

カラー情報と多重解像度を用いた動き推定法

1U-8

黒川雅人

日本アイ・ビー・エム(株)東京基礎研究所

1 はじめに

動画理解において連続するフレーム間の動きを推定することは、最も基本的な処理である。本稿では勾配法を基本とし、オプティカルフローの拘束式を、カラーのRGB 3次元の値について独立に算出し、Multiple Constraints[1]として用いる方法を示す。勾配法に関しては、数多くの手法が提案されている。しかし、その大部分は実際に撮影された動画像に対しては必ずしも良好な結果を示さない。これは、実写画像ではフレーム間での動きが1ピクセル程度ではなくより大きな変位を持っていることが一つの要因である。ここでは、ピラミッド型の階層構造を持った画像を作成し、大きな変位は粗い画像、小さな変位は細かい画像で抽出する戦略を用いる。

2 処理の概要

入力データとしては、ピラミッド型の階層構造を持った画像データを使用する。各階層のデータは、ガウシアンピラミッド(5X5のマスクを使用)の要領で作成する。各階層の画像に対して以下の処理を行う。

(1) 各ピクセルのRGB値の各々に関してオプティカルフローの拘束式(1)を立てる。

$$E_x u + E_y v + E_t = 0 \quad (1)$$

ここで (u, v) は動きベクトル、 E_x, E_y, E_t はそれぞれx軸方向、y軸方向、t軸方向の濃度微分である。次に、それらを局所的なウィンドウ($n \times n$)に関して集め、線形連立方程式を立てる。その結果2変数に対して $3 \times n^2$ 個の式のあるover-constrainedな式ができる。これらを最小二乗法を用いて解き、 (u, v) を得る。実際には、 E_x, E_y の共分散行列Dを作り、その逆行列を求めることになる。

またここで算出された動きベクトルの信頼度を表すために、共分散行列Dの条件数を算出し、それが大きい値を越えるもののみを採用する。

算出された動きベクトルに対してSmoothingの処理を行う。ここでのSmoothingは u, v を求めするための拘束条件としてではなく、後処理としてのSmoothingとな

る。本稿で用いるSmoothing[2]では、

$$\iint \lambda^2 (\bar{u}_x^2 + \bar{u}_y^2 + \bar{v}_x^2 + \bar{v}_y^2) + (u - \bar{u})^2 + (v - \bar{v})^2 dx dy$$

を最小化することを目的とし、

$$\bar{u}^{n+1} = (u^n + \lambda S(\bar{u}^n)) / (1 + 20\lambda^2)$$

$$\bar{v}^{n+1} = (v^n + \lambda S(\bar{v}^n)) / (1 + 20\lambda^2)$$

の式の繰り返し演算を行う。上記の式において $S(*)$ は、9ポイントのLaplacianマスクを用いた近似を示しており、 λ は、重みの定数である。

(1)の方法を評価する基礎実験として、人工的に作成した2枚の画像間での動きを推定した。使用したデータは、MPEG標準データのフラワーガーデン中の1枚であり、これをX方向に1, 2, 4, 6, 8画素ずらした画像を作成し、各処理階層毎に算出された動きベクトルの正答率を図1に示す。図1で横軸は与えた変位、縦軸は正答率、各ラインは処理階層を示している。また正答率の計算では、信頼度の高いベクトルのみを対象とした。

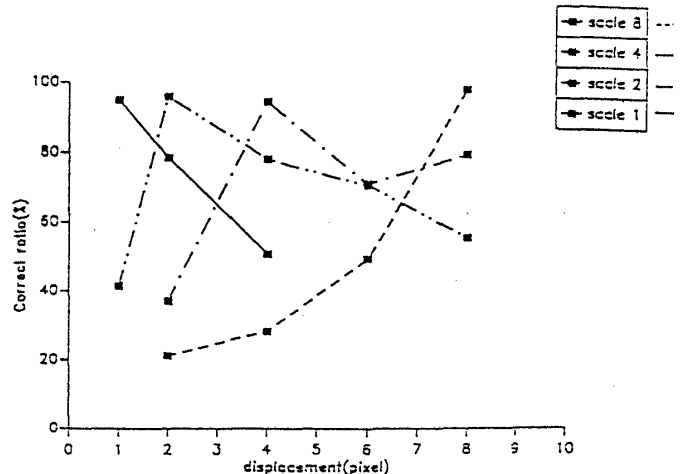


図1. 処理階層と正答率

この結果から、(1)の方法を用いる場合、動きの大きさに応じて処理階層を分ける必要があることがわかる。ここでは、逆に各処理階層毎にそこで算出されるべき動

"Motion estimation of color image by using multi-constraints with hierarchical framework"

Masato Kurokawa

IBM Research Tokyo Research Laboratory, 1623-14, Shimot-suruma, Yamato, Kanagawa, 242, Japan

きベクトルの大きさを限定するアプローチをとる。そのために(1)で算出された各階層中の動きベクトルに対し、以下の処理を行う。

(2) 適当な粗さの処理階層から出発して、以下の2つの基準を満たすベクトルを選択していく。

(a) 信頼性指標

行列Dの条件数 $> Th1$

(b) 大きさの限定

$Thmin < \text{動きベクトルのノルム} < Thmax$

上記の2条件に適合した動きベクトルの値を、最も細かい階層まで伝搬させる。(2)の処理はCoarse to fineで最も細かい階層まで行う。

3 実験結果

上記の処理を実際の動画像に適用する実験を試みた。使用したデータはMPEGの標準画像であるフラワーガーデンであり360 X 240画素RGB8ビットのデータである。各画像に対して4階層のピラミッド構造を作成した(用いたガウシアンマスクは5 X 5)。また共分散行列は3 x 3のウィンドウを単位として計算している。Smoothingの処理は5回反復計算したものを用いている。動きベクトルの選択の際に使用したしきい値は $Thmin = 0.6$, $Thmax = 1.8$ である。原画像を図2に、抽出された動きベクトルを図3に示す。対象画像では、カメラが平行移動するために、手前の樹木の部分と花畑、後方の家ではそれぞれ距離に応じて動きの大きさが異なる。図3の結果から動きベクトルの大きさに合わせて処理階層が正しく選択され、抽出された動きベクトルも比較的正しい事がわかる。ただし、木の幹のように色が類似している場所及び動きの境界では結果は不安定となっている。

4 おわりに

本稿では、カラー入力された動画像から動きを抽出する手法について述べた。今後は、上位階層の結果を用いて大きな粗い動きを抽出し、次に下位階層を用いてより精度良く動きを算出する方法[4]等を導入していく予定である。

参考文献

- [1] V.Markandey, B.E.Flinchbaugh, "Multispectral Constraints for Optical Flow Computation", 2nd ICCV, pp.38-41, 1990.

- [2] B.G.Shunk, "Image flow Segmentation and Estimation by Constraint Clustering.", IEEE PAMI, Vol.11, No.10, pp1010-1027, 1989.
- [3] B.K.P.Horn, B.G.Shunk, "Determining Optical Flow", Artificial Intelligence, vol.17, pp.185-203, 1981.
- [4] P.Anandan, "A computational framework and an algorithm for the measurement of visual motion." Computer Vision, 2:pp283-310, 1989.

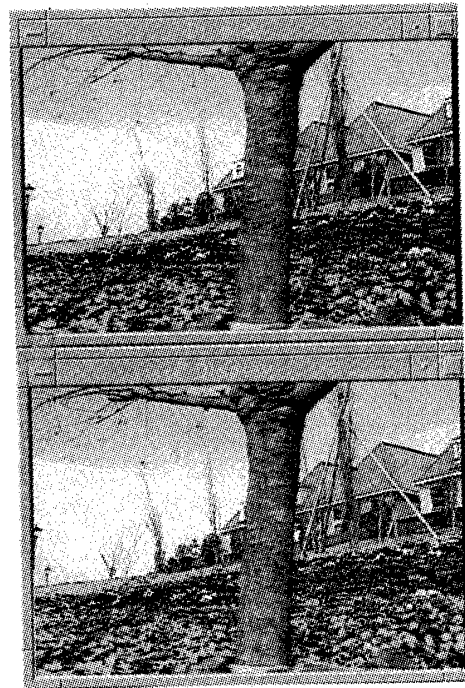


図2. 原画像

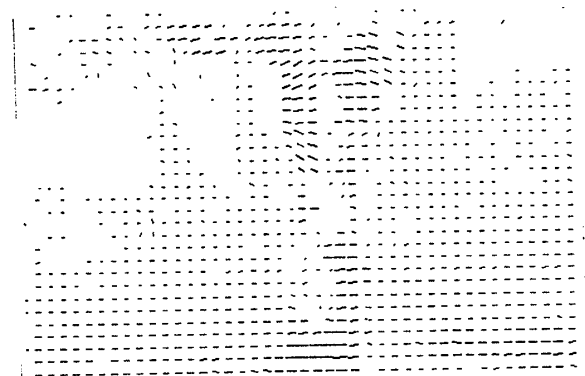


図3. 処理結果 (8 pixel おきに表示)