

4C-9

動画像から剛体形状の抽出

陳 謙

辻 三郎

(財) イメージ情報科学研究所

大阪大学

はじめに

動画像から物体形状を抽出する研究は、コンピュータビジョンの重要な研究課題の一つである。従来の方法は獲得した物体の形状情報の精度が低く、適用するのに色々な制約条件があるので、一般的に適用することが難しい。本稿では物体の形状を抽出する際、まず、物体の特徴点からカメラへの投影式を多項式で近似した後、各フレームの画像上の物体の特徴点の座標値で構成される行列を、線形代数のSVDのアルゴリズムを用いて分解し、動画像に含まれる物体の形状情報とカメラの移動情報を分離する新しいアルゴリズムを提案する。最後に、得られた結果をカメラの投影方程式に代入することにより物体の形状を求めることができる。

1 カメラの投影式及びその多項式近似

図1に示すように、カメラの運動が水平面内に限定されている場合、ピンホールカメラモデルによって、カメラの投影方程式は次のように表現できる。

$$x' = \frac{x}{f} \tag{1}$$

ここで $x = \frac{X_c}{Y_c}$

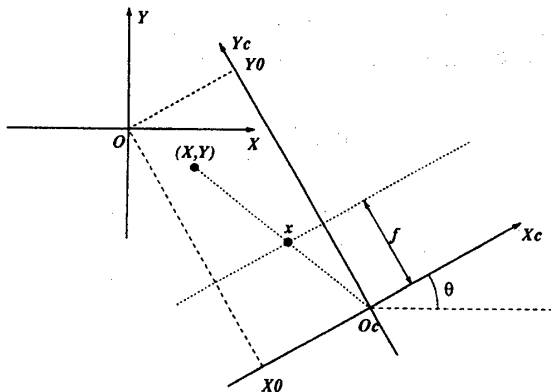


Figure 1. The camera and object coordinate systems.

⁰Extracting the Shape of a Rigid Object from Moving Pictures
¹Qian Chen, Lab. Image Information Sci. & Tech.
²Saburo Tsuji, Osaka University

x は物体の特徴点の像の X 座標、 f はカメラの焦点距離で、 (X_c, Y_c) は物体の特徴点のカメラ座標系での座標値である。

$$\begin{cases} X_c = X \cos \theta + Y \sin \theta + X_0 \\ Y_c = -X \sin \theta + Y \cos \theta + Y_0 \end{cases}$$

物体中心の座標系の原点を、物体の特徴点の重心に設定すると、 Y_c は次のように表現できる。

$$Y_c = Y_0 + (-X \sin \theta + Y \cos \theta) = Y_0 + \Delta Y \tag{2}$$

ここで $\Delta Y = -X \sin \theta + Y \cos \theta$ である。

$\frac{1}{Y_c}$ を Taylor 展開して、必要の精度に応じてその近似式を求める。例えば、相対誤差が 0.4% 以下で、 $\frac{\Delta Y}{Y_0} < 0.25$ の場合、前の 5 つの項を取れば十分である。この場合、式 (1) は次のように変形する。

$$x' = \frac{X_0}{Y_0} + f_1 X + f_2 Y + f_3 X^2 + f_4 XY + \dots + f_{19} XY^4 + f_{20} Y^5 \tag{3}$$

ここで $f_1, f_2, f_3, \dots, f_{19}, f_{20}$ はカメラの運動パラメータ X_0, Y_0, θ の関数である。

2 行列分解による形状情報と運動情報の分離

画像列から物体の特徴点を抽出し、特徴点のフレーム間の対応付けを行なった後、各フレームの画像内の特徴点の座標値を用いて次のような行列を作る。

$$I = \begin{pmatrix} x'_{1,1} & x'_{1,2} & \dots & x'_{1,n} \\ x'_{2,1} & x'_{2,2} & \dots & x'_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ x'_{m,1} & x'_{m,2} & \dots & x'_{m,n} \end{pmatrix}$$

この行列の中の i 行目、 j 列目の要素は画像列の i 番目のフレームにおける j 番目の特徴点の X 座標値である。カメラ投影方程式の多項式近似式 (3) により、次の式が得られる。

$$I = MS \tag{4}$$

ここで

$$M = \begin{pmatrix} \frac{X_{0,1}}{Y_{0,1}} & f_{1,1} & f_{2,1} & \cdots & f_{20,1} \\ \frac{X_{0,2}}{Y_{0,2}} & f_{1,2} & f_{2,2} & \cdots & f_{20,2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{X_{0,m}}{Y_{0,m}} & f_{1,m} & f_{2,m} & \cdots & f_{20,m} \end{pmatrix}$$

$$S = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ X_1 & X_2 & \cdots & X_n \\ Y_1 & Y_2 & \cdots & Y_n \\ X_1^2 & X_2^2 & \cdots & X_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ Y_1^5 & Y_2^5 & \cdots & Y_n^5 \end{pmatrix}$$

で、 M にはカメラの運動情報を含め、 S には物体の形状情報を含む。

SVD(Singular Value Decomposition)を用いて、行列 I を $m \times 21$ の行列 M' と $21 \times n$ の行列 S' に分解する。

$$I = M'S' \quad (5)$$

この分解は唯一ではないが、全ての答えの間に次のような関係式が成立する。

$$M'S' = M'C^{-1}CS' = (M'C^{-1})(CS') = M''S''$$

ここで C は 21×21 の行列である。

従って、

$$S = CS' \quad (6)$$

を成立させる行列 C が存在する。式 (6) により次の結果が得られる。

$$\begin{cases} X_i = c_{2,1}s'_{1,i} + c_{2,2}s'_{2,i} + \cdots + c_{2,21}s'_{21,i} \\ Y_i = c_{3,1}s'_{1,i} + c_{3,2}s'_{2,i} + \cdots + c_{3,21}s'_{21,i} \end{cases} \quad (7)$$

ここで $s'_{1,i}, s'_{2,i}, \dots, s'_{21,i}$ は S' の要素で、SVD を用いて行列 I を分解することにより求められる。従って、 $c_{2,1}, c_{2,2}, \dots, c_{2,21}$ 及び $c_{3,1}, c_{3,2}, \dots, c_{3,21}$ を決定すれば、式 (7) によって物体の形状を求めることができる。

3 物体の形状の獲得

物体の形状を、求めるのに必要な式 (7) の中の係数 $c_{2,1}, \dots, c_{2,21}, c_{3,1}, \dots, c_{3,21}$ を決定するために、式 (7) を式 (1) に代入し、整理すると、

$$\begin{cases} s'_{1,1}a_1 + \cdots + s'_{1,21}a_{21} - x'_{1,i}s'_{1,1}b_1 - \cdots \\ \quad - x'_{1,i}s'_{1,21}b_{21} + X_{0,i} = x'_{1,i}Y_{0,i} \\ s'_{2,1}a_1 + \cdots + s'_{2,21}a_{21} - x'_{2,i}s'_{2,1}b_1 - \cdots \\ \quad - x'_{2,i}s'_{2,21}b_{21} + X_{0,i} = x'_{1,i}Y_{0,i} \\ \vdots \\ s'_{n,1}a_1 + \cdots + s'_{n,21}a_{21} - x'_{n,i}s'_{n,1}b_1 - \cdots \\ \quad - x'_{n,i}s'_{n,21}b_{21} + X_{0,i} = x'_{n,i}Y_{0,i} \end{cases} \quad (8)$$

を得る。ここで

$$\begin{cases} a_k = c_{2,k} \cos \theta_i + c_{3,k} \sin \theta_i \\ b_k = -c_{2,k} \sin \theta_i + c_{3,k} \cos \theta_i \end{cases} \quad k = 1, 2, \dots, 21 \quad (9)$$

である。この方程式の中の未知数は $a_1, a_2, \dots, a_{21}, b_1, b_2, \dots, b_{21}$ 及び $X_{0,i}, Y_{0,i}$ である (i はフレームナンバである)。1 フレームの画像に当たりの方程式の個数は特徴点の数である。従って特徴点が多すぎると、この方程式を解くことができる。

我々は画像列の最初フレームのデータを用いて方程式 (8) の解を求め、そして最初フレームの画像が撮像された時のカメラの角度 θ_i を 0 とし (これは物体座標系の Y 軸とカメラ座標系の Y 軸が平行であることを意味する)、その時の $Y_{0,1}$ を 1 とすることにより、式 9 を用いて $c_{2,1}, \dots, c_{2,21}, c_{3,1}, \dots, c_{3,21}$ を計算する。そして式 (7) より、物体の形状を求めることができる。

4 実験

本論文で紹介したアルゴリズムの有効性を示すために、我々はこのアルゴリズムを実験室内で撮像した画像列に適用して実験をした。実験画像の解像度は 512×480 画素で、カメラの焦点距離は 382 画素に相当する。各フレームの画像から特徴点を抽出し、対応付けを行なった後、我々のアルゴリズムで物体の形状を求めた。

5 おわりに

本論文で紹介したアルゴリズムを利用して、物体の特徴点が広く分布している場合において、物体の形状を画像列から抽出することができる。このアルゴリズムは、一般的な環境において、一般的な運動をするカメラに適用でき、得られた物体の形状は安定で、正確である。我々が行った実験はこのアルゴリズムの有効性を示している。

参考文献

1. Qian Chen and Saburo Tsuji "A Hierarchical Method that Solves the Shape and Motion from an Image Sequence Problem", Proc. IROS'92, pp 2131-2138.
2. Qian Chen and Saburo Tsuji "A Generalized Shape and Motion From Image Sequences", Proc. First Korea-Japan Joint Conference on Computer Vision, pp 343-349.
3. C. Tomasi and T. Kanade, "Shape and Motion without Depth", Proc. of ICCV'90, pp 91-94.