

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

佐藤 智和[†] 神原 誠之[†] 横矢 直和[†] 竹村 治雄^{††}

Camera Parameter Estimation from a Long Image Sequence by Tracking Markers and Natural Features

Tomokazu SATO[†], Masayuki KANBARA[†], Naokazu YOKOYA[†], and Haruo TAKEMURA^{††}

あらまし 動画像からのカメラパラメータの復元は、三次元モデル復元、物体認識、ロボットナビゲーション、複合現実感など、様々な分野への応用が可能である。しかし、屋外環境のように複雑で広大な環境では、動画像上の特徴点が頻繁に出現・消失するために、安定した特徴点の追跡が困難となりカメラパラメータ推定の誤差が累積する。この問題を回避するため、本研究では三次元位置関係が既知の複数個の基準マーカを利用する。提案手法では、まず基準マーカを動画像の初期フレームの画像上で指定し、第2フレーム目以後で自然特徴点を画像上で自動追跡することによって、撮影時のカメラの外部パラメータと自然特徴点の三次元位置を逐次推定する。次に、最終フレームを含む少数の画像上で基準マーカを指定して動画像全体での最適化を行うことにより、累積的な外部パラメータの推定誤差を最小化する。また、自然特徴点の追跡と三次元復元の処理を統合的に扱うことで、安定した自然特徴点の追跡を行う。実験では、屋外環境を撮影した動画像を入力とした復元結果を示し、評価を行うことで本手法の有効性を示す。

キーワード 三次元復元, 単眼動画像, カメラパラメータ, 自然特徴点, マーカ

1. はじめに

動画像からのカメラパラメータの復元は、三次元モデル復元、物体認識、ロボットナビゲーション、複合現実感など、様々な分野への応用が可能である。そのため動画像内の見かけの特徴点の動きから、特徴点の三次元位置とカメラパラメータを自動的に復元する手法 (shape from motion) に関する研究が盛んに行われている [1] ~ [8]。

それらの代表的な手法に、カメラモデルを線形近似し、線形解法によって撮影対象の形状とカメラの運動パラメータを同時に求める因子分解法 [1], [2] がある。この手法は安定かつ高速に対象の復元を行うことがで

きるが、カメラモデルを線形近似するためにカメラパラメータの復元精度に問題がある。さらに、動画像の全てのフレームに撮影されている特徴点を用いることを前提としているため、特徴点が全てのフレームに渡って連続して観測できない一般的な動画像からの三次元復元は困難である。近年、不完全なデータに同手法を適用する手法 [9] や透視投影モデルへの適用を試みる手法 [10] が提案されているが、これらの問題を同時に解決することは難しく、複雑な屋外環境データへの手法の適用は困難である。

また、射影的な復元を利用する手法が提案されている [4] ~ [7]。これらは、各フレームにおいて、逐次的に射影的なカメラパラメータを復元し、射影座標系からユークリッド座標系への変換を行うことでカメラパラメータを復元する。しかし射影的な復元では、多くの自由度を残したまま逐次的な三次元復元を行うため、特徴点の誤対応による累積的な推定誤差が大きくなり、数十枚程度の画像からの復元を行うにとどまっている。また、これらの手法では現実世界と復元されるデータ

[†] 奈良先端科学技術大学院大学 情報科学研究科, 奈良県

Graduate School of Information Science, Nara Institute of Science and Technology, 8916-5 Takayama, Ikoma, Nara, 630-0101 Japan

^{††} 大阪大学 サイバーメディアセンター, 大阪府

Cybermedia Center, Osaka University, 1-32 Machikaneyama, Toyonaka, Osaka, 560-0043 Japan

との位置関係およびスケールの情報が失われるため、複数回に分けて撮影した動画像間の復元データを統合して広域な環境を復元するのは難しい。

一方、複数の奥行き画像を三次元空間内で統合することによって、撮影対象の三次元モデルと撮影時のカメラパラメータを同時に復元する手法が提案されている [11], [12]。しかし、カメラパラメータの良い初期値が必要であることや、重複領域を持つ高精度な奥行き画像を多数必要とするためにレーザレンジファインダなどの特殊な機材が必要であることが問題である。また、一般にレーザレンジファインダによる奥行き画像計測には時間がかかるため、広域な環境を密に撮影するためには撮影コストが膨大となる。

そこで我々は、屋外環境を動画像として撮影し、三次元位置が既知の特徴点（以下、基準マーカ）と自然特徴点を画像上で同時に追跡することでカメラの外部パラメータを推定する手法 [8] を提案し、推定されたカメラパラメータを用いてマルチベースラインステレオ法によってシーンの三次元モデルを復元する手法 [13] に応用している。本論文では、文献 [13] の前提となっていた基準マーカと自然特徴点の追跡によるカメラの外部パラメータの復元手法に関して詳述し、定量的評価を行う。

本手法では、基準マーカをトータルステーションと呼ばれる三次元測量機材を用いて計測し、初期フレームと最終フレームを含む少数の画像上で基準マーカを指定し、全体での最適化処理を行うことで、カメラの外部パラメータの累積的な推定誤差を最小化する。また、追跡中の自然特徴点の三次元位置を逐次更新し、次のフレームでの特徴点の探索に利用することで、自然特徴点の安定した追跡を実現している。これにより、複雑で広域な環境を撮影した千枚程度から成る動画像のカメラの外部パラメータを効率的かつ精度良く推定することが可能となる。ただし、本論文で用いる基準マーカは、自然特徴点または環境中に人工的に設置される色・形状が既知のマーカを想定しており、基準マーカの三次元位置および画像上の追跡座標は常に信頼できるものとする。

以下、2章では、特徴点（基準マーカと自然特徴点）を画像上で追跡し、カメラの外部パラメータと自然特徴点の三次元位置を復元する手法について述べる。3章では、実際に屋外環境を撮影した動画像を用いて実験と評価を行い、本手法の特徴と有用性を考察する。最後に4章で、本稿のまとめと今後の課題を述べる。

2. 特徴点の追跡によるカメラの外部パラメータの復元

本章では、動画像中の特徴点を追跡し、カメラの外部パラメータを推定する手法について述べる。ただし、カメラの内部パラメータは既知であるとする。提案手法では、図1に示すように、まず初期フレームにおいて三次元位置が既知の6個以上の基準マーカを画像上で指定する。これにより、初期フレームにおけるカメラの外部パラメータが推定される。次に、フレーム毎の処理（図1中A）を初期フレームから最終フレームまで繰り返すことにより、全てのフレームにおけるカメラの外部パラメータと自然特徴点の三次元位置を逐次的に推定する。この手法では蓄積誤差が含まれるため、最終フレーム、中間フレームなどの少数の画像上で基準マーカを指定し、推定されたカメラの外部パラメータと自然特徴点の三次元位置を入力動画像全体で最適化する（図1中B）。これにより、累積する外部パラメータの推定誤差を最小化する。また、本手法では、自然特徴点の三次元位置と信頼度を逐次更新し、次のフレームでの自然特徴点の追跡に利用することで、マーカが撮影されていないフレームにおいても精度良くカメラの外部パラメータを復元する。以下、各処理について詳述する。

2.1 逐次処理によるカメラの外部パラメータの推定

本節では、フレーム毎の処理（図1中A）によって、カメラの外部パラメータと自然特徴点の三次元位置を推定する手法の各処理について順に述べる。本手法は、追跡中の自然特徴点の三次元位置を逐次更新し、次のフレームでの特徴点の探索に利用することで、安定した自然特徴点の追跡を実現する。また、自然特徴点の信頼度や再投影誤差などの複数の評価尺度を用いることで、フレーム毎に自動で自然特徴点を追加・削除することができる。

2.1.1 マーカと自然特徴点の追跡

一般に、画像上の情報のみを用いた自然特徴点の追跡手法 [14], [15] における問題点として、(a) フレームが進むに従って増加するテンプレートの中心位置のずれ、(b) 周辺類似パターンとの誤対応、が挙げられる。本研究では、(a) を解決するために、入力画像上のエッジのコーナなどの追跡の容易な点を Harris のインタレストオペレータ [16], [17] を用いて検出し、この出力が極大値となる点を自然特徴点の追跡位置の候補とす

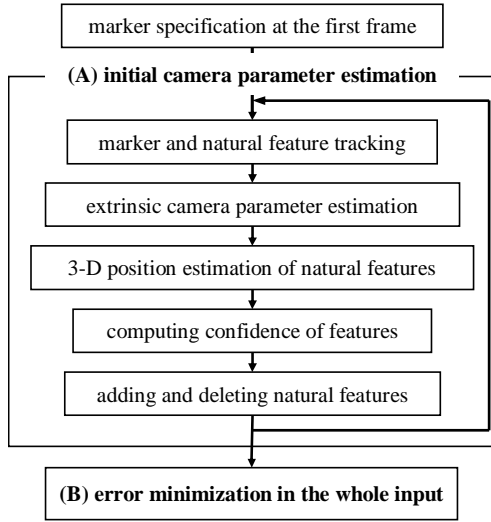


図 1 提案手法の処理の流れ

Fig. 1 Flow diagram of camera parameter estimation and 3-D reconstruction .

る。また、テンプレートをフレーム毎に更新することで、自然特徴点の見え方の変化に対応する。また (b) に関しては、既に推定されている自然特徴点の三次元位置とロバスト推定 [18] によって算出される暫定的なカメラパラメータを用いて、追跡すべき自然特徴点の探索範囲を限定することで解決する。

以下に、第 f フレーム ($f \geq 2$) における特徴点の追跡処理手順を述べる。

(1) 第 $(f-1)$ フレームにおける各基準マーカに対して、画像上での追跡を行なう。基準マーカとして、人工的に設置される色・形状が既知のマーカを利用する場合には、色・形状の情報を用いた一定範囲の窓内での自動追跡を行なう [8]。また、自然特徴点の一部を基準マーカとして利用する場合には、他の自然特徴点の三次元位置が精度良く推定されるまでの数十フレームの間、手動での追跡を行なう。

(2) 第 $(f-1)$ フレームで追跡されている各自然特徴点に対し、その特徴点の画像上の位置を中心とする探索窓内で SSD を評価尺度としたテンプレートマッチングを用いることで第 f フレームにおける仮の追跡を行う。ただし、テンプレートマッチングは Harris のインタレストオペレータ [16] の出力が極大値となる点に対してのみ行う。

(3) 以下、ロバスト推定の繰り返し処理を開始する。 i 回目の繰り返し処理においては、まずステップ (2) で

追跡された特徴点からランダムに n 個 $P_i = \{p_{i1}, p_{i2}, \dots, p_{in}\}$ を選択し、次節に述べる手法によって仮の外部パラメータ \hat{M}_i を推定する。次に、追跡された全ての特徴点に対して、図 2 に示すような、特徴点 p が追跡された画像上の座標 x_{fp} と特徴点 p の三次元座標 S_p を \hat{M}_i によって画像上に投影した座標 \hat{x}_{fp} との距離の二乗誤差 R_{ifp} (以下、再投影誤差) の中央値 RM_i を算出する。ただし、ここでは S_p として第 $(f-1)$ フレームで推定されている三次元位置を用いる。なお、再投影誤差 R_{ifp} とその中央値 RM_i は以下の式によって表される。

$$R_{ifp} = |x_{fp} - \hat{x}_{fp}|^2 \quad (1)$$

$$RM_i = \text{med}(R_{if1}, R_{if2}, \dots, R_{ifm}) \quad (2)$$

ただし、 m はステップ (2) で仮に追跡された特徴点の数である。

(4) ステップ (3) を g 回繰り返して得られる複数の仮の外部パラメータ $(\hat{M}_1, \dots, \hat{M}_g)$ に対して、以下に示す LMedS 基準を満たす外部パラメータを選択し、暫定パラメータ \bar{M}_f とする。

$$\text{LMedS} = \min(RM_1, RM_2, \dots, RM_g) \quad (3)$$

(5) 暫定パラメータ \bar{M}_f を用いて第 $(f-1)$ フレームにおける全ての特徴点 p の三次元座標 S_p を画像上に投影し、その座標を中心とする探索窓内でステップ (2) と同様の対応づけを行い追跡位置を確定する。ただし、ここで用いる探索窓はステップ (2) で用いた探索窓より小さくする。

2.1.2 カメラの外部パラメータの推定

本節では、特徴点の画像上の二次元座標と、特徴点の世界座標における三次元座標を用いてカメラの外部パラメータを推定する。本手法においては、式 (1) で示した再投影誤差 R_{fp} を誤差尺度として利用する。これにより、第 f フレームにおけるカメラの外部パラメータ M_f は次の誤差関数 E_f を最小化することで推定される。

$$E_f = \sum_p W_p R_{fp} \quad (4)$$

ここで、 W_p は特徴点 p における重みであり、後述する特徴点 p の信頼度により決定する。また、外部パラメータ M_f の自由度は 6 とし、カメラ座標系の基

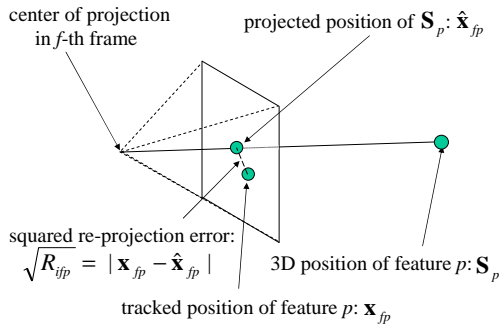


図 2 再投影の誤差
Fig. 2 Reprojection error.

基底ベクトルは単位直交条件を満たすものとする。このような誤差関数 E_f の最小化は非線形最小化問題であり、局所解や計算量の問題があるため、本手法では出口らの手法 [19] を利用し、最小自乗法によって外部パラメータの初期値を線形演算によって推定し、補正を行った後に、非線形最適化を行う。以下に外部パラメータの推定手順を示す。

(1) 出口らの手法 [19] を利用し、線形最小自乗法によってカメラの外部パラメータの初期値を推定する。ここで推定される外部パラメータは、カメラ座標系の基底ベクトルが単位直交条件を満たしておらず、自由度が 12 となるため、推定されるカメラの位置がノイズに敏感となる。

(2) ステップ (1) で推定された初期値を補正し、外部パラメータの自由度を 6 に補正する。ステップ (1) で用いた手法では、特徴点の画像上での追跡誤差が、推定されるカメラ位置の光軸方向に大きく影響する。このため、本手法では真のカメラ位置が推定された光軸上に存在すると仮定し、線形演算によって外部パラメータの自由度を 6 に補正する。

(3) ステップ (2) で求めた外部パラメータを初期値とし、最急降下法により E_f を最小化する。先に行った線形演算によって求めた初期値は真値に近い推定値であると考えられるため、局所解に陥ることを回避し、少ない演算回数で大域最小解を求めることができると考えられる。これにより、再投影誤差の重み付きの和である E_f を最小化する M_f を求める。

2.1.3 自然特徴点の三次元位置の推定

各フレームにおいて推定されたカメラパラメータと自然特徴点 p の画像上の座標 x_{fp} から、その自然特徴点の三次元位置 S_p を推定し、毎フレーム更新する。ここでは、カメラの外部パラメータの推定と同様に、

式 (1) で示した再投影誤差 R_{fp} を誤差尺度として利用する。自然特徴点の三次元位置 S_p は、次の誤差関数 E_p を最急降下法によって最小化することで推定する。

$$E_p = \sum_f R_{fp} \quad (5)$$

S_p の初期値は、自然特徴点の画像上の座標とカメラの投影中心を結ぶ直線群との距離の二乗和が最小となる三次元位置を、線形最小自乗法によって算出することで得る。

2.1.4 特徴点の信頼度の算出

第 f フレームにおいて算出される自然特徴点の信頼度は、第 $(f+1)$ フレームにおけるカメラの外部パラメータ推定の重みと、第 f フレームにおける自然特徴点の追加・削除の指標として用いられる。本節では第 f フレームにおける自然特徴点の信頼度の算出について述べる。

先に求めた第 f フレームにおける特徴点 p の画像上の座標 x_{fp} は、追跡誤差により一般的に、 p の三次元位置 S_p を画像上に投影した座標 \hat{x}_{fp} と一致しない。そこで、この追跡誤差が \hat{x}_{fp} 周辺に分散 σ_p^2 をもつガウス分布を成すものとしてモデル化を行う。このとき、 x_{fp} が真値である確率密度関数 $p(x_{fp})$ は次式のようになる。

$$p(x_{fp}) = \frac{1}{2\pi\sigma_p^2} \exp\left(-\frac{|x_{fp} - \hat{x}_{fp}|^2}{2\sigma_p^2}\right) \quad (6)$$

これを用い、第 f フレームにおける全ての特徴点 p の確率密度関数の積 P_f を考える。

$$P_f = \prod_p p(x_{fp}) \quad (7)$$

このような確率密度関数の積 P_f を最大とする外部パラメータ M_f は、

$$EM_f = \sum_p \frac{|x_{fp} - \hat{x}_{fp}|^2}{2\sigma_p^2} \quad (8)$$

を最小化することにより得られる。本手法では、 σ_p^2 として第 $(f-1)$ フレームまでの再投影誤差の分散を用い、式 (4) と式 (8) を比較することにより、第 $(f-k)$ フレームから第 $(f-1)$ フレームまで追跡されている特徴点 p の信頼度 W_p を次式のように定義する。

$$W_p = \frac{1}{2\sigma_p^2} = \frac{k}{2} \left\{ \sum_{i=f-k}^{f-1} |x_{ip} - \hat{x}_{ip}|^2 \right\}^{-1} \quad (9)$$

2.1.5 自然特徴点の追加・削除

本論文では、広域な屋外環境を撮影対象とした動画画像を扱うため、カメラの動きによって画像上での自然特徴点の出現・消失が起こる。そこで、複数の評価尺度を用いることで、フレーム毎に自動で自然特徴点の追加・削除を行なう。

自然特徴点の追加は、画像上の情報のみを用いて追跡された自然特徴点の候補から、以下の条件を全て満たすものを選択することで実現する。

- 信頼度が一定値以上
- 各画像間のマッチング誤差が一定値以下
- Harris オペレータによる特徴量が一定値以上
- 推定される三次元位置と各フレームのカメラの投影中心を結ぶ直線群の成す最大角が一定値以上

また、以下のいずれかの条件を満たす自然特徴点を削除する。

- 信頼度が一定値以下
- 各画像間のマッチング誤差が一定値以上
- 再投影誤差が一定値以上

2.2 動画全体での最適化

これまでの処理は逐次的に行うため各フレームにおける処理は短時間でできるが、誤差が蓄積するという問題がある。そこで、2.1 節で用いた基準マーカに加えて、最終フレームを含む少数のフレームの画像上で基準マーカを指定し、全体での最適化を行うことで、累積的な誤差を最小化する。ここでは、外部パラメータ M_f と自然特徴点 p の三次元位置 S_p をパラメータとし、最後に全フレームに対して最適化を行う。最適化の誤差関数には次式を用い、 M_f と S_p の初期値には逐次処理によって得られた推定結果を用いる。

$$E = \sum_f \left(A_f \sum_p W_p |x_{fp} - \hat{x}_{fp}|^2 \right) \quad (10)$$

ここで W_p は、逐次処理で得た自然特徴点 p の信頼度である。 A_f は、フレームの信頼度による重みであり、ここでは、基準マーカが存在しないフレームでは 1 を、基準マーカが存在するフレームでは 1 に対して十分に大きな値を設定する。これにより、マーカが存在するフレームでの推定結果を真値とみなして、動画全体の再投影誤差を最小化することができる。この最適化は非線形最小化問題となるが、初期値が真値に近ければ解は大域最小解に収束しやすくなり、また高速に求めることができる。

3. 実験

提案手法の有効性を確認するため、実環境における建物として奈良・朱雀門を、広域で複雑な環境として大阪・住道の町並みをビデオカメラで撮影し、カメラパラメータの復元実験を行った。両実験ともに、ワイドレンズ (Sony VCL-HG0758) を取り付けられたビデオカメラ (Sony DSR-DP-150) を用いて撮影を行い、得られた動画画像を入力とした。ただし、カメラの内部パラメータは Tsai の手法 [20] によりあらかじめ得た値を用いた。

3.1 建物を対象としたカメラパラメータの復元

本実験においては、手持ちのビデオカメラを用いて建物を中心とする円弧を描くように撮影を行い、図 3(a) を含む画像 982 枚 (720×480 画素、プログレッシブ撮影) を得た。また、あらかじめ、図 3(b) に印で示す点を含む複数の自然特徴点を基準マーカとし、その三次元位置関係を、三次元測量機器トータルステーション (Leica TCR1105) を用いて計測した。基準マーカの画像上の位置は初期フレーム、第 500 フレーム、最終フレームの画像上で指定し、第 2 フレーム以降、第 100 フレームまでの基準マーカの画像上の座標を手動で追跡した。また、その他のフレームにおいては基準マーカを利用しなかった。

図 3(b) に特徴点の追跡結果を示す。図中の印は指定した基準マーカを、白点は追跡された自然特徴点を表している。同図に示すように、多数の自然特徴点が追加・削除を伴って安定に追跡された。本実験においては、各フレームで平均 190 点の自然特徴点がそれぞれ平均 100 フレーム間自動で追跡された。実験には、PC (Pentium4 2GHz, Memory 1GB) を用い、カメラの外部パラメータの復元処理に要した時間は、自然特徴点の自動追跡を含む逐次的な復元処理に約 2800 秒、全体での最適化処理における 100 回の繰り返し計算に約 4700 秒であり、全体最適化後の自然特徴点の平均投影誤差は、0.95 画素であった。

図 4 に復元された自然特徴点の三次元位置とカメラパラメータの推定結果を示す。曲線は推定されたカメラパスを、錘台は 50 フレーム毎のカメラの姿勢を表す。本実験において復元されたカメラパスの長さは約 76m であった。同図から分かるように、カメラパスは滑らかに推定されており、また画像上で基準マーカを指定しなかったフレームにおいても、自然特徴点を手掛かりにカメラの位置・姿勢を復元することができた。



(a) Input images (b) Result of tracking features

図 3 入力画像と特徴点の追跡結果 (建物)

Fig. 3 Input images and result of feature tracking (building).

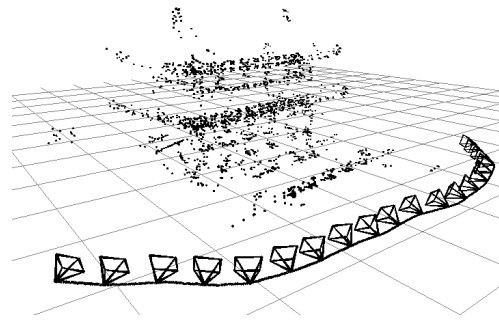


図 4 カメラパラメータの推定結果 (建物)

Fig. 4 Result of camera parameter estimation (building).

3.2 町並みを対象としたカメラパラメータの復元

本実験では、ビデオカメラを車の上に三脚で固定し、町並みを撮影しながら走行することで、図 5(a) を含む画像 500 枚 (720×480 画素、プログレッシブ撮影) を得た。また、前節の建物の復元実験と同様に、図 5(b) に印で示す自然特徴点を基準マーカとし、三次元位置関係をトータルステーションを用いて座標系を統一して計測した。基準マーカの画像上の位置は、初期フレーム、第 250 フレーム、最終フレームの画像上で指定し、第 2 フレーム以降、第 50 フレームまで基準マーカの画像上の座標を手動で追跡した。また、その他のフレームにおいては基準マーカを利用しなかった。

図 5(b) に特徴点の追跡結果を示す。図中の印は指定した基準マーカを、白点は追跡された自然特徴点を表している。本実験において追跡された自然特徴点の数は各フレーム平均 110 点であり、それぞれ平均 20 フレーム間自動で追跡された。また、実験には前節と同じ計算機を用い、カメラの外部パラメータの復元処理に要した時間は、自然特徴点の自動追跡を含む逐次的な復元処理に約 1000 秒、全体での最適化処理における 100 回の繰り返し計算に約 1800 秒であり、全体最適化後の自然特徴点の平均投影誤差は、0.82 画素であった。

図 6 に復元された自然特徴点の三次元位置とカメラパラメータの推定結果を示す。前節と同様に、曲線は推定されたカメラパスを、錘台は 50 フレーム毎のカメラの姿勢を表す。本実験において復元されたカメラパスの長さは約 130m であった。同図から分かるように、オクルージョンの発生する複雑な町並みを撮影した動画像を入力として用いた場合にも、カメラ位置の軌跡は滑らかに推定されており、また画像上で基準



(a) Input images (b) Result of tracking features
 図 5 入力画像と特徴点の追跡結果 (町並み)
 Fig. 5 Input images and result of feature tracking (street scenery).

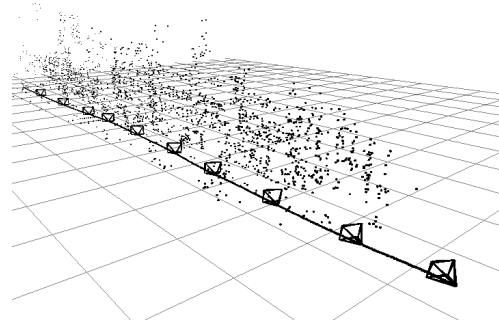


図 6 カメラパラメータの推定結果 (町並み)
 Fig. 6 Result of camera parameter estimation (street scenery).

マーカを指定しなかったフレームにおいても、自然特徴点を手掛かりにカメラの位置・姿勢を復元することができた。

3.3 カメラの外部パラメータの推定精度の評価

上述の実験により推定されたカメラの外部パラメータの定量的な評価を行う。一般に、屋外環境においてカメラの外部パラメータの真値を得ることは困難であるため、本実験では全ての入力画像上で基準マーカを手動で追跡することで誤追跡による誤差が累積しない外部パラメータを求め、この外部パラメータとの差を推定誤差とみなすことで評価を行った。ここでは基準マーカとして、3.1 節における建物を対象とした復元実験では各フレームにおいて平均 30 点、3.2 節における町並みを対象とした復元実験では各フレームにおいて平均 20 点を指定した。

建物を対象とした復元実験により推定された外部パラメータと基準マーカのみで復元された外部パラメータの各フレームにおける位置の推定誤差を図 7 に示す。また各フレームにおける光軸ベクトルの角度の推定誤差をグラフで図 8 に示す。図より、2.2 節で述べた全体最適化によって、外部パラメータ推定の累積誤差が軽減されていることが分かる。また、建物を対象とした復元実験においては、全体最適化前のカメラ位置の推定誤差は平均 240mm、光軸方向の推定誤差は平均 0.20°であり、全体最適化後にはそれぞれ、平均 110mm、平均 0.09°となった。

同様に、図 9、図 10 に、町並みを対象とした実験で復元した外部パラメータと基準マーカのみで復元された外部パラメータとの推定誤差を示す。町並みを対象とした復元実験においては、全体最適化前のカメラ位置の推定誤差は平均 220mm、光軸方向の推定誤差は

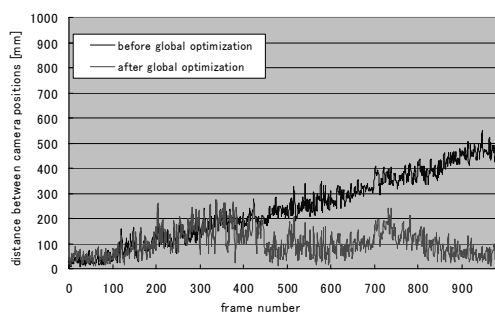


図 7 カメラ位置の推定誤差 (建物)

Fig. 7 Estimation errors of camera positions (building)

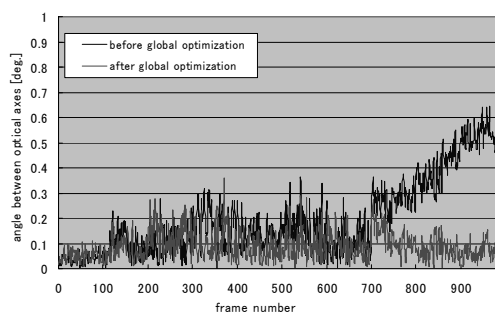


図 8 カメラ光軸の推定誤差 (建物)

Fig. 8 Estimation errors of camera rotation (building)

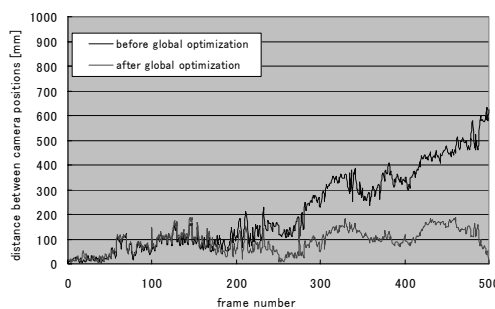


図 9 カメラ位置の推定誤差 (町並み)

Fig. 9 Estimation errors of camera positions (street scenery)

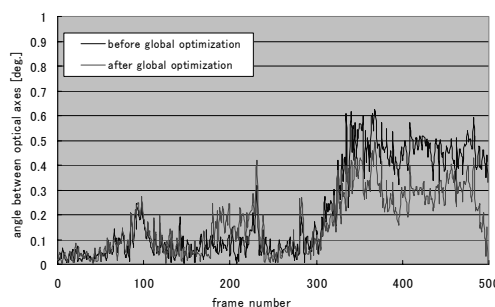


図 10 カメラ光軸の推定誤差 (町並み)

Fig. 10 Estimation errors of camera rotation (street scenery).

平均 0.22° であり, 全体最適化後にはそれぞれ, 平均 90mm , 平均 0.17° となった. 建物の復元実験と同様に, 全体最適化によってカメラの外部パラメータ推定の累積誤差が小さくなっていることが分かる. ただし, 図 10 に示したカメラの光軸方向の推定誤差に関して, 全体最適化前には, 中間フレームと最終フレームでの推定誤差が周辺のフレームの推定誤差に比べて小さくなっていたため, 他の結果と比較して全体最適化の明確な効果が現れていない.

4. まとめ

本論文では, まず基準マーカと自然特徴点を画像上で追跡することにより逐次的な三次元復元を行い, 次に全体最適化によって累積的な推定誤差を最小化することで, 精度良くカメラの外部パラメータを推定する手法を提案した. 実験により, 提案手法では多くのオクルージョンを含む千枚程度の画像から成る動画像に対しても累積誤差を抑えてカメラパラメータの復元を行うことが可能であることを確認した. 具体的には, カメラの外部パラメータの復元精度について定量的な評価を行い, 100m 程度の移動を伴う屋外環境の撮影においては, カメラ位置に関して平均誤差 100mm 程度, カメラの光軸方向に関して平均誤差 0.2° 程度で外部パラメータを得ることが可能であることを確認した.

今後は, 高解像度全方位カメラ [21] を用いた動画像からの三次元復元手法や, 推定したカメラパラメータを用いた三次元モデル獲得手法の検討を行う予定である.

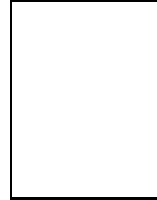
文献

- [1] C. Tomasi and T. Kanade: "Shape and Motion from Image Streams under Orthography: A Factorization Method," *Int. J. of Computer Vision*, Vol. 9, No. 2, pp. 137–154, 1992.
- [2] J. Poleman and T. Kanade: "A Paraperspective Factorization Method for Shape and Motion Recovery," *Tech. Rep. CMU-CS-93-219*, Carnegie-Mellon Univ., 1993.
- [3] R. Szeliski and S. B. Kang: "Recovering 3D Shape and Motion from Image Streams Using Non-linear Least Squares," *J. of Visual Communication and Image Representation*, Vol. 6, No. 1, pp. 10–28, 1994.
- [4] P. Beardsley, A. Zisserman and D. Murray: "Sequential Updating of Projective and Affine Structure from Motion," *Int. J. of Computer Vision*, Vol. 23, No. 3, pp. 235–259, 1997.
- [5] H. S. Sawhney, Y. Guo, J. Asmuth and R. Kumar: "Multi-view 3D Estimation and Application to Match Move," *Proc. IEEE Workshop on Multi-view*

- Modeling and Analysis of Visual Scenes, pp. 21–28, 1999.
- [6] G. Roth and A. Whitehead: “Using Projective Vision to Find Camera Positions in an Image Sequence,” Proc. 13th Int. Conf. on Vision Interface, pp. 87–94, 2000.
- [7] M. Pollefeys, R. Koch, M. Vergauwen, A. A. Deknuydt and L. J. V. Gool: “Three-dimensional Scene Reconstruction from Images,” Proc. SPIE, Vol. 3958, pp. 215–226, 2000.
- [8] T. Sato, M. Kanbara, H. Takemura and N. Yokoya: “3-D Reconstruction from a Monocular Image Sequence by Tracking Markers and Natural Features,” Proc. 14th Int. Conf. on Vision Interface, pp. 157–164, 2001.
- [9] D. W. Jacobs: “Linear Fitting with Missing Data: Applications to Structure-from-Motion and to Characterizing Intensity Images,” Proc. IEEE Conf. on Computer Vision and Pattern Recognition, pp. 206–212, 1997.
- [10] 出口: “透視投影条件下での因子分解法による多視点画像からの形状復元”, 情処研報 CVIM106-6, 1997.
- [11] 増田: “符号付距離場の一致による複数距離画像からの形状モデル生成”, 画像の認識・理解シンポジウム (MIRU 2002) 講演論文集, Vol. I, pp. 1–10, 2002.
- [12] 西野, 池内: “大規模距離画像群の頑健な同時位置合わせ”, 電子情報通信学会論文誌 (D-II), Vol. J85-D-II, No. 9, pp. 1413–1424, 2002.
- [13] 佐藤, 神原, 竹村, 横矢: “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”, パーチャルリアリティ学会論文誌, Vol. 7, No. 2, pp. 275–282, 2002.
- [14] C. Tomasi and T. Kanade: “Detection and Tracking of Point Features,” Tech. Rep. CMU-CS-91-132, Carnegie-Mellon Univ., 1991.
- [15] 藤井, 神原, 竹村, 横矢: “拡張現実のためのジャイロセンサを併用したステレオカメラによる位置合わせ”, 信学技報, MVE99-59, 2000.
- [16] C. Harris and M. Stephens: “A Combined Corner and Edge Detector,” Proc. Alvey Vision Conf., pp. 147–151, 1988.
- [17] C. Schmid, R. Mohr and C. Bauckhage: “Comparing and Evaluating Interest Points,” Proc. 6th Int. Conf. on Computer Vision, pp. 230–235, 1998.
- [18] 栗田, 坂上: “ロバスト統計とその画像理解への応用”, 画像の認識・理解シンポジウム (MIRU 2000) 講演論文集, Vol. I, pp. 65–70, 2000.
- [19] 出口: “射影幾何学による PnP カメラ補正問題の統一的解法”, 情処シンポジウム, Vol. 90, pp. 41–50, 1990.
- [20] R. Y. Tsai: “An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision,” Proc. IEEE Conf. on Computer Vision and Pattern Recognition, pp. 364–374, 1986.
- [21] S. Ikeda, T. Sato and N. Yokoya: “An Calibration Method for an Omnidirectional Multi-camera Sys-

tem,” Proc. SPIE, Vol. 5006, 2003.

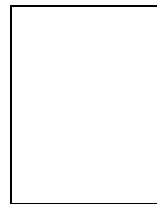
(平成 × 年 ×× 月 ×× 日受付)



佐藤 智和 (学生員)

1999年阪府大・工・情報工卒。2001年奈良先端科学技術大学院大学情報科学研究科博士前期課程修了。現在,同大博士後期課程に在学中。コンピュータビジョンの研究に従事。2001年電子情報通信学会学術奨励賞受賞。情報処理学会会員,IEEE各

会員。



神原 誠之 (正員)

1997年岡山大・工・情報工卒。2002年奈良先端科学技術大学院大学情報科学研究科博士後期課程修了。現在,同大情報科学研究科助手。コンピュータビジョン,拡張現実感の研究に従事。1999年電子情報通信学会学術奨励賞受賞。博士(工)。情報処理学会,日本バーチャルリアリティ学会,IEEE各会員。

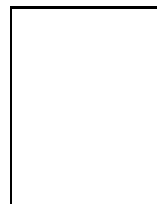
会員。



横矢 直和 (正員)

1974年阪大・基礎工・情報工卒。1979年同大大学院博士後期課程了。同年電子技術総合研究所入所。以来,画像処理ソフトウェア,画像データベース,コンピュータビジョンの研究に従事。1986~87年マツギル大・知能機械研究センター客員教授。

1992年奈良先端科学技術大学院大学・情報科学センター教授。現在,同大情報科学研究科教授。1990年情報処理学会論文賞受賞。工博。情報処理学会,日本バーチャルリアリティ学会,人工知能学会,日本認知科学会,映像情報メディア学会,IEEE各会員。



竹村 治雄 (正員)

1982年阪大・基礎工・情報工卒。1987同大大学院博士後期課程単位取得退学。同年(株)ATR入社。3次元ユーザインタフェース,CSCW,仮想現実の研究に従事。1994年奈良先端科学技術大学院大学・情報科学研究科助教授。2001年大阪大学サイバーメディアセンター教授。工博。情報処理学会,日本バーチャルリアリティ学会,IEEE,ACM,HFES,映像情報メディア学会,ヒューマンインタフェース学会各会員。

abstract

Camera parameter reconstruction from an image sequence is very useful for many applications such as 3D model reconstruction, object recognition, robot navigation, and mixed reality. However, there is a problem of the precision in the camera parameter reconstruction from many images because feature tracking errors and estimation errors are accumulated. In this paper, we propose a camera parameter reconstruction method which is based on using a number of markers with known position, color and shape, as well as natural features. The camera parameters and 3D positions of natural features are estimated efficiently in every frame by tracking these markers and natural features. Finally the accumulation of estimated error is minimized by specifying the markers in the last frame and optimizing parameters globally in the whole input. We demonstrate and evaluate experimental results of 3D reconstruction from real image sequences for showing the validity of the proposed method.

Key words: 3D reconstruction, Monocular image sequence, Camera parameters, Natural features, Markers