

[Original article]

(2008年4月7日 Accepted)

包囲カメラ群を用いた立体の全周復元法

守 啓祐¹, 曾根 俊昌², タン ジュークイ², 金 亨燮², 石川 聖二²

¹九州工業大学大学院工学研究科, ²九州工業大学工学部

要約: ステレオ法を用いた立体の3次元形状復元では、カメラで共通に観察される部分しか復元されないため、全周モデルを作成するには部分的に復元された形状をつなぎ合わせる必要がある。このため厳密なカメラキャリブレーションが必要であり、3次元復元の効率化を妨げている。本稿では、因子分解法の適用により、簡便なカメラキャリブレーションに基づく、つなぎ合わせを行わない全周復元法を提案する。まず物体を囲むように複数のカメラを配置し、全カメラで共通に観察される物体上の特徴点に因子分解法を適用して、全カメラの方向および特徴点の3次元座標を求める。次に前面カメラの方向および前面カメラでしか観察されない物体上の特徴点の画像座標から、それらの特徴点の3次元座標を計算し、また後面カメラの方向および後面カメラでしか観察されない物体上の特徴点の画像座標から、それらの特徴点の3次元座標を計算する。この手続きにより、物体の全周の特徴点が復元される。実験により提案法の有効性が示された。

キーワード: 3次元モデル, 形状復元, 因子分解法, 全周, 全形状, コンピュータビジョン

A Method of Entire Shape Recovery of a Solid by Surrounding Cameras

Keisuke Mori¹, Toshimasa Sone², Joo Kooi Tan², Hyoung Seop Kim², Seiji Ishikawa²

1) Graduate School of Engineering, Kyushu Institute of Technology

2) Faculty of Engineering, Kyushu Institute of Technology

Abstract: The existent 3-D modeling technique such as the stereoscopic vision recovers 3-D shape of part of a solid which can be observed commonly from the set cameras. This inevitably needs registration among recovered partial shapes in order to obtain an entire 3-D model of the solid. For this purpose, strict camera calibration is indispensable, resulting in not very efficient modeling procedure. In this paper, a technique based on a simple camera calibration is proposed for recovering entire shape of a solid without registration by the application of the factorization. In the first place, multiple cameras are set around an object interested and the feature points observable from all the cameras recover their 3-D locations by the factorization. By employing the information on the orientations of the cameras observing the front part of the object, those feature points observable only from the frontal cameras recover their 3-D locations. In the same way, the feature points observable only from the cameras looking at the rear part of the object recover their 3-D positions by the employment of the orientations of the rear cameras. In this way, all the feature points on the object recover the 3-D positions, providing entire shape of the object. Experimental results show satisfactory performance of the proposed technique.

Keywords: 3-D modeling, shape recovery, factorization, entire shape, computer vision

Seiji ISHIKAWA

Sensuicho 1-1, Tobata, Kitakyushu 804-8550, JAPAN

Phone: +81-(0)93-884-3183, Fax: +81-(0)93-884-3183, E-mail: ishikawa@cntl.kyutech.ac.jp

1. はじめに

3次元物体のモデル化は、カメラとコンピュータを使った画像情報処理研究のなかでも特に近年注目されている技術であり、仮想現実世界の生成や、ビデオゲームやマンマシンインタラクティブシステムにおける人の立体モデル生成等に利用されている。またスポーツやリハビリテーションにおける人の動作解析などへの応用も期待されている。

立体モデル化は物体の全周または全形状のモデル化（復元とも言う）である。従来法はステレオ法[1]を基本とし、3次元対象物体の周囲に設置されたカメラで物体の部分ごとの特徴点を復元（3次元座標を求めること）し、それをまとめて全周形状を得る。部分ごとの復元の積み重ねであるから、物体を取り巻くカメラ群は高精度にキャリブレーションがなされていなければならない。このため立体器具を用いたカメラキャリブレーションを事前に行う必要があるが、これは手間のかかる作業のため、立体モデル化システム（光学式モーションキャプチャシステムとして製品化されている）が、この分野の専門家でない一般ユーザの間に、広く普及しない要因のひとつとなっている。上述のように立体モデル化の需要は大きいため、簡便なキャリブレーションに基づく立体モデル化システムが求められている。

因子分解による立体モデル化法[2]は、移動撮影するカメラはキャリブレーション不要であるが、移動中常に観察される物体上の部分しか復元されないため、立体モデル化には復元された部分形状のつなぎ合せが必要である。しかし因子分解法は結像系に線形モデルを用いるため、対象とカメラとの距離が十分大きくないと少なからぬ復元誤差が混入し、つなぎ合せアルゴリズム[4]を適用することができない。

タン等[3]は因子分解法を応用して、つなぎ合せを行わずに物体（運動する人のような変形体を含む）の全周形状を復元する方法を提案した。これは、立体器具を使ったカメラキャリブレーションが不要という特徴を持つ方法であるが、対象の回りに設置する包囲カメラ群において、相対するカメラの光軸を一致させる必

要があった。

本稿では、タン等の方法を改良し、簡便なカメラキャリブレーションに基づいて、物体の周囲に設置したカメラの得る映像に因子分解法等を適用し、物体の全周復元、すなわち立体モデル化を行う方法を提案する。本法は、ステレオ法と異なり立体器具を使ったカメラキャリブレーションが不要な方法であり、タン等の方法とは違い、光軸を一致させた2対のカメラを相対する位置に設置する必要がない。そのかわり、前面のカメラ群と後面のカメラ群の間（すなわち全カメラ）で、物体上の少なくとも3特徴点が共通に観察される場合に利用できる方法である。

2. 因子分解法

提案法は因子分解法を応用した方法であるので、本節では因子分解法の概要を述べる。

図1に示すように、物体 O をカメラで移動しながら撮影する。撮影位置 f における物体 O 上の特徴点 p の画像面（撮像面 VP_f ）への投影位置を (x_{fp}, y_{fp}) で表す。 L_f はカメラレンズである。 P 個の特徴点($p=1,2,\dots,P$)および F 個の撮影位置($f=1,2,\dots,F$)に対し次の $2F \times P$ 行列 W を定義する。

$$W = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1P} \\ x_{21} & x_{22} & \cdots & x_{2P} \\ \vdots & \vdots & \ddots & \vdots \\ x_{F1} & x_{F2} & \cdots & x_{FP} \\ y_{11} & y_{12} & \cdots & y_{1P} \\ y_{21} & y_{22} & \cdots & y_{2P} \\ \vdots & \vdots & \ddots & \vdots \\ y_{F1} & y_{F2} & \cdots & y_{FP} \end{pmatrix} \equiv \begin{pmatrix} \boxed{x_{fp}} \\ \boxed{y_{fp}} \end{pmatrix} \quad (1)$$

ここで行列 W の fp 成分は x_{fp} 、 $(F+fp)$ 成分は y_{fp} である。 W を式(1)最右辺のように略記する。

物体 O 上の P 個の特徴点の重心 G を世界座標系（3次元座標系）の原点とする。結像系に平行投影を仮定すれば、撮影位置 f におけるカメラ画像上の原点 G の投影点 (x_f, y_f) は次式で与えられる。

$$(x_f, y_f) = \left(\frac{1}{P} \sum_{p=1}^P x_{fp}, \frac{1}{P} \sum_{p=1}^P y_{fp} \right) \quad (2)$$

これは式(1)の各行の平均値として求められる。

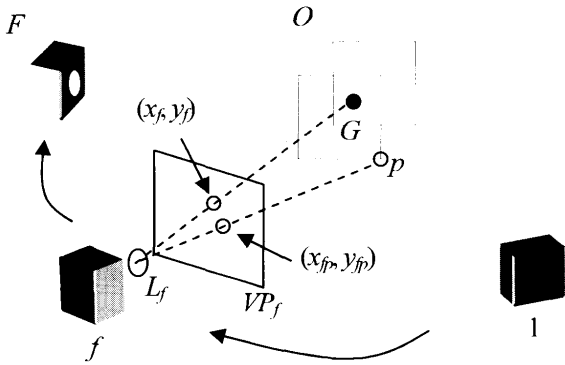


図1. 撮影位置 f におけるカメラレンズ L_f および撮像面 VP_f と物体上の特徴点 p および重心 G の関係。

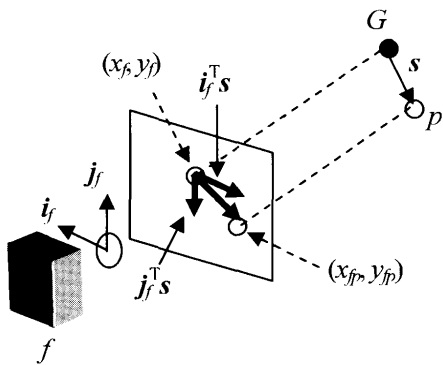


図2. 点 p の位置ベクトル s とその投影像の大きさ。

世界座標系において、原点 G からの点 p の変位 (これを s_p と表せば点 p の3次元空間内の位置は s_p で与えられる) は、画像面上では、原点 G の投影点 (x_f, y_f) からの点 p の投影点 (x_{fp}, y_{fp}) の変位となる。この関係を図2に示す。一方、撮影位置 f におけるカメラレンズ座標系を、 x 軸、 y 軸方向の単位ベクトル i_f, j_f で表せば (光軸方向は $i_f \times j_f$)、次式が成立する。

$$x_{fp} - x_f = i_f^T s_p, \quad y_{fp} - y_f = j_f^T s_p \quad (3)$$

本法では結像系に平行投影を仮定するが、式(3)は一般

にアファインカメラモデルについて成立する[2]。

式(2), (3)を用いて、式(1)より式(4)で与えられる $2F \times P$ 行列 \tilde{W} が得られる。行列 \tilde{W} は、物体上の特徴点の重心を原点とする世界座標系で表記した式(1)の行列 W であり、カメラの方向を与える行列 M と特徴点の3次元座標を与える行列 S の積として表される。

$$\tilde{W} = \begin{bmatrix} x_{fp} - x_f \\ y_{fp} - y_f \end{bmatrix} = \begin{bmatrix} i_f^T s_p \\ j_f^T s_p \end{bmatrix} \quad (4)$$

$$= \begin{bmatrix} i_f^T \\ j_f^T \end{bmatrix} \begin{bmatrix} s_p \end{bmatrix} \equiv MS$$

提案法では式(4)の関係 $\tilde{W} = MS$ を用いる。

3. 撮影システムの構成

提案する立体モデル化法は、複数のカメラを対象物体の周囲 (前面と後面) に置いて物体を撮影する。簡単のためにカメラ対3組の撮影システムを図3に示す。同図で物体 O の前面のカメラ群 C_i^F ($i=1,2,3$) には物体前面の特徴点 P_k^F ($k=1,2,\dots,K^F$) が観察され、物体後面のカメラ群 C_j^R ($j=1,2,3$) には物体後面の特徴点 P_l^R ($l=1,2,\dots,L^R$) が観察されるものとする。特徴点の集合をそれぞれ $S(P^F) \equiv \{P_k^F \mid k=1,2,\dots,K^F\}$, $S(P^R) \equiv \{P_l^R \mid l=1,2,\dots,L^R\}$ と表せば、物体 O の全特徴点の集合 S は $S = S(P^F) \cup S(P^R)$ で与えられる。 $S(P^F)$ と $S(P^R)$ の共通要素を P^{FR}_m と表記すれば、その集合 $S(P^{FR})$ は $S(P^{FR}) = \{P^{FR}_m \mid P^{FR}_m \in S(P^F) \cap S(P^R); m=1,2,\dots,M\}$ で与えられる。点 P^{FR}_m は、前面カメラ群 C_i^F と後面カメラ群 C_j^R で共通に観察される O 上の特徴点である。カメラ C の画像面上に投影される特徴点 P の位置を画像座標 (x, y) で表す。

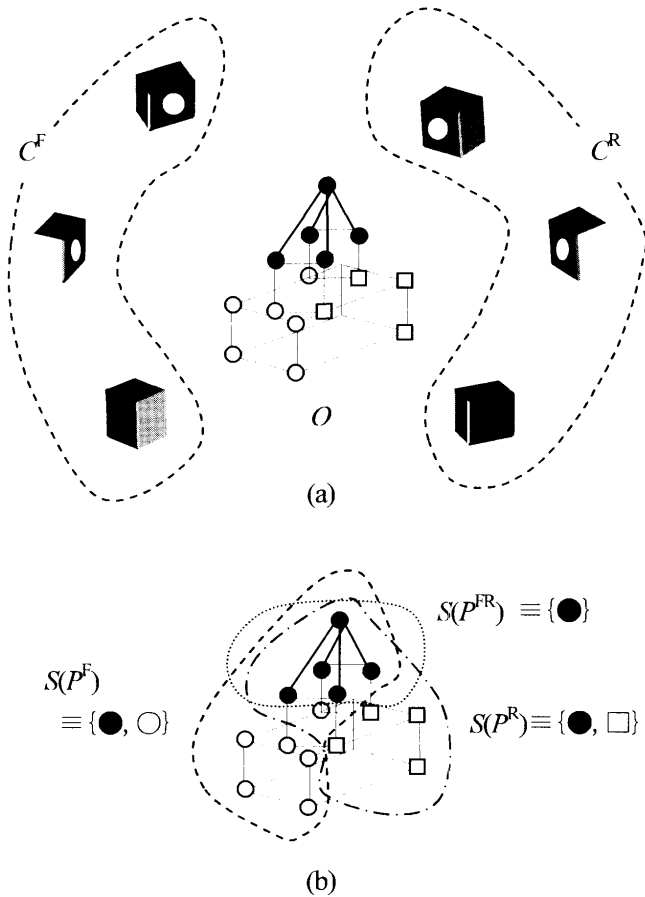


図3. 撮影システム. (a) 物体 O は前面カメラ群 C^F および後面カメラ群 C^R によって撮影される. (b) O 上の特徴点は, 前面カメラ群で観察される特徴点群 $S(P^F)$, 後面カメラ群で観察される特徴点群 $S(P^R)$, 両者に共通に観察される特徴点群 $S(P^{FR})$ に分類される.

4. 全周復元の方法

図3において, 共通特徴点の集合 $S(P^{FR})$ に因子分解法[2]を適用すれば, 次式が得られる.

$$W_{\{\bullet\}} = M_{CFR} S_{\{\bullet\}} \quad (5)$$

ここで $W_{\{\bullet\}}$ は, 集合 $S(P^{FR})$ に含まれる特徴点の画像座標が作る計測行列, $S_{\{\bullet\}}$ はそれらの特徴点の3次元座標を与える行列, M_{CFR} は全カメラ群の方向を与える行列である.

式(5)により全カメラの方向が求められるので, 前方カメラの方向 M_{CF} , および前面カメラ群でのみ観察される特徴点 (図3で○の特徴点) の集合 $S(P^F) \equiv S(P^F) \setminus S(P^{FR})$ を用いれば次式が成立する. (ただし集合

A, B に対し A/B は差集合を表す.)

$$W_{\{\circ\}} = M_{CF} S_{\{\circ\}} \quad (6)$$

ここで $W_{\{\circ\}}$ は, 集合 $S(P^F)$ に含まれる特徴点の画像座標が作る計測行列, $S_{\{\circ\}}$ はそれらの特徴点の3次元座標を与える行列である. 式(6)より

$$S_{\{\circ\}} = (M_{CF}^T M_{CF})^{-1} M_{CF}^T W_{\{\circ\}} \quad (7)$$

が得られるから, 前面カメラ群でのみ観察される特徴点が復元される.

同様に後面カメラの方向 M_{CR} , および後面カメラ群でのみ観察される特徴点 (図3で□の特徴点) の集合 $S(P^R) \equiv S(P^R) \setminus S(P^{FR})$ を用いれば

$$W_{\{\square\}} = M_{CR} S_{\{\square\}} \quad (8)$$

が成立する. ここで $W_{\{\square\}}$ は, 集合 $S(P^R)$ に含まれる特徴点の画像座標が作る計測行列, $S_{\{\square\}}$ はそれらの特徴点の3次元座標を与える行列である. 式(8)より

$$S_{\{\square\}} = (M_{CR}^T M_{CR})^{-1} M_{CR}^T W_{\{\square\}} \quad (9)$$

となるから, 後面カメラ群でのみ観察される特徴点が復元される.

式(5), (7), (9)から得られる行列 $S_{\{\bullet\}}, S_{\{\circ\}}, S_{\{\square\}}$ が, 物体 O 上で観察される全周の特徴点の3次元座標を与えるから, これにより全周復元が実現する.

5. 実験結果

提案法を用いて多面体の全周復元を行った. 2節で述べた因子分解法では, 少なくとも3方向の撮影方向 ($F \geq 3$) が必要である[2]. すなわち因子分解法を適用する場合, カメラを固定するなら少なくとも3台のカメラが必要である. 図4に示すように, 本実験では簡単のために, 複数のカメラを設置する代わりに1台のカメラを設置し, カメラから約2m離れた位置にある, 対象立体を載せたターンテーブルを回転させて立体の画像を取得する. 式(6)が成立するように立体の前面3方向 ($F=3$) の画像を取得し, また式(8)が成立するように後面3方向の画像を取得する. 式(5)は全6方向の画像に対し適用する. カメラの俯角は45度, ターンテーブルは前後面とも, 0度を中心として左右約30度回転して撮影する. ただしこれらの角度の値は, 復元計算には用いない.

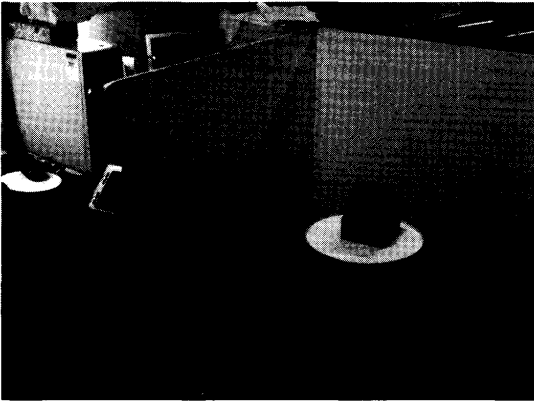


図4. 実験環境.

対象となる多面体の全面, 後面の形状を図5(a)に示す. 辺の長さは短辺5 cm, 長辺10 cm, 各頂点に集まる辺のなす角度は90度である. 特徴点は多面体の各頂点, および各長辺の midpoint に球状の小マーカを張り付けている. 総特徴点数は21点, すなわち $n(S)=21$, このうち共通点は13点, すなわち $n(S^{FR})=13$, 前面カメラ群でのみ観察される特徴点は4点, すなわち $n(S^F)=4$, 後面カメラ群でのみ観察される特徴点は4点, すなわち $n(S^R)=4$, である. ただし, 集合 A の要素数を $n(A)$ で表す.

復元計算の手順は以下のとおりである.

- 1: 図4の実験システムで撮影された多面体の前面3画像, 後面3画像から, 特徴点を, ①全カメラで共通に観察される特徴点, ②前面カメラ群でのみ観察される特徴点, ③後面カメラ群でのみ観察される特徴点に分類する.
- 2: 各特徴点の画像座標を求める.
- 3: ①の特徴点の3次元座標と6台のカメラの方向を式(5)より求める.
- 4: ②および③の特徴点の3次元座標を, それぞれ式(7)および式(9)から求める.

本稿の目的は全周復元法の提案とその有効性の検証であるから, 画像上での特徴点抽出と画像間の特徴点対応付け, および上記手順1の特徴点分類はマニュアルで行っている.

全周の復元結果を図5(b)に示す. ただし復元されるのは特徴点の3次元位置であるが, わかりやすいように面を貼り付けて表示している. 全周形状が復元されている.

復元精度の比較評価のために, 上記の多面体を使っ

た実験と同様の実験をシミュレーションでも行った. すなわち, グラフィックスソフトウェアである3次元ポリゴンモデラー (Metasequoia[©]) を用いて, 上記実験と同形の多面体 (ポリゴンモデル) をコンピュータ内に生成し, 俯角45度の仮想カメラで6方向からそれを撮影したとして得られる映像上の特徴点位置を求め, 提案する全周復元法を適用する. 多面体 (実体) を用いた場合とシミュレーションによる多面体 (ポリゴンモデル) を用いた場合との主な違いは, 後者の場合の方が特徴点の画像座標が丸め誤差の範囲内で正確に求められるという点である.

なお, 前述の復元計算の手順のうち, ステップ3の式(5)を用いた計算では $W_{(i)}$ に特異値分解が適用される. この計算には OpenCV[©] の関数 SVD を使用している. 関数 SVD は 1000 列程度の大きなサイズの行列でも計算可能であるため, 本研究の用途には十分である.

特異値分解の計算時間は, 本実験 (特徴点数13個) では, CPU3.0GHz の PC で 0.16 ms と高速であった.

6. 復元精度の評価

復元精度の評価は辺の長さ (21 辺) と頂点の角度 (42 角) の相対誤差の計算によって行う. 復元後の短辺の長さを x_i ($i=1, 2, \dots, n$; n は短辺の総数), 復元後の長辺の長さを X_j ($j=1, 2, \dots, N$; N は長辺の総数) で表せば, 真値ならば $X_j=2x_i$ である. そこで短辺を基準として平均辺長 \bar{x} を次式で定義する.

$$\bar{x} = \frac{1}{n+2N} \left(\sum_{i=1}^n x_i + \sum_{j=1}^N X_j \right) \quad (10)$$

式(10)を用いて短辺長の相対誤差 E_x (%), 長辺長の相対誤差 E_X (%) をそれぞれ次式で定義する.

$$E_x = \frac{x - \bar{x}}{\bar{x}} \times 100 \quad (11)$$

$$E_X = \frac{X - 2\bar{x}}{2\bar{x}} \times 100 \quad (12)$$

同様に, 多面体の各頂点の角度 θ の相対復元誤差 E_θ (%) を次式で定義する.

$$E_\theta = \frac{\theta - 90}{90} \times 100 \quad (13)$$

式(10)~(13)を用いて, 平均相対誤差を求めた. $n=21$, $N=42$ である. 結果を表1に示す.

A

7.おわりに

本稿では、立体器具を用いず画像ベースでカメラキャリブレーションを行うことによって、簡便に立体の全周形状を復元する方法を提案し、多面体の全周復元を行って提案法の有効性を示した。復元精度は辺長で4.83%、角度で3.58%であり、従来の因子分解法を用いた剛体や変形体の復元結果[5]と同等の精度である。

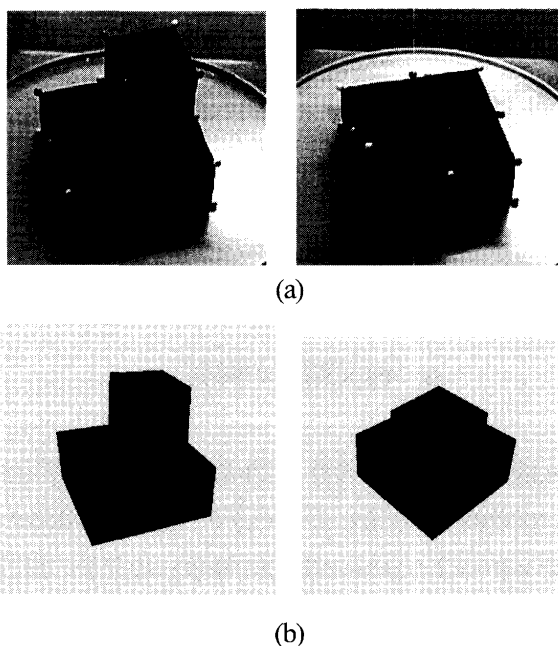


図5. 対象と復元結果. (a) 多面体の前面と後面およびマーカの位置, (b) 全周の復元結果.

表1. 多面体の相対復元誤差.

		多面体	ポリゴンモデル
辺長	平均	4.83 %	2.06 %
	標準偏差	2.78 %	0.83 %
角度	平均	3.58 %	1.57 %
	標準偏差	4.19 %	1.88 %

この値の許容性は応用分野で異なるであろうが、マン・マシン・インタラクティブ・システム等で使われる人モデル（アバター）の生成には十分利用できると思われる。

因子分解をベースとした復元に含まれる主な誤差の要因は、①結像系のアファインカメラモデル近似（線形近似）による誤差と、②特徴点の画像座標抽出時に混入する位置誤差がある。ポリゴンモデルを用いたシミュレーションによる復元では、特徴点の画像座標は計算で求められるから、その誤差は丸め誤差程度と考えられる。よって表1では、多面体の復元誤差は主に①、②によるもの、ポリゴンモデルの復元誤差は主に①によるものと考えられる。よってその差に注目すれば、多面体の復元では、辺長の場合2.77%、角度の場合2.01%が特徴点の画像座標抽出時に混入する位置誤差の影響、残り（辺長の場合2.06%、角度の場合1.57%）が結像系のアファインカメラモデル近似による誤差の影響と考えられる。

提案法は、結像系を線形近似しているため、従来のステレオ法よりも復元精度がやや低い（4~5%程度）、カメラキャリブレーションが簡素化されている（立体器具を用いた事前のキャリブレーションが不要。取得画像を用いたキャリブレーションを行う）点、また立体の周囲に配置するカメラ群も厳密な位置設定が必要ない点で、従来法よりも使いやすい方法である。対象の映像さえ撮影すればよいから、屋内外等、撮影場所も自由度が増える。ただし、立体上の少なくとも3点が全カメラから共通に観察されなければならない（共通点の条件）、これはカメラの俯角を大きめにとれば実現しやすい条件と思われる。無論、共通点が多いほど式(1)から得られるカメラ方向行列 M_{CTR} の精度が上がり、したがって復元精度も向上する。よって撮影時においては、全周をカバーしつつ共通点の多い撮影位置を確保するのがよい。

本実験では、立体（剛体）の全周復元を行ったが、サンプル時刻ごとに本法を反復適用すれば人の運動のような変形体の全周復元[3]も可能である。この検証を次の実験の課題とする。また本実験では、画像上での特徴点抽出と画像間の特徴点対応付け、および共通点等の特徴点分類をマニュアルで行っている。これらのマニュアル操作の自動化を今後検討する。

参考文献

[1] 井上康男, 桜井伸二, 矢部京之輔: “DLT法”, Japanese

- Journal of Sports Science, **10**, 3, 191-195 (1991 年).
- [2] 金出武雄, コンラッドポールマン, 森田俊彦: “因子分解法による物体形状とカメラ運動の復元”, 電子情報通信学会論文誌, **J76-D-II**, **8**, 1497-1505 (1993 年).
- [3] タン ジュークイ, 石川聖二, 有田哲志: “全周の特徴点情報を含む計測行列を用いた 3 次元物体モデルの一括生成法”, 映像情報メディア学会誌, **53**, **3**, 433-436 (1999年3月).
- [4] Chen, Y., Medioni, G: “Object modeling by registration of multiple range images”, Image and Vision Computing, **10**, **3**, 145-155 (1992).
- [5] タン ジュークイ, 石川聖二, 加藤清史: “因子分解を利用したモーションキャプチャ法”, 計測自動制御学会論文集, **36**, **11**, 980-984 (2000年11月).