

大局的幾何学制約を用いた特徴点追跡に基づく 画像列からの3次元モデル生成

保田和隆^{†*} 椋木雅之[†] 浅田尚紀[†]

本論文では、位置・姿勢が未知のカメラを用いて物体の全周囲を撮影した画像列から、3次元モデルを生成する一連の手順を提案する。3次元モデルの生成には、画像間に対応する特徴点を利用する。この場合、相関法のような局所情報のみによる特徴点追跡では、対応誤りが発生する。本手法ではまず、アフィン変換を用いて対応付けた画像全体の位置関係という大局的幾何制約（アフィン制約）を導入し、対応誤りが少ない特徴点追跡を行う。次に、再投影誤差最小化手法を直接適用することにより、対応点の誤差にロバストに形状復元を行う。さらに、復元物体の全体構造に基づく大局的幾何制約（奥行き制約）を用いた特徴点追跡と形状復元を繰り返す。これにより、特徴点の3次元位置を安定して復元でき、テクスチャを付加した高精細な3次元モデルが生成できることを示す。

3D Model Generation from Image Sequences Using Global Geometric Constraint

KAZUTAKA YASUDA,^{†*} MASAYUKI MUKUNOKI[†] and NAOKI ASADA[†]

This paper describes a method for generating a three-dimensional model from an uncalibrated image sequence taken around an object. Our method is based on feature tracking and minimization of re-projection errors. To cope with mis-matches in the result of feature tracking, we introduce two types of global geometric constraints. The one is "affine constraint" which imposes the positional relationship between pixels on the images. The other is "depth constraint" which imposes the three-dimensional structure of the object. First, we use the affine constraint to reconstruct the object roughly and then we refine the feature tracking and shape reconstruction using the depth constraint. Experimental results show that our method can automatically generate accurate three-dimensional models from real image sequences.

1. はじめに

実物体の3次元モデル化は、仮想空間におけるリアルな物体表現として利用されるだけでなく、文化財のデジタルアーカイブなど多様な目的に用いられる重要な技術である¹⁾。最近では、レーザレンジセンサを用いて高精細な3次元モデルを生成することが行われているが、専用の計測装置が必要であり、大きさや表面特性などいくつかの条件を満足する対象に限られるという性質がある。このため従来から、迅速かつ手軽に形状を生成できるシステムが渴望されており、コンピュータビジョン技術の応用が試みられている。その1つとして、多視点画像からカメラと対象物の相対運

動および対象物の3次元形状を復元する画像列からの3次元形状復元（Structure from Motion）を用いた3次元モデル生成が研究されている^{2)~8)}。

画像列からの形状復元を用いた3次元モデル生成では、まず、視点の異なる画像間に対応する点を追跡し、カメラの位置・姿勢と3次元形状の復元を行った後、最終的な3次元モデル生成が行われる。

特徴点追跡においては、Tomasiら⁹⁾がビデオカメラで撮影した視差の微小な連続画像間に対し、輝度勾配を用いた特徴点追跡手法を提案している。また、Zhangら¹⁰⁾や金澤ら¹¹⁾は、視差の大きな画像間に対し、画像情報のみからエピポラ方程式を求め、特徴点のランダム選択と局所対応の投票などにより、特徴点の対応付けを行う手法を提案している。しかし、これらの手法では、物体の見え方の変化や隠蔽、フレームアウトなどにより、大きな特徴点の対応誤りを生じたり、特徴点の追跡結果が得られなかったりする場合がある。

[†] 広島市立大学情報科学部

Faculty of Information Sciences, Hiroshima City University

^{*} 現在、ソニー株式会社

Presently with Sony Co.

最終的な3次元モデルの生成には、マルチベースラインステレオ法¹²⁾や距離推定¹³⁾、多視点解析¹⁴⁾などの手法が用いられる。これらの処理は、カメラ位置・姿勢が正確に求められていることが処理の前提となっているが、一般に、画像列からの形状復元により求められるカメラ位置・姿勢には誤差が含まれている。特に、特徴点の対応誤りに起因する誤差の影響は大きく、3次元モデルを安定して生成することは困難であった。

このように従来の研究では、個々の技術については多くの研究がなされているが、大きな視差を含む数十枚の全周画像列から3次元モデルを生成する一連の処理を構築するという観点での議論は十分なされていない。この一連の処理では、各段階に従来手法を適用することが考えられるが、前提としている条件が異なるため、それを組み合わせただけでは適切に3次元形状が復元できない。そのため、途中段階の処理に含まれる誤差を許容しつつ、新しい制約を追加していくことで段階的にモデル形状を復元する方法を考案する必要がある。

本研究は、物体の全周囲を撮影した画像列から、大局的幾何制約により特徴点追跡を行い、再投影誤差最小化原理に基づき安定に3次元モデルを生成する一連の手法を提案する。本手法では、まず、2画像間の変形をアフィン行列で表せると仮定し、変形した画像における画素の近傍領域に特徴点の対応付けを限定するという大局的幾何制約（アフィン制約）により、初期的な特徴点追跡を行う。この追跡結果をもとに、非線形最適化計算により、特徴点位置とカメラ位置・姿勢を推定する。得られた特徴点位置とカメラ位置・姿勢をもとに、特徴点の存在する奥行き範囲を限定するという大局的幾何制約（奥行き制約）を導入し、限定した範囲内を画像上で探索することにより、より信頼性の高い特徴点追跡を行う。以上の処理を繰り返すことにより、明らかな誤対応を確実に排除し、精度を向上させ、幾何学的に正しい3次元形状復元を行う。最終的に、信頼できる特徴点のみから構成されたテクスチャマッピングを行った3次元モデルを生成する。

2. 再投影誤差最小化による形状復元

形状復元においては、物体は P 個の特徴点で表現される（図1）。各特徴点は、物体とカメラを含む共通の座標系（世界座標系）上にあり、固有の3次元座標 s_p ($p = 1 \dots P$) を持つ。これらの点は最大 F 枚の画像上に投影され観測される。第 f 画像 ($f = 1 \dots F$) 上で観測された、第 p 特徴点の座標を u_{fp} とする。特徴点 p が画像 f で観測されていない場合 u_{fp} は定

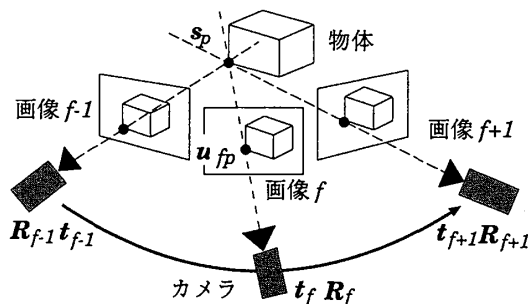


図1 形状復元とカメラ位置・姿勢の推定

Fig.1 Shape Reconstruction.

義されない。

各画像を撮影したカメラの位置・姿勢は、世界座標系における3次元の回転と並進 (R_f, t_f) で表される。投影モデルは透視投影とし、焦点距離は全画像を通じて共通の既知の値 l とする。

これらのパラメータを用いて、形状復元問題は次のように定式化される^{15),16)}。

$$\arg \min_{\mathbf{x}} E(\mathbf{x}) \quad (1)$$

$$E(\mathbf{x}) := \sum_{(f,p)} |P[R_f s_p + t_f] - u_{fp}|^2$$

$$\mathbf{x} := (s_p, \mathbf{q}_f, t_f \text{ を並べたベクトル})$$

$$\text{ただし, } P[x, y, z]^T := \frac{l}{z} [x, y]^T$$

ここで R_f は自由度3の回転行列であり、 \mathbf{q}_f は R_f のパラメータを並べたベクトルである。また、特徴点の画像座標 u_{fp} とその画像間での対応はあらかじめ与えられているものとする。

この式は、推定した形状 s_p を推定したカメラパラメータ R_f, t_f で投影した場合の投影座標 $P[\cdot]$ と、観測値 u_{fp} との誤差（これを再投影誤差とよぶ）の2次ノルムの総和が最小となることを要求するものであり、観測座標に等方性正規分布の誤差を仮定した場合の最尤推定を与える式である。すなわち、正規分布と見なせるような微小な誤差に対しては、ロバストに形状復元が行えることになる。

式(1)は、非線形最適化問題となる。一般に、非線形最適化計算では、繰返し処理により解を逐次更新する反復法が用いられる。計算法としては、Levenberg-Marquardt法などの任意の汎用非線形最適化手法が利用できるが、問題の規模が大きい場合、より計算量の少ない、前処理付き共役勾配法¹⁷⁾が適している。

非線形最適化計算においては、初期値の設定が重要となる。本研究では、物体の周囲を物体とほぼ同じ水平面内で1方向に移動しながら撮影した画像列を用いることを想定する。この撮影方法を、物体を中心にカメラが等速円運動をしながら物体を観測する注視回転

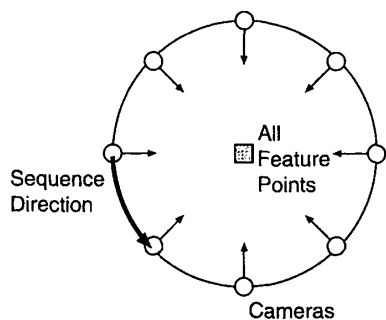


図2 注視回転による初期値
Fig.2 Initialization method.

によりモデル化する (図2)。この場合、初期値を数式で表すと次式ようになる。

$$\begin{cases} R_f = R_{\omega_f}^y & t_f = (0, 0, r)^T \quad \text{for all } f \\ s_p = (0, 0, 0) & \quad \text{for all } p \end{cases} \quad (2)$$

ここで、 $R_{\omega_f}^y$ は、 y 軸 (図2の紙面鉛直方向) 回りの角度 ω_f の回転を表す行列とする。また、 r はカメラの円軌道の半径であるが、計算空間の3次元座標系には物理単位が与えられていないため、 $r = 1$ としてよい。対応点追跡が正しく行われており、物体のほぼ全体が画像内に撮影されている場合、この初期値からほとんどの場合で最適な形状復元が行える¹⁸⁾。

3. 大局的幾何制約の導入

再投影誤差最小化による形状復元は、正規分布と見なせるような微小な誤差に対してはロバストであるが、統計的に外れ値と見なされる大きな特徴点の誤対応には対処できない。

一般に誤対応の除去には、Fischlerら¹⁹⁾が提案したRANSACやRousseeuw²⁰⁾が提案した最小メジアン法(LMedS推定)が用いられる。これらの手法では、選択する対応候補の中に十分な数の正しい対応が存在しなければならない。しかし、視差の大きな画像間や、類似したテクスチャの多い画像が対象の場合は、特徴点の対応付け結果の正解率が下がり、これらの手法で推定可能な精度を満たさない。

一方、画像間で対応するエピポーラ線を求め、特徴点の探索範囲をエピポーラ線上に限定することで、誤対応を軽減する手法も用いられている。しかし、画像列からの形状復元により推定されたカメラ位置・姿勢を利用する場合、推定誤差が含まれるため、必ずしも対応点がエピポーラ線上に存在しているとは限らない。これに対処するには、エピポーラ線の近傍に対しても対応点探索の範囲を拡大する必要がある^{21),22)}が、探索範囲を広げすぎると誤対応が生じる可能性が高まる。

本研究では、数十枚の画像列を処理するため、明ら

かな誤対応を確実に排除する必要がある。そのために、2種類の大局的幾何制約を導入する。

特徴点追跡の最初の段階では、利用できるのは画像情報のみであるため、まず画像そのものを直接対応させることで大局的制約を与える。すなわち、2画像間の見え方の変化を幾何学的変換により近似し、画像全体が最もよく対応付けられる変換を求める。画像内の各画素は、この幾何学的変換により変換された位置の近傍に対応点が存在すると仮定し、特徴点の対応付けをその範囲に限定する。画像対応のための幾何学的変換は、視差による見え方の変化をある程度表現できる自由度が必要であるが、過度の自由度は計算が不安定となる欠点がある。そこで、本研究では自由度6のアフィン変換を用いる。アフィン変換を用いた画像間の対応付けに基づく幾何制約をアフィン制約とよぶ。

アフィン制約により、特徴点対応における大きな外れ値は除外できるが、アフィン変換は画像間の変形を完全には表現できないため、検出できる特徴点数が少ないうえ、多少の誤差も存在する。より精度良く3次元形状を復元するために、特徴点追跡結果に基づき、いったん形状復元を行う。得られた3次元形状とカメラ位置・姿勢を用いると、画像内の各点における物体までの奥行きを求めることができる。この結果は誤差を含むが、一定の範囲を考慮することにより、誤差を許容しつつ特徴点の対応範囲を限定することができる。この制約を奥行き制約とよぶ。

RANSACやLMedS推定を用いた方法やエピポーラ線上の特徴点探索は、数点の特徴点の局所的な対応付け結果に基づいている。これに対して、本手法では画像全体あるいは物体の3次元構造全体をさらに利用することにより、適切に特徴点の対応範囲を限定することができ、より安定して形状復元を行うことが可能である。金澤ら²³⁾は、検出した特徴点に対して、局所相関、対応点のフロー、射影変換の一致度を確信度として逐次的に追加しながら、大域的な整合がとれるよう対応付けを更新し、最終的にエピポーラ拘束により対応付けを行う手法を提案している。一方、本手法は、アフィン制約を用いた初期復元の結果を利用し、形状復元と対応点探索を交互に行いながら全体を最適化することにより、モデル形状に含まれる誤りを除去する方法となっている。

4. 3次元モデル生成

4.1 処理の流れ

3次元モデル生成の処理の流れを、図3に示す。物体の全周囲を20~40枚程度で撮影した画像列が与え

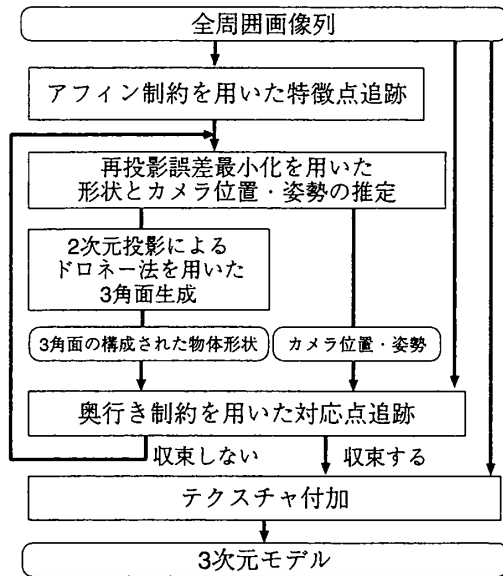


図3 提案手法の処理の流れ

Fig. 3 Flow diagram of the proposed method.

られる。撮影したカメラの焦点距離は既知とする。まず、アフィン制約により、特徴点の追跡を行う。次に、追跡結果から再投影誤差最小化により形状とカメラ位置・姿勢を復元する。この際、復元精度を向上させるために、誤対応と見なせる点の除去を行う。得られた復元結果を利用し、奥行き制約に基づき特徴点追跡を行う。この際、画像の各点の奥行きを求めるために、点群で表現された物体形状に、ドロネー法により3角面を生成する。奥行き制約を更新しながら、形状復元と特徴点追跡を繰り返すことにより、段階的に3次元形状の精度を向上させていく。特徴点追跡において、追跡される特徴点数がほぼ一定に収束した段階で、得られた3次元形状に対してテクスチャ画像を選択、付加することにより3次元モデルを生成する。

4.2 アフィン制約を用いた特徴点の対応付け アフィン変換行列

$$A = \begin{bmatrix} a & b & t_u \\ c & d & t_v \end{bmatrix}$$

を用いると、画像 I 上の座標 (u, v) と画像 I' 上の座標 (u', v') の対応は、次式により表される。

$$\begin{bmatrix} u' & v' \end{bmatrix}^T = A \begin{bmatrix} u & v & 1 \end{bmatrix}^T \quad (3)$$

2 画像間の最適アフィン変換行列 A は、対応する画素値の差の2乗和が最小となるように次式によって求める。

$$\arg \min_A \sum_{(u,v)} |I(u, v) - I'(u', v')|^2 \quad (4)$$

ここで、 $I(u, v)$ 、 $I'(u', v')$ はそれぞれ、画像 I 、 I' の座標 (u, v) 、 (u', v') での画素値とする。上記の計

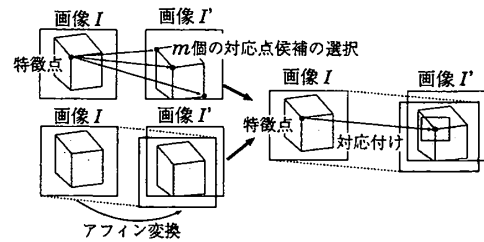


図4 アフィン制約を用いた特徴点の対応付け

Fig. 4 Feature point matching with affine transform constraint.

算は、まず並進成分 $[t_x, t_y]^T$ の最適値を1画素ごとの全探索により求め、次にその値を初期値として、6個すべてのパラメータを勾配法により求める2段階の探索により行った。なお、局所解を避けるため画像を平滑化し低周波成分によるマッチングを行う。

実際の手順は、下記のとおりである(図4)。

- (1) 画像 I 、 I' の両方において、Harris らの手法²⁴⁾、 \star を用いて特徴点を検出する。
- (2) アフィン変換前の画像において、画像 I の1つの特徴点 P に対して、画像 I' の特徴点の中で、周囲画素同士の画素値の残差2乗和(SSD)を評価基準として、最も類似しているものから順に m 個の点 $Q_1 \sim Q_m$ を選択しておく。
- (3) 2画像間で求めたアフィン変換行列を用いて、画像 I の特徴点 P を画像 I' 上に変換する。変換した点を P' とする。
- (4) $Q_1 \sim Q_m$ のうち、 P' の近傍領域に含まれるものがあれば、それを対応点とする。複数あれば、その中でSSDの評価基準により最も類似している点と対応付ける。近傍領域に含まれる点がない場合は、点 P の対応点はないものとし、特徴点から除外する。

以上の(1)~(4)の処理を、画像 I のすべての特徴点に対して順に行う。この際、画像 I の特徴点 P は、対応する特徴点の第1候補 Q_1 のSSDによる評価値が最も良い(類似している)ものから順に選択する。また、特徴点 P と対応付けられた特徴点 Q_{i^*} ($i^* \in 1 \dots m$) は、以降の対応付けからは除外する。この手法は、従来から行われている貪欲法に、アフィン制約を追加したものとなっている。これにより、大域的な幾何制約を考慮した特徴点追跡を行う。

4.3 再投影誤差最小化による形状復元

前節の処理によって得られた画像列全体の特徴点対応を用いて、物体の形状復元と各画像を撮影したカメ

\star <http://www.img.tutkie.tut.ac.jp/programs/index.html> のプログラムを利用した。

ラの位置・姿勢パラメータ推定を行う。これは「運動からの形状復元」問題として多くの研究がなされているが、本研究では文献 17) の再投影誤差最小化原理に基づく方法を用いて物体形状とカメラ位置・姿勢の同時推定を行う。

ただし、画像情報のみを利用して対応付けた特徴点は、一般に追跡誤りを含んでいるため、誤った特徴点に対処する必要がある。そのために、ロバスト推定の一形態である LMedS 推定や M 推定²⁵⁾ が用いられることがあるが、非線形最適化では収束性が悪くなる問題や計算の負荷が大きくなる問題がともなう。このため、まず誤差を含む特徴点の追跡結果に対して非線形最適化を利用して、特徴点の 3 次元座標とカメラ位置・姿勢を推定する。次に、特徴点の追跡誤差によるカメラ位置・姿勢の復元誤りを補正するため、復元後に 3 次元上の特徴点を画像上へ再投影し、2 乗誤差を計算する。この評価尺度において誤差が大きな特徴点を除去する。残りの特徴点に対して再度式 (1) を適用して形状復元を行うことにより、特徴点の数は減るが、より信頼性の高い復元結果を得る。

4.4 奥行き制約による形状復元の改善

前章で得られたカメラ位置・姿勢と物体構造を利用して、奥行き制約を用いてより信頼性の高い特徴点追跡を行う。そのためにまず、点群として得られた物体形状に 3 角面を生成する。

3 次元上の特徴点データを円筒面に投影し、投影した平面上で 2 次元ドロネー法を用いて 3 角面を構成する。なお、円筒面に投影するため、復元物体の重心を世界座標の原点に、上下方向を z 軸として定めておく。また、円筒面座標 (R, θ) 上に投影した際、適切な 3 角面を構成するために、各軸は正規化されているとする。

まず、3 次元空間上の物体を形成する i 番目の特徴点 (x_i, y_i, z_i) を、円筒面座標 (R_i, θ_i) で表す。

$$\theta_i = \tan^{-1} \frac{y_i}{x_i} \quad R_i = z_i \quad (5)$$

ここで、 θ_i は $[-\pi, \pi]$ の値をとるが、3 次元空間では $\theta_i = -\pi$ と $\theta_i = \pi$ は同じ位置を表しているため、 $\theta = -\pi$ 付近に投影された特徴点を $\theta = \pi$ 付近に、 $\theta = \pi$ 付近に投影された特徴点を $\theta = -\pi$ 付近にも投影する。次にこの平面においてドロネー法により 3 角面を作成する。この際、 $\theta_i = -\pi, \pi$ 付近の重複した 3 角面は削除する。以上の処理により構成された特徴点の接続関係を、もとの 3 次元空間での点群にあてはめることにより 3 角面の構成を行う。

3 角面が生成されると、各画像上の画素の奥行きは、

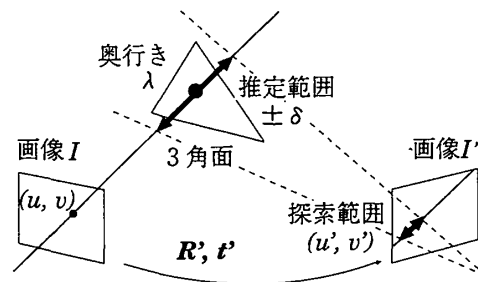


図 5 奥行き制約を用いた特徴点の対応点探索
Fig. 5 Feature tracking by depth constraint.

図 5 に示すようにカメラ位置とその画素を通る直線が 3 角面と交差する奥行き λ の周辺に位置すると考えられる。このとき、画像 I の座標 (u, v) に対応する画像 I' の探索位置 (u', v') は、次式の範囲に制限できる。

$$\begin{bmatrix} u' \\ v' \end{bmatrix} = \mathcal{P} \left[R' \hat{\lambda} \begin{pmatrix} u/l \\ v/l \\ 1 \end{pmatrix} + t' \right] + \begin{bmatrix} 0 \\ \alpha \end{bmatrix} \quad (6)$$

ここで $\hat{\lambda} \in [\lambda - \delta, \lambda + \delta]$

ここで、 l は焦点距離、 R' 、 t' は画像 I に対する画像 I' のカメラ位置・姿勢を表す。また、 δ は λ を中心とする探索する奥行き範囲、 α は誤差補正のための垂直軸方向への探索範囲である*。 $\mathcal{P}[\cdot]$ は透視投影による投影座標を表す。

画像 I 上で特徴点検出を行い、式 (6) で求められる図 5 に示す画像 I' 上の探索範囲を SSD により対応点探索する。これを全画像間で行い、得られた特徴点の追跡結果に対して、4.3 節の処理により形状復元する。以上の処理は、追跡される特徴点数が収束するまで繰り返す。

なお、本手法では、対象物は、特徴点を円筒面に投影したときに投影方向での奥行きの重なりがないような凸形状をしていると仮定している。また、円筒面に投影しているため、物体の上面、下面では 3 角面が大きく歪んで投影され、ドロネー法で適切な面が生成できない場合がある。本研究では、物体とほぼ同じ水平面内から撮影を行っているため、物体の上面、下面は、もともとほとんど観測されていないため、この点については、特に対処していない。

4.5 テクスチャの付加

特徴点数が収束した段階で、最終的に得られた形状復元結果に対しテクスチャを付加する。まず、構成された各 3 角面を画像列中の各画像上に投影し、可視

* α は、本来エッジ極面に垂直な方向の対応誤差を表すパラメータであるが、本研究ではカメラが水平移動し、光軸回りの回転がないため、式 (6) を用いた。

判定を行う。次に可視である画像の中で投影結果が最大面積となる画像を、3角面の最適なテクスチャ画像として選択する。この際、同一平面上にあると判定された点群からなる3角面には、同一の画像をテクスチャ画像として用いることにより、違和感のないテクスチャマッピングを行う。

5. 実験と評価

5.1 シミュレーション実験

仮想的な空間に、1辺10 unitと5 unitのテクスチャのある立方体を置き、図6に示す環境を設定した。ここで、1 unitは、仮想的な空間における長さの基準単位とする。全周囲を10度刻みに撮影し、36枚のシミュレーション画像(720×480画素)を得て実験を行った。

アフィン制約を用いて特徴点追跡を行い、形状復元した結果を図7に示す。この実験では、各画像から特徴点を300点検出し、アフィン変換で対応を求めた画素から15画素以内の範囲で特徴点对応付けを行った。中央の密集した点が3次元空間中の特徴点を表し、周囲の円弧上の点がカメラの位置、カメラ位置から中央の特徴点群に向かう線がカメラ光軸を表している。

図7では、特徴点位置は一部不正確であるが、全体の大まかな形状は、復元されている。また、カメラ位置は円弧状になっており、安定に復元できている。

次に、復元された物体形状とカメラの位置・姿勢を用いて、奥行き制約により特徴点追跡を行う。奥行き制約を用いない一意に指定した奥行きの移動範囲と、

奥行き制約を用いた場合の奥行きの移動範囲を画像上に投影した領域を図8を示す。奥行き制約を用いない手法では、物体すべてを含むように、奥行き λ' を26.5~38.4 unitとした。画像上では101画素程度の探索範囲となる。また、本手法では各点ごとの推定された奥行き λ に対し、経験的に $\delta = \pm 1.52$ unitの範囲(画像上では14画素程度)とした。なお、カメラ位置・姿勢の誤差を軽減させるため、式(6)中の α を ± 5 画素として、画像の垂直軸方向へも探索を行っている。

以上の設定の下で、両手法で特徴点の対応点探索を行う。これより得られた特徴点の3次元座標を、図9に示す。

図9(a)が奥行き制約を用いない場合の結果であり、図9(b)が奥行き制約を用いた場合の結果である。なお、特徴点は一方の画像のみから3,000点検出し、もう一方の画像では各特徴点の探索範囲内で対応点探索を行う。奥行き制約を用いない図9(a)は、物体の前に点群が存在している。これは誤った特徴点对応付けられたためである。本手法では探索範囲が各特徴点ごとに制限されているため、誤対応が軽減され、このような誤りが存在しない。以上により、エピポーラ線上の探索において、探索範囲を限定した対応点探索が有効であることが確認できる。

奥行き制約により得られた特徴点の追跡結果を利用した形状復元結果を図10に示す。図7と比較すると、カメラの位置や特徴点の3次元座標がより安定して推

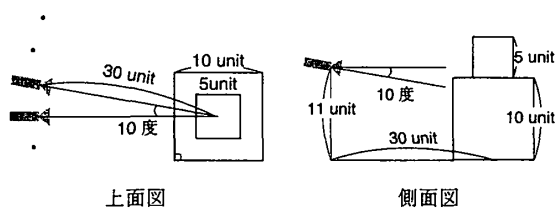


図6 物体形状とカメラの位置関係
Fig. 6 Configuration of synthetic data.

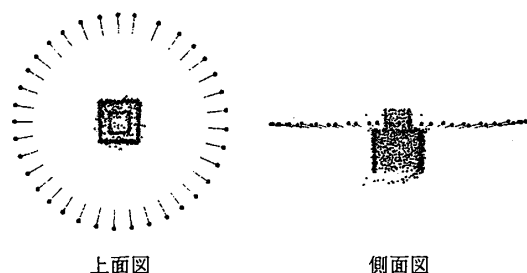
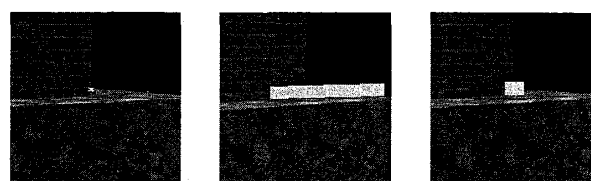
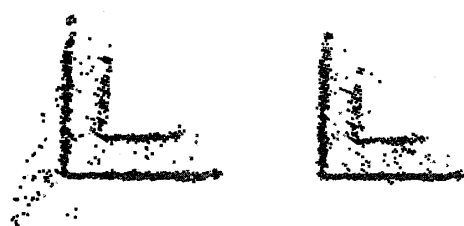


図7 アフィン制約を用いた初期復元結果
Fig. 7 Initial reconstruction result using affine constraint.



第13番目画像 第14番目画像
特徴点検出位置 奥行き制約なし 奥行き制約あり

図8 探索範囲の比較
Fig. 8 Search area.



(a) 奥行き制約なし (b) 奥行き制約あり
図9 ステレオ法により求めた特徴点の3次元座標
Fig. 9 Result of shape reconstruction.

表 1 シミュレーション画像による形状復元精度
Table 1 Accuracy of shape reconstruction for synthetic data.

	特徴点数	誤差 2 乗和	カメラ位置誤差	平均角度	距離誤差範囲
I 段階	1,436 点	0.967 画素	0.815 unit	86.3 度	-0.78~1.34 unit
II 段階	9,986 点	0.337 画素	0.126 unit	90.0 度	-0.35~0.37 unit
最終結果	11,277 点	0.332 画素	0.137 unit	90.0 度	-0.32~0.37 unit

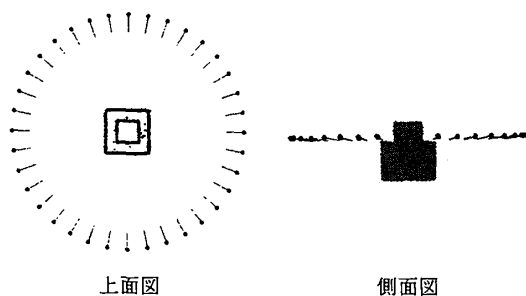


図 10 奥行き制約を用いた最終的な形状復元結果
Fig. 10 Reconstruction result using depth constraint.

定されている。

表 1 および図 11 に、アフィン制約を用いた初期形状復元結果 (I 段階) と、奥行き制約を 1 回適用した形状復元結果 (II 段階), 奥行き制約を 2 回適用して得られた最終結果を示す。さらに、復元された 3 次元モデルの精度評価を行うために、3 次元モデルを構成している点群から 1 平面上に存在する点群のみ抜き出し、点群にあてはめた平面の法線ベクトルを求めることで、各面間の平均角度および平面と特徴点との距離誤差の評価を行った結果も表 1 および図 12 に示した。ここで、物体が存在する仮想的な空間と、形状復元結果の空間では、そのままでは、位置およびスケールが一致していない。これを合わせるために、復元されたカメラ位置と正しいカメラ位置との誤差が最小となる 3 次元アフィン変換を求め変換を行っている。

I 段階目と比較して、II 段階目と最終結果では、特徴点数が大きく増えているにもかかわらず、誤差 2 乗和、カメラ位置誤差、直交性の精度が向上しており、高精度に点群が復元できていることが分かる。

特徴点と平面との位置誤差も、I 段階目では、-0.78~1.34 unit となっていたものが、II 段階目と最終結果では、より狭い範囲となっており、誤差が小さな特徴点の頻度も増えたことから、精度が向上したことが分かる。これは、図 11 において、I 段階目では物体から外れた点が現れているのに対して、II 段階目以降ではそのような点がなくなっていることから確認できる。

なお、奥行き推定において経験的に決定した値 $\delta = \pm 1.52$ unit は、I 段階目における位置誤差範囲よりも大きな値となっており、妥当な値だったと考えられる。

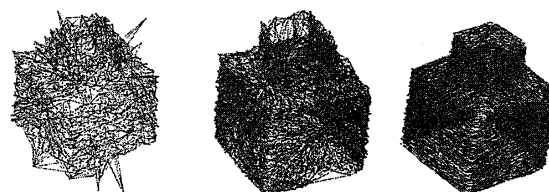


図 11 各段階の形状復元結果
Fig. 11 Result of shape reconstruction in each step.

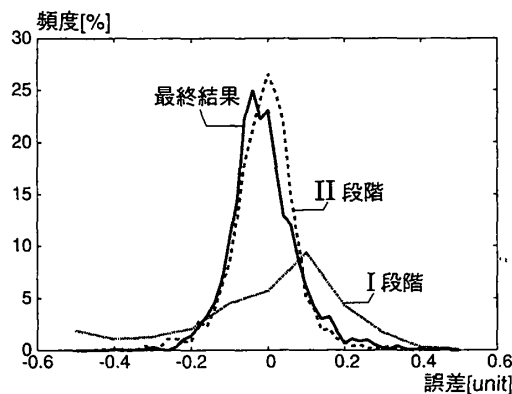


図 12 平面と特徴点との距離誤差の分布
Fig. 12 Distribution of error distance between the object plane and the reconstructed feature points.

5.2 実画像実験

大きさ $10 \times 8 \times 12$ cm の家の模型の全周囲をほぼ等間隔で撮影し、36 枚の画像列を得た。そのうち、11 枚を間引いた 25 枚の画像を用いて実験を行った。画像サイズは 720×480 画素とした。

図 13 に、画像列の各画像における、Zhang らの手法¹⁰⁾と金澤らの手法²³⁾および本手法で得られた対応特徴点の数を示す。なお特徴点は各画像から 300 点検出した。平均対応特徴点数は、Zhang らが 78 点、金澤らが 87 点、本手法が 113 点であり、本手法が最も多くの特徴点を対応付けることができた。

各手法の比較のために、図 14 に示すように視差が大きく対象物体が画像中心から外れた画像間で特徴点の対応付けを行った。得られた追跡結果は Zhang ら

☆ <http://www.sop.inria.fr/robotvis/personnel/zhang/software.html> を利用した。

☆☆ <http://www.img.tutkie.tut.ac.jp/programs/index.html> を利用した。

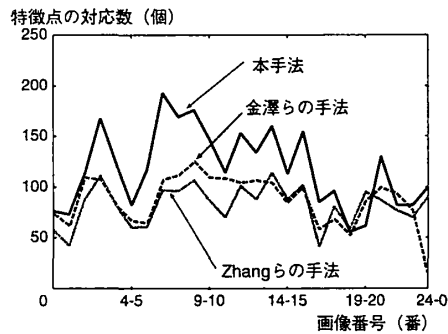


図 13 Zhang ら, 金澤ら, および本手法による対応特徴点数の比較

Fig. 13 Number of matched points by Zhang's method, Kanazawa's one and the proposed one.

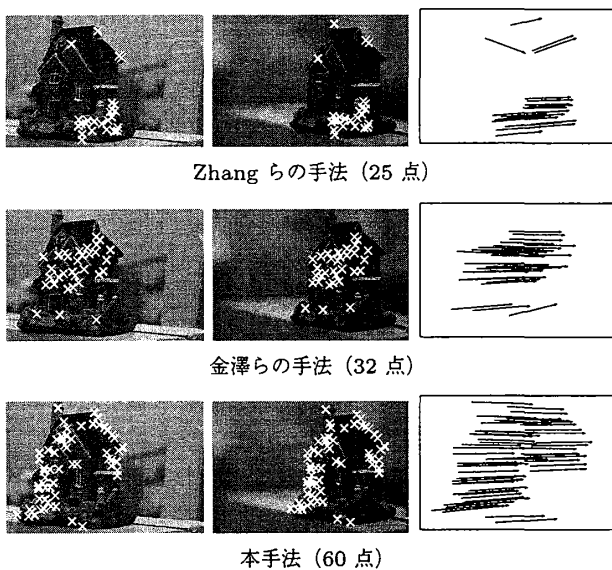
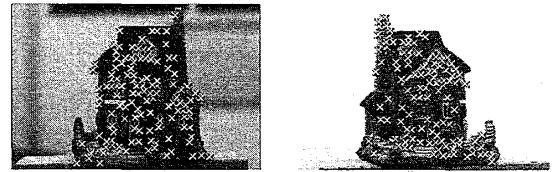


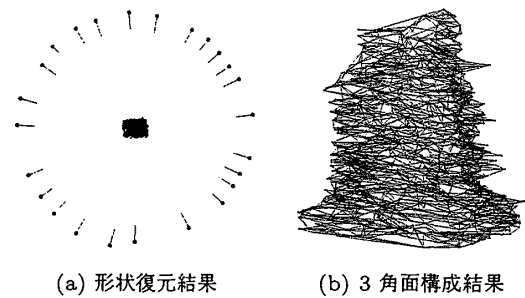
図 14 Zhang ら, 金澤ら, および本手法による特徴点の対応付け
Fig. 14 Matched feature points obtained by Zhang's method, Kanazawa's one and proposed one.

が 25 点, 金澤らが 32 点, 本手法が 60 点であった。図 14 右は, 各手法で得られた対応特徴点のオプティカルフローである。Zhang らの手法では大きな誤対応が存在し, 得られる対応特徴点数も少ないことが分かる。Zhang らの手法は, 画像間が局所的にアフィン変換により対応付けられると仮定し, 得られた特徴点候補に対して, LMedS 推定によりロバストにエピポーラ線を推定している。しかし, 図 14 のような画像では, 局所的な制約では正しい特徴点候補が十分得られず, LMedS 推定が有効に機能しないため, このような結果が得られたと考えられる。金澤らの手法では比較的正しい対応がとられているが, 得られる対応特徴点数は本手法より少ない。

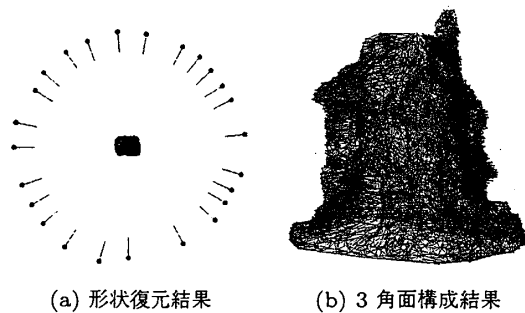
アフィン制約を適用して対応付けを行った特徴点を図 15 に示す (図中の "x" が特徴点を表す)。なお, 対応付けには 23×23 画素のテンプレートを使用し, アフィン変換で対応する画素を中心に 15 画素以内の



第 1 画像 第 18 画像
図 15 特徴点の追跡を行った一例
Fig. 15 Result of feature tracking.



(a) 形状復元結果 (b) 3 角面構成結果
図 16 アフィン制約を用いた初期復元結果
Fig. 16 Initial reconstruction result using affine constraint.



(a) 形状復元結果 (b) 3 角面構成結果
図 17 奥行き制約を用いた復元結果
Fig. 17 Reconstruction result using depth constraint.

領域で SSD を適用して対応特徴点を求めた。その結果, 形状を構成する特徴点として 1,294 点が得られた。式 (1) を用いて形状復元とカメラ位置・姿勢の推定を行った結果を図 16 (a) に示す。物体の復元形状とカメラ位置・姿勢は実際の形状や撮影状況にほぼ合致していることを確認した。この結果は, 画像数が少ないにもかかわらず, カメラ位置・姿勢が対象物体の全周囲を取り囲むように適切に推定されていることを示している。また, 3次元上の特徴点群に対し, 3角面を与えた結果を図 16 (b) に示す。特徴点数が少ないため, 物体形状は精細でない。しかし, 物体の大まかな形状を 3角面により表現できている。

次に, 奥行き制約を 2 回適用した後の形状復元結果を図 17 (a) に示す。なお, 特徴点は 15,751 点で構成されており, 図 16 (b) より 3角面が密に構成されたモデルとなっている。



図 18 3次元モデルの生成
Fig.18 Generated 3D model.

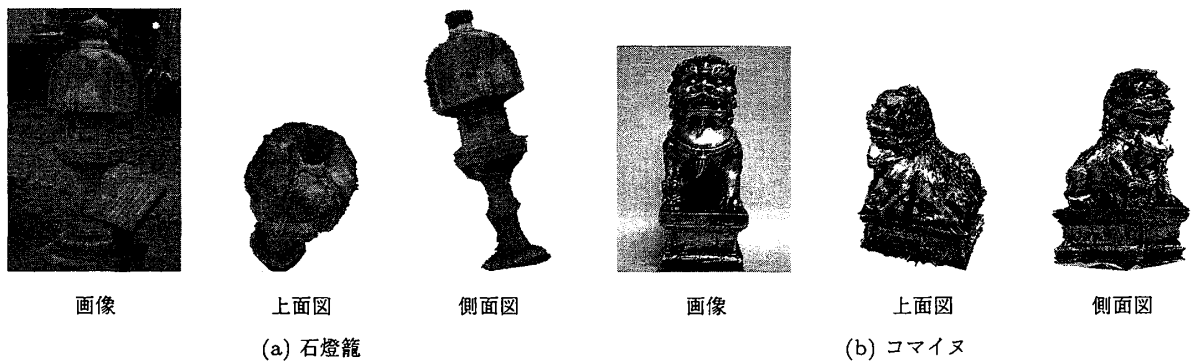


図 19 他の生成例
Fig.19 Other results.

表 2 生成に関する諸データ
Table 2 Data on the results.

画像列	撮影画像	使用画像	特徴点	処理時間
家の模型	36 枚	25 枚	15,751 点	36 分
石燈籠	35 枚	35 枚	4,587 点	34 分
コマイヌ	70 枚	30 枚	10,210 点	18 分

最後に、各 3 角面にテクスチャマッピングを行った結果を図 18 に示す。また、他のモデル生成結果を図 19 に、生成に関する諸データを表 2 に示す。処理は、CPU：Pentium4 2.0 GHz、メモリ：1 GBytes、OS：Vine linux 2.5 の計算機上で行った。また、奥行き制約の繰返しは、いずれの例でも 2 回であった。

撮影は、いずれの例でもほぼ等間隔で行った。得られた画像を間引きながら形状復元を行い、画像数が最小となる画像枚数を使用し復元した結果を示した。図 18 の家の模型では、側面の壁の部分が平面に近いので、11 枚の間引いても全周形状の 3 次元モデルが生成できた。一方、図 19 の石燈籠については、回転体に近い形状をしており、撮影した 35 枚の画像すべてを使わなければ 3 次元モデルが生成できなかった。コマイヌでは、70 枚の画像を撮影したが、結果的に

は 30 枚の画像から 3 次元モデルが生成できた。3 次元モデルの生成に必要な画像数は、対象に依存するため理論的な考察は不可能であるが、経験的には、物体の全周囲を 36 枚程度で撮影した画像があれば、本手法により生成が可能であることが多い。

生成された 3 次元モデルは、全体として良好といえる。ただし、本研究では物体とほぼ同じ水平面内から撮影を行っているため、特に物体の上面では正しく形状やテクスチャが得られていない。たとえば石燈籠の復元結果では、上面は本来尖形をしているが、特徴点が適切にとれず、平面に置き換えられている。

モデル生成の処理時間は、18～36 分程度となっている。このうち多くの部分は、特徴点の対応付けの処理時間である。アフィン制約による特徴点の対応付けでは、連続するすべての 2 画像間で勾配法によりアフィン変換を求めているため、画像数や画像の複雑さによって処理時間が変化したが、1 回あたり 6～17 分程度かかる。奥行き制約による特徴点の対応付けには、3～10 分程度かかり、これを 2 回繰り返している。特徴点の対応付けが与えられた後、非線型最適化により形状復元を行う処理は、特徴点数にもよるが、1 回あ

たり3~40秒程度である。処理時間の短縮は今後の課題であるが、アフィン制約では、解像度を下げた画像でアフィン変換のパラメータを求めた後、もとの画像でパラメータを改良する段階的手法の導入により、処理時間が短縮可能と考えられる。また、奥行き推定では、現在、3角面とカメラ位置との距離を求める処理に時間がかかっていることから、3次元的な幾何関係を高速に計算できるGPU処理などを導入することで、高速化が可能と考えられる。

6. おわりに

本論文では、対象物体を撮影した画像列から大局的幾何制約を用いることで、テクスチャ情報を保有した3次元モデルを生成する手法を示した。本手法は、まず大局的幾何制約として画像そのもののアフィン変換を利用することで、特徴点追跡を行う。次に、再投影誤差最小化により形状復元とカメラ位置・姿勢の推定を行う。さらに、復元された点群に対し3角面を構成し、その奥行きを大局的幾何制約とすることで、より多くの特徴点の追跡を行う。以上の処理を繰り返し行い、得られた3角面に対しテクスチャマッピングを行うことで、高精細な3次元形状モデルを生成する。

実画像を用いた実験により、本手法は画像間の視差が比較的大きい画像列においても、アフィン変換による大局的幾何制約によって特徴点の追跡が行え、特徴点の3次元座標とカメラ位置・姿勢が推定できることを確認した。さらに、3角面を用いて、その奥行き制約により画像上の探索範囲を限定することで、特徴点の対応付けの誤差を軽減し、安定した3次元モデル生成が可能であることを示した。

本研究では、前段階での結果に含まれる誤差を許容する機能を、後段階の処理が持つ手法となっている。このような段階的な方法は、ロバスト性や実用性の観点から重要であり、提案手法により、一貫した手順に従ったモデル生成が可能となる。特に、最終的な形状復元は非線形最適化により求めるため、画像全体で誤差を許容した復元結果を得ることができるという利点がある。

ただし、提案手法により任意の画像列から3次元モデルが生成できるわけではない。提案手法で3次元モデルが生成できるかどうかは、画像列の撮影位置、撮影枚数、対象物のテクスチャ、形状、光源状況など多様な要素が影響するため、明確ではない。経験的には、奥行き推定の範囲 δ や、4.3節の2乗誤差の大きい特徴点の除去において何点除去を行うかというパラメータを調整することで、全周囲を等間隔に36枚程度で

撮影した画像列から、3次元モデルが生成できることが多い。ただし、これらのパラメータの設定自体も経験的に行う必要があり、本手法の適用可能範囲を明らかにすることは今後の検討課題である。

また、円筒面への投影を用いた3角面の生成は、奥行き制約として利用するには十分であるが、最終的な3次元モデルとしては、特徴点の接続関係や3角面の形状が視覚的に好ましくない場合がある。この問題を解決することも今後の課題である。

謝辞 本研究の一部は、NEDO「産業技術研究助成事業費」の助成を受けて行った。また、本研究には、本文脚注に示した各プログラムを利用した。ここに、感謝の意を表します。

参考文献

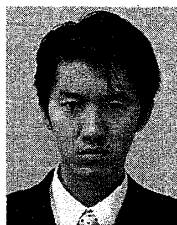
- 1) Oishi, T., Masuda, T. and Ikeuchi, K.: Digital Restoration of the Cultural Heritages, *Proc. VSMM*, pp.934-941 (2002).
- 2) Beardsley, P., Torr, P. and Zisserman, A.: 3D Model Acquisition from Extended Image Sequences, *Proc. 4th ECCV'96*, Vol.2, pp.683-695 (1996).
- 3) Koch, R., Pollefeys, M. and Van Gool, L.: Multi Viewpoint Stereo from Uncalibrated Video Sequences, *Proc. ECCV'98*, LNCS, pp.55-71, Springer-Verlag (1998).
- 4) 佐藤智和, 神原誠之, 横矢直和, 竹村治雄: マルチベースラインステレオ法を利用した動画像からの屋外環境の三次元モデル化, *日本VR学論*, Vol.7, No.2, pp.275-282 (2002).
- 5) Lhuillier, M. and Quan, L.: Quasi-Dense Reconstruction from Image Sequence, *Proc. ECCV'02*, pp.119-139 (2002).
- 6) 中川雅朗, 斎藤英雄, 小沢慎治: ハンディカメラを用いて撮影された画像列からのEPIの補正に基づく室内環境3次元再構築, *信学論D-II*, Vol.J84-D-II, pp.266-275 (2002).
- 7) Torr, P. and Zisserman, A.: MLESAC: A New Robust Estimator with Application to Estimating Image Geometry, *Computer Vision and Image Understanding*, Vol.78, No.1, pp.138-156 (2000).
- 8) 佐藤智和, 神原誠之, 横矢直和, 竹村治雄: マーカと自然特徴点の追跡による動画像からのカメラ移動パラメータの復元, *信学論D-II*, Vol.J86-D-II, pp.1431-1440 (2003).
- 9) Tomasi, C. and Kanade, T.: Detection and Tracking of Point Features, *CMU Tech. Rep. CMU-CS-91-132* (1991).
- 10) Zhang, Z., Deriche, R., Faugeras, O. and Luong, Q.-T.: A robust technique for matching

- two uncalibrated images through the recovery of the unknown epipolar geometry, *Proc. 3rd Artif. Intell.*, Vol.78, pp.87-119 (1995).
- 11) 金澤 靖, 金谷健一: 段階的マッチングによる画像モザイク生成, 信学論 D-II, Vol.J86-D-II, pp.816-824 (2003).
 - 12) Okutomi, M. and Kanade, T.: A multiple-baseline stereo, *IEEE PAMI*, Vol.15, No.4, pp.353-363 (1993).
 - 13) Matthies, L., Kanade, T. and Szeliski, R.: Kalman filter-based algorithms for estimating depth from image sequences, *IJCV*, Vol.3, pp.209-239 (1989).
 - 14) Gimel'farb, G.L. and Haralick, R.M.: Terrain reconstruction from multiple views, *Proc. CAIP'97*, LNCS 1296, pp.694-701, Springer (1997).
 - 15) Szeliski, R. and Kang, S.B.: Recovering 3D Shape and Motion from Image Streams using Non-Linear Least Squares, *Proc. CVPR*, pp.752-753 (1993).
 - 16) Triggs, B., McLauchlan, P.F., Hartley, R.I. and Fitzgibbon, A.W.: Bundle Adjustment — A Modern Synthesis, *Vision Algorithms: Theory and Practice*, LNCS 1883, pp.298-375, Springer (2000).
 - 17) 右田剛史, 天野 晃, 浅田尚紀: 3次元形状・運動復元のための高速非線形最適化計算法, 情報処理学会論文誌, Vol.44, No.11, pp.2864-2872 (2003).
 - 18) 椋木雅之, 右田剛史, 青山正人, 浅田尚紀: 非線形最適化による建物画像列からの全周形状一括復元のための初期値設定法, 情報処理学会論文誌：コンピュータビジョンとイメージメディア, Vol.45, No.SIG13(CVIM10), pp.64-73 (2004).
 - 19) Fischler, A. and Bolles, R.C.: Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography, *Comm. ACM*, Vol.24, pp.381-395 (1981).
 - 20) Rousseeuw, R.J.: Least median of squares regression, *J. American Stat. Assoc.*, Vol.79, pp.871-880 (1984).
 - 21) 内藤洋師, 岡谷貴之, 出口光一郎: 画像系列上の幾何学的拘束を用いた3次元形状復元, *MIRU 2002*, pp.II-79-84 (2004).
 - 22) 保田和隆, 右田剛史, 青山正人, 椋木雅之, 浅田尚紀: 疎な全周囲画像列からの密な3次元形状モデルの生成, 情報処理学会研究報告, CVIM-138, pp.73-80 (2003).
 - 23) 金澤 靖, 金谷健一: 大域的な整合性を保証するロバストな画像のマッチング, 情報処理学会論文誌：コンピュータビジョンとイメージメディア, Vol.44, No.SIG17(CVIM 8), pp.70-77 (2003).
 - 24) Harris, C. and Stephens, M.: A Combined Corner and Edge Detector, *Proc. Alvey Vision Conf.*, pp.147-151 (1988).
 - 25) Zhang, Z.: Determining The Epipolar Geometry And Its Uncertainty: A Review, *IJCV*, Vol.27, pp.161-195 (1998).

(平成 17 年 9 月 20 日受付)

(平成 18 年 3 月 20 日採録)

(担当編集委員 杉本 晃宏)



保田 和隆

2003年広島市立大学情報科学部知能情報システム工学科卒業。2005年同大学院情報科学研究科知能情報システム工学専攻前期課程修了。現在、ソニー株式会社勤務。2003年情報処理学会 CVIM 研究会第3回卒業論文セッション最優秀賞。



椋木 雅之 (正会員)

1991年京都大学工学部情報工学科卒業。同大学院博士課程、同工学研究科助手を経て、2002年広島市立大学情報科学部助教授、現在に至る。博士(工学)。画像認識、映像メディア処理の研究に従事。電子情報通信学会会員。



浅田 尚紀 (正会員)

1979年京都大学工学部電気工学科卒業。同大学院博士課程、福井医科大学助手、京都大学工学部助手、岡山大学工学部助教授を経て1995年から広島市立大学情報科学部教授、2006年から広島市立大学長、現在に至る。1989年シカゴ大学客員研究員。工学博士。コンピュータビジョン、コンピュータグラフィクス、文書構造解析、医用画像診断支援の研究に従事。1993年情報処理学会論文賞。IEEE CS, 電子情報通信学会, 人工知能学会, 医用画像工学会各会員。