

4 C-5 動画像におけるアフィン変換を用いたカメラワークの正規化

岩成 英一 有木 康雄
 龍谷大学 理工学部

1 はじめに

動画像が生成される過程においては、3次元世界に存在する対象物の動きと共に、観測者の動き、すなわちカメラワークがある。動画像解析においては、対象物の動きの検出、追跡や解析を目的とするため、このカメラワークを取り除くこと(正規化)が必要不可欠である。本研究では、動画像解析の前処理としてカメラワークの成分を抽出すると共に、カメラワークのない正規化した動画像を生成することを目的としている。ここでは、文献[1]に基づき、カメラワークが2次元画像のアフィン変換によって近似できると考え、動画像中からこの変換パラメータを推定して正規化を行なっている。以下にアフィン変換のパラメータ推定法[1]、および実験に関して述べる。

2 カメラワーク成分の抽出

2.1 アフィン変換による対応づけ

動画像において t 番目のフレーム画像を $y_t(u', v')$ 、 $t+1$ 番目のフレーム画像を $y_{t+1}(u, v)$ とする。このとき、画素 (u, v) がアフィン変換によって (u', v') に移るとすると、対応する画素 $(u, v), (u', v')$ のあいだには式(1)のような関係が成り立つ。

$$\begin{pmatrix} u' \\ v' \end{pmatrix} = \begin{pmatrix} x_5(u \cos x_2 - v \sin x_2 + x_3) \\ x_6(u \sin x_2 + v \cos x_2 + x_4) \end{pmatrix} \quad (1)$$

ただし、パラメータ x_2 から x_6 の意味は表(1)に示す通りである。これらのパラメータが求まると t 番目のフレーム画像から $t+1$ 番目のフレーム画像の間で、カメラワークのパンは $-x_3$ 、チルトは $-x_4$ 、回転は $-x_2$ 、ズームは $1/x_5, 1/x_6$ として推定することができる。表中、 x_1 は輝度の変化率を表すもので、式(2)が成り立つ。

$$y_{t+1}(u, v) = x_1 y_t(u', v') \quad (2)$$

ただし、変換の順序は、画像の中央を原点とし、回転、平行移動、拡大、輝度変化の順に固定している。

2.2 アフィン変換のパラメータ推定法

アフィン変換のパラメータを非線形最小二乗推定法により、反復計算で求める。[1]

表1: パラメータの意味

パラメータ	意味
x_1	輝度変化率
x_2	回転角度 (ラジアン、左まわり)
x_3	x 方向の平行移動量 (ドット)
x_4	y 方向の平行移動量 (ドット)
x_5	x 方向の拡大率
x_6	y 方向の拡大率

$y_t(u', v')$ からはじめて $y_{t+1}(u, v)$ に近づくよう、 k 回のアフィン変換を反復実行したのあとの推定画像を $\hat{y}(u, v)^{(k)}$ とする。

このとき、式(2)より式(3)が成り立つ。

$$\hat{y}(u, v)^{(k)} = x_1 y_t(u', v') \quad (3)$$

この k 回目の反復後におけるパラメータの推定値 $\hat{x}^{(k)} = [\hat{x}_1^{(k)}, \dots, \hat{x}_6^{(k)}]$ の微小変化 $\Delta \hat{x}^{(k)} = [\Delta \hat{x}_1^{(k)}, \dots, \Delta \hat{x}_6^{(k)}]$ に対する推定画像の変化量 $\Delta \hat{y}(u, v)^{(k)}$ は式(4)のように表すことができる。

$$\Delta \hat{y}(u, v)^{(k)} = \begin{bmatrix} \frac{\partial \hat{y}(u, v)^{(k)}}{\partial x_1} & \dots & \frac{\partial \hat{y}(u, v)^{(k)}}{\partial x_6} \end{bmatrix} \begin{bmatrix} \Delta \hat{x}_1^{(k)} \\ \vdots \\ \Delta \hat{x}_6^{(k)} \end{bmatrix} = a(u, v)^{(k)T} \Delta \hat{x}^{(k)} \quad (4)$$

$a(u, v)^{(k)T}$ の要素の値は式(1),(3),(4)より求めることができる。

一方、時刻 $t+1$ のフレーム画像上の画素 (u_i, v_j) の濃度値 $y_{t+1}(u_i, v_j)$ と k 回目の推定画像上の画素 (u_i, v_j) の濃度値 $\hat{y}(u_i, v_j)^{(k)}$ との差を求め、これを式(4)の変化量 $\Delta \hat{y}(u_i, v_j)^{(k)}$ として設定することにより、パラメータの微小変化量 $\Delta \hat{x}^{(k)}$ を推定することができる。

すなわち、すべての画素について式(5)の値を求め、それを縦に平べることにより式(6)のような行列を求める。

$$\Delta \hat{y}(u_i, v_j)^{(k)} = y_{t+1}(u_i, v_j) - \hat{y}(u_i, v_j)^{(k)} \quad (5)$$

$$\Delta \mathbf{y}^{(k)} = \begin{bmatrix} \frac{\partial \hat{y}(u_1, v_1)}{\partial x_1} & \dots & \frac{\partial \hat{y}(u_1, v_1)}{\partial x_6} \\ \vdots & & \vdots \\ \frac{\partial \hat{y}(u_n, v_m)}{\partial x_1} & \dots & \frac{\partial \hat{y}(u_n, v_m)}{\partial x_6} \end{bmatrix} \begin{bmatrix} \Delta x_1 \\ \vdots \\ \Delta x_6 \end{bmatrix} = A^{(k)} \Delta \hat{x}^{(k)} \quad (6)$$

表 2: 実験データ

変換	x_1	x_2	x_3	x_4	x_5	x_6
ズーム	1.0	0.0	0.0	0.0	1.1	1.1
回転	1.0	0.1	0.0	0.0	1.0	1.0
平行移動	1.0	0.0	5.0	0.0	1.0	1.0
混合	1.0	0.05	5.0	-3.0	0.9	0.9

$\Delta \hat{y}^{(k)}$ と $A^{(k)}$ はフレーム画像から求めることができるので、一般化逆行列によって k 回目の反復計算のあとのパラメータの微小変化量 $\Delta \hat{x}^{(k)}$ を次式で求めることができる。

$$\Delta \hat{x}^{(k)} = (A^{(k)T} A^{(k)})^{-1} A^{(k)T} \Delta \hat{y}^{(k)} \quad (7)$$

求めた $\Delta \hat{x}^{(k)}$ を現在の推定ベクトル $\hat{x}^{(k)}$ に加えることによって、式 (8) により $k+1$ 回目の推定値を導くことができる。

$$\hat{x}^{(k+1)} = \hat{x}^{(k)} + \Delta \hat{x}^{(k)} \quad (8)$$

実験では、 $k = 50$ になるまでこの繰り返し計算を続けている。初期値は、画像に何の変換も起こっていない状態、すなわち、

$x_1 = 1.0, x_2 = 0.0, x_3 = 0.0, x_4 = 0.0, x_5 = 1.0, x_6 = 1.0$ としている。また、時刻 t と $t+1$ のフレーム画像において求めたパラメータは、次の時刻 $t+1, t+2$ 間での推定パラメータの初期値としている。

3 実験および結果

3.1 人工生成した画像系列に対する実験

ビデオカメラにより取り込んだ一枚の画像からアフィン変換によって画像系列を生成し、そのパラメータが正しく推定されるかどうかを実験した。生成した画像系列は 5 フレームで、各フレーム画像は大きさが 320×240 画素、白黒 256 階調である。表 (2) に実験に用いたデータと、画像系列を生成した時のパラメータ値を示す。表中、「ズーム」、「回転」、「平行移動」は各々ズーム、回転、平行移動が生じるように生成した画像系列である。「混合」はズーム、パン、チルト、回転のすべてが生じるように生成した画像系列である。表 (3) にパラメータ推定の結果を示す。「混合」においては回転と平行移動において誤差を生じているが、その他の合成画像系列では正しくパラメータが推定されている。

3.2 ビデオ映像に対する実験

テレビで放映されているニュース番組から、パン、パンとズーム、チルトと輝度変化の生じている 3 種

表 3: パラメータ推定の結果

変換	x_1	x_2	x_3	x_4	x_5	x_6
ズーム	1.0	0.0	0.0	0.0	1.1	1.1
回転	1.0	0.1	0.0	0.0	1.0	1.0
平行移動	1.0	0.0	5.0	0.0	1.0	1.0
混合	1.0	0.04	3.9	-2.9	0.9	0.9

類の画像系列を取り込み、パラメータ推定の実験を行なった。このうち、パンの生じている画像系列に対して推定されたパラメータを図 1 に示す。

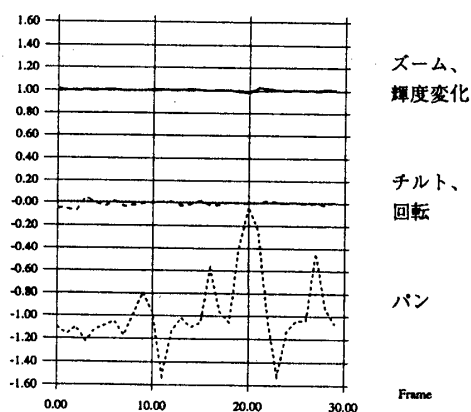


図 1: ビデオ映像で推定されたパラメータ値

フレーム数は 30 フレーム、フレーム画像の大きさは 320×240 画素、白黒 256 階調である。図 1 より、輝度変化およびズームのパラメータは 1.0 と推定され、これらのカメラワークは生じていないことがわかる。同様に、チルトと回転も、生じていないことがわかる。パンのパラメータは平均 -1.0 と推定されており、パンが生じていることがわかる。

4 まとめ

動画像におけるカメラワークがアフィン変換のパラメータによって検出できることを示した。これによってカメラワークが起こっていない正規化された画像系列を作ることが可能となる。今後はその方法を用いて動物体の抽出などを行なう予定である。

本研究は、科学技術庁の平成 4 年度科学技術振興調整費による「創造的研究支援のための自己組織型情報ベースシステムの構築に関する研究」の一環として行なわれた。

参考文献

- [1] 野村 由司彦, 原田 裕次郎, 藤井 省三: "画素データレベルでのアフィン変換画像のマッチング", 信学論 (D-II), j75-D-II, 9, pp1498-1503(1992-09).