画素単位のマッチング情報に基づく動きと領域分割の同時推定

正会員 大 網 亮 磨[†], 正会員 相 澤 清 晴^{**}, 正会員 羽 鳥 光 俊^{**}

Simultaneous Motion Estimation and Segmentation Based on Pixel-wise Matching Information

Ryoma Oami[†], Kiyoharu Aizawa^{**} and Mitsutoshi Hatori^{**}

Abstract The method proposed in this paper uses the area of the region around each pixel with small matching errors as a criterion for motion estimation and segmentation. This area is calculated for each displacement vector through directional distance transformation of a binary image obtained by thresholding the matching error at each pixel. On the basis of this area, the proposed method selects candidate vectors at each pixel and determines one of them to be the motion vector by majority rule in a local region. As a result, an image is divided into regions with uniform motion. This method has the advantage of being insensitive to surface textures since it uses only matching errors. The simulation results show that this method can perform well both motion estimation and segmentation even along motion boundaries when regions with coherent motion are relatively large.

1. ま え が き

領域分割動画像符号化や動き補償フレーム内挿などの応用では、精度の高い動き推定、領域分割が要求される。動き推定の精度向上をはかるには、何らかのセグメンテーションが必要となる。一方、動画像のセグメンテーションをより良く行うためには、精度の高い動きの情報が必要である。このように、動画像において動き推定と領域分割の2つの処理は、互いに密接な

関連を有していることが従来から指摘されてきた^{2)~4)}。このため、動き推定と領域分割の両方の処理を同時に行うことが望ましい。

両方の処理を同時に行う手法として、マルコフ確率場(MRF)に基づく手法⁵⁾、クラスタリングに基づく手法²⁾、従来のブロックマッチングを拡張した手法⁷⁾などが提案されている。文献5)では、領域を表すラベルの場をMRFでモデル化し、同一のラベルを持つ領域のみ動き場の平滑化を行い、動きと領域を同時に

キーワード：動き推定，領域分割，マッチング誤差，距離変換

1995年3月28日，1995年電子情報通信学会総合大会，D-322に発表

1995年10月16日受付，1995年12月4日再受付

[†] NEC 情報メディア研究所 (〒216 川崎市宮前区宮崎 4-1-1, TEL 044-856-2264)

^{**} 東京大学 工学部電気工学科 (〒113 東京都文京区本郷 7-3-1, TEL 03-3812-2111)

[†] Information Technology Research Laboratories, NEC Corp. (4-1-1, Miyazaki, Miyamae-ku, Kawasaki-shi, Kanagawa 216, Japan)

^{**} Department of Electrical Engineering, University of Tokyo (7-3-1, Hongo, Bunkyo-ku, Tokyo 113, Japan)

推定している．この手法では良好な結果が得られるが，処理の複雑さが問題となる．一方，文献2)では，色，位置，時空間輝度勾配の8次元特徴量に基づく競合学習により，色，動きが類似した近傍画素をクラスタとして求め，動き推定，領域分割を行っている．この手法では，色の類似性を領域分割の指標として用いるため，細かなテクスチャを有する物体に対処しにくいと考えられる．また，文献7)では，1つのブロックに複数の動きを割り当てることを許容するようにブロックマッチング法を拡張し，動き推定とブロック内の領域分割を同時に行っている．しかし，ブロック内での処理であるためアパーチャ問題が存在する⁸⁾．

本論文では，画素単位のマッチング誤差情報を用いて，相互に影響し合う動き推定，領域分割の2つの処理を同時に行う手法を提案する．提案手法では，画素単位のマッチング誤差より，各画素の周囲に広がるマッチング誤差の小さい領域の面積を算出し，この値を指標として用いて動きと領域を同時に推定する．本手法はMRFに基づく手法に比べ処理は単純である．また，マッチング誤差のみを用いて動き推定，領域分割を行うため，領域分割に画素値の類似性を利用する手法と比して，物体表面の模様などの影響を直接的には受けにくいという特長を有する．さらに，マッチング結果に基づき推定を行うため，画素値の時空間勾配の関係を用いて推定する勾配法に比べ，大きな動きに対しても安定した推定が可能であり，オクルージョンにもある程度対処できる．

以後，2章では画素単位でマッチングを行う手法について述べ，3章ではマッチング誤差の小さい領域の面積を求める手法について述べる．そして，4章では得られた面積値に基づき動き推定，領域分割を行う手法について述べ，5章で提案手法の特徴をまとめ，6章では本手法を実画像に適用した場合の結果を示す．

2. 1画素マッチング

ブロックマッチング法がブロック単位でマッチング誤差を計算するのに対し，1画素マッチング手法は，画素単位でマッチング誤差を計算し，この結果を積極的に利用する手法である¹⁾．(1)式で定義される関数(以後，1画素マッチングの誤差関数と呼ぶ)を各画素で計算し，この関数を用いて動き推定などの処理を行う．

$$e_x(\Delta x) \triangleq |I_r(x + \Delta x) - I_c(x)| \quad (1)$$

ここに， $I_c(x)$ ， $I_r(x)$ はそれぞれ現フレーム，参照フレームの画素 x の輝度値である．

(1)式の $e_x(\Delta x)$ は，変位置量 Δx に対する画素 x の

マッチング誤差を表す．このため，画素 x の動きベクトルである変位置量 Δx に対しては， $e_x(\Delta x) \approx 0$ となる．一方， $e_x(\Delta x) \approx 0$ となる変位置量 Δx は，動きベクトルのみとは限らず一般には複数存在する．この画素単位での誤差情報をもつことで，ブロック内での平均(すなわちブロックマッチング)以外に非線形演算などによる多様な推定を可能にする¹⁾．

3. マッチングの良好な領域とその面積の算出

本章では，動き推定，領域分割の際に用いるマッチングの良好な領域の面積値の算出手法について述べる．

前述のように，(1)式に示す $e_x(\Delta x)$ の値が小さくなる変位置量 Δx は複数存在し得る．しかしながら，動き場は物体境界以外では滑らかに変化することを考慮すると，ある画素 x_0 での動きベクトルが Δx_0 であるならば， $e_{x_0}(\Delta x_0)$ だけでなく， x_0 の周囲の画素 x においても， $e_x(\Delta x_0)$ の値は小さくなる．すなわち，画素 x_0 を含むある広がりを持つ領域に対して，マッチング誤差が小さくなる．逆に，ある変位置量 Δx_0 に対して，画素 x_0 だけでなく， x_0 の周囲の多くの画素 x で同時に $e_x(\Delta x_0)$ の値が小さくなるならば，この変位置量 Δx_0 は画素 x_0 での動きベクトルである可能性が高いと考えられる．このことに基づき，本手法では，各画素の周囲に広がる $e_x(\Delta x)$ の値が小さくなる領域(以後，マッチングの良好な領域と呼ぶ)の面積値を各変位置量 Δx に対して求め，この大小により動き推定，領域分割を同時に行う．

本手法では，このマッチングの良好な領域の面積を以下の手順で算出する．

- (1) 各画素での誤差関数値 $e_x(\Delta x)$ を閾値処理して得られる2値画像(以後，閾値画像と呼ぶ)を作成する．
- (2) (1)で得られる閾値画像に対し，方向性のある距離変換を行う．
- (3) (2)で得られる距離値よりマッチングの良好な領域の面積を算出する．

上述の処理を各変位置量 Δx に対して行う．以後の節では，この各ステップに関して詳述する．

3.1 閾値画像の作成

マッチング誤差が閾値以下の画素には1，そうでない画素に0を割り当てて得られる2値画像を閾値画像と呼ぶことにする．例えば，ある変位置量 Δx に対する誤差関数値 $e_x(\Delta x)$ が図1(a)の場合には，この変位置量に対する閾値画像は図1(b)のようになる．図1(a)では，白いほど誤差関数値が大きいことを表し

ている。

変位量 Δx に対する閾値画像を $L_{\Delta x}$ で表すことにすると、各画素 x での閾値画像の値 $L_{\Delta x}(x)$ は(2)式で与えられる。

$$L_{\Delta x}(x) = \begin{cases} 1 & (e_x(\Delta x) \leq T_m(x)) \\ 0 & (e_x(\Delta x) > T_m(x)) \end{cases} \quad (2)$$

ここに $T_m(x)$ は閾値を表し、位置 x の関数である。この閾値 $T_m(x)$ を、最適な動き推定に対するマッチング誤差と雑音の影響を考慮して決定する。動き推定の精度が $\Delta x \pm \delta x$ であるときには、最適な推定ベクトルに対して、 $|\nabla I_c(x) \cdot \delta x| (\leq |\nabla I_c(x)| \delta x)$ 程度のマッチング誤差を生じ得る。そこで、 $T_m(x)$ として(3)式の関数を用いる。

$$T_m(x) = |\nabla I_c(x)| \delta x + T_n \quad (3)$$

ここに T_n は雑音の大きさを考慮した項である。

3.2 閾値画像に対する方向性のある距離変換

本節では、3.1 節で求めた閾値画像に対して方向性のある距離変換を行い、マッチングの良好な領域の広がり具合を求める手法について述べる。

距離変換は、2 値画像において各画素に背景(値が 0

の画素)までの最短距離を与える変換である⁶⁾。この変換は、2 値画像に対して最小値フィルタ処理を繰り返し行うことで実現できる。この処理で、画素値が 1 から 0 になるまでの処理の反復回数が、その画素での距離値となる。これに対し、方向性のある距離変換とは、特定の方角に関して背景までの最短距離を求めるものであり、処理に用いる最小値フィルタのサポートに方向性を持たせることで実現できる。

本手法では、図 2 に示すサポート領域をもつ最小値フィルタを用い、閾値画像に対して方向性のある距離変換を行う。閾値画像では、値が 0 となるのはマッチング誤差の大きい画素であるため、この方向性のある距離変換で得られる距離値は、その方角に存在するマッチング誤差の大きい最近隣画素までの距離となる。すなわち、この距離値は、その方角に関するマッチングの良好な領域の広がり程度を表す。したがって、図 2 の各最小値フィルタによる結果から、全方位を 8 等分した各方位に対するマッチングの良好な領域の広がり程度が求まることになる。

例えば、図 1(b)の閾値画像を図 2 の 1 番目の最小値フィルタにより距離変換することで、図 3 に示す距離変換結果を得る。この図では、黒いほど距離値が大きいことを表している。このような変換を図 2 の 8 通

(a)

(b)

図 1 閾値画像の例

An example of thresholded DFD (displaced frame difference) images.

図 2 方向性のある距離変換で用いる最小値フィルタのサポート領域

Support regions of minimum filters used for directional distance transformation on thresholded DFD images.

図 3 閾値画像に対する方向性のある距離変換の例

An example of directional distance transformation on the thresholded DFD image shown in Fig. 1 with the first minimum filter in Fig. 2.

図4 方向性のある距離変換により得られる距離値
Values of distance at the pixel P obtained by directional distance transformation with the eight minimum filters in Fig. 2.

りの最小値フィルタに対してそれぞれ行い、8通りの距離変換結果を得る。図4(a)に示す閾値画像の例では、点Pにおいては、図2の*i*番目のフィルタにより、図4(b)の距離値 $d^{(i)}$ が求まることになる。

3.3 マッチングの良好な領域の面積値の算出

本節では、前節で求めた8つの距離値を用いて、マッチングの良好な領域の面積を近似し、算出する手法について述べる。ここで、図2の*i*番目のフィルタで閾値画像 $L_{\Delta x}$ を方向性のある距離変換し、その結果得られる画素 x での距離値を $d_x^{(i)}(\Delta x)$ で表すことにする。

本手法では、この $d_x^{(i)}(\Delta x)$ ($i=1, \dots, 8$) より、画素 x の周囲に広がるマッチングの良好な領域の面積 $a_x(\Delta x)$ を(4)式により求める。

$$a_x(\Delta x) = \frac{1}{2} \sum_{i=1}^8 \{d_x^{(i)}(\Delta x)\}^2 \quad (4)$$

例えば、図4(a)の画素Pの場合には、この面積値は、図4(b)の斜線部の領域の面積に相当する。

4. マッチングの良好な領域の面積値に基づく動きと領域分割の同時推定

本章では、3章で述べたマッチングの良好な領域の面積 $a_x(\Delta x)$ に基づき、動き推定、領域分割を行う手法について述べる。

一般に、 $a_x(\Delta x)$ の値が大きくなる変位量 Δx は、画素 x での動きベクトルである可能性が高いと考えられる。本手法ではこの仮定に基づき、 $a_x(\Delta x)$ を指標として用いて、以下の手順で動き推定と領域分割を行う。

- (1) 候補ベクトルの選出
- (2) 動きベクトルと領域の決定
- (3) 後処理と非対応領域の決定

以下に各ステップについて詳述する。

4.1 候補ベクトルの選出

各画素 x において面積値 $a_x(\Delta x)$ が大きくなる変位量 Δx を候補ベクトルとして選出する。この候補ベクトルの集合を $V(x)$ とすると

$$V(x) = \{\Delta x | a_x(\Delta x) \geq \max(a_{x_{\max}}, A)\} \quad (5)$$

とする。ここに、 $a_{x_{\max}}$ は

$$a_{x_{\max}} = \max_{\Delta x} a_x(\Delta x) \quad (6)$$

であり、 a, A はそれぞれ $0 \leq a \leq 1, A \geq 0$ を満たす定数とする。面積値 $a_x(\Delta x)$ が小さい場合には、雑音などの影響を受けやすくなり、指標としての信頼性が低下するため、ここでは候補の選択に A という閾値を用いる。

4.2 動きベクトルと領域の決定

局所的な多数決処理を反復することで、 $V(x)$ 内の1つの変位量を動きベクトルとして決定する。画素 x に対して割り当てられた動きベクトルを $MV(x)$ で表すことにする。

まず、初期値として、 $MV(x)$ に(7)式の値を割り当てる。

$$MV(x) = \begin{cases} \arg \max_{\Delta x} a_x(\Delta x) & (V(x) \neq \phi) \\ \text{unknown} & (V(x) = \phi) \end{cases} \quad (7)$$

次に、近傍画素に割り当てられているベクトル $MV(x)$ の頻度を調べ、その頻度が最大となる $V(x)$ 内のベクトルで $MV(x)$ の値を置き換える。この処理を繰り返し行うことで、動きの様な領域の形状は平滑化される。そして、 $MV(x)$ の値が等しい連結領域を1つの領域とする。この結果、画像は様な動きを有するいくつかの領域に分割される。

4.3 後処理と非対応領域の決定

4.2節までの処理では、微小な領域や動きの不定な

図5 非対応領域

Regions which have no correspondent region in a reference frame.

領域が残存するため、これらの領域を除去し、隣接領域の動きベクトルを割り当てる。ここでは、ある閾値 T_a 以下の面積の領域を微小領域と判定する。そして、これらの領域の周囲から内部へと動きベクトルを割り当てていく。この処理ですべての画素の $MV(x)$ に何らかのベクトルが割り当てられる。

オクルージョンなどにより、参照フレーム上に対応領域の存在しない領域（以後、非対応領域と呼ぶ）が生じることもあるが、これに対しては、動きベクトルと領域分割の整合性により判断する。例えば、図5のように動きベクトルと領域が求まっている場合、各領域を動きベクトルに従って動かすと、領域1と2は参照フレーム上で重なり合う。このことは動き推定と領域分割の結果が矛盾することを意味する。そこで、このような領域に対しては、割り当てられた変位量 $\Delta x = MV(x)$ に対する $a_x(\Delta x)$ の値が最大となるものを優先し、それ以外の領域は非対応領域とする。図5の例では、参照フレーム上で同一の画素に対応する画素 x_1 と画素 x_2 に対して、 $a_{x1}(\Delta x_1) < a_{x2}(\Delta x_2)$ ならば、画素 x_1 を非対応領域に属する画素とする。

5. 本手法の特徴

本手法の特徴をまとめると、以下ようになる。

- ・物体表面の模様の影響を直接受けにくい

正しい動きベクトルに対しては、物体表面の模様によらずマッチング誤差は小さくなる。このマッチング誤差情報のみより動き推定、領域分割を行うため、画素値の一様性を用いて領域分割を行う手法（例えば2））と比べ、模様の直接的な影響を受けにくい。このため、複雑な模様を有する動物体であっても、動きが同じであれば1つの領域に分割される。ただし、輝度値勾配によって、動きベクトルの推定誤差に対するマッチング誤差の変化の度合いが異なるため、本手法では、マッチングの良好度を判定する閾値を(3)式のように変化させることで対処している。

- ・アパーチャ問題を回避できる

従来の動き推定法の多くは、動き推定対象画素の

周囲にウィンドウを設定し、そのウィンドウ内で処理を行い動きを算出するため、エッジ近傍では、エッジに平行な動き成分が求まらないというアパーチャ問題が存在する⁸⁾。しかし、本手法では、推定にウィンドウを用いず、動き場の連続性を反映した面積値 $a_x(\Delta x)$ を推定の指標に用いるため、アパーチャ問題が存在しない。このため、エッジ近傍でも良好な推定結果が得られる。

- ・大きな動きに対しても安定した推定が可能である

輝度値の時空間勾配より動きを算出する勾配法とは異なり、基本的にはマッチングベースの手法であるため、探索範囲内に動きベクトルが存在すれば、大きな動きに対しても推定が可能である。

- ・動きの一様性の小さい領域には対処困難である

動きの一様性が小さいところでは、 $a_x(\Delta x)$ が急峻なピークを持たなくなり、雑音などの影響を受けやすくなる。本手法では、この限界を候補ベクトルを選ぶ際の閾値 A で規定している。

また、本手法では、1つのベクトルに対して8回の距離変換を行うため、距離変換を行う総数は探索ベクトル数の8倍となる。この1回の距離変換には、数回から数十回の最小値フィルタ処理が要求されるため、演算量は、従来のブロックマッチング法などと比較すると、かなり大きくなる。現段階では、これが問題である。しかし、面積値が求まった後は、比較的単純な反復処理で動き推定、領域分割が行われる。

6. 実験結果

本章では、実画像に対して本手法を適用した実験結果について述べる。実験は画像“Table tennis”、“Flower Garden”（ともに 352×240 ）に対して行った。なお、雑音の低減のために、 3×3 のローパスフィルタによる前処理を行っている。

1画素マッチングの誤差関数 $e_x(\Delta x)$ は、整数精度の変位量に対して計算した。このとき、(3)式の δx は、 $[0.5, 0.5]^T$ となる。(3)式において、 $|\nabla I_c(x)|$ は 3×3 の Sobel オペレータを用いて算出し、 T_n の値は5とした。また、(5)式の候補ベクトル算出時のパラメータは、 $\alpha=0.5$ 、 $A=75$ とし、動きベクトルの決定に用いた多数決処理の近傍領域は、処理対象画素を中心とする 17×17 の正方形領域、微小領域判定の閾値 T_a は、 A と同じく75とした。これらの値に関しては、実験により適切と思われるものを採用しているが、特に最適化は行っていない。これらのパラメータのうち、 T_n の値は、小さ過ぎると正しい推定に対してもマッチング誤差が大きいと判定し、 $a_x(\Delta x)$ の算

(a)

(b)

図 6 動き推定, 領域分割結果 (画像: “Table tennis” 2, 4 フレーム間)
Results of motion estimation and segmentation.
(Image: “Table tennis”)

(a)

(b)

図 7 動き推定, 領域分割結果 (画像: “Flower garden” 25, 27 フレーム間)
Results of motion estimation and segmentation.
(Image: “Flower garden”)

出に大きな影響を与えるため、雑音による輝度値の変化をほぼ吸収できる程度の値に設定している。また A は、小さくし過ぎると不要なベクトルが候補に選ばれ、多数決処理に影響を与えるため、これらの影響が排除できる程度の大きさに設定している。 α は、動きが連続的に変化する部分でも真の動きベクトルが候補から洩れない程度の値に設定すればよい。実際、模様がある程度存在する領域では、 $a_x(\Delta x)$ の値は、 Δx の変化に対して急峻なピークを示すため、 α の値は小さな値（例えば 0.2）に設定してもそれほど問題はない。また、多数決処理を行う領域の大きさに関しては、小さく過ぎると領域形状の平滑化が行われにくくなる傾向にあるため、今回は比較的大きな領域を用いている。

画像 “Table tennis” に対する動き推定, 領域分割結果を図 6 に示す。図 6(a) は本手法により求めた動き場, 同図(b) は(a) に対応する領域分割結果である。図 6(b) の白い領域は、非対応領域と判定された領域である。また、画像 “Flower garden” に対する結果を図 7 に示す。図 6, 図 7 より、動き推定, 領域

分割がともに精度良く行われていると推察される。特に、“Flower garden” に対する結果では、テクスチャの細かい花壇の部分においても、その影響を受けることなく動きの様な領域ごとに分割されているのわかる。

7. む す び

本論文では、画素単位のマッチング結果よりマッチングの良好な領域の面積を算出し、この値を指標として動き推定, 領域分割を同時に行う手法について述べた。本手法を実画像に対して適用し、動きの様な領域がある程度大きい画像に対して、動き境界でも良好な動き推定, 領域分割結果が得られることを確認した。

本手法では、動きの様な性の低い領域に対しては、推定が困難である。しかしながら、マッチングの良好な領域の面積値はこのような領域に対しては極端に小さい値となるため、このような領域を閾値を用いて判定し、他の動き推定, 領域分割手法を行うことで、対処できると考える。

本研究を行う上で有用な御助言を頂いた、株式会社グラフィックスコミュニケーションラボラトリーズの佐藤真樹氏に感謝致します。

〔参 考 文 献〕

- 1) 大網亮磨, 佐藤真樹, 相澤清晴, 羽鳥光俊: “一画素マッチングによる誤差平面を利用した動き推定”, テレビ誌, マルチメディア時代の画像処理特集, **49**, 3, pp. 319-329 (Mar. 1995)
- 2) 柴藤稔, 白井良明: “色, 位置, 輝度こう配に基づく領域分割による2次元動き推定”, 信学論D-II, **J76-D-II**, 11, pp. 2324-2332 (Nov. 1993)
- 3) H. G. Musmann, M. Hötter and J. Ostermann: “Object-oriented Analysis-synthesis Coding of Moving Images”, Signal Processing: Image Communication, **1**, 2, pp. 117-138 (Oct. 1989)
- 4) R. Thoma and M. Bierling: “Motion Compensating Interpolation Considering Covered and Uncovered Background”, Signal Processing: Image Communication, **1**, 2, pp. 191-212 (Oct. 1989)
- 5) M. M. Chang, M. I. Sezan, Z. M. Tekalp: “An Algorithm for Simultaneous Motion Estimation and Scene Segmentation”, Proc. IEEE ICASSP94, pp. V221-V224 (1994)
- 6) 鳥脇純一郎: “画像理解のためのデジタル画像処理(II)”, 昭晃堂
- 7) 小松隆, 斎藤隆弘: “複数の動きを推定可能なブロックマッチング法”, テレビ誌, マルチメディア時代の画像処理特集,

49, 3, pp. 328-335 (Mar. 1995)

- 8) A. Singh: “Optic Flow Computation”, IEEE Computer Society Press (1991)

おおあみ りょうま

大網 亮磨 1993年, 東京大学工学部電子工学科卒業。1995年, 同大学大学院修士過程修了。同年, NECに入社。在学中, 動画像動き推定の研究に従事。正会員。

あいざわ きよはる

相澤 清晴 1983年, 東京大学工学部電子工学科卒業。1988年, 同大学大学院博士課程修了。現在, 同大学電子情報工学科助教授。1990年から2年間, 米国イリノイ大学客員助教授。画像の符号化と処理, モデルベース符号化, 画像取得手法, イメージセンサなどの研究に従事。工学博士。正会員。

はとり みつとし

羽鳥 光俊 1963年, 東京大学工学部電気工学科卒業。1968年, 同大学大学院電気工学科博士課程修了。同年, 同大学工学部電気工学科専任講師を経て, 現在, 同大学電子情報工学科教授。通信方式, 放送方式, 適応信号処理, 画像通信, 移動通信, 光通信, ケーブルテレビなどの研究に従事。工学博士。正会員。