

局所的バンドル調整を用いた自己運動推定における累積誤差の軽減化

池田 友彦[†] 杉本 晃宏^{††}

[†] 千葉大学大学院自然科学研究科
〒 263-8522 千葉県千葉市稲毛区弥生町 1-33

^{††} 国立情報学研究所
〒 101-8430 東京都千代田区一ツ橋 2-1-2
E-mail: †sugimoto@nii.ac.jp

あらまし 搭載した 2 台の能動カメラそれぞれを注視点制御することにより、カメラの自己運動を推定する手法が提案されている。しかし、その手法は逐次的に運動を推定するため、運動が長くなると推定誤差が累積し、運動軌跡の推定精度が次第に悪くなるという問題を抱えている。そこで、本稿では、時々刻々に推定されたカメラ運動のパラメタがそれまでの入力画像系列と幾何学的に整合するように局所的にバンドル調整を適用し、推定結果をそのつど修正する手法を提案する。そして、この修正によって長い運動に対しても安定な運動推定を実現することができることを示す。
キーワード 自己運動推定, 局所的バンドル調整, 累積誤差, 注視点.

Reducing Accumulated Errors in Ego-Motion Estimation using Local Bundle Adjustment

Tomohiko IKEDA[†] and Akihiro SUGIMOTO^{††}

[†] School of Science and Technology, Chiba University
Inage-ku, Chiba 263-8522, Japan
^{††} National Institute of Informatics
Chiyoda, Tokyo 101-8430, Japan
E-mail: †sugimoto@nii.ac.jp

Abstract The binocular independent fixation control is a camera control of two mounted active cameras where each of them independently and automatically fixates its optical axis to its own fixation point. A method using this camera control was proposed to incrementally estimate ego-motion from two time-series frames. The method, however, has a problem that estimation accuracy gradually becomes worse as the motion trajectory becomes longer and longer. This is due to accumulation of estimation errors incurred in each estimation step. To keep estimation accuracy stable even for a long trajectory, we propose to locally apply the bundle adjustment to each estimated motion so that the modified estimation becomes geometrically consistent with time-series frames acquired so far. This modification realizes stable estimation of a long motion trajectory.

Key words ego-motion estimation, local bundle adjustment, accumulated errors, fixation point.

1. はじめに

コンピュータの急速な高性能化, 小型化, 低価格化にともない, コンピュータが我々の生活により身近なものになってきている。そして, そう遠くない将来に, コンピュータを装着したまま日常生活することが現実味をおびつつある [1], [3], [5], [13]. そこでは, 人間が装着したセンサから得られる情報をもとに, 装着者の意図や興味を理解し, 有益な情報を自律的に提供するシ

ステムが必要となる [16] が, そのためには, 装着者が存在する環境とその環境内における装着者の位置の情報を知ることが重要である [1], [5], [21]. 一方, ロボットビジョンの分野でも, ロボットのナビゲーションに関する研究が盛んであり, ロボット自身がどのような環境に存在し, どこに位置しているかを認識する技術が必要となっている [4], [6], [7], [17], [18], [24], [28]. このように, カメラを装着した人間やカメラを搭載したロボットの, 3次元空間中における運動および位置を推定することは, 重要な

問題の一つとなっている。

Gonçalves ら [9] は、ステレオ視により 3 次元空間中の点を復元し、それとオプティカルフローを利用することで運動推定を行う手法を提案している。また、Molton ら [17] は、ステレオ視により 3 次元空間中の複数の点を復元し、運動前後におけるそれらの点の画像上での対応を求めることにより、運動推定を実現している。一方、Davison ら [6] は、ロボットの運動が平面的であるという前提の下で、ステレオ視により復元された空間中の点の情報と、平面運動における幾何拘束を利用したロボットの運動推定手法を提案している。

しかし、ステレオ視を用いた場合、2 台のカメラの視野を共有させることが不可欠となる。その上で、運動推定のためには 2 枚の画像において複数の特徴を対応づける必要があるが、正確にしかも安定して対応づけを得ることは困難である [8]。また、ステレオ視を利用する場合、その推定精度は 2 台のカメラの基線長に依存する [15], [20] が、ロボットにカメラを搭載する場合や人間がカメラを装着する場合、基線長が長くなるように 2 台のカメラを配置するのは難しく、高精度な運動推定が期待できない。

これに対し、杉本ら [11], [12], [21] ~ [23] は、2 台のアクティブカメラを独立に注視点制御する 2 眼独立注視点制御 (図 1) という考え方を導入し、搭載したカメラの基線長に依存しない推定精度を保証する運動推定手法を提案した。そこでは、2 眼独立注視点制御を通じて得られる注視点対応と、注視点周辺で得られる直線対応 [21], [22] やオプティカルフロー [11], [12], [23] を利用することで、カメラの自己運動を推定している。しかし、この手法は、高精度な運動推定が期待できる反面、時間的に連続する 2 フレームの画像を用いて逐次的に運動を推定しているため、運動が長くなると推定誤差が累積し、運動軌跡の推定精度が次第に悪くなるという問題を抱えている。

一方、カメラパラメータのキャリブレーション [27] では、投影モデルに基づいて推定されたパラメータの値に対して、再投影誤差 (パラメータの推定値の下での特徴点の投影位置の予測と実際に観測される投影位置との差) を考え、それを最小化するように推定値を修正するというアプローチがとられている。特に、マルチカメラの外部パラメータのキャリブレーションの際には、順番に選ばれたカメラ対に対して外部パラメータを求め、最後に、全てのカメラが全体として整合するようにパラメータ修正を行い、最終的な外部パラメータを求めるという手法が一般的である。このような再投影誤差の最小化は写真測量の分野で古くから行われており、バンドル調整 [10], [26] とよばれている。バンドル調整は、キャリブレーションに限らず、運動からの形状復元でも用いられている。

本稿では、杉本ら [11], [12], [21] ~ [23] の 2 眼独立注視点制御を用いた運動推定にバンドル調整を適用することで推定誤差の累積を軽減する手法を提案する。バンドル調整は、もともと、パラメータの推定が全て終わった後に、最後に一括して全パラメータの修正をはかるものであるが、ここでは、局所時間内に得られた推定結果にバンドル調整を適用する [29], [30] という考え方を採用する。そして、時々刻々に推定された運動がそれまでの入力

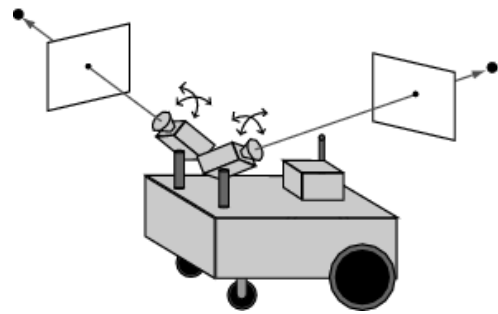


図 1 2 眼独立注視点制御

画像系列と幾何学的に整合するようにそのつどバンドル調整を適用し、誤差が累積しないように推定結果を随時修正する手法を提案する。また、この修正によって、長い運動に対しても安定な運動推定を実現することができることを実験によって示す。

2. 2 眼独立注視点制御を用いた運動推定

2 眼独立注視点制御 [21], [22] とは、2 台の装着型アクティブカメラを独立に注視点制御するカメラ制御のことである (図 1)。この制御は、各カメラの注視点から見れば、そのカメラを装着した人物、もしくは、そのカメラを搭載したロボットを対象として、ステレオ視を実行している状況であると解釈できる。この場合、注視点間距離がステレオ視における基線長に相当するので、注視点間距離が長くなるように 2 眼独立注視点制御を実行することで高精度な運動推定が可能となる。このように、2 眼独立注視点制御を用いた運動推定手法 [11], [12], [21] ~ [23] は、搭載した 2 台のカメラの基線長に依存せず、高精度な運動推定を実現する方法となっている。

以下では、2 眼独立注視点制御から得られる情報と注視点周辺のオプティカルフローから得られる情報とを統合することによってカメラの自己運動を推定する手法 [11], [12], [23] を概観する。

左右 2 台のカメラのうち右カメラを基準に考え、2 台のカメラを固定した台座の運動によって生じる右カメラの回転、並進運動をカメラ運動とよぶことにする。したがって、カメラ運動には、注視点制御のために行うカメラの首振り運動は含まれないとする。左右のカメラのカメラ座標系を一致させるための回転行列、並進ベクトルを R_{in}, T_{in} とする。また、時刻 t から $t+1$ の間におけるカメラ運動のパラメータを R, T とする。前者が回転を表す行列であり、後者が並進を表すベクトルである。なお、カメラの内部・外部パラメータは校正済みであるとする。

時刻 t において、右カメラの投影中心から右カメラの注視点に向かう単位ベクトルは、時刻 t の右カメラ座標系で v_r^t であるとする。このとき、右カメラの注視点対応からカメラ運動のパラメータに関する制約式

$$\det [R_0 v_r^t \mid R_0 R v_r^{t+1} \mid T] = 0$$

が得られる。ここで、 R_0 はワールド座標系の姿勢を、時刻 t の右カメラ座標系に一致させるための回転行列である。同様に、左カメラの注視点対応からは、

$$\det [R_0 R_{in}^T v_\ell^t \mid R_0 R_{in}^T v_\ell^{t+1} \mid \mathbf{T} - R_0(R-I)R_{in}^T \mathbf{T}_{in}] = 0$$

が得られる。ここで、 v_ℓ^t は、時刻 t において、左カメラの投影中心から左カメラの注視点に向かう単位ベクトルを表す（ただし、時刻 t の左カメラ座標系で表示）。

一方、時刻 t において、右カメラで撮影された画像の注視点周辺の領域内にある点 q_r^t のオプティカルフロー u_r^t が得られたとすると、それからカメラ運動パラメタに関する制約式

$$\det [R_0 R (M u_r^t + \tilde{q}_r^t) \mid R_0 \tilde{q}_r^t \mid \mathbf{T}] = 0$$

が得られる。ただし、 f_r を右カメラの焦点距離として、

$$M = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}^T, \quad \tilde{q}_r^t = ((q_r^t)^T, f_r)^T$$

とする。同様に、左カメラで画像の注視点周辺の領域内にある点 q_ℓ^t のオプティカルフロー u_ℓ^t が得られたとすると、それからカメラ運動パラメタに関する制約式

$$\det [R_0 R R_{in}^T (M u_\ell^t + \tilde{q}_\ell^t) \mid R_0 R_{in}^T \tilde{q}_\ell^t \mid \mathbf{T} - R_0(R-I)R_{in}^T \mathbf{T}_{in}] = 0$$

が得られる。ここに、 $\tilde{q}_\ell^t = ((q_\ell^t)^T, f_\ell)^T$ である (f_ℓ は左カメラの焦点距離)。

回転、並進ともに 3 自由度あるので、カメラ運動の自由度は 6 である。これに対し、注視点対応から得られるカメラ運動の制約式は 2 個である。また、オプティカルフローに基づく制約式は、1 点に対して 1 個ずつ独立に得られる。したがって、4 点以上に対してオプティカルフローに基づく制約式を得ることで不足分の制約式を補い、それらをまとめて非線形最適化問題として捉え、最適化手法を適用することで、カメラ運動のパラメタを推定することができる。

3. 局所的バンドル調整

前節に述べた自己運動推定手法は、時間的に連続する 2 フレームを用いて逐次的な運動推定を実現している。しかし、各時刻での運動推定には非線形最適化を行っているため、最適化の過程で局所解に陥る可能性があり、カメラ運動のパラメタを正しく推定できているとは限らない。たとえ大域的な最適解が求められたとしても、観測誤差に起因する推定誤差が生じている。したがって、運動の長さが長くなればなるほど推定誤差が累積し、運動軌跡の推定精度が次第に悪くなる。このため、各画像の注視点とそれ以前に得られた画像系列の注視点とを推定された運動パラメタによって幾何学的に関係づけようとする、そこに不整合が生じる。ここでは、バンドル調整の考え方を取り入れて、この不整合が生じないように運動パラメタの推定値を修正することを考える。

いま、離散的な時刻で画像系列が得られるとし、時刻 s から時刻 t ($t \geq s+2$) までのカメラ運動のパラメタが連続する 2 フレーム間で逐次的に求められているとする。そして、時刻 i ($i = s, \dots, t-1$) と時刻 $i+1$ の間で推定された運動パラメタ

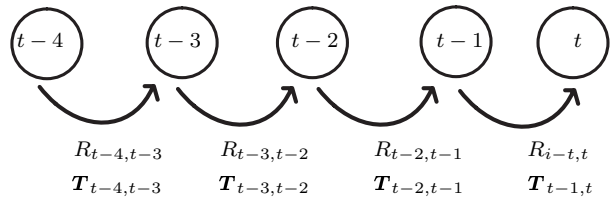


図 2 隣接する時刻間におけるカメラ運動のパラメタ

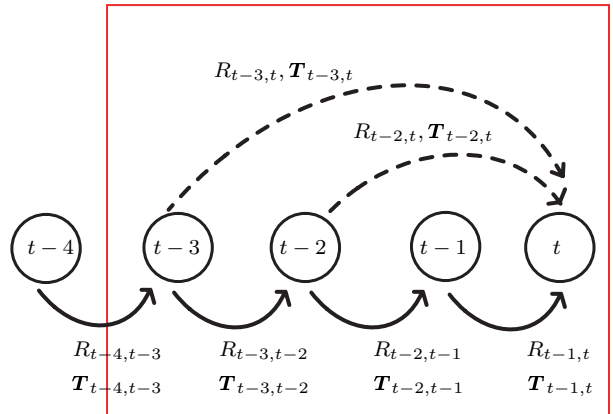


図 3 局所的バンドル調整を適用する範囲内のカメラ運動のパラメタの関係 (4 フレームの場合)

を $R_{i,i+1}, T_{i,i+1}$ とする (図 2)。ここに、 $R_{i,i+1}, T_{i,i+1}$ は、それぞれ、時刻 i から時刻 $i+1$ におけるカメラ運動の回転行列、並進ベクトルを表すとする。なお、注視点制御によって、時刻 s から時刻 t の間の各画像間で、注視点の像 p_i ($i = s, \dots, t$) の対応が得られているとする。

このとき、時刻 t の画像において、時刻 t を含む過去 n 時刻 ($3 \leq n \leq t-s+1$) に得られた画像系列に注目する。そして、この画像系列に対してバンドル調整を適用することを考える。時間的に連続する 2 フレーム間のカメラ運動のパラメタが推定されているので、これを合成することによって、時刻 j ($t-n+1 \leq j \leq t-1$) から時刻 t までのカメラ運動 $R_{j,t}, T_{j,t}$ を得ることができる (図 3 参照):

$$R_{j,t} = \prod_{i=j}^{t-1} R_{i,i+1}, \quad T_{j,t} = \sum_{i=j}^{t-1} T_{i,i+1}.$$

そこで、これを用いて、 p_j によって定まるエピポーラ線を時刻 t の画像上に考える。本来、このエピポーラ線は p_t を通るはずであるが、推定誤差やその累積のため、図 4 に示す例のように、 p_t を通らない。これは、各画像の注視点とそれ以前に得られた画像系列の注視点とが幾何学的に不整合である、すなわち、エピポーラ拘束を満たさないことを意味する。

p_j 、および、 $R_{j,t}, T_{j,t}$ によって得られる時刻 t の画像上のエピポーラ線は、時刻 t の画像上の点の同次座標を \tilde{x} とすると、

$$\tilde{x}^T E_{j,t} \tilde{p}_j = 0 \quad (1)$$

と表される。ただし、 \tilde{p}_j は p_j の同次座標であり、

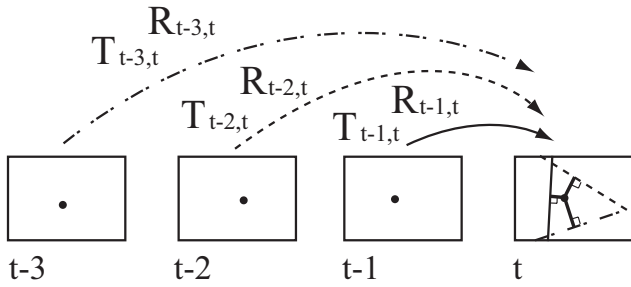


図4 注視点と推定したパラメタから得られるエピポーラ線

$$E_{j,t} = [T_{j,t}]_{\times} R_{j,t}$$

である(注1)。したがって、 \tilde{p}_t とエピポーラ線 (1) とのずれによって、 $R_{j,t}, T_{j,t}$ の誤差を評価することができる。そこで、 $(n-1)$ 枚の画像から得られるエピポーラ線群と \tilde{p}_t とのずれの和を評価し、それを最小化するように最新の運動パラメタ $R_{t-1,t}, T_{t-1,t}$ を修正する。幾何学的距離でエピポーラ線とのずれを評価(図5参照)すると、最小化するべき評価関数は、

$$\sum_{j=t-n+1}^{t-1} \left(\frac{\tilde{p}_t^T E_{j,t} \tilde{p}_t}{\sqrt{a_{j,t}^2 + b_{j,t}^2}} \right)^2 \quad (2)$$

となる。ただし、 $(a_{j,t}, b_{j,t}, c_{j,t})^T = E_{j,t} \tilde{p}_j$ とした。

式(2)を最小化するように運動パラメタ $R_{t-1,t}, T_{t-1,t}$ を修正することによって、過去 n 時刻に得られた画像系列において、注視点に関するエピポーラ拘束が満たされるような修正が実現される。ここで修正するパラメタは $R_{t-1,t}, T_{t-1,t}$ のみであるので、その数は n の値に依存しない。したがって、局所的なバンドル調整を適用する時間範囲(フレーム数)を変更しても計算の手間はほとんど変わらないことになる。

以上のように時々刻々に推定された運動パラメタに対して、局所的にバンドル調整を適用し、そのつど推定結果を修正することによって、入力画像系列の注視点と推定された運動パラメタが幾何学的に整合するようになる。その結果、修正前の運動パラメタに含まれていた誤差が軽減され、逐次的な推定を行っても誤差の累積が抑えられ、推定精度の向上が期待される。なお、式(2)は運動パラメタ $R_{j,t}, T_{j,t}$ に関して非線形方程式となっているので、Powell法[19]などの非線形最適化手法を用いることで評価関数の最小化を行うことができる。

4. 実験

提案手法を用いて、カメラの自己運動を推定する実験を行った。まず始めに、シミュレーションにより、提案手法を用いることでどの程度推定誤差が軽減されるかを定量的に評価した。また、何フレームに対して局所的バンドル調整を適用するのが運動推定に適しているかを評価した。次に、実環境下において、カメラの自己運動軌跡を推定する実験を行った。

(注1): $[T_{j,t}]_{\times}$ は、 $T_{j,t}$ から構成される 3×3 の歪対称行列である: 任意の 3 次元ベクトル y に対して、 $[T_{j,t}]_{\times} y = T_{j,t} \times y$ が成立する。

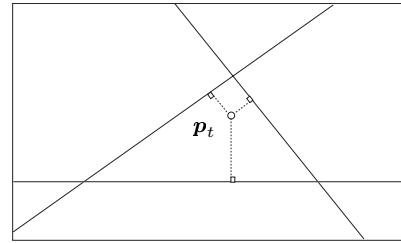


図5 注視点とエピポーラ線のずれの評価

4.1 シミュレーションによる定量的評価

ここでは、提案手法によって誤差がどの程度軽減されるかをシミュレーション画像を用いて定量的に評価した。比較のために杉本らの手法[11], [12], [23] (以下、従来手法とよぶことにする)を実装し、そこで生じる誤差に対する軽減の度合いを評価した。まず、前節において $n=3$ とした場合、すなわち、過去の3フレームに対して局所的バンドル調整を適用した場合において、提案手法と従来手法による運動パラメタの推定誤差を比較した。

カメラの基線長を距離の単位とした。また、カメラの画角は 21.0° とし、焦点距離を 0.025 とした。さらに、右カメラと左カメラの姿勢は一致しているとし、得られる画像のサイズを $512\text{pixel} \times 512\text{pixel}$ とした。一方、注視点の基線からの奥行きを 10.0 (一定) とした。

実験では、注視点間距離が 10.0 となるように左右のカメラの注視点が発生させた。さらに各画像中で、注視点の周りにオプティカルフローを算出する点を注視点の周辺領域内に3個ずつランダムに発生させた。なお、注視点の周辺領域は、画像上で注視点の像を中心とする $20\text{pixel} \times 20\text{pixel}$ の範囲に投影される領域とした。

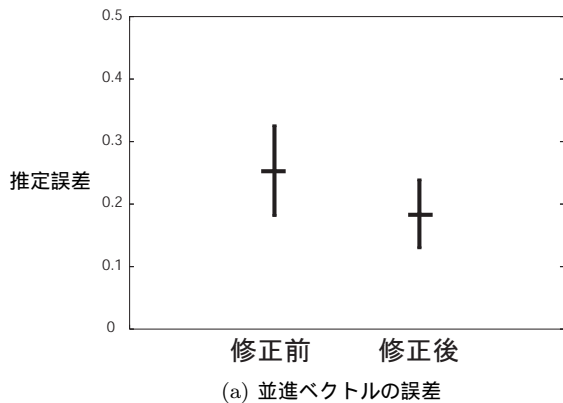
3フレームの画像を得るために2ステップの運動が発生させた。第1ステップにおける並進ベクトルは、ノルムが 0.25 になるようにランダムに発生させ、回転行列は、回転角がカメラ座標系の Z 軸周りに 1.0° となるように発生させた。第2ステップの並進に関しては、第1ステップで発生させた並進ベクトルをカメラ座標系の Z 軸周りに $-1.0 \sim 1.0^\circ$ の間でランダムに回転させた。また、第2ステップにおける回転に関しては、第1ステップにおけるカメラの回転角に $-1.0 \sim 1.0^\circ$ の間でランダムに発生させた回転角を加えることで発生させた。

カメラ運動の前後において、各点をそれぞれの撮像面に投影し、その投影された点に対してノイズを加えた。ここでは、平均 0.0pixel 、標準偏差 2.0pixel の正規分布に従って発生させた乱数をノイズとして加えた。そして、各ステップにおいて、カメラの運動の前後で画像上で各点の対応を算出し、カメラ運動の推定を行った。

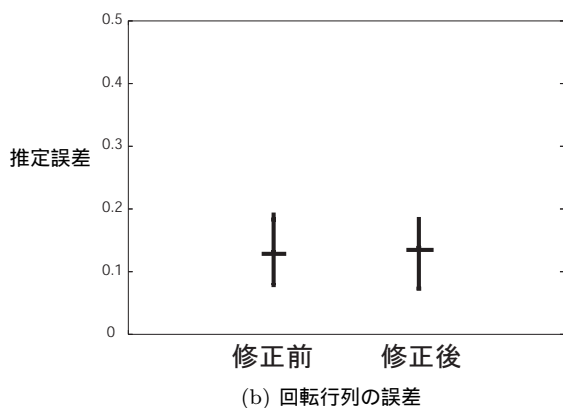
推定されたカメラの回転 \hat{R} 、および、並進 \hat{T} の誤差は、以下の式で評価した。

$$\frac{\|\hat{\theta}\hat{r} - \theta r\|}{|\theta|}, \quad \frac{\|\hat{T} - T\|}{\|T\|}$$

ここでは、誤差評価のため回転行列 \hat{R} を、回転軸 \hat{r} ($\|\hat{r}\| = 1$)



(a) 並進ベクトルの誤差



(b) 回転行列の誤差

図 6 注視点間距離を 10 基線長とした場合の推定誤差

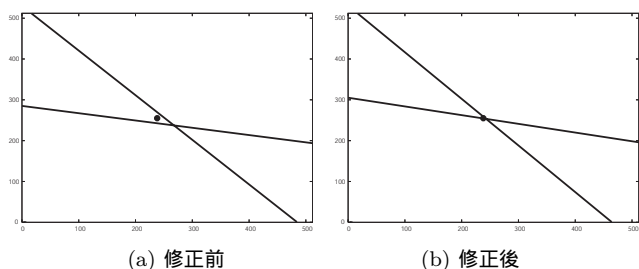


図 7 注視点と修正前後の推定パラメタから得られたエピポーラ線

と回転角 $\hat{\theta}$ に分解した。 r ($\|r\| = 1$) は回転軸の真値であり、 θ は回転角の真値である。また、 T は並進ベクトルの真値を表す。上記の過程を 200 回繰り返し、200 回の繰り返しにおける推定誤差の平均値と標準偏差を求めた。

図 6 にその実験結果を示す。これより、提案手法による推定誤差の軽減が読みとれる。まず、図 6(b) をみると、回転行列に関しては従来手法と提案手法とで推定誤差にあまり差異がないことが分かる。回転行列に関しては、もともと推定精度が十分高い [11], [12], [23] ので、それほどの改善が得られなかったと考えられる。しかし、図 6(a) をみると、並進ベクトルに関して、提案手法を用いた方が推定誤差の平均値、および、標準偏差がともに小さくなっていることが分かる。この差に意味があるかどうかを判定するために、有意水準 5% でウェルチの検定 [25] を

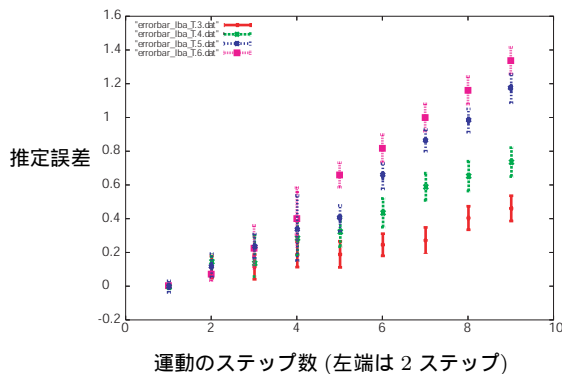


図 8 局所的バンドル調整を適用するフレーム数と並進運動の推定誤差

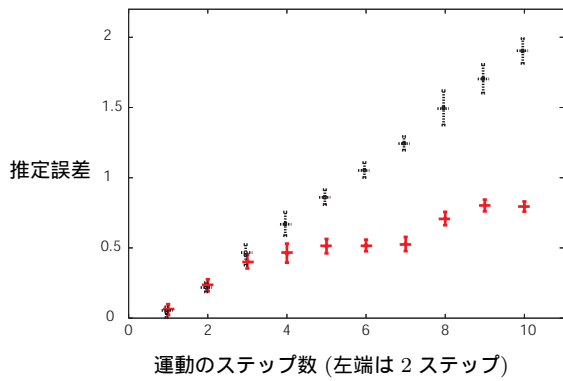
行ったところ、この差は統計的に有意であるとの結論を得た。

200 回のうちのある試行に対して、エピポーラ線を実際に引いた結果を図 7 に示す。図 7 から、修正によって、過去のフレームから引かれたエピポーラ線を注視点を通るようになっていことが分かる。この時の注視点と各エピポーラ線とのおよその距離は、修正の前後でそれぞれ、10pixel, 0.2pixel であった。200 回の試行における注視点とエピポーラ線との距離はほとんどの場合、この程度であった。

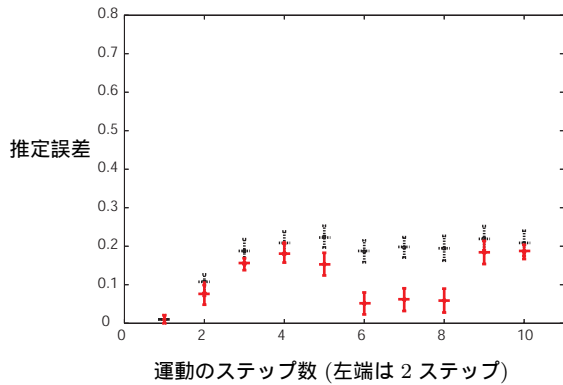
200 回全ての試行において、修正前後で評価関数の値が小さくなっていることを確認しているが、評価関数の値が小さくなることが必ずしも推定値が真の値に近づくことを意味するとは限らないため(回転と並進での相殺が起こり得るため)、推定精度の改善にはある程度の限界があつて然るべきである。この点を踏まえ、提案手法を用いることにより、逐次推定された運動パラメタが入力画像系列と幾何学的に整合し、推定精度は向上していると結論できる。

次に、局所的バンドル調整を適用するフレーム数と推定誤差との関係を見るために、 $3 \leq n \leq 10$ の場合に対して、上と同一の手順で実験を行った。ここでは、10 ステップの運動を発生させた。なお、回転運動ではあまり改善が見られていないので、ここではカメラの並進運動の推定のみに着目した。この時の実験結果のうち、 $3 \leq n \leq 6$ の場合を図 8 に示す。それぞれの場合において、必要なフレーム数が得られるまではバンドル調整による修正は行っていない。たとえば、 $n = 5$ の場合は、4 ステップの運動以降に対してのみバンドル調整を適用している。

図 8 から、 $n = 3$ の場合、最も推定精度がよいことがわかる。一見、局所的バンドル調整を適用するフレーム数をふやせばふやすほど推定誤差の軽減がよりはかられるように思うが、実際は、そうではない。なぜなら、フレーム数が多くなると、バンドル調整を適用するまでの累積誤差が大きくなってしまっていて、バンドル調整を適用する拘束条件の信頼性が低下していると考えられるからである。それゆえ、バンドル調整ではもはやその誤差を十分に軽減できなくなってしまうと考えられる。 $n = 3$ の場合の誤差が最も小さいというのはこのことを裏付けている。累積誤差が大きくならないうちに随時バンドル調整を



(a) 並進ベクトル



(b) 回転行列

図9 3フレームを用いた局所的バンドル調整を適用した場合の累積誤差 (赤は提案手法, 黒は従来手法を表す)

適用し, 誤差を軽減する必要がある。

誤差を最も抑えるには, 3フレームに対して局所的バンドル調整を適用すればよいことがわかったので, $n = 3$ の場合に対して, 累積誤差がどのように変化するかを評価した。このときの実験の手順も同一である。この実験で発生させた運動は, 11ステップである。

図9に結果を示す。比較のため, 従来手法で推定を行った場合についても結果を示した。回転運動に関しては, やはり, 提案手法, 従来手法ともに, あまり誤差の累積がみられない。しかし, 並進運動に関しては, 図9(a)から容易に読みとれるように, 従来手法では運動のステップ数がふえるにほぼ比例して誤差が累積しているのに対し, 提案手法では誤差の累積の割合は大幅に改善されている。これより, 提案手法によって, 長い運動に対しても安定に運動推定を実現することができると考えられる。

4.2 実画像を用いた運動軌跡推定

前節でのシミュレーション実験の結果をふまえて, 実環境下において, 提案手法を用いてカメラの自己運動軌跡を推定する実験を行った。また, 提案手法の有効性を確認するため, 従来手法との比較も行った。

本実験では, アクティブカメラとしてSONY製EVI-G20を用いた。そして, 基線長が約20cmとなるように2台のカメラを三脚に配置し, 注視点間距離が2~3m程度になるように, 2



図10 実験時のカメラ配置



図11 実験環境

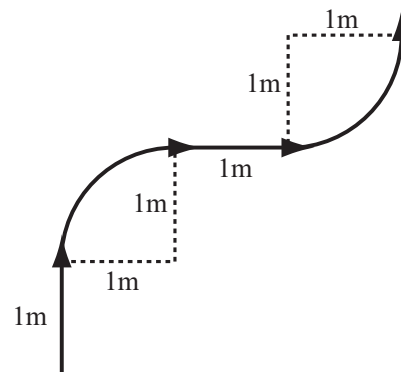


図12 カメラ運動の軌跡

台のカメラの向きを設定した(図10)。このとき, カメラ座標系のZ軸は, カメラの光軸と平行であるとし, ワールド座標系は, 初期位置における右カメラのカメラ座標系に一致しているとした。そして, 右カメラの投影中心の運動軌跡を推定した。

図11に示す軌跡(約5m)に沿って実験室内でカメラを運動させ, 提案手法, および, 従来手法を用いて, その運動を推定した。実験では, 図12に示したように, カメラ運動の軌跡を130ステップに刻んだ(2つの直線部分がそれぞれ20ステップ, 2つの円弧の部分がそれぞれ45ステップである)。そして, 各ステップごとに提案手法, および, 従来手法を適用した。なお, それぞれの手法では, 計6点を左右のカメラの画像から均等に選り, そのオプティカルフローを用いた。

表 1 130 ステップにおける推定誤差の平均と標準偏差

	修正前	修正後
平均 [cm]	0.268	0.0394
標準偏差 [cm]	0.102	0.00739

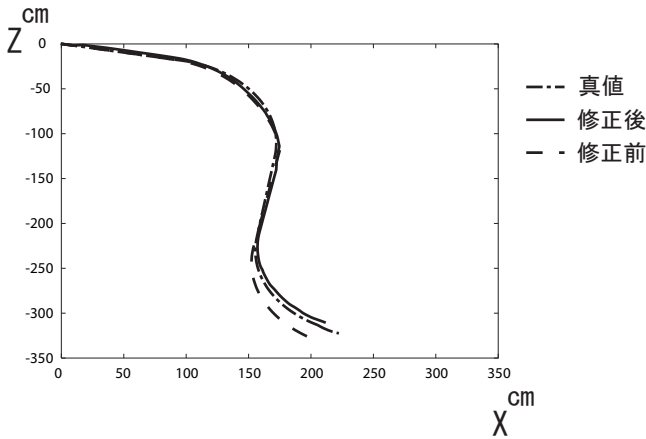


図 13 カメラ位置の推定結果

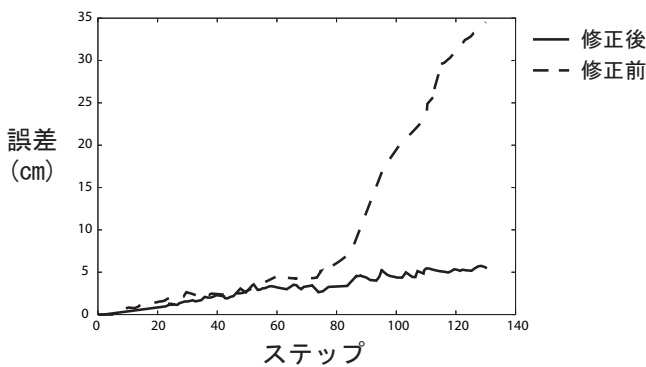


図 14 カメラ位置の推定誤差

注視点制御における注視点の対応付けは手動で行った。また、運動の過程において、注視点画像から外れてしまった場合には注視点の更新が必要となるが、この更新は手動で行った（更新回数は左右のカメラともに7回であった）。一方、各カメラで得られた画像ごとに、注視点を中心とする100pixel×100pixelの領域においてオプティカルフローを算出し[2], [14]、さらに、注視点を中心とする30pixel×30pixelの領域からランダムに3点ずつ選んだ。そして、運動推定には、この6点のオプティカルフローを利用した。また、提案手法では、時間的に連続する3フレームにおける左右のカメラそれぞれで、注視点対応を用いて局所的バンドル調整を行った。

提案手法、および、従来手法を利用してカメラ運動を推定した各ステップでの結果をつなぎ合わせることによって、運動軌跡の推定結果を得た。得られたカメラの推定位置とその真値をワールド座標系のXZ平面にそれぞれ投影し、重ね合わせた結果を図13に示す。また、提案手法および従来手法それぞれについて、各ステップにおけるカメラ位置の推定誤差を図14に示す。図13、図14において、実線が提案手法を用いた結果であり、点線が従来手法を用いた結果である。また、1点鎖線が真値を表している。表1に、130ステップの運動における位置の推定誤差の平均と標準偏差を示す。

図13をみると、従来手法に比べ、提案手法を用いた方が真値に近い結果が得られていることが分かる。図14をみると、特に、80ステップ目あたり（軌跡の2つ目の円弧部分）において、従来手法は推定に失敗して真の軌跡から大きなずれが生じ始めているが、提案手法では、修正によって誤差が著しく減少し、軌跡推定に成功していることがわかる。最終的な推定位置の誤差は、従来手法では約35cm、提案手法では約4cmと、大幅な改善がみられる。表1からも、提案手法の方が130ステップの位置推定における誤差の平均、および、標準偏差がともに小さくなっていて、高精度で安定に推定を行っていることがわかる。

5. おわりに

2眼独立注視点制御を用いた運動推定[11], [12], [21]~[23]は、時間的に連続する2フレームの画像による逐次推定手法であるため、運動が長くなると推定誤差が累積し、運動軌跡の推定精度が次第に悪くなるという問題を抱えていた。そこで、本稿では、バンドル調整を局所時間内に得られた運動パラメータの推定結果に適用し、そのつど推定結果を修正する手法を提案した。これによって、入力画像系列の注視点と運動パラメータが幾何学的な整合性を保つようになり、逐次的に推定された運動パラメータに含まれていた誤差が軽減され、逐次型運動推定の精度向上が確認された。また、提案手法により、長い運動に対しても安定な運動推定を実現することができることを実験によって示した。

本稿では、左右のカメラの注視点対応に対して局所的バンドル調整を適用したが、画像中のより多くの特徴点を用いて過去の画像系列との幾何学的な整合性を総合的に評価することで、さらなる推定精度の向上が期待される。この点に関しては、今後検討する予定である。また、局所的バンドル調整を選択的に適用するなどの工夫によって処理の高速化をはかることも今後の課題として残されている。

謝 辞

本研究の一部は、科学研究費補助金13224051、および、14380161の補助を受けて行った。

文 献

- [1] H. Aoki, B. Schiele and A. Pentland: *Realtime Personal Positioning System for Wearable Computers*, Vision and Medline Technical Report, TR-520, Media Lab. MIT, 2000.
- [2] J.-Y. Bouquet: *Pyramidal Implementation of the Lucas Kanade Feature Tracker Description of the Algorithm*, Intel Corporation, Microprocessor Research Labs, OpenCV Documents, 1999.
- [3] B. Clarkson, K. Mase and A. Pentland: *Recognizing User's Context from Wearable Sensors: Baseline System*, Vision and Modeling Technical Report, Vismod TR-519, Medial Lab. MIT, 2000.
- [4] A. J. Davison: *Mobile Robot Navigation Using Active Vi-*

- tion, Ph.D thesis, University of Oxford, 1998.
- [5] A. J. Davison, W. W. Mayol and D. W. Murray: Real-Time Localisation and Mapping with Wearable Active Vision, *Proc. of IEEE/ACM Int. Symposium on Mixed and Augmented Reality*, pp. 18-27, 2003.
- [6] A. J. Davison and D. W. Murray: Mobile Robot Localisation using Active Vision, *Proc. of ECCV*, Vol. 2, pp. 809-825, 1998.
- [7] G. N. DeSouza and A. C. Kak: Vision for Mobile Robot Navigation: A Survey, *IEEE Trans. on PAMI*, Vol. 24, No. 2, pp. 237-267, 2002.
- [8] D. A. Forsyth and J. Ponce: *Computer Vision : A Modern Approach*, Prentice Hall, 2003.
- [9] N. Gonçalves and H. Araújo: Estimation of 3D Motion from Stereo Images, *Proc. of 16th International Conference on Pattern Recognition (ICPR)*, Vol. I, pp. 335-338, 2002.
- [10] R. Hartley and A. Zisserman: *Multiple View Geometry in Computer Vision*, Cambridge Univ. Press, 2000.
- [11] 池田友彦, 杉本晃宏: アクティブカメラの注視点制御に基づく自己運動の推定, 画像の認識・理解シンポジウム (MIRU2004), Vol. 1, pp. 39-44, 2004.
- [12] 池田友彦, 杉本晃宏: アクティブカメラの注視点制御に基づく自己運動の推定, 電子情報通信学会論文誌, Vol. J88-D-II (2005), No. 8, to appear.
- [13] M. Kouroggi, T. Kurata and K. Sakaue: A Panorama-Based Method of Personal Positioning and Orientation and Its Real-Time Applications for Wearable Computers, *Proc. of Int. Symposium on Wearable Computers*, pp. 107-114, 2001.
- [14] D. D. Lucas and T. Kanade: An Iterative Image Registration Technique with an Application to Stereo Vision, *Proc. of Imaging Understanding Workshop*, pp. 121-130, 1981.
- [15] J. Matas, O. Chum, M. Urban and T. Pajdla: Robust Wide Baseline Stereo from Maximally Stable Extremeal Regions, *Proc. of BMVC*, 2002.
- [16] 松山隆司, 杉本晃宏, 佐藤洋一, 川嶋宏彰: 人間と共生する情報システムの実現を目指して, 人工知能学会誌, Vol. 19, No. 2, pp. 257-266, 2004.
- [17] N. Molton and M. Brady: Practical Structure and Motion from Stereo When motion is Unconstrained, *Int. J. of Computer Vision*, Vol. 39, No. 1, pp. 5-23, 2003.
- [18] T. Nakagawa, T. Okatani and K. Deguchi: Active Construction of 3D Map in Robot by Combining Motion and Percived Images, *Proc. of ACCV*, vol. 1, pp. 563-568, 2004.
- [19] W. H. Press, B. P. Flannery, S. A. Teukolsky and W. T. Vetterling: *Numerical Recipes in C*[日本語版], 技術評論社, 1993.
- [20] P. Pritchett and A. Zisserman: Wide Baseline Stereo Matching, *Proc. of 14th International Conference on Pattern Recognition (ICPR)*, pp. 754-760, 1998.
- [21] 杉本晃宏, 長友渉, 松山隆司: 装着型能動視覚センサを用いた人物の位置および運動の推定, 画像の認識・理解シンポジウム (MIRU2002), Vol. 1, pp. 215-222, 2002.
- [22] A. Sugimoto, W. Nagatomo and T. Matsuyama: Estimating Ego Motion by Fixation Control of Mounted Active Cameras, *Proc. of ACCV*, pp. 67-72, Jeju, Korea, 2004.
- [23] A. Sugimoto and T. Ikeda: Diverging Viewing-Lines in Binocular Vision: A Method for Estimating Ego Motion by Mounted Active Cameras, *Proc. of the 5th Workshop on Omnidirectional Vision, Camera Networks and Non-classical Cameras*, pp. 67-78, Prague, Czech Republic, 2004.
- [24] Y. Sumi, Y. Ishiyama and F. Tomita: 3D Localization of Moving Free-Form Objects in Cluttered Environments, *Proc. of ACCV*, vol. 1, pp. 43-48, 2004.
- [25] 東京大学教養学部統計学教室編: 統計学入門, 東京大学出版会, 1991.
- [26] B. Triggs, P. McLauchlan, R. Hartley and A. Fitzgibbon: Bundle Adjustment – A Modern Synthesis, *Vision Algorithms: Theory and Practice (B. Triggs, A. Zisserman and R. Szeliski eds.) Lecture Notes in CS*, Vol. 1883, pp. 298-372, Springer, 2000.
- [27] 植芝俊夫, 岡谷貴之, 佐藤智和: カメラキャリブレーション, 情報処理学会 CVIM 研究会, CVIM148, pp. 1-18, 2005.
- [28] M. Werman, S. Banerjee, S. Dutta Roy and M. Qiu: Robot Localisation Using Uncalibrated Camera Invariants, *Proc. of IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, Vol. II, pp. 353-359, 1999.
- [29] Z. Zhang and Y. Shan: *Incremental Motion Estimation through Local Bundle Adjustment*, Technical Report MSR-TR-01-54, Microsoft Research, May 2001.
- [30] Z. Zhang and Y. Shan: Incremental Motion Estimation through Modified Bundle Adjustment, *Proc. International Conference on Image Processing (ICIP)*, Vol. II, pp. 343-346, 2003.