

アフィン空間拘束とエピポラ拘束を利用した 2組の時系列画像における画像間対応軌跡推定

高橋 秀和^{†1,*1} 杉本 茂樹^{†1} 奥富 正敏^{†1}

本論文では、大きなベースライン長を持つ2組の時系列画像間の対応を求めるための新しい手法を提案する。提案手法では、それぞれの時系列画像中の特徴点をそれぞれ追跡し、画像間に存在するエピポラ拘束を利用して、一方の時系列画像のそれぞれの特徴点軌跡を他方の時系列画像の全特徴点軌跡が属するアフィン空間にあてはめる。これにより、一方の時系列画像から得られた特徴点軌跡に対する他方の時系列画像中の特徴点軌跡を推定することができ、画像間の特徴点の対応軌跡が得られる。この方法は、画素値を利用したステレオ画像対応点探索とは異なり、両方の画像で同時に観測されていない点に対しても特徴点の対応座標が得られるという特徴を持つ。さらに、得られた対応をもとに、対象の3次元形状を復元することもできる。合成画像および実画像を用いた実験を通じて、提案手法の有効性を示す^{*1}。

Image Correspondence Estimation from Affine Space Constraint and Epipolar Constraint on a Pair of Image Sequences

HIDEKAZU TAKAHASHI,^{†1,*1} SHIGEKI SUGIMOTO^{†1}
and MASATOSHI OKUTOMI^{†1}

In this paper, we propose a novel approach for image correspondence estimation using a pair of wide-baseline synchronized image sequences. In the proposed approach, after tracking the feature points in each image sequence over several frames, we utilize the consistent epipolar constraints on the image pairs for fitting each trajectory in one image sequence to the motion subspace derived from all trajectories in the other sequence. Then the stereo correspondence of each trajectory is obtained. Dissimilarly to the conventional stereo correspondence estimation based on matching using pixel values, the proposed approach enables us to obtain the image correspondences even though the trajectories are observed in only one image sequence. The validity of the proposed approach is shown by the experiments using synthetic and real images.

1. はじめに

複数の画像間で特徴点の対応座標を求めることは、コンピュータビジョンにおける基本的な問題であり、Shape from Motion、ステレオビジョン、物体認識などのアプリケーションに広く応用されている。

画像間の対応を求める場合には、一般に画像の画素値もしくは画素値から得られる特徴量を用い、その類似度を評価する方法が用いられる^{3),9)-11)}。時系列画

像中の特徴点をトラッキングする場合は、隣接フレーム間では画素値の変化が少ないため、このような画素値に基づいたアプローチによって比較的精度の良い対応推定が可能である。一方で、ベースライン長の大きな画像間の対応を推定する場合は、たとえエピポラ拘束が利用できる場合でも、(a) 大きな視点の違いによる画像の歪みや輝度の変化、(b) 小さい共通視野、(c) オクルージョンの発生、などが原因で対応推定が困難となる。特に、(b) や (c) により、一方の画像では観測された特徴点が他方の画像には観測されていない場合には、画素値に基づいたアプローチを用いて画像間の対応を推定することは原理的にできない。

そこで、本論文では、画素値に基づいたアプローチ

^{†1} 東京工業大学大学院理工学研究科機械制御システム専攻

Department of Mechanical and Control Engineering,
Graduate School of Science and Engineering, Tokyo Institute of Technology

^{*1} 現在、株式会社デンソーディーゼル噴射技術1部

Presently with Diesel Injection Engineering Department 1, Denso Corporation

^{*1} 本研究は第1著者が本大学院在学時になされたものである。

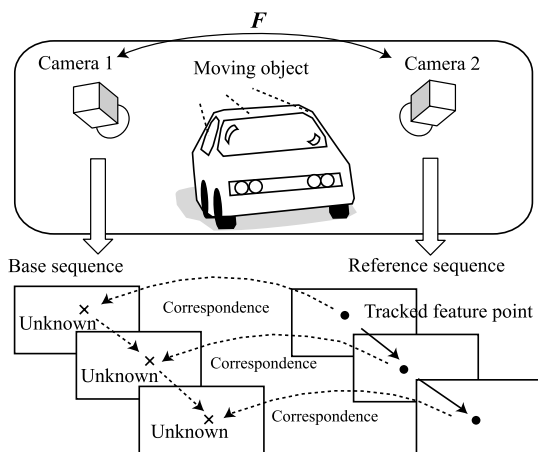


図 1 ベースライン長の大きな時系列画像間の対応

Fig.1 Correspondence between two wide-baseline image sequences.

における時系列画像中の特徴点トラッキングの容易さを利用して、大きなベースライン長を持つ2つのカメラから得られる時系列画像間の対応を、画素値を評価することなく求める手法を提案する。ここでは、図1に示すように、大きなベースライン長を持ち、かつ相対的に静止した2つのカメラから同一の移動物体を観測し、同図右の時系列画像においてトラッキングされた特徴点軌跡に対応する、同図左の時系列画像中の未知の軌跡を求める。このとき、その未知の軌跡は、左の時系列画像において観測されている必要はない。

この問題を解くための基本的な考え方は次のとおりである。まず、2組の時系列画像の特徴点をそれぞれ独立にトラッキングする。そして、図1左の時系列画像の軌跡からアフィン部分空間^{8),12)}を求める。アフィン部分空間をあてはめた時系列画像を基準画像列、他方の時系列画像を参照画像列とする。このとき、2組の時系列画像が同じ移動物体を観測したと仮定すると、参照画像列の軌跡に対応する基準画像列の未知の軌跡は、そのアフィン部分空間に属するはずである。さらに、2つのカメラ間の基礎行列が事前にキャリブレーションされていると仮定すると、その未知の軌跡の各時刻における座標は、参照画像列の軌跡の各時刻の座標と基礎行列とが定めるエピポーラ線上に存在するはずである。本論文では、この2つの拘束を用いることにより未知の軌跡を求められることを示す。

提案手法は、未知の軌跡が基準画像列中に観測されているか否かにかかわらずその軌跡を求めることができるため、提案手法の適用後に、探索範囲を限定して画素値に基づくアプローチを適用したり、その結果からオクルージョンを検知したりするなどの応用が考え

られる。また、提案手法の結果を利用して3次元復元を行うことも可能である。本論文では、応用の1つとして3次元復元を取り上げ、提案手法によって得られたステレオ画像間の対応座標から3次元復元を行い、その結果を別のアフィン空間にあてはめることで復元精度を向上させる方法について述べる。さらに、そのアフィン空間にあてはめた際の誤差を利用して、提案手法における軌跡が張るアフィン部分空間の適切な次元についても検討を行う。

時系列ステレオ画像を利用して、ステレオ画像間の輝度を比較することなく対応座標を推定する枠組みは、これまでもいくつか提案されている。Dornaikaら¹⁾は、強キャリブレーション済みの2つのステレオカメラ間の回転と並進と、特徴点の対応座標から推定される時系列画像間のego-motionを利用すれば、ステレオ画像間の対応座標は2つのエピポーラ線の交点として得られることを示した。しかし、この場合、時系列画像間のego-motionは、微小なカメラ運動から推定する必要があり、そのego-motion推定はトラッキング誤差に対してきわめて脆弱であることが知られているため¹⁴⁾、この方法を用いたステレオ画像間の対応座標推定の精度は低い。

これに対し、Hoら^{4),5)}は、2つのアフィンカメラから得られるステレオ時系列画像に因子分解法¹²⁾を適用した手法を提案している。この方法は、因子分解法が特徴点軌跡をアフィン空間にあてはめることから⁷⁾、提案手法と類似した方法といえる。Hoらは、2組の時系列画像の特徴点軌跡とカメラ外部パラメータから導出される行列が4次元部分空間を張ることを利用し、その部分空間の基底を求めることによりステレオ画像間の特徴点の対応軌跡を求めている。ただし、この手法では、部分空間の基底を求める際に、ステレオ画像間の4組以上の特徴点の対応軌跡を得ることが不可欠であるのに対し、提案手法は、あらかじめステレオ画像間の対応軌跡を得る必要はない。すなわち、提案手法では、ステレオカメラが物体の表裏をそれぞれ観測し、2つのカメラに同時に観測された特徴点がまったく存在しない極端な場合でも、特徴点の対応軌跡を求めることが可能である。

本論文の構成は以下のとおりである。まず、2章では提案手法、すなわち、参照画像列中の軌跡に対応する基準画像列中の未知の軌跡を、アフィン部分空間による拘束とエピポーラ拘束とを利用して求める手法について説明する。次に、3章では、提案手法の1つの応用として、2章で得られた画像間の対応座標を利用して3次元復元を行い、その結果を別のアフィン空

間にあてはめることで精度を向上させる方法について述べる。4章では、合成画像と実画像を用いた実験を行い、提案手法の有効性を示し、最後に本論文をまとめる。

2. 2組の時系列画像間の対応座標推定

カメラに対し相対的に移動する剛体を、2台の互いに位置が固定されたカメラで撮影し、一方の基準カメラと他方の参照カメラから得られる2組の時系列画像のそれぞれにおいて移動物体上の特徴点をトラッキングする。ただし、2台のカメラは同期しているものとする。移動物体は剛体であるので、2組の時系列画像中の特徴点群は各フレームにおいて相対的に静止している。基準カメラから得られた画像（以下、基準画像とする）上の特徴点の履歴座標を $(u_{ij}, v_{ij})^T$ ($i = 1, \dots, M$), ($j = 1, \dots, N_1$) とし、参照カメラから得られた画像（以下、参照画像とする）上の特徴点の履歴座標を $(u'_{ik}, v'_{ik})^T$ ($i = 1, \dots, M$), ($k = 1, \dots, N_2$) とする。ただし、 M は画像列のフレーム数を表し、 N_1 と N_2 は、基準画像列と参照画像列においてトラッキングされた特徴点の数をそれぞれ表す。

以下では、参照画像上の k 番目の特徴点の履歴座標 $(u'_{ik}, v'_{ik})^T$ ($i = 1, \dots, M$) に対応する、基準画像上の特徴点の履歴座標 $(u''_{ik}, v''_{ik})^T$ ($i = 1, \dots, M$) を求める問題を考える。すなわち、参照画像上の k 番目の特徴点の履歴座標を並べることによって表される既知の軌跡ベクトル

$$p'_k = \left[u'_{1k}, v'_{1k}, u'_{2k}, v'_{2k}, \dots, u'_{Mk}, v'_{Mk} \right]^T \quad (1)$$

に対応した、基準画像側の未知の軌跡ベクトル

$$p''_k = \left[u''_{1k}, v''_{1k}, u''_{2k}, v''_{2k}, \dots, u''_{Mk}, v''_{Mk} \right]^T \quad (2)$$

を求める。ただし、カメラ間の基礎行列 (Fundamental Matrix) は既知とする。

2.1 アフィン空間生成

基準画像列から得られた既知の軌跡ベクトルを

$$p_j = \left[u_{1j}, v_{1j}, u_{2j}, v_{2j}, \dots, u_{Mj}, v_{Mj} \right]^T \quad (3)$$

とする。 $j = 1, 2, \dots, N_1$ の全軌跡ベクトルを並べて、次のような行列を作る。

$$W = \left[p_1, p_2, \dots, p_{N_1} \right] \quad (4)$$

W の各列からすべての軌跡ベクトルの重心ベクトルを引いた行列を \bar{W} とし、次のように表す。

$$\bar{W} = \left[\bar{p}_1, \bar{p}_2, \dots, \bar{p}_{N_1} \right] \quad (5)$$

ただし、

$$\bar{p}_j = p_j - p_G, \quad p_G = \frac{1}{N_1} \sum_{j=1}^{N_1} p_j \quad (6)$$

である。そして、 \bar{W} を主成分分析することで基底 e_1, \dots, e_{2M} を得る*1。

基準画像列の軌跡ベクトル p_j がすべて同一物体上にあり、かつカメラがアフィンカメラモデルを満たす場合、 p_j は3次元アフィン空間に含まれる⁸⁾。すなわち、 p_G と3つの基底 e_1, e_2, e_3 を用いて次式のように表される。

$$p_j = p_G + \alpha_{1j}e_1 + \alpha_{2j}e_2 + \alpha_{3j}e_3 \quad (7)$$

ただし、 $\alpha_{1j}, \alpha_{2j}, \alpha_{3j}$ は基底の係数である。

カメラがアフィンカメラモデルの場合は、前述のとおり基底数は3で十分であるが、実際のカメラはアフィンカメラモデルとは厳密には異なる。そのため、実際のデータは3次元より高い次元のアフィン空間にも含まれると考えられる。そこで、以下では、カメラモデルをより一般化し、必要に応じ次のように、より高次のアフィン空間を利用する。

$$p_j = p_G + \alpha_{1j}e_1 + \alpha_{2j}e_2 + \dots + \alpha_{\nu j}e_{\nu}, \quad (8)$$

ただし、 $3 \leq \nu$ とする。

2.2 対応軌跡推定

基準画像列の軌跡ベクトルは式(8)で示すアフィン部分空間に属するため、2台のカメラが同一の移動物体を観測したと仮定すると、参照画像上の k 番目の特徴点軌跡ベクトル p'_k に対応する基準画像上の軌跡 p''_k も、式(8)で表されたアフィン空間に属する。よって、次式が成立する。

$$p''_k = p_G + \beta_{1k}e_1 + \beta_{2k}e_2 + \dots + \beta_{\nu k}e_{\nu} \quad (9)$$

ただし、 β_{nk} ($n = 1, 2, \dots, \nu$) は、 n 番目の基底の係数である。すなわち、式(9)は、アフィン部分空間による拘束を表している。

次に、基準画像と参照画像との間のエピソード拘束を定式化する。参照画像上の k 番目の特徴点軌跡ベクトル p'_k と、それに対応する基準画像上の軌跡ベクトル p''_k は、画像間のエピソード拘束によって関係づけることができる。すなわち、 i 番目のフレームでは、カメラ間の基礎行列 F を用いて次式が成立する。

$$\begin{bmatrix} u'_{ik} & v'_{ik} & 1 \end{bmatrix} F \begin{bmatrix} u''_{ik} \\ v''_{ik} \\ 1 \end{bmatrix} = 0, \quad (10)$$

*1 モーメント行列 $M = \bar{W}\bar{W}^T$ を作成し、モーメント行列の固有ベクトルを求めると、その固有ベクトルが求める基底 e_1, \dots, e_{2M} に対応する。

$$\text{where } \mathbf{F} = \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix} \quad (11)$$

以下では，アフィン部分空間拘束による式 (9) と，エビポーラ拘束による式 (10) を用いて，未知の軌跡ベクトル \mathbf{p}_k'' を求めることができることを示す．

式 (9) における部分空間の基底 $\{\mathbf{p}_G, \mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_\nu\}$ は，基準画像列の特徴点軌跡より得られるため， \mathbf{p}_k'' を求めるためには，係数 $\{\beta_{1k}, \beta_{2k}, \dots, \beta_{\nu k}\}$ を求めればよい．そこで，まず，式 (9) から， i 番目のフレームに関する式を抽出する．

$$\begin{bmatrix} u_{ik}'' \\ v_{ik}'' \end{bmatrix} = \mathbf{p}_G^{(i)} + \beta_{1k} \mathbf{e}_1^{(i)} + \beta_{2k} \mathbf{e}_2^{(i)} + \dots + \beta_{\nu k} \mathbf{e}_\nu^{(i)} \quad (12)$$

ただし， $\mathbf{p}_G^{(i)}, \mathbf{e}_1^{(i)}, \dots, \mathbf{e}_\nu^{(i)}$ は， $\mathbf{p}_G, \mathbf{e}_1, \dots, \mathbf{e}_\nu$ のそれぞれのベクトルから $2i-1$ と $2i$ 番目の行を取り出したベクトルを表す．そして，式 (10) に式 (12) を代入すると， i 番目のフレームについて，次式が得られる．

$$\tilde{\mathbf{u}}_{ik}^T \left\{ \mathbf{f}_1 \begin{bmatrix} \mathbf{p}_G^{(i)} & \mathbf{E}^{(i)} \end{bmatrix} \begin{bmatrix} 1 \\ \beta_{1k} \\ \vdots \\ \beta_{\nu k} \end{bmatrix} + \mathbf{f}_2 \right\} = 0 \quad (13)$$

ただし，

$$\begin{aligned} \tilde{\mathbf{u}}_{ik}^T &= \begin{bmatrix} u_{ik}' & v_{ik}' & 1 \end{bmatrix}, \\ \mathbf{E}^{(i)} &= \begin{bmatrix} \mathbf{e}_1^{(i)} & \dots & \mathbf{e}_\nu^{(i)} \end{bmatrix}, \\ \mathbf{f}_1 &= \begin{bmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \\ f_{31} & f_{32} \end{bmatrix}, \quad \mathbf{f}_2 = \begin{bmatrix} f_{13} \\ f_{23} \\ f_{33} \end{bmatrix} \end{aligned}$$

とする．さらに，全フレーム ($i=1, \dots, M$) について式 (13) を縦に並べると，次式が得られる．

$$\begin{aligned} & \mathbf{M}_k^T \left\{ \mathbf{F}_1 \begin{bmatrix} \mathbf{p}_G & \mathbf{E} \end{bmatrix} \begin{bmatrix} 1 \\ \beta_{1k} \\ \vdots \\ \beta_{\nu k} \end{bmatrix} + \mathbf{F}_2 \right\} \\ &= \begin{bmatrix} \mathbf{0} \end{bmatrix} \end{aligned} \quad (14)$$

ただし，

$$\mathbf{M}_k^T = \begin{bmatrix} \tilde{\mathbf{u}}_{1k}^T \\ \tilde{\mathbf{u}}_{2k}^T \\ \vdots \\ \tilde{\mathbf{u}}_{Mk}^T \end{bmatrix},$$

$$\mathbf{E} = \begin{bmatrix} \mathbf{E}^{(1)} \\ \vdots \\ \mathbf{E}^{(M)} \end{bmatrix},$$

$$\mathbf{F}_1 = \begin{bmatrix} \mathbf{f}_1 & & \\ & \ddots & \\ & & \mathbf{f}_1 \end{bmatrix}, \quad \mathbf{F}_2 = \begin{bmatrix} \mathbf{f}_2 \\ \vdots \\ \mathbf{f}_2 \end{bmatrix}$$

とする．

式 (14) は， ν 個の未知数 β_{nk} ($n=1, 2, \dots, \nu$) に対して式数が M 個なので，アフィン空間の次元 ν を $3 \leq \nu \leq M$ の範囲で選択すれば，未知数 β_{nk} を求めることができる．

式 (14) を次のように書く．

$$\mathbf{A}_k \mathbf{x}_k = \mathbf{b}_k \quad (15)$$

ただし，

$$\mathbf{A}_k = \mathbf{M}_k^T \mathbf{F}_1 \mathbf{E}, \quad \mathbf{x}_k = \begin{bmatrix} \beta_{1k} \\ \vdots \\ \beta_{\nu k} \end{bmatrix},$$

$$\mathbf{b}_k = -\mathbf{M}_k^T \mathbf{F}_1 \mathbf{p}_G - \mathbf{M}_k^T \mathbf{F}_2$$

である．式 (15) を

$$\mathbf{x}_k = (\mathbf{A}_k^T \mathbf{A}_k)^{-1} \mathbf{A}_k^T \mathbf{b}_k \quad (16)$$

のように解き，未知数 β_{nk} ($n=1, 2, \dots, \nu$) を求める．これにより，参照画像上の k 番目の軌跡ベクトル \mathbf{p}_k' に対する基準画像上の軌跡ベクトル \mathbf{p}_k'' が得られる．

この方法を，すべての k についてそれぞれ求めることにより，参照画像にしか写っていない特徴点でも，その特徴点に対応する基準画像上の特徴点の位置を全フレームにわたって推定することができる．基準画像と参照画像の役割を入れ替えて同様の計算を行うことにより，一方のカメラにしか写っていない特徴点でも，両方のカメラで対応座標を求めることができる．すなわち，両方のカメラで得られた画像上のすべての特徴点の対応座標を求めることが可能である^{*1}．

3. 3次元再構成への応用

本章では，提案手法の応用として，3次元復元を取

*1 なお，Shape from Motion により，一方の時系列画像から3次元復元する方法では，復元の不定性が残るため，それを他方の画像に再投影して対応座標を求めることはできない．

り上げる．ここでは，2章で述べた提案手法を用いて3次元復元を行い，剛体としての一貫性を利用して，その復元結果の精度を向上させる方法について述べる．

本章で述べる方法では，全フレームにわたって全特徴点を3次元復元した結果が，剛体運動の一貫性を示す3次元アフィン部分空間に含まれることを利用し，復元結果をそのアフィン空間にあてはめることにより，3次元復元の精度を向上させる．また，そのあてはめを行った際の残差を用いて，式(9)における基底数 ν を決定する方法についても述べる．ただし，以下では，基礎行列に加え，ステレオカメラの内部パラメータを既知とする．

2章に述べた方法によって得られた第 i 番フレームにおける k 番目の特徴点の対応から，カメラの内部・外部パラメータを利用して，その特徴点の3次元座標を計算する．得られた3次元座標を $X_{ik} = (x_{ik}, y_{ik}, z_{ik})^T$ とし，全フレームについて並べたベクトル X_k を次式で表す．

$$X_k = \begin{bmatrix} X_{1k} & X_{2k} & \dots & X_{Mk} \end{bmatrix}^T \quad (17)$$

i 番目のフレームのワールド座標系に対する物体座標系の基底ベクトルを，それぞれ i_i, j_i, k_i とし，物体座標系の原点を τ_i とする． k 番目の特徴点の3次元座標を物体座標系で表したときの座標を， $(a_k, b_k, c_k)^T$ とすると， X_{ik} は次式で表される⁸⁾．

$$X_{ik} = \begin{bmatrix} x_{ik} \\ y_{ik} \\ z_{ik} \end{bmatrix} = \tau_i + a_k i_i + b_k j_i + c_k k_i \quad (18)$$

よって，上式を全フレームにわたって縦に並べることで X_k を得る．

$$X_k = \tau + a_k i + b_k j + c_k k \quad (19)$$

ただし， τ, i, j, k は， $d = [d_1^T, d_2^T, \dots, d_M^T]^T$ ($d = \tau, i, j, k$)で表される．

さらに，全特徴点をならべることで次式を得る．

$$X = \begin{bmatrix} X_1 & X_2 & \dots & X_N \end{bmatrix} \\ = \begin{bmatrix} \tau & i & j & k \end{bmatrix} \begin{bmatrix} 1 & 1 & \dots & 1 \\ a_1 & a_2 & \dots & a_N \\ b_1 & b_2 & \dots & b_N \\ c_1 & c_2 & \dots & c_N \end{bmatrix}$$

ただし N は全特徴点数である．

物体座標系の原点はどこに設定してもよいので，特徴点の3次元座標の重心を原点に設定すると，次式を得る．

$$\bar{X} = \begin{bmatrix} \bar{X}_1 & \bar{X}_2 & \dots & \bar{X}_N \end{bmatrix} \\ = \begin{bmatrix} i & j & k \end{bmatrix} \begin{bmatrix} a_1 & a_2 & \dots & a_N \\ b_1 & b_2 & \dots & b_N \\ c_1 & c_2 & \dots & c_N \end{bmatrix} \quad (20)$$

ただし，

$$\bar{X}_k = X_k - \tau, \quad \tau = \frac{1}{N} \sum_{k=1}^{N_2} X_k \quad (21)$$

である．

式(20)は，対象物体が3次元空間中で一貫した剛体であることを示す式(18)から導かれたものであり，全フレームの全特徴点の3次元復元結果を示す $3M \times N_2$ の行列 \bar{X} の各行ベクトルが，3次元アフィン空間に含まれることを意味している．本問題の仮定により式(18)は有効であるため，式(20)はつねに成立する．

2章で述べた手法により対応を推定し，その結果を用いて3次元復元を行うと，トラッキング誤差などの影響により， \bar{X} は式(20)の3次元アフィン空間には属さず，より大きな次元の空間を生成する．そこで， \bar{X} を主成分分析し，改めて \bar{X} を式(20)で示す3次元アフィン空間にあてはめることにより，3次元復元の精度を向上させることができる．これは，その3次元復元の結果を画像に再投影することにより，対応推定の精度を向上させることも意味する．

さらに，3次元復元結果を式(20)にあてはめた際の残差を利用して，適切な基底数 ν の値を決定することが考えられる．すなわち， \bar{X} を主成分分析し，第4番目以下の寄与率を用いてデータを3次元アフィン空間にあてはめた場合の残差を計算し，その残差が最小になるように基底数 ν を決定することができる．実験ではこの方法を用いることにより，より透視投影モデルに近いときや，トラッキングノイズのレベルによって適切な ν の値がどのように変化するかを解析する．

4. 実験結果

提案手法の有効性を確認するために，シミュレーション実験と実画像実験を行った．シミュレーション実験では，まず，2章で述べた提案手法の有効性をマッチング精度により確認するとともに，3章で述べた3次元復元を行うことにより， ν の適切な値を検討する．また，実画像実験では，実際に3次元復元を行った結果を示す．

4.1 シミュレーション実験

シミュレーション実験では，図2に示すようにカメ

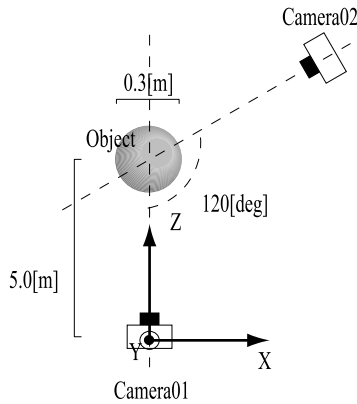


図 2 シミュレーション実験におけるカメラと物体の配置
Fig. 2 Arrangement of cameras and object.

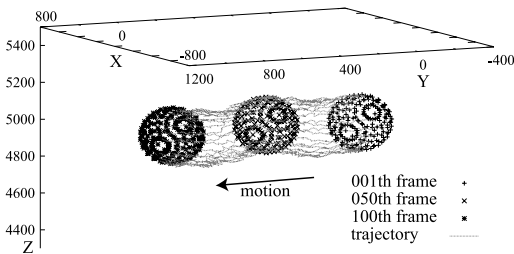


図 3 球の軌跡 (各軸の単位は mm)
Fig. 3 Trajectories of sphere's motion.

ラと対象物体を配置した．対象物体は球であり，2 台のカメラは球の中心点から 5.0 [m] 離れている．2 台のカメラはともに画角は 7.8 [deg]，焦点距離は 35 [mm] である．球はすべてのフレームにわたって 1 平面上をランダムに移動している．このときの球の移動軌跡を図 3 に示す．

球上に 200 点の特徴点を配置し，2 台のカメラそれぞれで特徴点を仮想的に撮影する．カメラ 1 では 71 点，カメラ 2 では 77 点の特徴点を 100 フレームにわたって追跡した．このときトラッキングした特徴点を図 4 に示す．各フレームにて得られた 2 枚の画像の各特徴点に対して，画像座標上で平均 0，標準偏差 σ [pix] の正規乱数によるトラッキング誤差を加えた．

2 章で述べた提案手法によって得られた対応点座標の推定結果を図 5 に示す．ただし，式 (9) のアフィン部分空間の基底数は $\nu = 3$ であり，トラッキング誤差は $\sigma = 1$ である．また，同図は，図 4 のカメラ 1 で撮影された画像の 1 フレームを示しており，カメラ 1 で撮影された特徴点を + で示し，提案手法によって図 4 のカメラ 2 で撮影された特徴点の対応推定結果を × で示している．

次に 2 章で述べた提案手法によって得られた対応

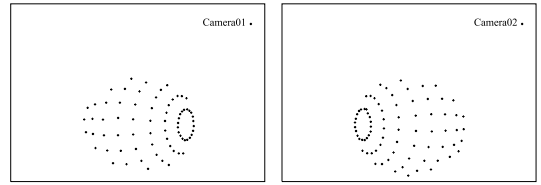


図 4 全フレームで観測された特徴点 (左: カメラ 1, 右: カメラ 2)

Fig. 4 Feature points observed over all frames (left: Camera1, right: Camera2).

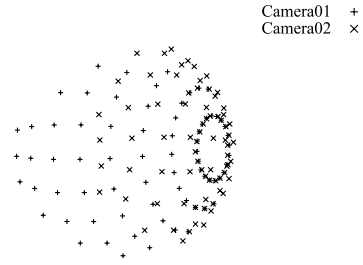


図 5 対応位置推定結果
Fig. 5 Inferred correspondences.

点座標の推定精度を図 6，図 7 に示す．ただし，アフィン部分空間の基底数は $\nu = 3$ とし，図 6 はカメラ 2 を基準画像とした場合の推定結果を示し，トラッキング誤差における σ (横軸) と，カメラ 1 に写っている特徴点をカメラ 2 の画像上に投影した際の，全フレームにわたる全特徴点の真の座標と推定した座標との RMSE (縦軸) [pix] を示している．同図の結果は，1 つの σ について 20 回対応点推定を行い，各 σ における平均誤差を計算したものである．また，図 7 は，カメラ 1 を基準画像とした場合の推定結果を示し，図 6 と同様に，カメラ 2 に写っている特徴点をカメラ 1 の画像上に投影した際の，真の座標と推定した座標との RMSE (縦軸) を示している．いずれにおいても， $\sigma = 0$ のときは，対応座標推定の RMSE はゼロであり，良好な対応座標推定ができています．また，トラッキング誤差がある場合は，その標準偏差 σ よりも，推定した対応の RMSE の方が低い．これは，提案手法では，未知の特徴点の軌跡を，基準画像列の特徴点軌跡が張る 3 次元アフィン空間にあてはめていることから，このあてはめにより，各特徴点座標に独立に付与したトラッキング誤差を低減させていると考えることができる．

次に，3 章で述べた 3 次元復元手法を利用して，アフィンカメラ効果とトラッキング誤差レベルを変化させた場合の適切な基底数 ν について調べた．ここでは，画像上での物体の大きさ，移動量がほぼ同じにな

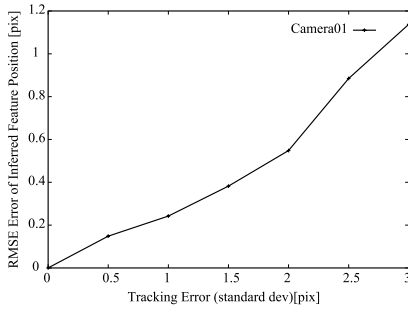


図 6 対応点推定精度 (カメラ 1)

Fig. 6 Correspondence accuracy (Camera1).

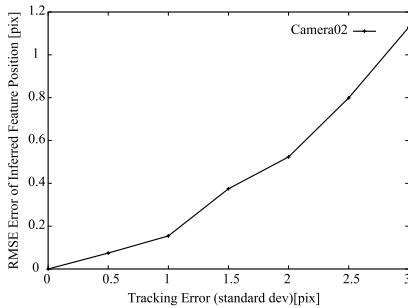


図 7 対応点推定精度 (カメラ 2)

Fig. 7 Correspondence accuracy (Camera2).

るように、3次元空間中の物体の位置と移動量、およびカメラの焦点距離(9.0 [mm] ~ 49.0 [mm])を変化させ、アフィンカメラモデルの効果(初期フレームにおける物体とカメラとの距離を、物体の厚みで割った値)を変化させながら、対応座標推定を行った。図8、図9は、カメラ2の画像とカメラ1の画像をそれぞれ基準画像としたとき、アフィンカメラモデルの効果(横軸)と、3次元復元結果を式(20)にあてはめた際の残差を最小にする基底数 ν (縦軸)との関係を示している。同図の各グラフは、トラッキング誤差の標準偏差 σ を0.2~2.0 [pix]の間で変化させた場合においてそれぞれ選択された基底数を示している。ただし、1つの σ について20回試行を繰り返し、選択された基底数 ν の平均を示したものである。

図8、9により、トラッキング誤差が小さいときは、アフィンモデルとの乖離が大きくなるほど(横軸の値が小さくなるほど)、大きな基底数が選択される傾向があることが分かる。トラッキング誤差が $\sigma = 0.5$ の場合には、 $\nu = 6$ 程度のアフィン部分空間を利用した際に3次元形状の剛体としての一貫性が良くなっており、 $\nu = 3$ よりも高い次元のアフィン部分空間を導入する意義が示されているといえる。一方で、トラッキング誤差が大きいときは、つねに小さな基底数が選択されていることが分かる。これは、高次元のアフィン

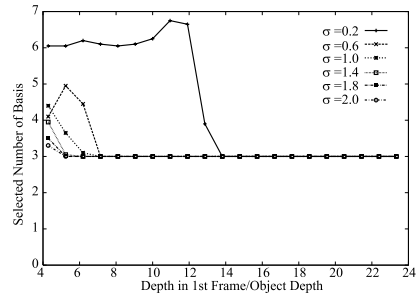


図 8 撮影条件と基底数 (カメラ 1)

Fig. 8 Selected number of basis with respect to shooting condition (Camera1).

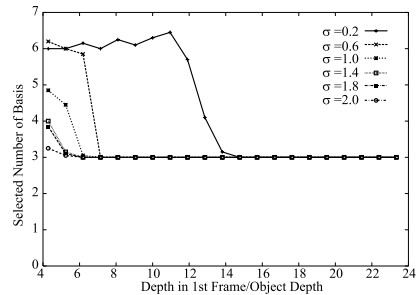


図 9 撮影条件と基底数 (カメラ 2)

Fig. 9 Selected number of basis with respect to shooting condition (Camera2).

空間は、必ずしも一般的なカメラモデルを表現するためのより良いモデルとはいえないため、トラッキング誤差が大きくなると、特定次元数のアフィン部分空間によってトラッキング誤差とモデル誤差を分離することはできず、結果として、より基底数が少ないほうが各フレームにおける3次元形状の剛体としての一貫性が高くなり、小さな ν が選択されるものと思われる。実画像を用いた場合、トラッキング誤差は比較的大きなものになると考えられる。よって、2章で述べた提案手法を実画像に適用することを考えると、カメラからの距離と比較して物体の厚みが十分小さいと判断できる場合は、 $\nu = 3$ でよいといえる。

次に、3章の手法を用い、式(20)にあてはめることによって得た3次元復元結果を図10に示す。ただし、 $\sigma = 1$ 、基底数は $\nu = 3$ である。図10は3次元復元結果に真値を重ねたものである。図10を見ると真値の中に推定した3次元位置の特徴点が含まれている。3次元形状における真値と推定位置とのRMSEは1.5 [mm]であった。式(20)にあてはめを行わず、2章の提案手法で求めた対応のみから3次元復元を行った際のRMSEは8.8 [mm]であったことから、3章の手法を用いて、3次元復元の精度を向上さ

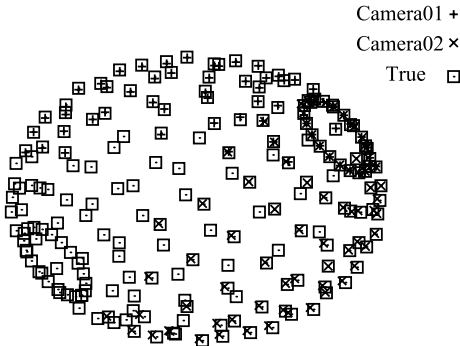


図 10 3次元復元結果
Fig.10 3D reconstruction result.

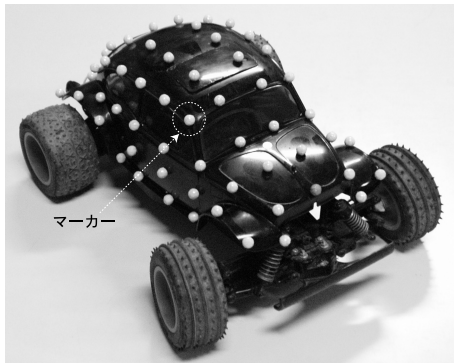


図 11 撮影対象
Fig.11 Real object.

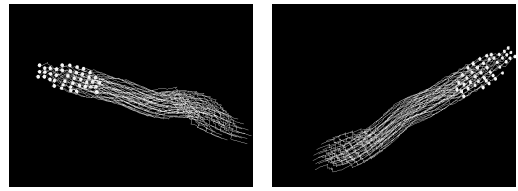


図 12 トラッキング結果 (左: カメラ 1, 右: カメラ 2)
Fig.12 Observed trajectories (Left: Camera1, Right: Camera2).

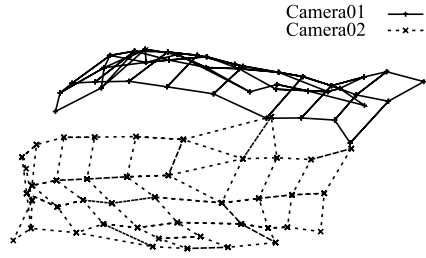


図 13 3次元形状復元結果 (ワイヤフレーム表示)
Fig.13 3D reconstruction result (wire-frame view).

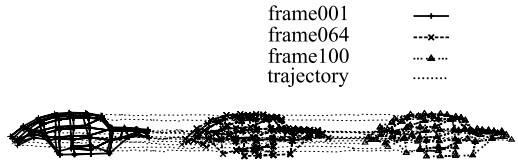


図 14 3次元軌跡復元結果 (横方向からのビュー)
Fig.14 3D trajecotries (side view).

せることができることが分かる。

4.2 実画像実験

実画像実験では、共通部分をほとんど設けずに物体を撮影して提案手法の有効性を確認した。ただし、真の対応を得ることが困難なため、ここでは、3章の手法を利用して物体の3次元形状復元が精度良く行えることを確かめた。

実験に用いた撮影対象は図 11 であり、この物体には特徴点追跡用に全部で 83 個のマーカーがついている。同図の物体を、平面上をジグザグに運動させマーカーをトラッキングした。図 12 はトラッキング結果を表しており、図中の線は、100 フレームにわたってトラッキングされた特徴点の軌跡である。カメラ 1 で撮影した画像上では 42 個のマーカーが追跡でき、カメラ 2 で撮影した画像上では 43 個のマーカーが追跡できた。

提案手法により対応座標推定を行い、3章の手法を用いて対象物体を3次元形状復元した。ただし、アフィン空間の次元は $\nu = 3$ とした。3次元形状復元した結果をワイヤフレームで示したものが図 13 である。図中右半分がカメラ 1 で見えている部分の復元結

果(カメラ 2 を基準画像とした結果)を示し、左部分がカメラ 2 で見えている部分の復元結果(カメラ 1 を基準画像とした結果)を示す。同図より 2 つのカメラで、共通で見えている特徴点(2点)が重なっていることが分かる。これにより、良好な 3次元形状が得られていることが分かる。

また、図 14 は、第 1 フレーム、第 64 フレーム、および第 100 フレームの復元結果と、物体の運動の軌跡とを示したものである。同図は、物体を真横から見たものをしており、物体の運動軌跡がほぼ一直線上になっていて、物体が一平面上を運動していることをよく表している。

ワイヤフレームによる復元結果に、三角パッチをあて、CG 表示したものが図 15 上段である。図 15 下段は撮影対象をそれぞれの CG の結果の視点にあうように、実物体を撮影したものである。いずれの復元結果も真の形状をよく表している。このことから、提案手法と 3章で述べた手法により、3次元復元が精度良く行われているといえる。

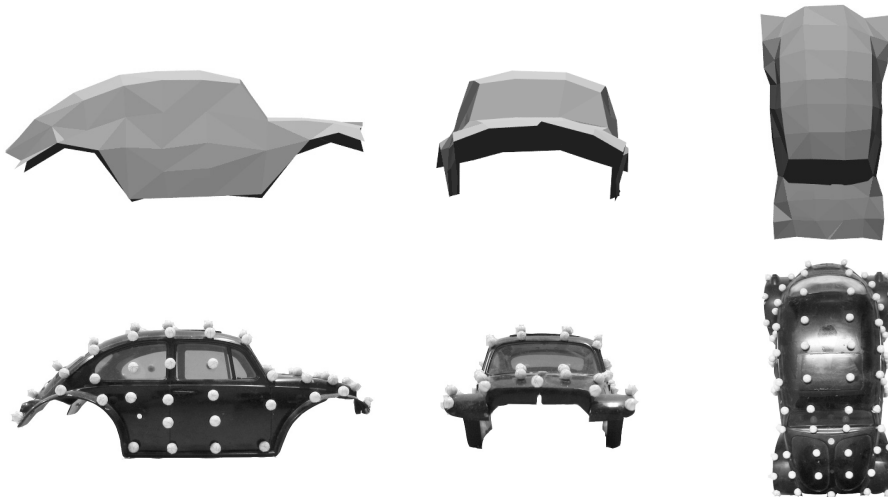


図 15 上段：CG 復元結果．下段：真の形状（左から，横，正面，鉛直方向からのビュー）
Fig.15 Upper row: 3D reconstruction by CG view. Bottom row: true shape
(side, front, and top view).

5. おわりに

本論文では，時系列ステレオ画像を使った新しいステレオ対応座標推定の方法を提案した．提案手法では，基準画像列から得られた特徴点軌跡を用いてアフィン空間を構成し，そのアフィン空間拘束と，エピポーラ拘束とを用いて，参照画像列の軌跡に対応する基準画像列の軌跡を推定した．また，提案手法の応用例として，得られた対応軌跡から 3 次元形状復元を行い，各フレームにおける 3 次元形状が 1 つの 3 次元アフィン空間に属することを利用して，画素値による対応を求めることなく，3 次元復元精度を向上させる方法を示すとともに，特徴点軌跡が張るアフィン空間の次元について検討した．

提案手法の最も大きな特徴は，輝度情報を用いた画像間の対応点を直接探索する方法とは異なり，カメラ間のエピポーラ拘束と，同一物体上の特徴点の軌跡ベクトルが持つ拘束条件とを利用して，幾何学的拘束条件のみから画像間の特徴点の対応座標を推定することである．これにより，両方の画像から同じ特徴点が観測されていない場合でも，一方の画像で観測された特徴点軌跡に対応する，もう一方の画像上での軌跡を求めることができた．

今後は，実画像を利用した手法において，マーカを使用せず，画像上の特徴点のみを利用して提案手法の有効性を示す方法について検討する予定である．また，実環境下では，シミュレーション実験とは異なり，トラッキング誤差が大きくなる傾向があるため，アウト

ライアの除去が必要となると考えている．その方法として，基準画像列のアフィン部分空間を推定する際に RANSAC²⁾ や GPCA¹³⁾ などを利用する方法が考えられる．また，3 次元復元を行わずに ν の値を推定する方法として，幾何学的 AIC⁶⁾ を利用する方法が考えられる．今後は，これらの方法を検討する予定である．

参考文献

- 1) Dornaika, F. and Chung, R.: Stereo Correspondence From Motion Correspondence, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol.1, pp.70–75 (1999).
- 2) Fischer, M.A.: Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography, *Comm. ACM*, Vol.24, No.6, pp.381–395 (1981).
- 3) Freeman, W. and Adelson, E.: The Design and Use of Steerable Filters, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.13, No.9, pp.891–906 (1991).
- 4) Ho, P.-K. and Chung, R.: Use of Affine Camera Model and All Stereo Pairs in Stereo-Motion, *IEEE International Conference on Intelligent Vehicles*, Vol.1, pp.323–328 (1998).
- 5) Ho, P.-K. and Chung, R.: Stereo-Motion with Stereo and Motion in Complement, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.22, No.2, pp.215–220 (2000).

- 6) 金澤 靖, 金谷健一: 幾何学的 AIC による画像モザイク生成の安定化, 電子情報通信学会論文誌, Vol.J83-A, No.6, pp.686-693 (2000).
- 7) 金谷健一, 菅谷保之: 因子分解法の完全レシピ, 信学技報, PRMU-2003-118 (2001).
- 8) 黒澤典義, 金谷健一: アフィン空間分離法による運動物体の分離, 情報処理学会研究報告, 2001-CVIM-125-3 (2001).
- 9) Lazebnik, S., Schmid, C. and Ponce, J.: Sparse Texture Representation Using Affine-invariant Neighborhoods, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp.319-324 (2003).
- 10) Lowe, D.G.: Distinctive Image Features From Scale-Invariant Keypoints, *International Journal of Computer Vision*, Vol.60, No.2, pp.91-110 (2004).
- 11) Lucas, B.D. and Kanade, T.: An Iterative Image Registration Technique with An Approach to Stereo Vision, *Image Understanding Workshop*, pp.121-130 (1981).
- 12) Tomasi, C. and Kanade, T.: Shape and Motion From Image Streams Under Orthography: A Factorization Method, *International Journal of Computer Vision*, Vol.9, No.2, pp.137-154 (1992).
- 13) Vidal, R., Ma, Y. and Sastry, S.: Generalized Principal Component Analysis (GPCA), *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.27, No.12 (2005).
- 14) Zhang, Z.: Determining The Epipolar Geometry and Its Uncertainty: A Review, *International Journal of Computer Vision*, Vol.27, No.2, pp.161-198 (1998).

(平成 19 年 5 月 8 日受付)

(平成 19 年 11 月 21 日採録)

(担当編集委員 川崎 洋)



高橋 秀和

2004 年東京工業大学制御システム工学科卒業。2006 年東京工業大学大学院理工学研究科機械制御システム専攻修士課程修了。同年(株)デンソー入社。2004 年日本機械学会

会畠山賞受賞。



杉本 茂樹

1995 年東京工業大学制御工学科卒業。1997 年東京工業大学大学院情報理工学研究科情報環境学専攻修士課程修了。2003 年東京工業大学大学院理工学研究科機械制御システム

専攻産学官連携研究員。コンピュータビジョン, 画像計測, ITS 関連の研究に従事。電子情報通信学会, IEEE 各会員。



奥富 正敏(正会員)

1981 年東京大学工学部計数工学科卒業。1983 年東京工業大学大学院理工学研究科制御工学専攻修士課程修了。同年キヤノン(株)入社。1987~1990 年カーネギーメロン大

学コンピュータサイエンス学科客員研究員。1994 年東京工業大学大学院情報理工学研究科情報環境学専攻助教授。2002 年同大学院理工学研究科機械制御システム専攻教授。コンピュータビジョン, 画像処理, 画像計測に関する研究に従事。工学博士。電子情報通信学会, 計測自動制御学会, 画像電子学会, IEEE 各会員。