

部分空間分離に基づく関節構造を持った物体の運動と形状の推定法

松尾 幸治^{†1} 小川原 光一^{†1}
倉爪 亮^{†1} 長谷川 勉^{†1}

本稿では、画像列上の特徴点軌跡からカメラ運動と対象の3次元形状の復元を目的とする運動からの形状復元 (Structure from Motion, SFM) において、対象が関節構造を持つ剛体であり、また全ての特徴点が同時に観測できない場合を扱う。本研究では、部分空間への分離と当てはめに基づき推定を行うが、このとき使用する特徴点の位置および分離の精度が高いことが非常に重要であり、本稿では逆に結果に悪影響を与える特徴点をなるべく除去することによって、安定した形状復元に寄与する手法を紹介する。

Articulated Structure from Motion using Subspace Separation

KOJI MATSUO^{†1} KOICHI OGAWARA^{†1}
RYO KURAZUME^{†1} and TSUTOMU HASEGAWA^{†1}

This paper considers the problem of articulated structure from motion where the target object is made of articulated rigid bodies and all the feature points are not always observed simultaneously. We deal with this problem by subspace separation method and it is quite important to use feature points whose accuracy in position and separation is high. In this study, we propose a method that rejects wrong feature points as much as possible so as to estimate structure and motion stably with the remaining good feature points.

^{†1} 九州大学
Kyushu University

1. はじめに

本稿では、画像列上の特徴点軌跡からカメラの運動と対象の3次元形状の復元を目的とする運動からの形状復元 (Structure from Motion, SFM) において、特に対象が関節構造を持つ剛体であり、また全ての特徴点が同時に観測できない場合を扱う。対象が単一の剛体でかつカメラがアフィンカメラモデルで近似できる場合には、この問題は画像上の時系列特徴点軌跡を表す縦ベクトルを特徴点の数だけ横に並べた行列 (観測行列, W) を、カメラの運動を表す行列 (運動行列, M) と対象の3次元形状を表す行列 (形状行列, S) の積に分解できることが知られており^{1),2)}、また運動行列および形状行列のランクが制限されることから、この問題は部分空間に対する当てはめ問題と等価であることが知られている³⁾。

SFM に関する研究はこの20年の間に非常に進歩を遂げ、対象も複数物体^{4),5)}、関節物体^{6),7)}、非剛体運動⁸⁾などが扱われる一方で、ある時点で不可視な特徴点に起因する欠損のある観測行列から精度よく運動と形状を推定するために、閉形式解^{9),10)}、交互推定法¹¹⁾、非線形最小化¹²⁾など多くの方法が提案されてきた。

これらの手法の性能は、観測された特徴点の位置の精度、および対象が複数の場合には分離の精度に大きく依存する。そのため、処理の自動化のためには高性能な特徴点の検出・追跡アルゴリズムが必要になるとともに、結果に悪影響を与える特徴点を除去することが重要になる。本研究では、部分空間への分離と当てはめに基づき繰り返し法によって物体の分離と運動・形状の復元を行うが、その過程で精度低下の要因となりうる特徴点を積極的に除去することによって、最終的な復元精度の向上を図る。

本稿では、復元性能低下の要因となりうる例として、(1)異なる物体を表す複数の部分空間に同程度に当てはまりのよい特徴点軌跡、および(2)特徴点の誤追跡によって一貫性がなくなった特徴点軌跡の2つの場合を考える。計算の過程でこれらが見つかった場合には適宜削除することによって、推定性能の向上に寄与する特徴点のみが残ることを意図する。本来有用な特徴点が除去される可能性もあるが、復元自体は後で述べるように再投影誤差の二乗和の最小化問題として定式化され誤分離や誤追跡は推定性能の大幅な低下を招くため、可能性のあるものはなるべく除去する方針を取る。

2. 関連研究

対象の幾何モデルを持たずに関節構造の推定を行う研究には、複数のカメラを使う例¹³⁾もあるが、単眼の場合には観測行列のランクによって判断することができる。一般の剛体の

運動は観測行列のランクが最大で4となり、 N 個の物体が同時に運動する場合には観測行列全体のランクは最大で $4N$ となる。関節物体の場合は一方の物体がもう片方の物体の運動を拘束するため、例えば球関節の場合はランクが1つ下がり、蝶番関節の場合はランクが2つ下がる。

Tresadernらはアフィンカメラモデルを仮定してランク推定に基づき関節構造を推定する方法を提案し⁶⁾、Yanらは透視カメラを仮定してやはりランク推定に基づき関節構造を推定する方法を提案した⁷⁾。Yanらの手法では、Local Space Affinity (LSA) と呼ばれる局所的な少数の特徴点群から構成される部分空間同士の類似度を用いて異なる物体に属する特徴点の分離を行い、隣接する2つの特徴点群全体のランクを推定することによって、関節構造の種類と関節中心位置を推定する。本稿でも、透視カメラを仮定し、物体の分離と関節構造の推定に関してはYanらの手法を用いる。

透視カメラモデルを仮定した運動と形状の復元については自己校正法^{14),15)}が知られており、本稿でもこれを使用する。

欠損の補完方法については、前説で述べたように様々な方法が提案されているが、そのうち交互推定法の1種であるPower Factorization¹¹⁾は欠損が無い場合に非常に高速かつ安定に2つの行列の積に分解することが可能であり、8割以上の要素が欠損している場合でも分解することができたという報告もある。またVidalらは、Power Factorizationによって5次元の共通部分空間に特徴点軌跡ベクトルを投影することによって物体の分離を行った¹⁶⁾。本稿でも、欠損した観測行列から分離のための共通の部分空間を生成し、また分離後の運動と形状の推定にも自己校正法とともにPower Factorizationを使用する。

3. 提案手法

本研究では、1台のカメラの前で未知の関節構造を有する物体が自由に運動する(典型的には人が物体を操作する)様子を撮影し、これから対象の関節構造および運動と形状を同時に復元することを目的とする。

図1に提案手法の概要を示す。基本的には、前章で述べたように従来法を組み合わせるが、本研究の主眼は、その際になるべく復元性能低下の要因となりうる特徴点を削除する点にある。

3.1 画像列上における特徴点追跡

任意の特徴点検出アルゴリズムと特徴点追跡アルゴリズムを用いて、時系列画像全体にわたって特徴点の追跡を行う。このとき、画像列の途中で現れたり消えたりする特徴点も等し

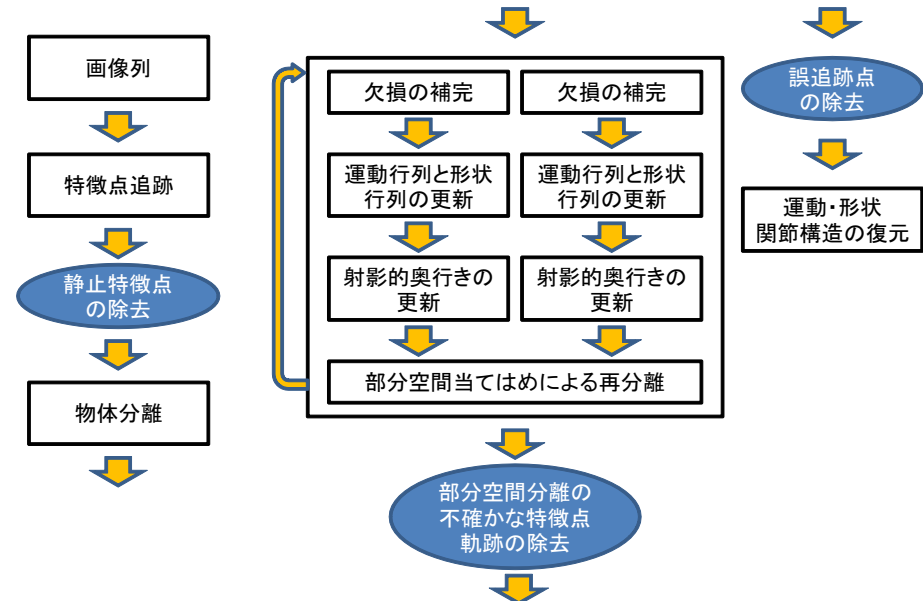


図1 手法の概要

く追跡する。

ただし、特徴点が現れる期間が短いものや、検出されてもほとんど移動しないものは、物体分離を行う際に誤分離される可能性が高いためこの時点で削除する。

3.2 物体分離

欠損のある観測行列 W の非欠損項 $W_{i,j}$ について、式(1)を最小化する行列積 MS^T を次に示すPower Factorizationによって求める。

$$\sum_{i,j} |W_{i,j} - (MS^T)_{i,j}|^2. \quad (1)$$

ただし、 M は $3F \times r$ の行列で F は画像列の数を表し、 S は $r \times N$ の行列で N は特徴点の数を表す。 r は手順によって使い分け、物体分離の場合は5に、運動・形状推定の場合は4にする。 W は $3F \times N$ の行列となり、欠損の無い要素には特徴点の投影点の同次座標が格納されている。

各行列の求め方は、まずいずれかの行列(例えば S_1)を乱数で初期化し、以下のステップ

を収束するまで繰り返す．

$$M_k = (S_{k-1}^T S_{k-1})^{-1} S_{k-1}^T W$$

orthonormalize M_k

$$S_k = (M_k^T M_k)^{-1} M_k^T W \quad (2)$$

次に，Local Space Affinity (LSA)⁷⁾ を用いて，上で5次元単位球に投影された各特徴点軌跡ベクトルについて，近傍の特徴点軌跡ベクトルを集めて局所的な部分空間を構成し，部分空間同士の類似度を正準角によって表現して，スペクトラルクラスタリング¹⁷⁾ によって物体分離を行う．

3.3 運動と形状の復元

Power Factorization¹¹⁾ と自己校正法^{14),15)} を併用し，以下の手順を収束するまで繰り返すことによって運動と形状の復元を行う．

3.3.1 欠損の補完

カメラが長時間運動し画像列が長くなる場合には，観測行列中で欠損項の割合が高くなることもあり，この状態で以下の手順に従い運動・形状・射影的奥行きを同時に求めると，収束に非常に時間がかかるか，もしくは解が不安定になりやすい．

そこで，正則化項として以下の手順により欠損部分を補完することにより，以降の推定の安定化を図る．

欠損補完の手順を図2に示す．図2(a)において，1と書いてある要素は非欠損項であり，それ以外の欠損項を補完したい．各非欠損項は同次座標 $x_{i,j} = (x, y, z)^T_{i,j}$ による表現であり，3次元座標は特徴点 j ごとに決まるスケールファクタ sc_j と，後の節で求める射影的奥行き $z_{i,j}$ をかけ以下のように表現できる．

$$X_{i,j} = sc_j z_{i,j} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{i,j} \quad (3)$$

スケールファクタ sc_j は， $X_{i,j}$ の z 座標が $(MS)_{i,j}$ の z 座標とそうように決定する．このとき， $X_{i,j}$ は以下のように，時刻 i の j を除いた4つ以上の特徴点 $X_{i,j'} (j' \neq j)$ の重み付き和として以下のように表現できる．

$$X_{i,j} = \sum_{j'} w_{i,j'} X_{i,j'} \quad (4)$$

ここで，3次元空間において対象が剛体運動をする限り，この重みは固定値であると考え

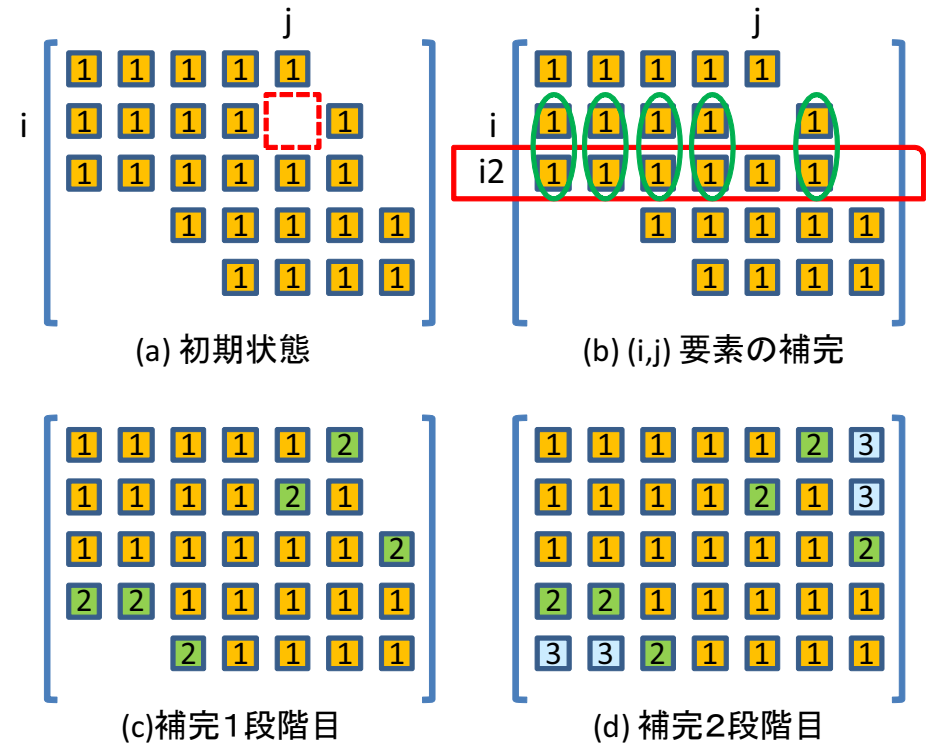


図2 段階的な欠損項の補完

られる．そこで，図2(b)のように，時刻 i 以外で，特徴点 j 以外のどちらも非欠損である特徴点の数が最も多い時刻 (この場合は $i2$) を探索し，以下の式に従って $X_{i2,j}$ の重みを求め，これを使って $X_{i,j}$ を内挿する．

$$X_{i2,j} = \begin{bmatrix} X_{i2,j1} & \cdots & X_{i2,jk} \end{bmatrix} w_{i2} = A w_{i2}, \quad (5)$$

$$w_{i2} = A^T (A A^T)^{-1} X_{i2,j}.$$

$$X_{i,j} = \begin{bmatrix} X_{i,j1} & \cdots & X_{i,jk} \end{bmatrix} w_{i2}.$$

これを，図2(c)に示すように，1以下の項のみを使用してできるだけ補完する．補完できる項が無くなれば，図2(d)に示すように2以下の項のみを使用して順に補完を行う．

3.3.2 運動行列と形状行列の更新

物体分離の節で説明したのと同様 Power Factorization によって運動行列と形状行列を更新する。ただし、前節で補完した3次元座標は、段階が進むほど近似精度が悪くなるため、以下の重み付き方程式を解いて更新する。

$$M_k = (S_{k-1}^T C^{-2} S_{k-1})^{-1} S_{k-1}^T C^{-1} W$$

orthonormalize M_k (6)

$$S_k = (M_k^T C^{-2} M_k)^{-1} M_k^T C^{-1} W$$

ただし、 C は図2で表示されている番号を対角成分に持つ行列である。

3.3.3 射影的奥行き更新

特徴点軌跡ベクトルが運動行列 M の4次元部分空間に最もよく当てはまるように、特徴点ごとの射影的奥行き $z_{i,j}$ を求める¹⁴⁾。

3.4 再物体分離

正規化された各特徴点軌跡ベクトル p_j が、各物体 s ごとに運動行列 M_s によって表現される4次元部分空間のどれと最も当てはまりがよいのかを以下の式によって決める。

$$s^* = \operatorname{argmax}_s \sum_{k=1}^4 (p_j, u_{s,k})^2. \quad (7)$$

ここで、 $(,)$ は内積を表し、 $u_{s,k}$ は4次元部分空間の直交基底ベクトルである。

ただし、最大値と2番目に最大の値の差が閾値以下の場合には、あいまいな特徴点とみなし削除する。

分離結果が安定するまで3.3節に戻って繰り返す。

3.5 誤追跡点の除去

各特徴点軌跡の各要素 $X_{i,j}$ ごとに再投影誤差を計算し、最大再投影誤差が閾値以上の場合は誤追跡とみなしてその特徴点を削除する。

3.6 運動・形状・関節構造の推定

残った特徴点を用い、3.3節と同じ方法で運動と形状を復元し、非線形最適化(バンドル調整)によって微調整を行う。最後に、文献⁷⁾の関節構造推定法を用いて関節構造の推定と2物体のスケール調整を行う。

4. 実験

チェッカーパターンを6面に貼付した2個の直方体が頂点で点関節を構成するように物体

を作成し、これを人間が操作して運動と形状の復元を行った。

カメラは、PointGrey社製のIEEE1394カメラGrasshopper(2M)を用い、解像度640×480で動画撮影した画像を実験に用いた。

また、バンドル調整にはLourakisらによって開発された疎行列に対応したバンドル調整プログラム¹⁸⁾を用いた。

図3(a)に、3.2節で説明した初期物体分離の結果を示す。黄色い数字が左の直方体に属していることを表し、赤い数字が右の直方体に属していることを表す。誤分離が生じており、このまま運動・形状を復元してもよい結果は期待できない。

図3(b)に、提案手法により復元性能低下の要因となる特徴点を削除した結果を示す。青い数字が静止特徴点、紫の数字があいまいな特徴点、水色の数字が誤追跡点を表す。

残った特徴点を利用して3次元復元した結果を図4に示す。

5. まとめ

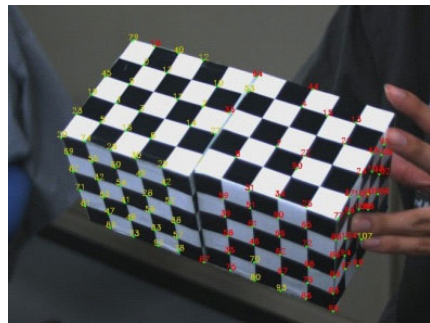
本稿では、関節構造物体を対象に欠損を含む観測行列から運動と3次元形状を推定する手法を提案した。提案手法では、部分空間への当てはめと分離に基づくが、その確度が低い特徴点を積極的に削除することにより、復元性能の向上に寄与できたことを実画像を使用した実験によって確認した。

謝辞

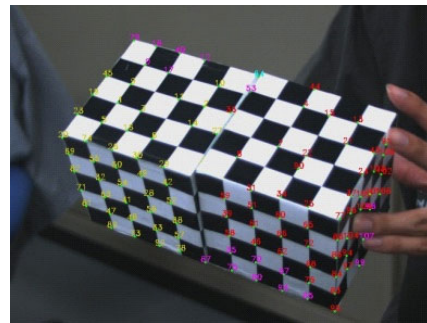
本研究は、科学研究費補助金若手(B)(21700224)および科学技術総合推進費補助金「若手研究者の自立的な研究環境整備促進」の補助を受けている。

参考文献

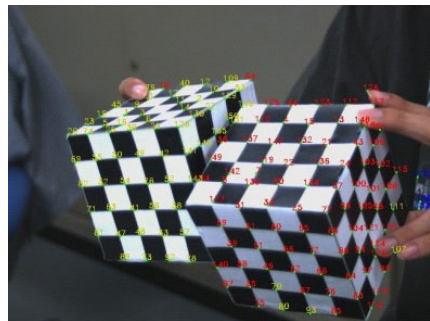
- 1) Tomasi, C. and Kanade, T.: Shape and Motion from Image Streams under Orthography: a Factorization Method, *International Journal of Computer Vision: IJCV*, Vol.9, No.2, pp.137–154 (1992).
- 2) Hartley, R. and Zisserman, A.: *Multiple View Geometry in Computer Vision*, Cambridge University Press (2000).
- 3) Kanatani, K.: Motion Segmentation by Subspace Separation and Model Selection, *Proc. of ICCV*, pp.301–306 (2001).
- 4) Costeira, J. and Kanade, T.: A multi-body factorization method for motion analysis, *Proc. of ICCV*, pp.1071–1076 (1995).
- 5) Rao, S., Tron, R., Ma, Y. and Vidal, R.: Motion Segmentation via Robust Sub-



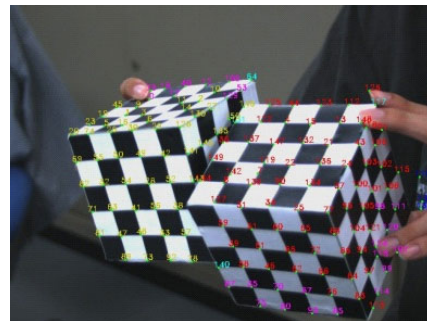
140 frame



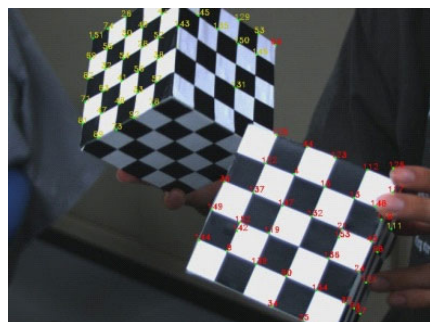
140 frame



170 frame

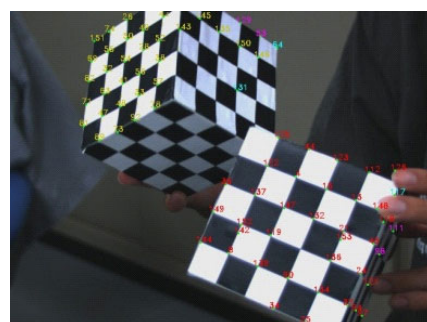


170 frame



200 frame

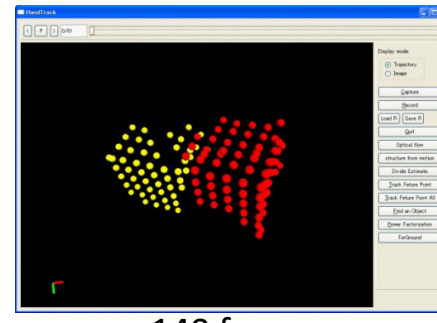
(a) 初期分離結果



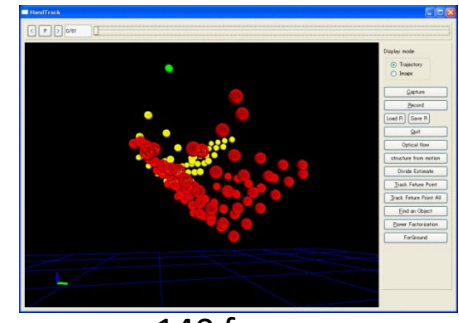
200 frame

(b) 最終分離結果

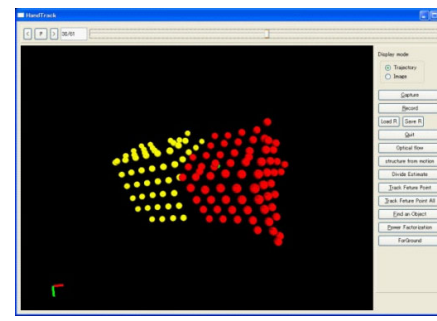
図 3 (a) 初期分離結果, (b) 最終分離結果. 数字の色が赤, 黄色はそれぞれの物体に属する特徴点. 青は静止点と判断し除去した特徴点. 紫色は部分空間分離で区別がつかず除去した点. 水色は誤追跡と判断し除去した特徴点.



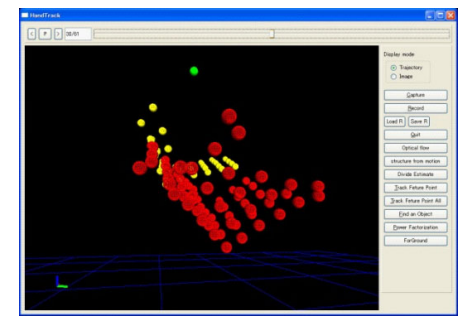
140 frame



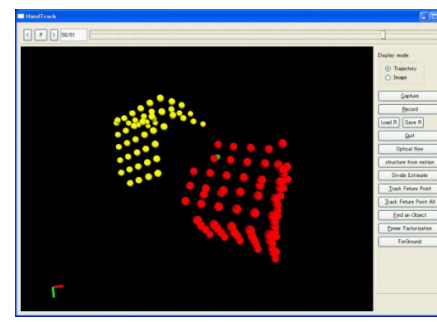
140 frame



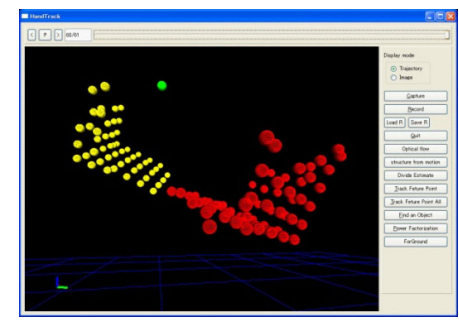
170 frame



170 frame



190 frame



190 frame

(a) 正面から見た復元結果

(b) 横から見た復元結果

図 4 (a) 3次元復元結果をカメラと同じ視点からレンダリングした結果, (b) 3次元結果を横からレンダリングした結果. 緑の点は推定された関節中心位置を表す.

- space Separation in the Presence of Outlying, Incomplete, or Corrupted Trajectories, *IEEE Conference on Computer Vision and Pattern Recognition: CVPR*, pp. 1–8 (2008).
- 6) Tresadern, P. and Reid, I.: Articulated Structure from Motion by Factorization, *IEEE Conference on Computer Vision and Pattern Recognition: CVPR*, pp.20–25 (2005).
 - 7) Yan, J. and Pollefeys, M.: A Factorization-Based Approach for Articulated Non-rigid Shape, Motion and Kinematic Chain Recovery From Video, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.30, No.5, pp.865–877 (2008).
 - 8) Bue, A.D.: A Factorization Approach to Structure from Motion with Shape Priors, *IEEE Conference on Computer Vision and Pattern Recognition: CVPR*, pp. 1–8 (2008).
 - 9) Martinec, D. and Pajdla, T.: Structure from Many Perspective Images with Occlusions, *7th European Conference on Computer Vision: ECCV*, pp.355–369 (2002).
 - 10) Jacobs, D.W.: Linear Fitting with Missing Data for Structure-From-Motion, *Computer Vision and Image Understanding: CVIU*, Vol.82, No.1, pp.57–81 (1997).
 - 11) Hartley, R. and Schaffalitzky, F.: Powerfactorization: 3D reconstruction with missing or uncertain data, *Proc. of Australia–Japan Advanced Workshop on Computer Vision*, pp.1–9 (2003).
 - 12) Buchanan, A. and Fitzgibbon, A.: Damped Newton algorithms for matrix factorization with missing data, *IEEE Conference on Computer Vision and Pattern Recognition: CVPR*, pp.316–322 (2005).
 - 13) G.Cheung, S.Baker, T.K.: Shape-From-Silhouette of Articulated Objects and its Use for Human Body Kinematics Estimation and Motion Capture, *IEEE Conference on Computer Vision and Pattern Recognition: CVPR* (2003).
 - 14) 金谷健一, 森昭延, 菅谷保之: 自己校正法の最新レシピ, 情報処理学会研究報告: コンピュータビジョンとイメージメディア (CVIM), No.2006-CVIM-153-31, pp.199–206 (2006).
 - 15) ハノアッカーマン, 新妻弘崇, 金谷健一: 自己校正法のための射影復元の計算量削減, 情報処理学会研究報告: コンピュータビジョンとイメージメディア (CVIM), No.2007-CVIM-160-11, pp.63–70 (2007).
 - 16) Vidal, R. and Hartley, R.: Motion Segmentation with Missing Data using PowerFactorization and GPCA, *IEEE Conference on Computer Vision and Pattern Recognition: CVPR*, pp.1–6 (2004).
 - 17) Shi, J. and Malik, J.: Normalized Cuts and Image Segmentation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.22, No.8, pp.888–905 (2000).
 - 18) Lourakis, M. and Argyros, A.: The Design and Implementation of a Generic Sparse Bundle Adjustment Software Package Based on the Levenberg-Marquardt Algorithm, Technical Report 340, Institute of Computer Science - FORTH, Heraklion, Crete, Greece (2004). Available from <http://www.ics.forth.gr/~lourakis/sba>.