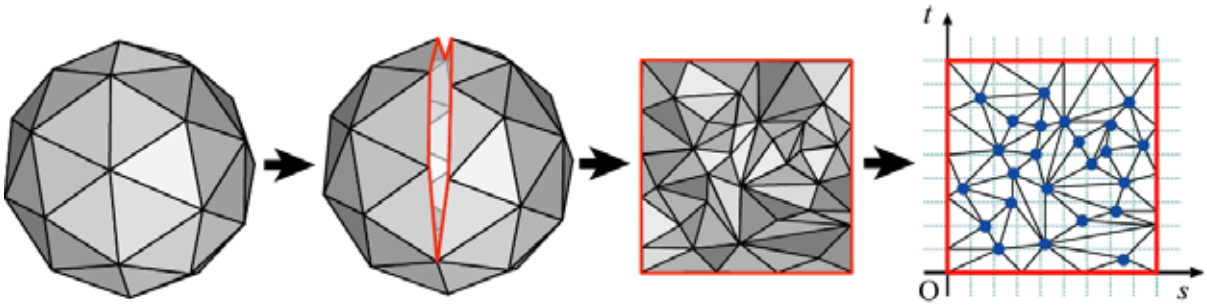
図１６　３次元ビデオと全方位パノラマビデオを編集した映像

【６】３次元ビデオの圧縮アルゴリズム

３次元ビデオは、従来の２次元ビデオと比べデータ量が膨大となり、そのままのデータ形式では、記録・伝送に大きな障害が生じ、データ圧縮法を考案することが実用化にとって不可欠となる。

本研究では、平成１７年度に３次元ビデオ圧縮の方法として skin-off と名付けたアルゴリズムを考案し、特許出願を行った。以下ではその基本的考え方を述べる。

３次元ビデオは、時間経過とともに動的に変化する３次元メッシュと、その表面の



(a) ３次元メッシュ　(b) カットパス　(c) 平面展開　(d) 格子を使ったサンプリング

図１７　skin-off の基本的考え方

テクスチャから構成されている。図１７は skin-off のアイデアを示したもので、以下のような処理が行われる。

Step1.　３次元メッシュの切り開き

任意の連結した１つの３次元メッシュ（図１７(a)）は、適当なエッジで切り開けば平面へ変換（展開）可能であるため、初期値として、切り開きに必要なエッジ集合（カットパスと呼ぶ。図１７(b)の赤線部分）を選択する。

Step2.　３次元メッシュの矩形平面上への展開

カットパスが２次元平面上の矩形の枠に射影され、他のメッシュ頂点が矩形内に射影されるように、３次元メッシュを切り開き、矩形に展開する（図１７(c)）。

Step3.　平面上の頂点位置の最適化

矩形平面内でのメッシュ頂点の配置を最適化する。このとき、3 次元メッシュ上での各パッチができるだけ歪むことなく平面上に配置されるよう、各頂点の矩形中での位置を最適化する。Skin-off では、この最適化において、３次元パッチの幾何学的歪とテクスチャ歪の２つの観点から最適化を行う。

Step4.　カットパスの更新

Step3 で頂点配置を最適化したあと、Step1 で選ばれたカットパスを更新する。カットパスの選び方によって、最適化された２次元平面上での歪みが変化するため、歪みを最小化するようにカットパスの更新、頂点位置の最適化を反復する。

Step5.　矩形領域内でのサンプリング

カットパスおよび平面上での頂点位置の最適化が終了した後、矩形平面上に適当な解像度の格子点を設定し、各格子点上で３次元ビデオデータをサンプリングする（図１７(d)）。これによって３次元ビデオデータを表現する２次元配列が得られる。具体的には、各格子点に記録する(x, y, z)および(R, G, B)の値を、その周囲に配置されている頂点の３次元座標および色情報から内挿によって求める。格子点同士の接続関係は、２次元配列の上下左右、対角線上の格子点が接続されていると考えれば、データ構造として明示的に表現する必要がない。

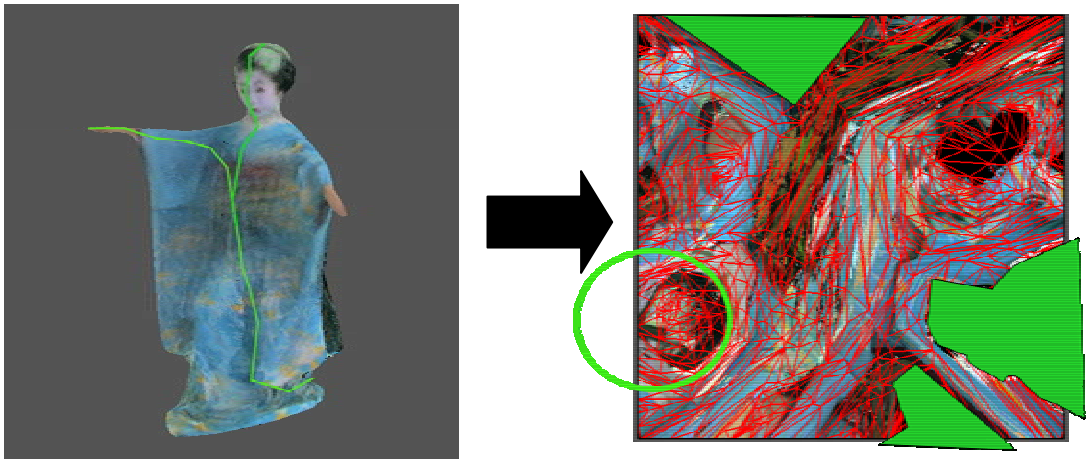


図１８　skin-off アルゴリズムで得られたカットパス（左図）と平面展開画像（○が顔の部分、緑の領域は撮影されていない部分）

Step6. 映像圧縮手法の適用

得られた２次元配列（デジタル画像）に対して、既存のビデオ圧縮手法を適用する。

図１８は、舞妓の３次元ビデオの１フレームを上記のアルゴリズムで平面展開して得られた２次元画像を表している。

Skin-off アルゴリズムは、データ圧縮に通常の２次元映像用のコーデックが利用でき、既存技術との親和性が高いという特徴を持っており、実用的な３次元ビデオ圧縮法としては有望であると考えられる。

現在は、以下の課題に対して研究を進めており、中間評価の際にはある程度の評価結果が得られるものと考えている。

- これまで開発したアルゴリズムはあくまでもプロトタイプであるため、実用化を目指した改良と定量的な圧縮効率の評価が必要である。
- 現在のアルゴリズムは１フレームの圧縮法であり、３次元ビデオを圧縮するには、時間軸方向での変化の符号化が必要であり、そのためのアルゴリズムを開発する必要がある。

（ｃ）まとめ

京都大学大学院情報学研究科松山研究室では、無形文化財の高精度デジタル化ソフトウェアの開発をめざして、多視点映像から運動対象の3次元形状とその表面情報を時系列データとして生成し記録する3次元ビデオの実用化を目指した研究開発を行っている。

本稿では、平成１６年度、１７年度の研究で開発した手法を統合化したベースライン・システムおよび改良版のレベル１・システムの概要を述べ、開発したソフトウェアの性能を実験によって示した。主なソフトウェアとしては、以下のものがある。

(1)　PCクラスタを用いた多視点ビデオ映像からの実時間3次元形状復元

(2)　弾性メッシュモデルを用いた高精度3次元形状・運動復元法

(3)　自然な3次元ビデオ生成のための高精度テクスチャマッピング法
(4)　3次元世界における多様なカメラワークが可能となる3次元ビデオ編集、インタラクティブな立体映像表示

平成１７年３月には、これらの課題に関する研究が３名の博士学位論文[7][8][9]としてまとめられ、人材育成の面でも大きな成果が得られた。

また、平成１７年度には、３次元ビデオ撮影用スタジオとして、従来の宮崎スタジオに加え、西陣スタジオを新たに構築した。

表２は、ベースライン・システムとレベル１・システムの性能比較および宮崎、西陣スタジオの比較を行ったものである。

さらに、上記(1)、(4)のソフトウェアに関しては、複数の企業へソフトウェアのライセンシング契約が成立し、産業界への技術移転も順調に進んでいる。

今後の展開としては、

・安定なシルエット抽出法の開発
・3 次元形状復元のさらなる高速化・高精度化
・鑑賞に耐えうる画質を持った 3 次元ビデオ生成法

の開発といった現在の手法の改良に加え、

(1) 首振りカメラを用いて、複数の人物を追跡、ズームアップ撮影して[4]3次元ビデオ化する技術の開発
(2) 複数人物によって演じられる舞踊などの同時３次元ビデオ撮影
(3) 撮影時の照明環境の推定[5]と、3次元ビデオに対する照明効果の導入
(4) 従来の2次元ビデオと比べ桁違いに大量となるデータの圧縮・符号化法[6]の改良とそのソフトウェア開発
(5) 3次元ビデオに映された対象の動作解析および、その結果に基づいた3次元ビデオの編集法の開発

があり、中間評価に向けてこれらの課題に関するソフトウェアの開発をすすめている。

平成 16 年度 ベースラインシステム	平成 17 年度 レベル 1 システム
並列視体積交差法の高速化・高精細化	
12fps / 10mm	60fps / 10mm, **30fps / 8mm,** 17fps / 5mm
弾性メッシュ変形による高精度３次元形状復元	
凹部の形状が不完全	**凹領域の復元精度が向上**
レンダリング	
頂点ベース、静止画背景	**面ベース、動画背景**

(宮崎スタジオ)　撮影スタジオ　(西陣スタジオ)	
VGA カメラ、12fps 空間解像度 10mm	**XGA カメラ、30fps** **空間解像度 2.5mm**
今後の取り組み	
自動追尾撮影による撮影可能領域の広範囲化	**複数人物**の同時撮影

表２　平成１６年度と平成１７年度の研究成果の比較

参考文献

[1] T. Matsuyama, X. Wu, T. Takai, and S. Nobuhara: "Real-Time 3D Shape Reconstruction, Dynamic 3D Mesh Deformation, and High Fidelity Visualization for 3D Video", International Journal on Computer Vision and Image Understanding, Vol.96, No.3 pp.393-434 ,Dec.2004

[2] 剣持雪子, 小谷一孔, 井宮淳:“点の連結性を考慮したマーチング・キューブ法”, 電子情報通信学会技術報告, PRMU98-218, pp. 197−204, 1999

[3] 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D-II, Vol. J81-D-II, No. 6, pp. 1182-1193, 1998

[4] 松山 隆司, 浮田 宗伯:″能動視覚エージェント群による協調追跡″, 日本ロボット学会誌, Vol. 19, No. 4, pp. 25-31, 2001

[5] T. Takai, K. Niinuma, A. Maki, and T. Matsuyama : "Difference Sphere: An Approach to Near Light Source Estimation", Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp.I98-I105 ,2004

[6] H. Habe, Y. Katsura, and T. Matsuyama: "Skin-off: Representation and Compression Scheme for 3D Video", Proc. Picture Coding Symposium (PCS) 2004, San Francisco , Dec.2004

[7] 仵小軍: “Parallel Pipeline Volume Intersection for Real-Time 3D Shape Reconstruction on a PC Cluster”, 京都大学大学院情報学研究科, 博士学位論文, 2005

[8] 高井勇志, “High Fidelity and Versatile Visualization of 3D Video”, 京都大学大学院情報学研究科, 博士学位論文, 2005

[9] 延原章平, “Deformable Mesh Model for 3D Shape and Motion Estimation from Multi-Viewpoint Video”, 京都大学大学院情報学研究科, 博士学位論文, 2005