

時系列視体積の統合による 3 次元復元形状の再現性の向上

豊浦 正広[†] 飯山 将晃^{††} 角所 考^{††} 美濃 導彦^{††}

[†] 京都大学大学院情報学研究科 〒 606-8501 京都市左京区吉田本町

^{††} 京都大学学術情報メディアセンター 〒 606-8501 京都市左京区吉田本町

E-mail: [†]toyo@mm.media.kyoto-u.ac.jp, ^{††}{iiyama,kakusho,minoh}@media.kyoto-u.ac.jp

あらまし 視体積交差法では、対象物体を観測するカメラ台数が多ければ多いほど、復元形状の再現性が向上する。しかし、カメラ台数の増加は空間的制約やキャリブレーションの難しさにより、その実現が難しいことがある。そこで我々は、時系列で得られる画像の情報を統合し、少ないカメラ台数で多くのカメラを用いたときと同様の復元形状を得ることを目指す。対象物体の剛体運動を各時刻の視体積から推定することができれば、複数の時刻でのカメラ画像の情報を統合して、1つの時刻で得る視体積よりも、より再現性の高い視体積を得ることができる。対象物体の運動推定は、特徴点を各時刻の視体積から抽出し行う。この際に、視体積が物体が占める領域以外に余分な領域を含み、視体積から安定して特徴点を抽出することが難しい。我々は視体積要素の画像への逆投影を用いて、特徴点を安定して抽出した。シミュレーション実験において、多くのカメラを用いたときと同等の復元形状を得られることを確認した。

キーワード 視体積交差法, 特徴点抽出, 運動推定, 時系列データ

An Accurate Shape Reconstruction Method by Integrating Visual Hulls in Time Sequences

Masahiro TOYOURA[†], Masaaki IIYAMA^{††}, Koh KAKUSHO^{††}, and Michihiko MINOH^{††}

[†] Graduate School of Informatics, Kyoto University, Yoshidahonmachi, Sakyo-ku, Kyoto, 606-8501 Japan

^{††} Academic Center for Computing and Media Studies, Kyoto University, Yoshidahonmachi, Sakyo-ku,
Kyoto, 606-8501 Japan

E-mail: [†]toyo@mm.media.kyoto-u.ac.jp, ^{††}{iiyama,kakusho,minoh}@media.kyoto-u.ac.jp

Abstract More accurate shapes can be reconstructed with many cameras in the volume intersection method. Generally, using many cameras is difficult due to physically restrictions and difficulty of calibration. In this paper, we propose a method to integrate information of images in time sequences to get more accurate shapes instead of using many cameras. In the case that rigid object motion of the target is estimated, virtual motion of the camera set can be calculated. The motion is estimated with the feature points of a visual hull at each time. The visual hull includes target object region and extra region. The extra region makes it difficult to extract feature points. Feature points are extracted by back projection of voxels of the visual hull. As a result, we can get more accurate shapes with virtual images of virtual camera set.

Key words Volume intersection method, extracting feature points, motion tracking, time series data.

1. はじめに

多種多様な昆虫や発掘物、建造物などを 3 次元デジタルコンテンツとして持つ電子博物館 [1] が注目されている。電子博物館のような博物学的用途に対しては、高い再現性で多種多様な物体を計測し、その 3 次元形状を復元することが求められている。

多種多様な物体の 3 次元形状復元の手法として視体積交差法 [2] がある。視体積交差法は、博物学のように、広範な物体を対象に形状復元を行う必要がある場合の手法として大変優れている。これは、視体積交差法が物体表面の詳細な色情報を必要としない性質を持つためである。

視体積交差法では、カメラ台数が増えれば増えるほど、より原形状に近い復元形状を得ることができる。近年、カメラは低

価格になり、大量カメラ環境の構築が容易に可能となった。これに伴い、より復元形状を原形状に近づけることができるようになってきた。

ここで、無限台数のカメラを用いて視体積を構築すると、復元形状として物体の最小凸包が得られる。つまり、物体の凹面を表現することはできない。これは視体積交差法による再現性の限界といえる。しかし、電子博物館のコンテンツへの適用を考えると、凹面が表現できないことは問題とはならない。なぜなら、物体の凹面は最小凸包上で陰影として表現されるため、復元形状にテクスチャを貼ることによって凹面を表現できないという欠点を補うことができるからである。以上の議論より、本稿では復元形状が物体の最小凸包に近づくことを再現性の向上と呼び、これを目標とする。

大量カメラが物理的には入手可能になったが、空間的に設置できるカメラ台数は限界があり、カメラのキャリブレーションの精度を十分に確保できるカメラ台数にも限界がある。そのため、視体積交差法で得られる復元形状の再現性の向上にもまた限界がある。シミュレーションデータを用いた予備実験により、数十台程度の現実的な台数のカメラを用いて視体積交差法を行ったとしても、十分な再現性が得られないことが確かめられた。

そこで、数十台程度の現実的な台数のカメラを用いて視体積の再現性を確保するために、時系列に並んだ画像情報を統合することで、カメラ台数を増加させた場合と等価な視体積を得る手法が考えられている。対象物体が観測している間にその形状が変化しないときには、物体が運動することによる対象とカメラの相対位置の時間的変化をカメラ位置の空間的な変化とみなすことができ、複数の時刻での視体積の情報を統合することで、1つの時刻で得る視体積よりもより再現性の高い視体積を得ることができる。

カメラの相対位置の時間的変化をカメラ位置の空間的な変化とみなす手法として、ターンテーブルを用いる手法が一般的に用いられる。ターンテーブルを用いる手法では、ごく少数のカメラにより物体の復元形状を獲得することができる。しかし、ターンテーブルを用いる手法では限られた方向からの画像しか獲得できず、また、対象物体をテーブルの上に固定する必要があることから、テーブルに接する部分の形状を獲得することが難しい。さらに、ターンテーブルに固定することのできない大きな物体や動く物体への適用はできない。

これに対し、Cheung ら [3] は画像上で物体の特徴点を追跡し、物体の運動を推定することによってカメラ台数を増加させた場合と等価な視体積を得る手法を提案した。しかし、この手法では特徴点の追跡に物体の色情報を用いており、視体積交差法の特長である物体表面の詳細な色情報を必要としない性質を損ねている。視体積交差法の特長を生かすためには、色情報を使わずに特徴点抽出を行うべきである。

これらの従来研究に対して、我々は色情報を用いず、各時刻の視体積の形状情報のみから特徴点をとらえて、運動推定を行う。このときに、各時刻の視体積は物体が占める領域以外に余分な領域を含むことが、特徴点抽出の上で問題となる。そこで、

視体積表面に現れる物体表面上のボクセルに注目する。これらのボクセルの中には、カメラと物体の位置関係の変化に依らず安定して抽出されるものが多数存在するため、これらを特徴点として抽出し、その特徴点を元に運動推定を行う。

2節ではまず、本研究の背景である視体積交差法について説明し、物体の運動を知ることによって仮想的なカメラ台数の増加が可能になることの原理を示す。3節では各時刻の視体積上で特徴点を求める手法を提案し、4節では特徴点からの物体の運動推定手法について説明する。5節では形状復元を含めた実験結果を示し、その精度について考察を行い、6節で全体についてまとめる。

2. 復元形状の再現性の向上

本節では視体積交差法の原理を説明し、復元形状の再現性向上のための手法を提案する。

2.1 視体積交差法

視体積交差法では、複数台のカメラで対象となる物体を観測する。物体はすべてのカメラの観測範囲内に収まるものとする。以下ではカメラ n 台で視体積交差法を行う原理の説明をする。

カメラ C_i の観測範囲に物体 O が存在するとき、物体 O は画像上に投影され、その投影像は2次元画像平面上の領域として得られる。この領域を物体投影領域 R_i と呼ぶ。 P_i はカメラ C_i に固有な投影変換行列である。

物体 O はカメラ C_i のレンズ中心を端点とし、物体投影領域 R_i 上の任意の点を通る半直線の集合からなる錐体状の開空間に内接して存在する。この開空間をカメラ C_i における視錐体と呼ぶ。これは C_i における画像から得られる物体 O の3次元形状に関する情報である。ここで、各カメラにおける画像から得られる視錐体 V_1, V_2, \dots, V_n について、式 (1) のような V を考える。

$$V = \bigcap_{i=1}^n V_i \quad (1)$$

この V を求めることで各カメラから得られる物体 O の3次元形状に関する情報を統合することができる。

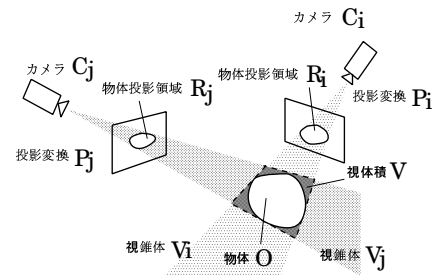


図1 視体積交差法

本研究では、視体積交差法の実現にボクセルモデルを用いた。3次元の位置情報を持つ1つ1つのボクセルが、それぞれのカメラが成す視体積の中に含まれるかどうかを判定する。1度でも含まれないと判定されたボクセルには0を与え、含まれると判定される限り1を与えた。1を与えられたボクセルの集合が

視体積となる。

2.2 カメラ台数と視体積の再現性の向上

視体積の持つ特徴として次の2つが挙げられる。

(1) 物体は視体積に常に内接する

(2) 視点が増加すればするほど視体積の体積は減少する

視体積は一般的に物体形状に一致しないが、これらの特徴より、カメラ台数が多ければ多いほど視体積は物体形状に近づくことと結論づけられる。つまり、復元形状の再現性の向上を図るためにはカメラ台数の増加を行えばよい。

そこで、1節に述べた限度以上の視点増加を実現するため時系列で得られた視体積を統合することを考える。複数の時刻での物体の姿勢と物体の視体積が得られ、かつ、物体が剛体ならば、各時刻間での物体の運動を推定し、視点を仮想的に増加させることができる。

2.3 標本化誤差による再現性向上の限界

原理的には、カメラ台数が増えれば増えるほど視体積の再現性は向上するが、入力データである画像が離散的なデータであることに起因する再現性向上の限界が存在する。そこで、再現性の向上がこの限界に達するカメラ台数に関する考察を行う。

多くの台数のカメラを用いて視体積交差法を行うことを考える。カメラ画像の1つの画素に含まれる物体投影領域の面積が、画像の標本化を行う際、画素面積の半分以上であるときには物体投影領域を示す画素であるとみなされ、半数に満たない場合には物体投影領域外の領域を示す画素であるとみなされる。画像と空間の標本化間隔が一致するように取られているときには、画像の量子化誤差により、視体積表面には ± 0.5 ボクセルの標本化誤差が現れる。これが画像の標本化に起因する再現性の向上の限界である。

視体積は各カメラ画像から得られる物体存在可能領域の積領域を求めると、無限台数のカメラで得られた離散化された画像を用いた視体積は、無限台数のカメラで得られた離散化されていない画像を用いた視体積と比較して、表面から0.5ボクセル削られた形状をとる。つまり、無限台数のカメラで得られた離散化された画像を用いた視体積と、その表面に1ボクセル分の厚みを加えた視体積との間に離散化の影響を受けない視体積が存在する。これが再現性の向上の限界となる。

この性質に基づき、再現性の向上が限界に達するカメラ台数をシミュレーション実験によって求めた。シミュレーションデータを用いて、カメラ中心をランダムに与え台数を増やすときの、視体積に含まれる余分な領域のボクセル数の減少を調べた。ほぼ凸物体である約430000ボクセルで表現される物体形状を対象とし、カメラ台数を1台ずつ増加させ視体積を構築した。40000台を用いて得られるときの視体積を無限台数のカメラで得られる形状とみなし、これとの差分のカメラ台数による変化を図2に示す。

約430000ボクセルからなる球を考えると、その表面に現れるボクセル数は約28000であると計算できる。いま、画像と空間の標本化間隔は一致するように取っているため、画像の量子化誤差により、視体積表面には ± 0.5 ボクセルの量子化誤差が現れる。つまり、約14000ボクセル以内の余分な領域は量子

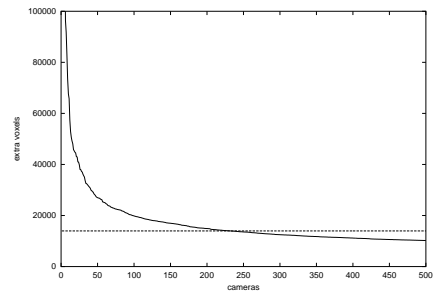


図2 カメラ台数と余分な領域の推移

化誤差に吸収されることになる。一般の物体では、球よりも表面積が大きくなり、これ以上の量子化誤差が起こることが予想される。

図2より、カメラ台数が230台程度を超えるときには余分な領域は量子化誤差に吸収され、再現性の向上の理論的限界に到達したと考えられる。つまり、再現性の向上のためには、230台程度以上のカメラ情報の統合が有効である。

再現性の向上に必要な230台程度以上のカメラ台数を用いて得られる視体積と同程度の再現性を、現実的なカメラ台数で実現するために、時系列で得られた視体積を統合することを考える。複数の時刻での物体の姿勢と物体の視体積が得られ、かつ、物体が剛体ならば、各時刻間での物体の運動を推定し、カメラ台数を仮想的に増加させることができる。我々は1節で述べたようなターンテーブルを用いる手法の問題点を解決するために、各時刻で得られる視体積から物体の運動を推定し、各時刻で得られる画像の情報を統合することを考える。

2.4 形状復元のための物体運動

実験環境として、校正済みの複数台のカメラが観測する空間内に、物体を糸で吊り、運動を加えるようなものを想定する。各時刻の視体積から運動推定が行われるので、各時刻の視体積はある程度の再現性が必要であり、本研究では10数台のカメラを持つとする。

このような実験環境下では、物体の運動のおおまかな大きさを観測者が与えることができる。物体が静止し続ける場合には本手法による再現性の向上は望めないが、カメラの設置間隔を超えるような大きな運動を与える必要はない。例えば、円周上に4台のカメラがあり、この円の軸の周りに物体が回転することを考えると、 x° 回転させたときのカメラ画像は、 $(x + 90)^\circ$ 回転させたときのカメラ画像と同じものとなる。

3. 特徴点抽出

本研究では、各時刻で求められる視体積から特徴点を抽出し、隣接する時刻間で対応する特徴点の運動を考えることにより、対象物体の運動推定を行う。本節ではまず、各時刻での視体積からの特徴点抽出を考える。

3.1 特徴点に要求される性質

各時刻で視体積から特徴点抽出を行い、隣接する時刻間で特徴点を対応づけることにより物体の運動追跡を行う。その際、特徴点として物体とカメラの位置関係が変化しても、安定して

抽出されるような点を用いることが必要となる。

一般的に、視体積のような3次元形状からの特徴点の抽出には3次元曲率が利用される。この手法では、物体表面上のすべての点において3次元曲率を計算し、局所的に最大になるような点を特徴点とする。これを視体積に適用することを考えたとき、各時刻の視体積が余分な領域を含むことが問題となる。各時刻での視体積の物体形状領域以外に余分な領域の影響により、視体積表面の3次元曲率が物体形状表面の3次元曲率と一致しないことがあるためである。

そこで、我々は各時刻の視体積に含まれる余分な領域の影響を受けない点を視体積表面から抽出することを考える。この点は、物体表面上の点の一部である。

物体表面上の点の抽出にあたり、物体表面上の点は余分な領域によって、視体積表面上に現れないことがあることを考慮する必要がある。視体積表面から物体表面上の点であることが保証されるような点を抽出することを考える。このような点を以降、特徴点と呼ぶ。

3.2 逆投影による特徴点抽出手法

特徴点を求めるために、カメラの画像のシルエットの境界領域を利用する。図3のように視体積表面上のボクセルを画像へ逆投影する。あるボクセルを通過し、かつ、シルエットの境界領域の画素を通る視線に注目し、その視線上に1つの特徴点候補のみが存在するとき、この視体積表面上のボクセルを特徴点に採用する。これを対象となる時刻の視体積を構築するすべての画像について行う。

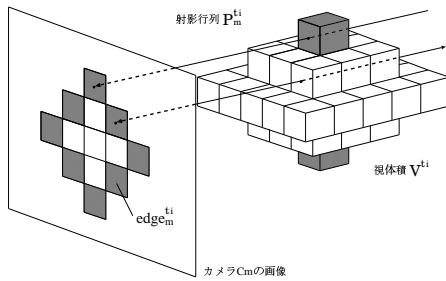


図3 境界要素の再投影

つまり、時刻 t_i における特徴点 F^{t_i} は以下の式で表される。

$$F^{t_i} = \{v \mid \forall m, P_m^{t_i} v \in \text{edge}_m^{t_i}, v \in V^{t_i}\} \quad (2)$$

ただし、 $P_m^{t_i}$ は時刻 t_i におけるカメラ C_m の射影行列、 $\text{edge}_m^{t_i}$ は時刻 t_i におけるカメラ C_m のシルエットの境界領域、 V^{t_i} は時刻 t_i の視体積である。さらに、

このようにして得られた特徴点は視体積表面上のボクセルであることはもちろんであるが、物体領域に含まれるボクセルであることも保証ができる。その理由は視体積交差法の原理より明らかになる：視体積交差法の原理からシルエットは視体積の投影像に一致することが保証される。もし、斜線で示したボクセルが物体領域に含まれないとすれば、図3で示す視体積の投影像はシルエットに一致しないことになり、この保証が成り立たなくなる。

また、このような点は、各時刻で安定抽出されることが言える。なぜなら、物体形状で突出しているような部分が特徴点に抽出されやすいためである。図4にこれを示す。図4(a), (c)では物体形状を持つような部分ではカメラと物体の相対的な位置変化によって、特徴点となるボクセルが変化しない。一方、図4(b), (d)のように物体形状を持つような部分では、カメラと物体の相対的な位置変化によって、特徴点となるボクセルが変化する。これは他のボクセルからの自己隠蔽の起こりやすさに起因するものであり、物体上の突出した部分ほど安定した点の特徴点として抽出されやすく、逆に谷間にある部分は特徴点として抽出されることがない。

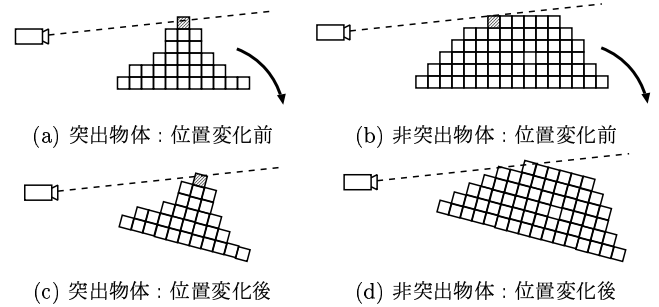


図4 相対位置の変化による特徴点の変化

運動推定を行う時刻間の物体の運動変化が小さい場合には、多数の突出した部分に対応するボクセルが安定して抽出されることが予想される。しかし、すべての特徴点が、突出した部分に対応するボクセルとはならない。カメラ配置によっては突出していない部分も特徴点として抽出され得るからである。これは視体積が余分な領域を含むためである。そこで運動成分推定の際に、外れ値を扱うことのできるロバスト推定を行い、突出した部分を表現する特徴点のみを推定に用いる工夫を行っている。これは4節で扱う。

4. 運動推定

特徴点を用いて剛体運動を推定するとき、この運動を表す剛体変換行列を用いる。運動追跡はこの剛体変換行列を収束法を適用して求めることで行う。

4.1 剛体変換行列

剛体変換行列 $D^{t_i t_j}$ は、時刻 t_i と時刻 t_j の間の物体の運動変化を示す斉次座標行列で、 3×3 の回転移動行列 $R^{t_i t_j}$ と平行移動成分ベクトル $\mathbf{t}^{t_i t_j}$ を用いて、次のように表すことができる。

$$D^{t_i t_j} = \begin{pmatrix} R^{t_i t_j} & \mathbf{t}^{t_i t_j} \\ \mathbf{0}_{1 \times 3} & 1 \end{pmatrix} \quad (3)$$

剛体変換行列は回転、平行移動の変数を決定することで得られる。回転行列には4元数表示を用い、線形性を高めた。また、 $D^{t_i t_j} = (D^{t_j t_i})^{-1}$ が常に成り立つ。

剛体変換行列を用いて異なる時刻の視体積を統合する手法を説明する。まず、物体の画像への投影を考える。図5に示すよ

うに記号を定義する.

まず, 時刻 t_i で物体 O が占める領域を O_{t_i} とし, o_{t_i} を O_{t_i} の要素とする. また, P_n はカメラ C_n の射影行列であり, 時間によって不変であるとする. さらに, r_{t_i} は o_{t_i} をカメラ n のスクリーンに投影した点であり, r_{t_i} が成す投影領域を R_{t_i} とする.

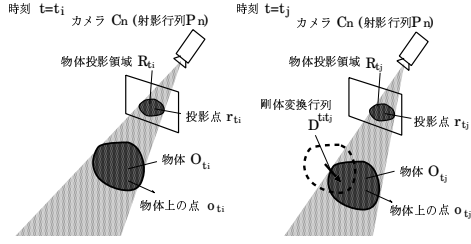


図 5 記号の定義

このときに, o_{t_j} を $(D^{t_i t_j})^{-1}$ により移動した点を P_n で画像に投影すると, これは R_{t_i} の要素となる.

つまり式で表すと,

$$o_{t_j} \in O_{t_j} \text{ ならば, } r_{t_i} = P_n((D^{t_i t_j})^{-1} \cdot o_{t_j}) \in R_{t_i}$$

となる. 同様にして,

$$\begin{aligned} o_{t_i} \in O_{t_i} \text{ ならば, } r_{t_j} &= P_n((D^{t_j t_i})^{-1} \cdot o_{t_i}) \\ &= P_n(D^{t_i t_j} \cdot o_{t_i}) \in R_{t_j} \end{aligned}$$

であるといえる. この式は, $D^{t_i t_j}$ と R_{t_j} が与えられれば O_{t_i} の要素である可能性を持つ o_{t_i} の集合が求められることを示す. この集合は時刻 t_i において, 実際とは異なる視点配置で求めた視体積に等しい. つまり, 時刻 t_i で求めた視体積と $D^{t_i t_j}$ と R_{t_j} から求めた視体積の積領域を求めることは, 時刻 t_i での視体積と時刻 t_j の視体積を統合することと等価である.

4.2 運動成分の求め方

誤差関数を最小にするような剛体変換行列 $D^{t_i t_j}$ を Powell の収束法 [4] を用いて求める. Powell の収束法は他の収束法に比べ, 多くの変数を持つ関数に対して収束が速く, 局所解に陥りにくいという性質を持つ.

時刻 t_i で求めた特徴点は, 剛体変換行列 $D^{t_i t_j}$ により移動させると, 時刻 t_j で求めた特徴点に一致するはずである. しかし実際には, 量子化誤差や特徴点のうち物体の突出した部分に相当しないボクセルが抽出されることがあり, これらは完全には一致しない. そこで誤差関数 E を定め, これを最小化する剛体変換行列 $D^{t_i t_j}$ を Powell の収束法を用いて求める. 時刻 t_i で求めた特徴点 $p_s^{t_i}$ を剛体変換行列 $D^{t_i t_j}$ により移動させたとき, 時刻 t_j における特徴点のうちで最も近くにある特徴点を $p_s^{t_j}$ に対応する点と考え, この点との二乗距離を求める. 誤差関数はこの 2 乗距離が M_{th} 以下であるものの平均をとった値に相当する. 誤差関数 E を以下のように定める.

$$E = \sum_u f \left(\min_s (p_u^{t_j} - D^{t_i t_j} p_s^{t_i})^2 \right) \quad (4)$$

$$\text{ただし, } f(x) = \begin{cases} x & x \leq M_{th} \\ M_{th} & x > M_{th} \end{cases}$$

ここで, $p_s^{t_i}$, $p_u^{t_j}$ はそれぞれ時刻 t_i , t_j での特徴点の座標を表し, M_{th} は時刻 t_i , t_j の間の移動距離の最大量を定めた閾値である. f は, 大きく離れた対応点を持つような特徴点が収束に与える影響を小さくしている. このように, 外れ値の影響を小さくする関数を用いて収束法を適用する推定方法は, 一般にロバスト推定と呼ばれる.

隣接する時刻間で物体の運動量が微小であると仮定し, 剛体変換行列をまず, 隣接する時刻間で求める. 任意の時刻間での剛体変換行列 $D^{t_i t_j}$ は以下のように求める.

$$D^{t_i t_j} = D^{t_j-1 t_j} D^{t_j-2 t_j-1} \dots D^{t_i+1 t_i+2} D^{t_i t_i+1} \quad (5)$$

5. 実験結果

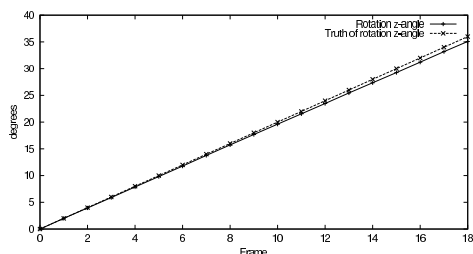
実験はシュミレーションデータを用いて行った. 3次元データは VRML 形式で与えられ, 仮想空間に配置したカメラから撮影を行い, 視体積の構築を行った. 3次元データには, 恐竜の形状を用いている. 恐竜の形状データは 2 節の再現性の理論的限界を考えるとときに扱ったものと同一形状であり, 約 430000 ボクセルで表される形状である. VRML 形式のデータを座標値調整により物体を仮想的に移動させ, それぞれの位置に対して撮影した画像を, それぞれのフレームで撮影した画像とした. 対象物体を各フレーム間で x, y, z 軸方向に 1 ボクセル平行移動, z 軸の周りに 2° 回転移動させ, 正 12 面体の頂点に配置した 20 台のカメラを用いて, 提案手法を適用した.

物体を z 軸周りにのみ回転させるときに 72° 周期で同じ画像が現れること, また, カメラが物体に近接していない場合には原点に対して対照に位置するカメラからはほぼ同じ画像が得られると考えられることから, 実験では $18(=72^\circ \div 2^\circ \div 2)$ フレーム分の画像情報を用いて, 統合視体積を構築した.

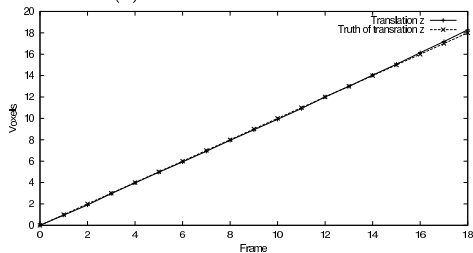
まず, 提案手法で特徴点群から得られた運動推定結果を与えた運動とともに示す. 図 6(a) は特徴点群から推定された z 軸周りの回転運動成分であり, 図 6(b) は z 軸方向の平行移動成分である.

回転移動成分では絶対値の平均で 0.26° , 絶対値の最大で 0.82° の推定誤差が生じた. 平行移動成分では絶対値の平均で 0.34 ボクセル, 絶対値の最大で 1.03 ボクセルの推定誤差が生じた. 平行移動成分については, ほとんどが 1 ボクセル以内の誤差に収まり量子化誤差に吸収される程度の精度の運動推定が行われているといえる. 回転移動成分については, 数値だけでは量子化誤差で吸収される範囲にあるかどうかは判別できないので, 統合視体積の結果を見ることで判断する.

次に, 1 フレームで得られる視体積と統合視体積の差を図 7 を用いて説明する. このときには $360(=20 \times 18)$ 台分のカメラ情報を用いて, 視体積を構築したことになる. 図 7(a) は正解とみなすことのできるカメラ 40000 台を用いたときに構築される視体積, 図 7(b) は 1 フレームの情報のみから構築される視体積, 図 7(c) は図 7(b) で示した視体積から得られる特徴点



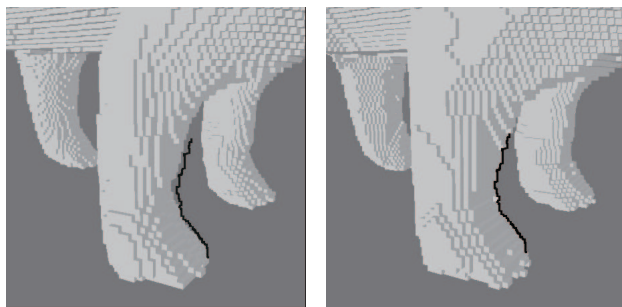
(a) z 軸周りの回転移動成分



(b) z 軸方向の平行移動成分

図 6 特徴点からの運動推定結果

群, さらに, 図 7(d) は 18 フレーム統合時の視体積である. 太線で 1 フレームの情報のみから構築される視体積の輪郭の一部を示す.



(a)40000 台での視体積

(b)1 フレームでの視体積



(c) 特徴点群

(d) 統合視体積

図 7 18 フレームでの視体積統合結果

図 7(c) では, 3 節で示した突出した部分に多く特徴点が抽出されていることが確認できる. 図 7(d) で示す足の前後のくびれの部分について, 特に再現性の向上が顕著に見られる.

また, 数値的な評価を行うために, 40000 台を用いて得られた視体積を正解の視体積とし, 正解の視体積と提案手法によって統合された視体積との違いを検証した. 評価には, 提案手法で得られた視体積に含まれ, 正解の視体積に含まれないボクセル (extra voxels), 提案手法で得られた視体積に含まれず, 正解の視体積に含まれるボクセル (missing voxels), 両者の和

(extra + missing voxels) を用いた. これを図 8 に示す.

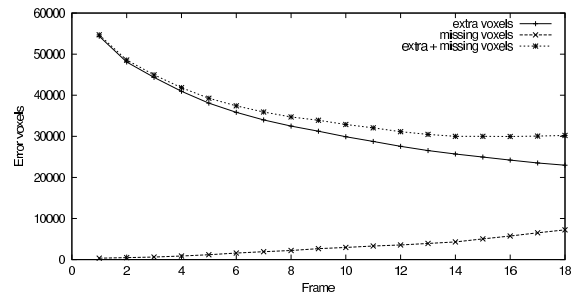


図 8 視体積の再現性の向上

統合するフレームが増加するにつれ, 単調にエラーである統合視体積に含まれる余分なボクセルと統合視体積に含まれない正解の視体積中のボクセルの和が減少した.

提案手法で得られた視体積に含まれず, 正解の視体積に含まれるボクセルの増加は運動推定誤差によって引き起こされるものと考えられる. このボクセルの数の増加は, 現実のカメラ増加によっては起こらない問題であり, 提案手法の問題点であるといえる. しかし, このボクセルの数は 18 フレーム統合時において, 2 節で述べた量子化誤差に吸収される範囲内にとどまっております, 再現性の向上を阻害するものではないと考えられる.

以上の結果より, 提案手法で得られる特徴点群は画像情報の統合に十分な運動推定を可能にするといえる. また, 得られた推定運動によって時系列の画像情報を統合し, 視体積の再現性を向上することができた.

6. おわりに

本研究では, 視体積交差法において時系列で得られる画像の情報を統合し, 現実的なカメラ台数で理論的境界の再現性を示す台数のカメラを用いたときと同様の復元形状を得ることを目指した. 各時刻で特徴点を求め, これを追跡することにより対象物体の運動推定を行い, 別の時刻での画像の情報を統合して視体積の構築を行った.

今後の課題として, 実環境での適用を試みることを挙げる. 実環境下では, 物体のシルエットが正確に得られない場合がある. そのような場合の特徴点抽出方法について, 考察を行う必要がある. 同一点を観測するカメラが複数存在することを用いて, シルエットの修正を行い, 視体積の修正を行うことを考えている.

文 献

- [1] APAN DIGITAL ARCHIVES ASSOCIATION(JDAA) WEB SITE, "デジタルアーカイブ推進協議会," <http://www.jdaa.gr.jp/index.htm>
- [2] Aldo Laurentini, "How Far 3D Shapes Can Be Understood from 2D Silhouettes," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol.17, No.2, pp.188-195, 1995
- [3] "Visual Hull Alignment and Refinement Across Time : A 3D Reconstruction Algorithm Combining Shape-From-Silhouette with Stereo," Proc. CVPR, pp.77-84, 2003.
- [4] NUMERICAL RECIPES in C, 技術評論社, William H.Press, Saul A.Teukolsky, William T.Vetterling, Brian P.Flanney