

5 A B - 3

3次元線形結合を用いた因子分解法による 物体形状とカメラ運動の復元

黄 國彰 山澤 一誠 岩佐 英彦 竹村 治雄 横矢 直和

奈良先端科学技術大学院大学 情報科学研究科

1 はじめに

近年、ステレオ画像や時系列画像から物体形状やカメラ運動を復元する研究が盛んに行なわれている[1~3]. 前者は、複数カメラを必要とするのに対して、後者は、単一のビデオカメラにより撮影された時系列画像から物体形状とカメラ運動を復元する。後者の代表的な手法の一つに因子分解法がある[1].

従来の因子分解法は、特異値分解を用いて計測行列と呼ばれる画像内の特徴点の時系列データを分解することにより物体形状とカメラ運動を復元する。本研究では、3次元線形結合を用いた因子分解法により計測行列を分解する手法を提案し、因子分解法の高速度を試みる。rank-theoremによると計測行列の階数が3であるならば必ず3つの独立ベクトルが存在するため、計測行列をそれらの線形結合で表現できる。このようにして、直接3つの独立ベクトルを見つければ、比較的簡単に分解ができ、高速化が可能である。以下では、提案手法の概要と有効性を検証した実験について述べる。

2 3次元線形結合を用いた因子分解法

本研究では、以下の前提条件をおく。

- カメラモデルが正射影で近似できる(対象物を遠距離から望遠レンズで撮影する様な場合)。
- 時系列画像の各フレームで同じ特徴点がすべて抽出され、対応がとれている。
- 少なくとも4つの特徴点を3フレームにわたって追跡する[4].

2.1 計測行列とその階数

第1フレームからP個の特徴点を抽出した後、時系列画像F枚において特徴点を順次追跡して式(1)のような計測行列Wを得る。

$$W = \begin{bmatrix} X_{fp} \\ Y_{fp} \end{bmatrix}, f = 1 \dots F, p = 1 \dots P \quad (1)$$

ここで、 X_{fp}, Y_{fp} は、各フレームの特徴点の座標からなる行列である。

A New Factorization Method Based on 3D Linear Combination for Shape and Motion Recovery
Kuo-chang Hwang, Kazumasa Yamazawa, Hidehiko Iwasa, Haruo Takemura and Naokazu Yokoya
Nara Institute of Science and Technology (NAIST)
8916-5 Takayama, Ikoma, Nara 630-01, Japan.

次にWから各フレームの特徴点の重心($\bar{X}_f = \frac{1}{P} \sum_{p=1}^P X_{fp}, \bar{Y}_f = \frac{1}{P} \sum_{p=1}^P Y_{fp}$)を引き、各フレームの特徴点位置をその重心に対する相対位置で表す行列 \tilde{W} (式(2))を求める。

$$\tilde{W} = W - \begin{bmatrix} \bar{X}_f \\ \bar{Y}_f \end{bmatrix} e_p^T \quad (2)$$

ここで、 e_p は $1 \times P$ の行列であり、 $e_p = [1 \dots 1]^T$ である。

以下で、カメラモデルが正射影の場合は、計測行列 \tilde{W} の階数が3以下であることを示す。

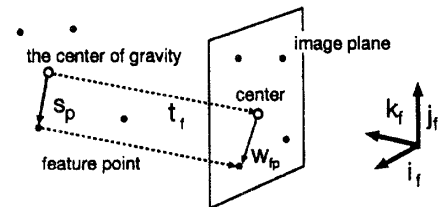


図1: カメラの正射影モデル

図1では、f番目の画像のx,y成分の単位ベクトルを i_f, j_f 、空間中の特徴点を s_p 、特徴点の重心から画像面の中心へのベクトルを t_f とする。正射影の場合、画像への投影線は、 $k_f = i_f \times j_f$ と平行である。このとき、各画像上への投影は式(3)になる。ここで、特徴点の重心をワールド座標の原点、すなわち、 $\sum_{p=1}^P s_p = 0$ とし、式(3)を式(2)に代入すれば、式(4)が得られる。式(4)を整理すれば、式(5)になる。

$$x_{fp} = i_f^T (s_p - t_f), y_{fp} = j_f^T (s_p - t_f) \quad (3)$$

$$\tilde{x}_{fp} = i_f^T s_p, \tilde{y}_{fp} = j_f^T s_p \quad (4)$$

$$\tilde{W} = \begin{bmatrix} i_f^T \\ j_f^T \end{bmatrix} [s_p] = MS \quad (5)$$

Mは各フレームの姿勢を、Sは特徴点の座標を表す行列である。つまり、 \tilde{W} をカメラの運動Mと物体の形状Sに分解できる。MとSの階数が高々3であるため、 \tilde{W} の階数は3以下である。

2.2 3つの独立ベクトルによる分解

\tilde{W} の階数が3以下であるということは、 \tilde{W} は、必ずある3つの独立ベクトルで構成される空間に存在することを意味する。即ち、 \tilde{W} の各行は、3つの独立ベクトルで構成された空間に射影され、係数を表す

行列 \hat{M} と3つの独立ベクトルを表す行列 \hat{S} の3次元線形結合方程式で表せる(式(6)).

$$\hat{W} = \hat{M}\hat{S} \quad (6)$$

ここで, 3つの独立ベクトルを $\hat{S} = [\hat{s}_k]$ と定義する($k = 1, 2, 3$). \hat{W} は \hat{S} で構成された空間に射影されている. 言い換えれば, \hat{W} から適当にある行 \hat{w}_{k_1} を1つの独立ベクトル \hat{s}_1 と定めることができる(式(7)). そして, \hat{s}_1 との内積が最も小さい \hat{w}_{k_2} を \hat{W} から選び出し, \hat{s}_2 を \hat{s}_1 と直交させるために式(8)のように \hat{s}_2 を定める. 同様に式(9)のように \hat{s}_3 を定める.

$$\hat{s}_1 = \hat{w}_{k_1} \quad (7)$$

$$\hat{s}_2 = \hat{w}_{k_2} - \frac{\hat{s}_1^T \cdot \hat{w}_{k_2}}{\|\hat{s}_1\|^2} \hat{s}_1 \quad (8)$$

$$\hat{s}_3 = \hat{w}_{k_3} - \frac{\hat{s}_1^T \cdot \hat{w}_{k_3}}{\|\hat{s}_1\|^2} \hat{s}_1 - \frac{\hat{s}_2^T \cdot \hat{w}_{k_3}}{\|\hat{s}_2\|^2} \hat{s}_2 \quad (9)$$

\hat{w}_f は, \hat{s}_k で構成された空間に射影されているため, $\hat{w}_f = \hat{m}_{f_1}\hat{s}_1 + \hat{m}_{f_2}\hat{s}_2 + \hat{m}_{f_3}\hat{s}_3$ となるような3次元線形結合で表現できる. 従って $\hat{M} = [\hat{m}_{fk}]$ と定義すれば, 次のように求めることができる.

$$\hat{m}_{fk} = \frac{\hat{s}_k^T \cdot \hat{w}_f}{\|\hat{s}_k\|^2}, \quad k = 1, 2, 3 \quad (10)$$

このように, 提案手法は, 従来の特異値分解を用いた手法と比べて, 因子分解に複雑な計算をしないため, 高速化が可能である.

2.3 拘束条件による解の決定

上で述べたように3次元線形結合により, 計測行列を2つの行列の積に分解できたが, 分解は一意ではない. そこで, 以下のような拘束条件を設定し, 解を一意に決定する. まず, 可逆な 3×3 行列 A を用いて, M, S を式(11)のように表す.

$$M = \hat{M}A, \quad S = A^{-1}\hat{S} \quad (11)$$

行列 M は, カメラの姿勢を表す単位ベクトル i_f, j_f から構成されているため,

$$\|\hat{m}_f^T A\| = 1 \quad (12)$$

$$\hat{m}_f^T A \cdot A^T \hat{m}_{f+F} = 0 \quad (13)$$

のような拘束条件を満たす必要がある. この条件から行列 A が求まれば, 式(11)に代入して物体の形状 S とカメラの運動 M を復元することができる.

3 実験結果と評価

本実験は, 図3のような模型を用いて連続して取り込んだ70フレームの画像を使用し, 図中の白枠で示した16個の特徴点(手動で設定)をテンプレートマッチングによって自動追跡して模型の形状とカメラ運動の復元を試みた. 実験の方法は, 物体とカメラ

の位置関係をほぼ一定に保ち, 手で模型を載せた台車を回転させて撮影した. その復元結果を図4に示す.

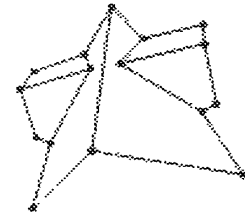
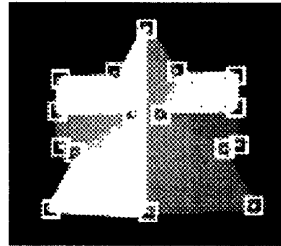


図3: 1枚目の画像と特徴点 図4: 模型の形状復元結果

また, 提案手法が高速であることを示すため, 図3の時系列画像を用いて特異値分解を用いた因子分解法と提案手法の計算時間(追跡時間を含まない)を表1に示す. 提案手法は, 約 $\frac{1}{5}$ の時間で解を得ることができた.

表1: 計算時間の比較

特異値分解	3次元線形結合
266 msec	149 msec

4 むすび

本稿では, 3次元線形結合を用いた因子分解法による物体形状とカメラ運動の復元手法を提案した. この手法は3次元空間に射影されている3つの独立ベクトルを直接求め, 計測行列を物体の形状とカメラの姿勢の2つの行列に分解して復元する. 実験により, 提案手法によって物体の3次元形状を復元できることを確認した. また, 従来の特異値分解を利用した因子分解法よりも高速であることを確認できた. 拘束条件を変えれば, カメラモデルが近中心投影への拡張も可能である. 復元精度の定量的な評価も必要である. また, ノイズに対する提案手法のロバスト性の検証も今後の重要な課題である.

謝辞: 本研究を進めるにあたり, 有益な討論を賜った本学の唐元生氏, カーネギーメロン大学のDaniel Morris氏に深く感謝します.

参考文献

- [1] C. Tomasi and T. Kanade: Shape and Motion from Image Streams: a Factorization Method, *Technical Report CMU-CS-91-172*, Carnegie Mellon University, Pittsburgh, PA, Sep. 1991.
- [2] 黄, 木下: “モーメントによるアフィンパラメータの推定法”, 1997年信学総大, No. D-12-9, Mar. 1997.
- [3] C.V. Stewart, R.Y. Flatland and K. Bubna: Geometric Constraints and Stereo Disparity Computation, *Int. Journal of Computer Vision*, 20(3), pp.143-168, 1996.
- [4] S. Ullman: *The Interpretation of Visual Motion*, The MIT Press, Cambridge, MA, 1979.