

ズームアップ撮影による高精細3次元ビデオ生成のための パン・チルトカメラワーク

稲積 慶人 浮田 宗伯 木戸出 正繼^{†1}

無形文化財のデジタルアーカイブを行うために、本研究では複数台のカメラで対象を撮影するだけで対象の形状とテクスチャの3次元情報が得られる3次元ビデオという情報メディアを用いる。従来の3次元ビデオでは、常に動く撮影対象の全身を全てのカメラが完全に捉えられるまでズームアウトし、カメラを静止させていた。しかし、この従来手法では全てのカメラが対象の全身を完全に捉える必要性により各カメラに対してズーム制限が生じる。その結果、撮影画像の高精細化に上限が発生してしまうという欠点を有する。そこで、本研究では従来よりもズームアップ撮影をすることで対象の一部分のみを捉えたカメラを能動的にパン・チルト回転させることで対象全体の形状の正確さの低下を抑制しつつ対象全身の高精細画像を取得する手法に着目する。それにより、従来よりも高品質な3次元ビデオを作成することを目的とする。本稿では(1)実環境下で動的な対象をズームアップ撮影するためのシステム構築、(2)そのシステムを用いてどのカメラが対象のどの部分を撮影するかを決定する手順であるカメラワーク決定手法について述べる。

Zoom-In Pan-Tilt Camera-Work for High-Fidelity 3D Video

YOSHITO INAZUMI, NORIMICHI UKITA and MASATSUGU KIDODE^{†1}

In this paper, we propose a real-time human-body shooting system with zoom-in pan-tilt cameras. The task of the system is to observe the whole body of a moving person as detail as possible for making high-fidelity 3D video with high-resolution shape and texture. With a zoom-in camera, the high-resolution image of only a part of the body can be captured. The system must therefore control each camera towards a body-part that can be captured well by the camera so that the whole body is captured by all the cameras. We developed the following camera architecture and camera assignment algorithm: (1) a pan-tilt camera architecture in which its pan-tilt angle and capturing timing are fully synchronised and (2) the target body-part of each camera is determined so that 3D reconstruction errors in the whole body are suppressed as much as possible. Experimental results with our prototype system demonstrated that the system could capture the whole body of a moving/dancing person and the resolution of captured videos was much better than that captured by fixed zoom-out cameras.

1. はじめに

近年、能や伝統舞踊に代表される無形文化財をデジタルアーカイブする試みが盛んになされている。そこで、本研究では3次元ビデオ映像という3次元情報を持った情報メディア¹⁾を用いて、無形文化財の実世界における姿・形・色の時間的変化を3次的にそのまま記録することで無形文化財のデジタルアーカイブを行う。

3次元ビデオ映像は複数台のカメラで対象を撮影し、表面のテクスチャ情報と形状に関する情報を取得して組み合わせることで作成される。このうち、テクスチャ

情報は撮影画像から手に入れることが可能であるが、複数の2次元画像から3次元形状情報を獲得する手法については、数多くの研究が行われている。

それらの中の手法の一つとして、Shape-From-Silhouette(以下SFSとする)と呼ばれるものがある。これは、以下の様な手順で撮影対象の3次元形状情報を得る。

- (1) 複数台のカメラで多視点同期撮影をし、撮影対象の全身を隈なく撮影する
- (2) 得られた同期画像群から撮影対象のシルエットを抽出する
- (3) 取得したシルエットからSFSによりvisual-coneを構成する(図1)
- (4) 各visual-coneの共通領域をVisual Hull(以下VHとする)とする(図1)

^{†1} 奈良先端科学技術大学院大学情報科学研究科
Graduate School of Information Science, Nara Institute
of Science and Technology

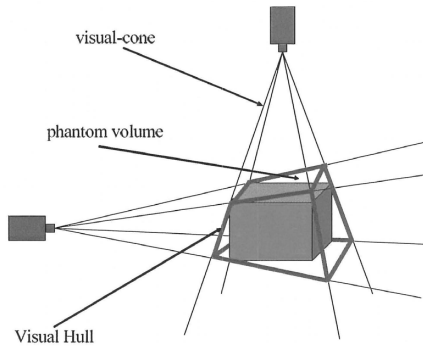


図 1 Shape-From-Silhouette の原理
Fig. 1 Principle of Shape-From-Silhouette

また、この手法の利点として、特定方向からの画像だけでなく対象の全周囲画像が得られる、カメラ間の距離が離れている環境下でも使用可能、カメラ台数を増やすほど正確な形状が取得できるなどがある。

一方で、原則としてそれぞれのカメラが対象全体を画像内に捉える必要がある(図 2) ためにズーム制限が発生し、図 1 に示すファントムボリュームが発生するなどという欠点を有する。特にズーム制限が生じることによって、3次元形状の空間解像度が低く留まるという問題点が存在する。この問題点を解決するために、本研究では SFS の枠組みを前提とした、実画像から VH を作成する際の高空間解像度化手法を提案する。高空間解像度化によって、指先などの細かな形状や、より高解像度なテクスチャが獲得可能になる。特にテクスチャが 3次元ビデオの見栄えに及ぼす影響は大きいために、3次元ビデオ映像の品質を大きく改善できることが期待される。

過去の研究³⁾⁴⁾においてもズーム制限を取り払い VH 作成における高空間解像度化手法は提案されてきたが、実環境下での動的対象に対する撮影手法については未だ研究がなされていないので、本研究にてその撮影手法を提案する。

2. 本研究で用いる 3次元ビデオ映像作成の流れ

本章では(1)撮影環境とカメラワークの定義、(2)人体の3次元形状解析手法について先に述べた後、実環境下で動的な対象をズームアップしながら撮影し、(1)(2)を利用する事で3次元ビデオ映像として記録する手法の流れを示す。

2.1 3次元ビデオ撮影環境とカメラワーク

まず、3次元ビデオ撮影環境について述べる。撮影に用いられるカメラ群は図 3 に示すように、撮影対



図 2 ズーム制限を有するカメラで撮影した画像例
Fig. 2 Image taken by zoom-out camera

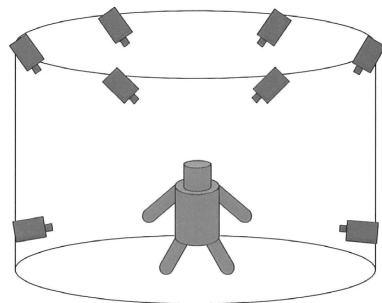


図 3 3次元ビデオ撮影環境
Fig. 3 3D video studio

象を取り囲むように配置されている。そこで、このカメラ群のカメラワークをカメラの視線方向、カメラ位置、ズーム値と定義する。

2.1.1 ズーム制限を有するときのカメラワーク

ズーム制限を有するカメラワークは以下のようになる。

- ・ 視線方向は固定
- ・ カメラ位置は固定
- ・ ズーム値は固定

ズーム制限を有する場合においては、カメラワークを変更しなくても常に対象を捉えることが可能であるという仮定のおとで動的対象を撮影する。

2.1.2 ズームアップ撮影時における問題点

部分撮影問題 現状では、図 3 に示す 3次元ビデオ撮影環境下において SFS により VH を作成する際、それぞれのカメラが対象の全身を隈なく撮影する必要がある。本研究ではその制限を取り払い対象にズームアップしながら撮影する。すると、ズーム制限がある場合と異なり、それぞれの画像には対象の一部分しか撮影されない(図 4) ことが分かる。

そのため各カメラを能動的に動かし、どのカメラが



図4 ズーム制限を取り払ったカメラで撮影した画像例
Fig. 4 Images taken by zoom-up camera

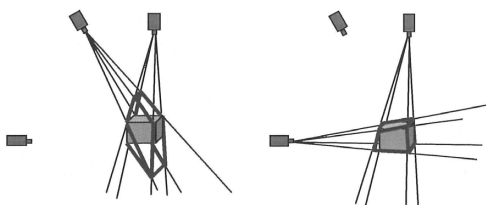


図5 カメラワークがファントムボリュームに及ぼす影響
Fig. 5 Influence of camerawork on phantom volume

どの部分を撮影するかを定めるために、適切なカメラワーク決定を行わなければならない。また、対象は常に動くのでカメラワーク決定をオンラインにて行う必要がある。

ファントムボリューム

図5に示すように、同じ対象を撮影しているにも関わらず、カメラワークによってファントムボリュームの量が変わることが分かる。したがって、上記の部分撮影問題だけでなく、ファントムボリュームを考慮したカメラワークを決定しなければならない。

2.1.3 ズームアップ撮影時におけるカメラワークの定義

2.1.2で述べたとおりズームアップ撮影時にはオンラインでカメラワークの変更を行う必要がある。そのために、それぞれのカメラを能動的に動かさなければならないので、カメラワークを以下のように変更する。

- ・ 視点を中心としたパン・チルト回転により視線方向を決定可能な機構をカメラに持たせる²⁾
- ・ カメラ位置は固定
- ・ ズーム値は固定

カメラ位置とズーム値に関しては問題の単純化のために固定にし、カメラの視線方向のみをパン・チルト動作により変更することでカメラワーク決定を行う。したがって、本稿ではカメラ位置とズーム値を固定とするのでカメラワークの定義をパン・チルト回転角の2

変数に置き換えることにする。

2.2 人体の3次元形状解析の必要性

本稿ではカメラワーク候補数が膨大なるのを防ぎ効率良くカメラワーク決定を行うために、人体の3次元形状解析を行うことで、撮影対象の頭や手などの各パーツがどこにあるかを把握する必要がある。その理由として、次の例を挙げる。有効なパン・チルト角の絶対角度が $-\theta \sim \theta$ の範囲内であり、また $\Delta\theta$ 刻みで指定可能であるとする、パン・チルト角度の組み合わせは一つのカメラに対して $N_{cw} = 4.0 \times (\theta/\Delta\theta)^2$ 通り存在することになる。さらにそれをカメラを n 台で使用すると、全てのカメラワーク候補数は N_{cw}^n 通りとなる。 $\theta = 15\text{degree}$, $\Delta\theta = 0.1\text{degree}$, $n = 8$ 台とすると、約 10^{35} 通りの膨大なカメラワーク候補が存在することとなり、オンラインでのカメラワーク決定を行う事が非常に困難になる。

そこで、人体パーツである頭・右手・左手・胴・下半身の位置を解析し、各パーツのボクセル重心のみカメラに注目させることで、一つのカメラに対するパン・チルト角度の組合せを5通りにまで減らすことが可能になる。すると、全カメラワーク候補数は約 10^5 通りになる。

以上の例のように、人体の3次元形状解析を行い、各パーツ位置を知ることができると、(1)カメラワーク候補を絞ることが可能であるということだけでなく、(2)カメラワークの定義をどのパーツにどのカメラを割当てるかに変更することができる。

2.3 ズーム制限を取り払ったときの3次元ビデオ映像作成全体の流れ

以上の議論を踏まえた3次元ビデオ撮影の全体の流れは図6のようになる

まずは、オンラインにて各カメラが常に画像を取り続けるのと並行して、ある固定時間間隔(本稿では0.5sec間隔で行う)で保存した同期画像群のシルエット抽出、SFSによるVHの作成、人体の3次元形状解析、カメラワーク決定を行う。ただし、このときのSFSによるVHの作成は人体の3次元形状解析とカメラワーク決定を行うための必要最小限の空間解像度で行う。

そして、オフラインにて出来る限り高空間解像度でSFSを行い、作成されたVHにオンラインで保存した高解像度画像を用いてテクスチャマッピングを実行し、3次元ビデオを完成させる。

2.4 本研究で用いる人体の3次元形状解析の従来手法

カメラワーク決定で用いる人体の3次元形状解析手法に関しては、様々な研究が行われている。本研究では次の様な条件を満たす3次元形状解析手法が必要となる。

- ・ 着物などの緩い着衣に対応可能できること
- ・ 形状解析に要する時間が目標時間間隔 500msec

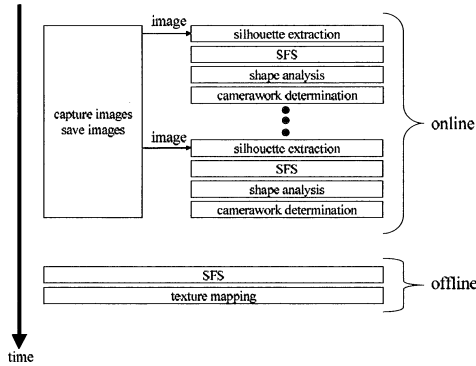


図 6 ズーム制限を取り払ったときの 3 次元ビデオ映像作成全体の流れ

Fig. 6 Flowchart of 3D video creation

よりも短いこと

過去の研究³⁾では剛体モデルを入力 VH とマッチングさせることで形状解析を行っている。形状モデルはパーツ自体の形状が全く変化しない剛体とし、各パーツ間の関節の回転によって動きを表現している。パーツ間の接続関係を木構造で表現する。そしてこの木構造の根にあたるパーツから順にマッチングを行う。マッチングはそのパーツの関節回転角度を調節することで行う。また、マッチングに要した計算時間は約 15sec である。したがって、この形状モデルは非剛体ではなく、さらにマッチング時間が長いなどといった欠点を有するために、本研究においてこの 3 次元形状モデルは使用することは困難である。

そこで、本研究では上記条件を満たす文献⁵⁾に記載されている 3 次元形状解析手法を部分的に用いる事にする。この手法では非剛体形状モデルを主成分分析により低次元空間に記録し、低次元空間中で入力 VH とのマッチングを行う。以下に (1) 3 次元形状モデルの作成方法、(2) 主成分分析を用いて 3 次元形状モデルを低次元空間で記録する方法、(3) 低次元空間中でのマッチング手法の詳細について述べる。

2.4.1 3 次元形状モデルの作成手法

まずは、各パーツごとに着色された着物を着た撮影対象の画像を全てのカメラに対して得る。そして、これらの画像から SFS によって VH を生成する。次にこの VH に対して各パーツを割当てて考える。そのために、まずは、色検出により各色に対応したラベルが付けられたパーツ領域が存在する画像を作る。このラベル付き画像と生成した VH とを用いてパーツ分割された 3 次元形状モデルを作成する。まず、各カメラの視点からラベル付き画像のパーツ領域の画素を通る半直線を描き、その半直線と一番初めに交わった VH のボクセルがそのパーツとなる。これで VH の表面ボクセルのみがパーツを割当てられたことになる。

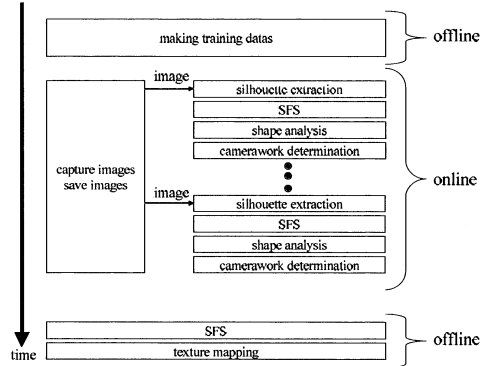


図 7 本研究で用いる 3 次元ビデオ映像作成全体の流れ
Fig. 7 Proposed flowchart of 3D video creation

本研究ではこのような手法で生成された表面ボクセルのみを用いて各パーツ重心を求めることとする。このことを全フレームに対して行い、生成した VH とパーツ分割された 3 次元形状モデルを一つの踊りに対する時系列学習データとして記録し、二つのデータを対応させておく。

2.4.2 3 次元形状モデルを主成分分析により低次元空間で記録する手法

まず、図 7(b) の VH (以下学習 VH と呼ぶ) を主成分分析して低次元空間に射影することを考える。主成分分析を行うためには各フレームにおける次元数を統一する必要があるために、学習 VH の重心を中心とした直方体であるバウンディングボックスを作る。そして、このことを 2.4.1 で記録した全フレームの VH に対して行うことで、全フレームのバウンディングボックス内に存在するボクセルデータを主成分分析し、低次元空間を作成し、その低次元空間に各フレームの学習 VH を射影し、時系列学習データとして記録する。

2.4.3 低次元空間にて入力 VH とパーツ分割された 3 次元形状モデルとをマッチングさせる方法

入力 VH を 2.4.2 で作成した低次元空間へ射影する。そして、記録してある全学習データ点とのユークリッド距離を調べ、最も距離が小さいものを選び出す。つまり、入力 VH と最も形状に近い学習 VH がどれであるかが分かるので、2.4.2 で記したように入力 VH と同じ形状を持ちながらパーツ分割された 3 次元形状モデルを選ぶことにもなる。よって、これにより入力 VH とパーツ分割された 3 次元形状モデルとをマッチングさせることが可能である。

2.5 本研究で用いる 3 次元ビデオ映像作成全体の流れ

本研究では図 7 に示す流れで 3 次元ビデオ映像作成を行う。図 6 で示す 3 次元ビデオ映像作成の流れと異なる点は、2.4 で述べたように、オフラインにて人

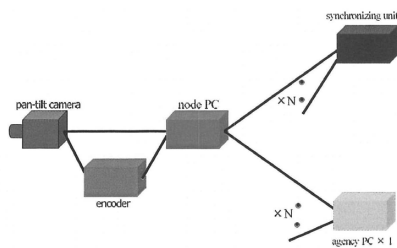


図 8 提案したシステム構成
Fig. 8 Proposed System

体を各パーツごとに分割した学習データを作成する既存手法を使用する点である。この学習データ作成手法を用いる事により、着物などの緩い着衣に対応する事が可能になる。

次章からは、まずは本研究で採用する SFS の実装手法の詳細について述べた後、実環境下にて図 7 の撮影の流れでオンラインカメラワーク決定を行うために、(2) 必要なシステム構成、(3) カメラワークの評価手法、(4) カメラワーク決定アルゴリズムに関して記す。

3. オンラインカメラワーク決定のためのシステム構成

2.5 で示した流れで 3 次元ビデオ撮影を行うために、本稿では図 8 の様なシステムを提案する。

このシステムはカメラ、パン・チルト角度を取得するためのエンコーダ、ノード PC がそれぞれ N 台、エージェント PC が 1 台で構成される。本節では提案したシステム構成の処理の流れについて述べる。

3.1 撮影画像・角度情報・撮影時刻の取得

まずは、ノード PC において以下の様な事柄を受信する。

- ・ 撮影画像
- ・ エンコーダから取得した正確なパン・チルト角度情報
- ・ 画像の撮影時刻を刻んだタイムスタンプ

ただし、これらは全て同期している。

3.2 撮影画像からのシルエット抽出

3.1 で述べた撮影画像から背景差分・色情報を用いた手法により撮影対象のシルエットを抽出する。

3.3 エージェント PC における処理

- (1) 各ノード PC から送信された撮影対象シルエット、パン・チルト角度情報、タイムスタンプを受信する。
- (2) タイムスタンプを利用して、それぞれのノード PC から獲得した対象シルエットを同期させる。
- (3) SFS を実行する。
- (4) 得られた VH を主成分分析により低次元空間に

射影し、Nearest Neighbor 法により最もユークリッド距離に近い学習データを選び出す。

- (5) 選び出した学習データに対して、後述の手法により適切なカメラワーク決定を施す。
- (6) 決定したカメラワークをノード PC に送信する

4. カメラワークの評価手法

本節ではそのカメラワークがどれくらい適切かを定量的に評価するために、カメラワークの評価基準を定め、それらに基づく評価関数を考案した。それぞれの評価基準を次に示す。

- (1) カメラ全体に関する評価基準
 - ・ 対象の全表面に対する被撮影率
- (2) パン・チルト動作の機械的な事柄に関する評価基準
 - ・ 移動距離の短さ
 - ・ 可動範囲内であるか
- (3) 個々のカメラに関する評価基準
 - ・ パーツの見え方
 - ・ 2 つのカメラ間角度の直角への近さ
- (4) 個々のパーツに関する評価基準
 - ・ パーツの重要度

以降それぞれの評価基準に基づく評価関数について述べる。

4.1 カメラ全体に関する評価基準

4.1.1 対象の全表面に対する被撮影率

デジタルアーカイブされた無形文化財を閲覧する際には必ず全身が見れるようにしなければならない。例えば、3 次元ビデオが手が欠けている人体の立体映像であってはならない。その事を避けるためには、撮影時に全てのカメラを用いて対象の全身を隈なく捉える必要がある。そこで、本稿では VH の全表面ボクセル数に対する被撮影ボクセル数の割合を計算し、その割合がある閾値以上の場合のみ評価値を与える事とする。これを実現するために閾値を N_{th} 、被撮影ボクセル数を n 、全表面ボクセル数を N とすると、評価関数は以下ようになる。

$$E_1(n) = \frac{1}{1 + \exp(a(n - N_{th}N))} \quad (1)$$

4.2 パン・チルト動作の機械的な事柄に関する評価基準

4.2.1 移動距離の短さ

カメラをパン・チルト回転させると機械的振動やモーションブラーなどにより正常な画像が得られない可能性がある。その可能性を抑制するためには、なるべく低速でカメラを回転させなければならない。そこで、回転時間を一定時間に固定すると回転角度が短いほどカメラ回転をより低速にできることから、本稿では回転角度の小ささを評価基準として用いることにする。そこで、回転角度を r_m とし、以下の様な評価関数を

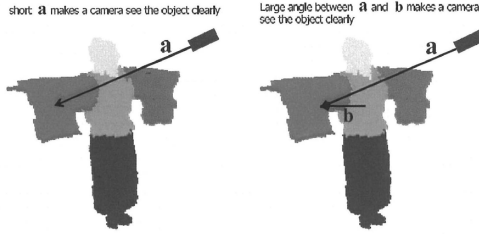


図9 パーツの見え方を評価する原理
Fig. 9 The method of evaluation of parts occlusion

定めることにより上記のような評価基準を満たすカメラワークに評価値を与えるようにする。ただし、 r_m が大きすぎると一定時間内に目標角度まで到達できない可能性があるため、 R_{th} を回転角度の閾値とし、 r_m が R_{th} を超えると評価値を与えないように考慮した。

$$E_2(r_m) = \frac{1}{1 + \exp(a(r_m - R_{th}))} \quad (2)$$

4.2.2 目標パン・チルト角度が機械的限界値を超えない事

カメラワーク決定をしても、パン・チルト角度が機械的限界を超えていれば実環境下において、そのカメラワークを実現することは不可能である。そのため、目標パン・チルト角度が機械的限界を超えていないかを評価する必要がある。本稿では目標角度を a_t 、機械的限界角度の負値の限界と正值の限界をそれぞれ A_{min} 、 A_{max} とする事により、次のような評価関数を定めた。

$$E_3(a_t) = \min\left(\frac{1}{1 + \exp(-a(a_t + A_{min}))}, \frac{1}{1 + \exp(a(a_t - A_{max}))}\right) \quad (3)$$

4.3 個々のカメラに関する評価基準

4.3.1 パーツの見え方

一つのカメラが一つのパーツを撮影するとき、遮蔽が起こらずに見えている度合い、つまり見え方を定量化する関数を定める事にする。図9に示すように、カメラからパーツまでのベクトル \mathbf{a} 、胴重心からパーツ重心へのベクトル \mathbf{b} を定義し、次のような評価基準を決める。

- \mathbf{a} が短い
- \mathbf{a} と \mathbf{b} のなす角度が大きい

これらの評価基準に基づき、本稿では以下の様な評価関数を定義した。ただし、 \mathbf{m} をカメラから最も近いパーツまでのベクトルとする。

$$E_4(\mathbf{a}, \mathbf{b}) = \exp\left(-\frac{(|\mathbf{a}| - |\mathbf{m}|)^2}{2\sigma_1^2}\right) \exp\left(-\frac{(\theta_{ab} - \pi)^2}{2\sigma_2^2}\right) \quad (4)$$

4.3.2 2つのカメラ間角度の直角への近さ

一般にSFSでは2つのカメラ間角度が直角に近い

とファントムボリュームが小さくなるとされている。したがって、2つのカメラ間角度が直角に近いほど評価値を大きくする評価関数を定める。本稿では、次の様な評価関数を定義した。ただし、 θ_{ij} をカメラ i, j 間角度と定義する。

$$E_5(\theta_{ij}) = \exp\left(-\frac{(\theta_{ij} - \pi/2)^2}{2\sigma^2}\right) \quad (5)$$

4.4 個々のパーツに関する評価基準

4.4.1 パーツの重要度

パーツによってテクスチャの細かさやファントムボリュームの存在量が異なる。そこで、本稿では(1)テクスチャが細かく、(2)ファントムボリュームの存在量が多いほど重要度 I が高いとし、重要度が高いほど評価値が高くなる関数を定義する。

まずは重要度 I の決め方についてであるが、本稿では人手によって I の値を定める。人が最も重要度が低いと判断したパーツから昇順に正の整数値を I に代入する。

そこで、本稿では重要度 I に応じた評価関数を次のようにした。

$$E_6(I) = I \quad (6)$$

5. カメラワーク決定アルゴリズム

本研究で要求するカメラワーク決定アルゴリズムの条件は以下ようになる。

- 目標時間である500msec未満であること
- 4章で定めた全ての評価関数を用いてカメラワーク決定が可能であること

そこで、まずはカメラワーク決定アルゴリズムに関する従来研究について述べる。文献3)で考案されたアルゴリズムは、(1)パーツの見え方、(2)カメラの視線方向のばらつき度合いを評価した関数を二つ定義した場合のみに適応可能であり、パン・チルト機構の機械的な事柄を考慮した評価関数が利用できないため、動的な実環境下でこのアルゴリズムを用いることはできない。また、文献4)で定められたアルゴリズムは、図10に示すエラー分布を定義し、エラー分布の最大値を最小にするようなカメラワークを分枝限定法により探索する。このアルゴリズムはシミュレーション上で行うことを想定したもので、文献3)のアルゴリズムと同様にパン・チルト機構の機械的な事柄を考慮していない。さらに、場合によっては実行時間が数時間にも及ぶために確実なオンライン性がない。したがって、このアルゴリズムも本研究では利用不可能である。

したがって、上記条件を満たすカメラワーク決定アルゴリズムの従来手法が存在しないので、本稿では以下に述べるアルゴリズムを提案した。

5.1 提案したカメラワーク決定アルゴリズム

本稿では4章で定めた個々の評価関数値をある程度大きくするためにカメラワーク決定アルゴリズム

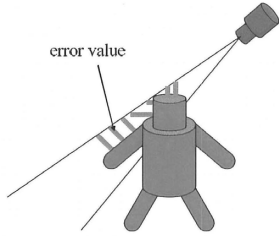


図 10 エラー分布
Fig. 10 Error Value

を提案した。このアルゴリズムはパーツの重要度を定量化した評価関数 E_6 の値が高いパーツから順に、そのパーツを見るカメラを一つずつ決定していくものである。

提案したアルゴリズムの詳細を以下に述べる。ただし、カメラを C_i とし、添え字はカメラの番号を表すとする。またパーツを P_j とし、添え字は重要度が高いほど小さくなるパーツ番号とする。さらに、カメラの個数を N_c 、パーツの個数を N_p とする。

- (1) $j = 1$ とする。
- (2) C_i にて P_j を見たときの評価値 $E = \sum_{i=0}^{N_c} a_i E_i$ を $j = 1, 2, N_c$ について計算する。ただし、 a_j は評価関数 E_j の係数とする。
- (3) (2) で最も評価値 E が大きかったカメラ C_i をパーツ P_j に割当てる。
- (4) パーツに割当てていないカメラがなくなり次第アルゴリズムを終了する。
- (5) $i = N_p$ の場合 $j = 1$ とする。
- (6) i に 1 を加算し、(2) に戻る。

このカメラワーク決定アルゴリズムの計算量は次のようになる。まず、始めにパーツ P_1 に割当てるために探索するカメラ数は N_c 通りある。そして、次にパーツ P_2 に割当てるために $N_c - 1$ 通りのカメラを探索する。割当てる全カメラがなくなるまで同様に繰り返すので、全計算量は以下のようになる。

$$\sum_{i=1}^{N_c} (N_c - i) = \frac{1}{2} N_c (N_c - 1) [\text{times}] \quad (7)$$

したがって、上式により計算量はカメラ数だけに依存し、カメラ数が決まれば一定の量になることが分かる。このアルゴリズムにおけるカメラ数と探索回数との関係を表したグラフを図 11 に示す。

6. 実験の結果と評価

実環境下で 7 台のカメラを用いてズームアップ撮影を行った実験結果を以下に記す。(1) このシステムの処理性能、(2) 取得画像と VH に関する実験結果と評

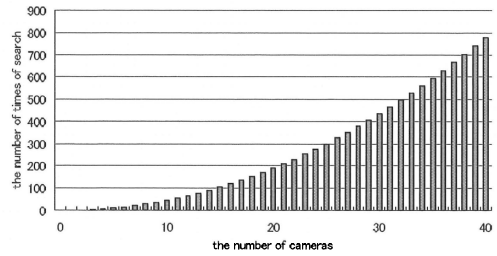


図 11 カメラワーク決定アルゴリズムの計算量
Fig. 11 Amount of calculation of algorithm

価について述べる。

6.1 システム性能

本節では使用するシステムの処理性能について以下に示す。

ノード PC 内での各処理性能は次のようになる。

- ・ キャプチャレート 約 7.5fps
 - ・ パン・チルト動作に要する時間 200msec に固定
- また、エージェント PC 内での各処理性能は
- ・ SFS にかかる時間 約 30msec
 - ・ マッチング速度 約 50msec
 - ・ カメラワーク決定に要する時間 10msec 未満

となる。ただし、マッチングは寄与率が約 90 に対して行われる。

この結果より、設定したカメラワーク決定間隔 500msec 以内に全ての処理が行われていることがわかる。

6.2 取得画像

ズームアウトしたときとズームアップしたときの撮影画像を比較することにより評価を行う。一つの姿勢に対するズームアウトとズームアップによる撮影画像をそれぞれ図 12 と図 13 に示す。

図 12 と図 13 からズームアップによる画像の方がズームアウトによる画像よりも、対象表面の詳細なテクスチャが得られていることが示される。その一方で、図 14 の VH を構成するボクセル数はズームアウト時が 1074 個、ズームアップ時が 1382 個となる。また図 15 も同様にズームアウト時が 961 個、ズームアップ時が 1215 個となる。よって、ズームアップによる VH の方がよりファントムボリュームが多く存在していることが分かる。

7. おわりに

本研究においては無形文化財のデジタルアーカイブを 3 次元ビデオという情報メディアを用いて行おうと

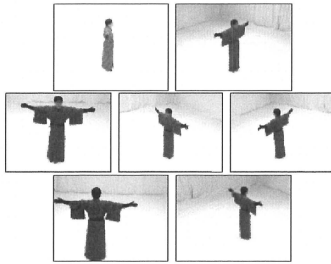


図 12 ズームアウトしたときの撮影画像
Fig. 12 Images taken by zoom-out camera

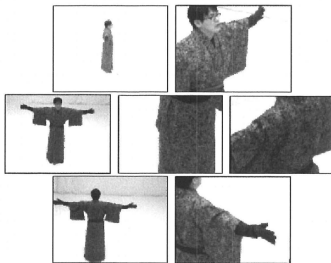


図 13 ズームアップしたときの撮影画像
Fig. 13 Images taken by zoom-up camera

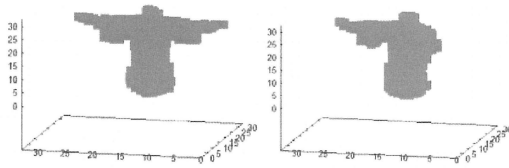


図 14 ズームアウトしたときの VH
Fig. 14 VH created with zoom-out images

した。本研究では一般的な3次元ビデオ撮影が有するズーム制限を取り払うことで撮影画像を高精細化し、また高精細化と引き換えに撮影対象の一部しか捉えられなくなったカメラを能動的にパン・チルト回転させることでカメラワーク決定をし、対象形状の正確さがある程度保持しながら高精細なテクスチャを取得する手法に着目した。そして、それを初めて実環境下で動的な対象に対して撮影を行うために、新たな3次元ビデオ映像作成の流れを考案し、その流れを実現するためのシステム構成、カメラワーク評価法、カメラワーク決定アルゴリズムを提案した。その結果、実環境下で動的対象を高精細にとりつつ、形状の正確さを維持しながら撮影することに成功した。

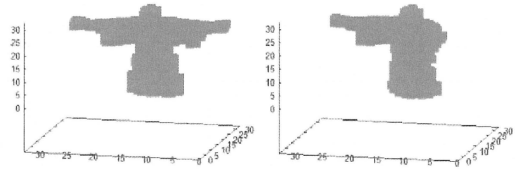


図 15 ズームアップしたときの VH
Fig. 15 VH created with zoom-up images

謝辞 本研究の一部は、文部科学省プロジェクト「知的資産の電子的な保存・活用を支援するソフトウェア基盤技術の構築」における「大型有形・無形文化財の高精度デジタル化ソフトウェアの開発」の補助を受けた。また、色検出については、和歌山大学和田俊和教授にソフトウェアを提供して頂いた。深謝致します。

参考文献

- 1) 松山隆司, 高井勇志, ウ小軍, 延原章平: 3次元ビデオ映像の撮影・編集・表示, TVRSF Vol.7 No.4(2002)
- 2) 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- Vol.J81-D- No.6 pp.1182-1193(1998)
- 3) 西山学, 松山隆司: 3次元形状モデルを用いたカメラワークプランニング, コンピュータビジョンとイメージメディア, pp.49-56(2002)
- 4) 西山学, 牧淳人, 松山隆司: 高精細3次元形状復元のための多視点カメラワーク最適化, 京都大学大学院 修士論文(2004)
- 5) 辻良介, 浮田宗伯, 木戸出正継: 時系列形状の主成分分析に基づく人体・着衣の形状解析, 奈良先端科学技術大学院大学 修士論文(2007)
- 6) Wolfgang Niem, Marcus Steinmetz: Camera Viewpoint Control for the Automatic Reconstruction of 3D Objects, ICIP(1996)
- 7) ウ小軍, 和田俊和, 東海彰吾, 松山隆司: 平面間透視投影を用いた並列視体積交差法, 情報学会論文誌: コンピュータビジョンとイメージメディア, Vol.42(2001)
- 8) Norimichi Ukita: Real-Time Cooperative Multi-Target Tracking by Communicating Active Vision Agents, pp.73-112
- 9) 日浦慎作, 村瀬健太郎, 松山隆司: ダイナミックメモリを用いた実時間対象追跡, 情報処理学会論文誌, Vol.41, No.11(2000)
- 10) German K. M. Cheung, Simon Baker, Takeo Kanade: Shape-From-Silhouette of Articulated Objects and its Use for Human Body Kinematics Estimation and Motion Capture, IEEE Computer Society Conference on CVPR(2003)