

## 色と輝度勾配に基づく領域分割による2次元動き推定

栄藤 稔† 白井 良明‡

†松下電器 中央研究所  
†570 守口市 八雲中町 3-15

‡大阪大学 工学部  
‡565 吹田市 山田丘 2-1

E-mail: etoh@crl.mei.co.jp

あらまし 動画像から動きを推定する手法を提案する。色、画素の位置、動きの特徴に基づくクラスタリングにより、画像を領域断片に分解する。各断片を、色、位置、輝度の時空間勾配の分布パラメータで記述する。クラスタリングは、逐次、画素の分類と同時にパラメータ推定を行なう競合学習の考えに基づいて行なわれる。この競合学習により、セグメンテーションと各断片の2次元の動き推定を同時に行なう。色、画素の位置、動きを結合することにより、従来の輝度勾配に基づく動き推定の手法に比べて、異なる動きを持つ領域の境界や、テクスチャの情報が少ない領域についても、より安定に2次元動き推定を行なうことができる。

和文キーワード 動き推定, 領域分割, 色, 輝度勾配, クラスタリング, 競合学習

## 2D Motion Estimation by Region Partitioning According to Color and Intensity Gradients.

Minoru ETOH† Yoshiaki SHIRAI‡

†Central Research Laboratories of Matsushita Electric Ind.

‡Osaka University

†3-15, Yakumonaka-machi, Moriguchi 570, Japan

‡2-1, Yamadaoka, Suita 565, Japan

E-mail: etoh@crl.mei.co.jp

Abstract This paper presents a new 2D motion estimation method. Combining color, motion and pixel position, an image is decomposed into region fragments by a clustering process. Each fragment contains distribution parameters of pixel positions, color, and spatial and temporal intensity gradients. Segmentation and motion estimation is simultaneously performed on the fragmental description by parameter estimation and sample assignment for each fragment. The clustering is achieved by a competitive learning in which the description is dynamically updated by the result of classification. As a result, regions are segmented more precisely, motion boundaries are not blurred, and 2D motion is obtained even in non-textured area.

英文 key words motion estimation, segmentation, color, intensity gradient, clustering, competitive learning

# 1 まえがき

本研究は、動画像中の物体の2次元な動き(画面上の移動量)を正確かつ安定に求めることを目的としている。想定している応用は、物体追跡、符号化、画像合成などであり、動き推定だけでなく領域の境界を保存したセグメンテーション(領域分割)が必要となっている。この報告では次の2つの課題を議論する。

1. 動き推定とセグメンテーションの相互依存.
2. ノイズに頑健な実時間処理システムの実現.

ここで、動きベクトルとは画面上の2次元動きを意味するものとする。最初の課題に関する問題として「窓問題(aperture problem)」がある。これは、局所的なエッジからは、真の動きベクトルに対して、エッジに対する法線方向成分のみ得られ、接線方向成分が未知となる問題である[1]。この問題を解くために様々なレベルで拘束条件が用いられている(例えば[2])。対象物体やカメラの動きパラメータに仮定を置かない一般的な拘束条件には、画像の局所的近傍で動きベクトルが一様である(一様性)か、あるいは、滑らかに変化する(連続性)の2つが考えられる[3][4][5][6]。これらの仮定により、異なる方向のエッジ成分が近傍にあれば、動きベクトルを求めることができる。この一様性または連続性を可能な限り広い領域で仮定できれば、ノイズを含む実際の画像に対しても、安定して動きベクトルを得ることができる。従って、ノイズに対して頑健かつ正確な動きベクトルを求めようとすれば、一様性または連続性を仮定する近傍領域を得るためのセグメンテーションが重要となる。一方、逆に、動きベクトルが正確に求まれば、セグメンテーションが可能である[7]。このように、動き推定とセグメンテーションは相互に強く依存している。

この報告では、動き推定とセグメンテーションを同時に行なう方法として色と輝度勾配の領域分割による2次元動き推定法を提案する。この方法の概略を、図1に示す。各画素について色(colors)、位置(positions)、時空間輝度勾配(intensity gradients)の8次元の特徴が与えられている。そして、競合学習の考えに沿ったクラスタリングにより、色が類似である近傍画素を集合として求める。これにより、セグメンテーションと動きベクトルの推定を同時に行なう。筆者らは既に静止画を画素の色と位置の分布により領域断片の集合として記述する手法を提案しているが[8]、本報告では、物体追跡、符号化、画像合成に適した動画像の記述を提案することになる。一つのクラスに属する画素の集合を領域断片(region fragment)あるいは単に断片とよぶことにする。この断片の中では動きベクトルが一様であると仮定する。図2に従来の連続性の拘束を用いたHornとSchunk[3]の正則化手法と本手法の比較を示す。

セグメンテーションと動き推定を同時に行なう典型的な手法には、大きく分けて、ラインプロセス付き弛緩法とクラスタリングの2つがある。ラインプロセスでは、画素単位の動きベクトル推定を行なえるが、対象画像は比較的単純な画像に限られる[9]。想定している応用では、画像中の小領域単位の動きベクトル推定で十分であることもあり、

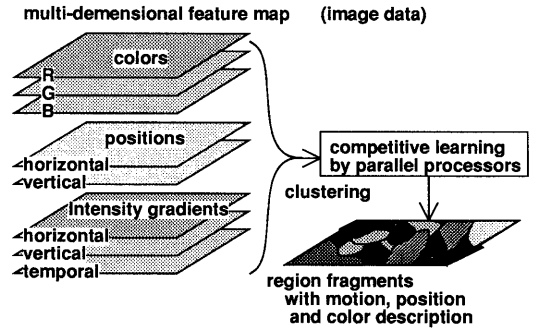


図1 領域分割による2次元動き推定の流れ

筆者らはパラメトリックなクラスタリングの手法を用いた。推定の粒度は犠牲となるが、領域に基づく推定は精度、ノイズに対する頑健性の点で優れている。クラスタリングに基づく手法には、ハフ変換などパラメータ空間への投票を行なう手法を含めて、従来研究があるが[10][11][12]、筆者らの提案手法は以下の点で異なる。

**色と動きの統合:** セグメンテーションと動き推定は、各領域断片に属する画素の色、位置、動きの結合確率密度関数のパラメータ推定として統一的に扱い、クラスタリングにより求める。これにより、動領域の境界を保存した動き推定が可能である。テクスチャの情報が少ない領域についても、色情報の類似性から動きベクトルの推定が可能である。

**並列処理と動的なパターン表現:** 実時間処理システムの実現に関して、クラスタリング処理を競合学習の考えに沿って実現した。これは並列処理が可能であり、また動画像の時間的に緩やかに変動する非定常パターンを表現するよい手段となっている。

以下、2章では提案手法の仮定と定義、3章では競合学習に基づく動き推定アルゴリズム、4章では実画像への適用結果を述べてく。

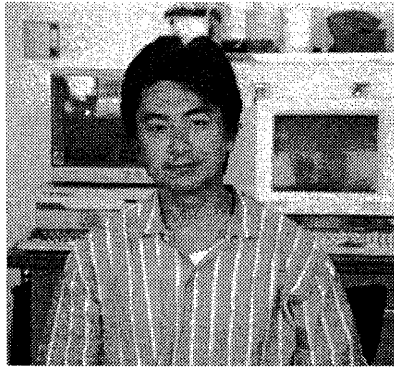
## 2 動画像記述のモデル

各画素における画像特徴は、色  $x$  (RGBの値)、位置  $p$  (水平、垂直) と輝度の時空間勾配  $g = (\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y}, \frac{\partial I}{\partial t})^t$  からなっている ( $I$  と  $(\cdot)^t$  は輝度と行列の転置を表す)。ここで、画像  $R$  は  $c$  個の領域断片  $\omega_i, 1 \leq i \leq c$  に分解されると考える。すなわち  $R = \{\omega_1, \omega_2, \dots, \omega_c\}$  とする。以下、この画像特徴の分布の仮定を行ない、画素がどの断片に属するかを決定する距離関数を定義する。

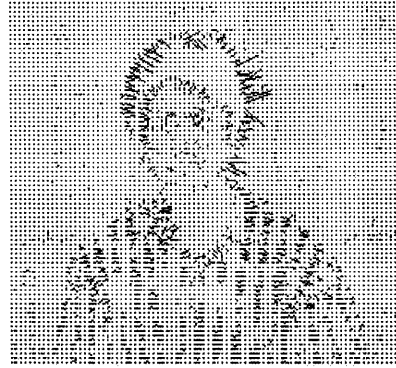
### 2.1 画像特徴の分布に関する仮定

任意の領域断片内の画像特徴の分布について、以下に示す色、位置、動きの確率密度関数を独立に仮定する。

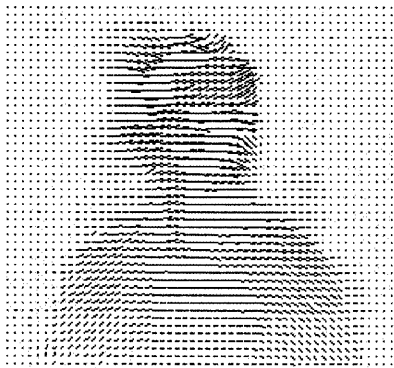
<sup>†</sup>窓問題は原理的な問題として残る。



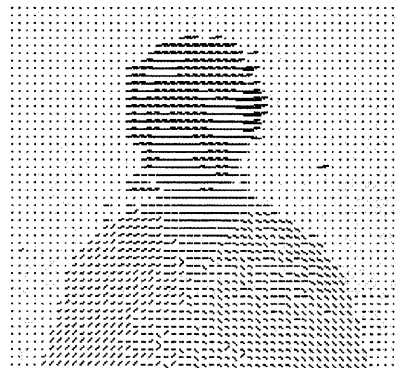
(a) A swinging person.



(b) Normal flow.



(c) Smoothed optical flow.



(d) 2D motion flow obtained by our method.

図2 従来手法との比較

(a) 画面上、人が左から右へ身体を振っている。全ての動き成分は時空間輝度勾配から求めている。(b) 平滑化を行なう前のオプティカルフロー。人物輪郭、服の縦縞部分、顔の陰影部分で得られているが、エッジの法線方向成分のみである。(c) は Horn と Schunk の正則化手法により得られたオプティカルフローである。反復回数は 1000 回。動領域境界がぼけている。(d) は、提案手法の結果である。

色：

断片  $\omega_i$  内の色  $\mathbf{x}$  は 3 次元正規分布  $N(\bar{\mathbf{x}}_i, \mathbf{X}_i)$  に従うものとする。すなわち、

$$p(\mathbf{x}|\omega_i) = \frac{1}{(2\pi)^{3/2} |\mathbf{X}_i|^{1/2}} \exp\left[-\frac{1}{2}(\mathbf{x} - \bar{\mathbf{x}}_i)' \mathbf{X}_i^{-1} (\mathbf{x} - \bar{\mathbf{x}}_i)\right] \quad (1)$$

であり、 $\bar{\mathbf{x}}_i$  と  $\mathbf{X}_i$  は色の平均と共分散である。単一光源下で単一絶縁物体の観測される色は RGB 空間の平面に分布することが知られている (例えば、[14][15][16])。従って、一般的な正規分布を仮定することにより、 $\mathbf{X}_i$  の 3 番めの固有値に相当する固有ベクトルが張る部分空間で異なる物体の領域が分割できることが期待できる。

位置：

画素の位置  $\mathbf{p}$  は各断片内で、2次元正規分布  $N(\bar{\mathbf{p}}, \mathbf{P}_i)$  に従うと仮定する。すなわち、

$$p(\mathbf{p}|\omega_i) = \frac{1}{(2\pi)^2 |\mathbf{P}_i|^{1/2}} \exp\left[-\frac{1}{2}(\mathbf{p} - \bar{\mathbf{p}})' \mathbf{P}_i^{-1} (\mathbf{p} - \bar{\mathbf{p}})\right] \quad (2)$$

であり、 $\bar{\mathbf{p}}, \mathbf{P}_i$  は位置の平均 (中心) と共分散である。明らかに、これは近似のための仮定である。正規分布の仮定は記述の簡略化と計算効率の点で都合が良い。共分散  $\mathbf{P}_i$  より、等価楕円領域を考えることができる。その楕円内で画素が一様分布するとすると、その楕円半径の平方は、固有値の 3 倍に等しい。

動き：物体表面の輝度一定を条件として、オプティカルフローの基本式 [3] が導き出される。

$$u \frac{\partial I}{\partial x} + v \frac{\partial I}{\partial y} + \frac{\partial I}{\partial t} = 0 \quad (3)$$

ここで、 $u, v$  は動きベクトルの各々水平、垂直成分である。動きベクトルを  $\mathbf{m}_i = (u_i, v_i, 1)'$  と表現する。ここで、時空間輝度勾配との内積は、1次元正規分布  $N(0, s_i^2)$  に従うと仮定する。すなわち、

$$p(g|\omega_i) = \frac{1}{\sqrt{2\pi} s_i} \exp\left[-\frac{1}{2s_i^2} (\mathbf{m}_i' \mathbf{g})^2\right] \quad (4)$$

であり、 $(\mathbf{m}_i' \mathbf{g})^2$  と  $s_i^2$  は式 3 に対する 2 乗誤差とその分散である。式 4 の仮定には、領域断片内で、動きベクトルが一様であるとの前提がある。

## 2.2 距離関数

領域断片  $\omega_i$  について、色、位置、輝度勾配の結合確率密度関数が次のように記述できる。

$$p(\mathbf{x}, \mathbf{p}, \mathbf{g} | \omega_i) = p(\mathbf{x} | \omega_i) p(\mathbf{p} | \omega_i) p(\mathbf{g} | \omega_i) \quad (5)$$

ここで、 $p(\omega_i)$  を領域断片  $\omega_i$  の事前確率とすると、Bayes の類別規則 [17] により、識別関数を定義することができる。

$$f_i(\mathbf{x}, \mathbf{p}, \mathbf{g}) = \ln p(\mathbf{x}, \mathbf{p}, \mathbf{g} | \omega_i) + \ln p(\omega_i), i = (1, 2, \dots, c) \quad (6)$$

これにより、特徴  $(\mathbf{x}, \mathbf{p}, \mathbf{g})$  を持つ画素を式 6 を最大化する領域断片  $\omega_k$ ,  $k = \arg \max_{0 \leq i \leq c} f_i(\mathbf{x}, \mathbf{p}, \mathbf{g})$  へ最小の誤り率で分類することができる。式 6 の定数項を除き、符号を換えて単純化すると距離関数が式 7 として得られる。

$$i = (1, 2, \dots, c) \text{ について, } d_i(\mathbf{x}, \mathbf{p}, \mathbf{g}) = (\mathbf{x} - \bar{\mathbf{x}}_i)^T \mathbf{X}_i^{-1} (\mathbf{x} - \bar{\mathbf{x}}_i) + (\mathbf{p} - \bar{\mathbf{p}}_i)^T \mathbf{P}_i^{-1} (\mathbf{p} - \bar{\mathbf{p}}_i) + (\mathbf{m}_i^T \mathbf{g})^2 / s_i^2 + \ln |\mathbf{X}_i| + \ln |\mathbf{P}_i| + \ln s_i^2 - \ln p(\omega_i) \quad (7)$$

これは画像特徴  $\mathbf{x}, \mathbf{p}, \mathbf{g}$  を持つ画素を領域断片  $\omega_k$ ,  $k = \arg \min_{0 \leq i \leq c} d_i(\mathbf{x}, \mathbf{p}, \mathbf{g})$  に最小誤り率で分類する尺度となる。次章で述べるクラスタリングにはこの距離関数を用いる。事前確率  $p(\omega_i)$  に関して特に情報が無いときは、全ての  $\omega_i$  について等確率と考えて、 $\ln p(\omega_i)$  の項は、式 6 及び式 7 で無視される場合が多い。この場合、通常の最尤推定法と同じになる。本手法では、応用に応じて先験的情報を事前確率  $p(\omega_i)$  で記述し、式 7 で用いる。

## 2.3 領域記述

ここで、 $c$  個のパラメータの集合を簡単のため、「領域記述」と呼ぶことにする。

$$\text{領域記述: } \theta = \{\theta_1, \theta_2, \dots, \theta_c\} = \{(\bar{\mathbf{x}}_1, \mathbf{X}_1, \bar{\mathbf{p}}_1, \mathbf{P}_1, \mathbf{G}_1), (\bar{\mathbf{x}}_2, \mathbf{X}_2, \bar{\mathbf{p}}_2, \mathbf{P}_2, \mathbf{G}_2), \dots, (\bar{\mathbf{x}}_c, \mathbf{X}_c, \bar{\mathbf{p}}_c, \mathbf{P}_c, \mathbf{G}_c)\} \quad (8)$$

ここで、 $\mathbf{G}_i = E(\mathbf{g}\mathbf{g}^t)$ ,  $\mathbf{g} \in \omega_i$  である。動きベクトル  $\mathbf{m}_i$  と分散  $s_i^2$  の代りに、 $\mathbf{g}$  の共分散行列を用いる。これは、 $\mathbf{m}_i$  と  $s_i^2$  が、

$$\mathbf{m}_i = \frac{1}{u_{33(i)}} \mathbf{u}_{3i} \quad (9)$$

$$s_i = \sqrt{\lambda_{3(i)} / u_{33(i)}} \quad (10)$$

として表現できることによる。 $\mathbf{u}_{3i}$  は  $\mathbf{G}_i$  の 3 番めに大きな固有値に対応する固有ベクトルであり、 $u_{33(i)}$  は  $\mathbf{u}_{3i}$  の第 3 要素である。すなわち、

$$\mathbf{G}_i = \lambda_{1(i)} \mathbf{u}_{1i} \mathbf{u}_{1i}^t + \lambda_{2(i)} \mathbf{u}_{2i} \mathbf{u}_{2i}^t + \lambda_{3(i)} \mathbf{u}_{3i} \mathbf{u}_{3i}^t, \quad \lambda_{1(i)} > \lambda_{2(i)} > \lambda_{3(i)} \quad (11)$$

である。式 4 の対数を  $\mathbf{m}_i$  について最大化を図ることを考えると、式 12 が必要条件として得られる。

$$\sum_{\mathbf{g} \in \omega_i} \frac{\partial \ln p(\mathbf{g} | \omega_i)}{\partial \mathbf{m}_i} = \sum_{\mathbf{g} \in \omega_i} (\mathbf{g}\mathbf{g}^t \mathbf{m}_i) = 0 \quad (12)$$

式 12 の最小化のためには、 $\mathbf{m}_i$  は  $\mathbf{G}_i$  の 3 番めに大きな固有値に対応する固有ベクトルとならなければならない。式 4 の正規分布の仮定から、輝度勾配  $\mathbf{g}$  の共分散 (平均値は零

と仮定) を求めることにより、動きベクトルが式 9 として求められることになる。この計算は、 $\mathbf{g}$  の標本行列を求めて特異値分解 (SVD) などにより最小 2 乗推定を行なう計算に比べて非常に簡便である。

## 3 競合学習によるパラメータ推定

Kohonen の文献 [13] にならって、ランダムに標本を得て、クラスへの割り当てと同時にそのクラスパラメータを逐次変更する処理を競合学習と呼ぶことにする。

### 3.1 処理の概要

処理は以下のステップで表せる。

1.  $t$  を時間軸の変数として、画像特徴の標本時系列を  $(\mathbf{x}(t), \mathbf{p}(t), \mathbf{g}(t))$ 、領域記述を  $\{\theta_i(t) : \theta_i = (\bar{\mathbf{x}}_i(t), \mathbf{X}_i(t), \bar{\mathbf{p}}_i(t), \mathbf{P}_i(t), \mathbf{G}_i(t)), i = 1, 2, \dots, c\}$  とする。
2.  $\theta_i(0), i = 1, 2, \dots, c$  を初期化する。
3. 時間軸  $t = 1, 2, 3, \dots$  の上で、ランダムに選んだ標本  $(\mathbf{x}(t), \mathbf{p}(t), \mathbf{g}(t))$  を順次入力し、パラメータ  $\theta_i(t)$  について照合比較する。最も照合の良い  $\theta_i(t)$  を選び、その値が  $(\mathbf{x}(t), \mathbf{p}(t), \mathbf{g}(t))$  に、より近くなるように変更する。この照合は式 7 の距離関数  $d_i(\mathbf{x}(t), \mathbf{p}(t), \mathbf{g}(t))$ ,  $i = 1, 2, \dots, c$  により、最小距離の領域断片を選ぶことにより行なわれる。最小距離となる領域断片を  $\omega_k$  とすると、 $d_k(\mathbf{x}(t), \mathbf{p}(t), \mathbf{g}(t))$  が減少するようにパラメータ  $\theta_k(t)$  を変更し、他の  $i \neq k$  なるパラメータ  $\theta_i$  は変更しない。この変更規則を「デルタ規則」と呼ぶ。

以上に示したように、各領域断片のパラメータが、ランダムに選ばれた標本により調整 (学習) されて行く。

このような競合学習を用いた理由は、以下の 2 点である。

- 競合学習によりリアルタイム処理に必要なオンライン処理系を構築することがきである。図 3 に、並列処理系の構成図を示す。プロセッサ  $i$  に  $\omega_i$  のパラメータ  $\theta_i$  を持たせ、距離  $d_i(\cdot)$  を並列に計算させる。距離計算後、比較回路により、最小距離となるプロセッサを選び、パラメータを変更することが可能となる。

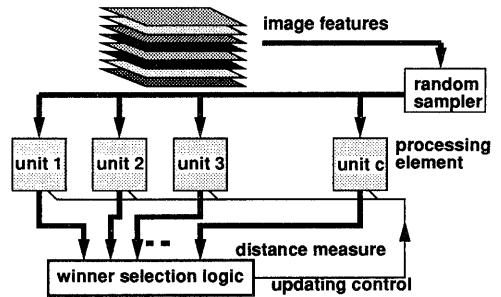


図 3 競合学習を行なう並列アーキテクチャ

- 動画像は、フレームを一つのパターンとすると、緩やかに変動する非定常パターンであり、一般的に前後するフレーム間では大きな変化はない。ランダムに選んだ標本からデルタ規則により 逐次、領域記述を変更することは、記述の連続的な更新を保証し、領域の追跡に利用できる。

### 3.2 デルタ規則

各領域断片について距離の総和  $E = \sum_R \sum_{\omega_i} d_i(\mathbf{x}, \mathbf{p}, \mathbf{g})$  が小さくなる方向に、逐次パラメータを修正していくデルタ規則を考える。ここで、次に示すデルタ規則は離散時間 ( $t=0,1,2,\dots$ ) におけるランダムな標本に対して動作する。 $(\mathbf{x}(t), \mathbf{p}(t), \mathbf{g}(t))$  は時刻  $t$  に得られた標本画像特徴、 $\omega_k$  は距離関数  $d_i(\mathbf{x}(t), \mathbf{p}(t), \mathbf{g}(t))$  が最小となる断片を表すものとする。 $E$  の降下ステップを考えると (付録: デルタ規則の導出)、

色:

$$\bar{\mathbf{x}}_k(t+1) = (1 - \gamma_x)\bar{\mathbf{x}}_k(t) + \gamma_x \mathbf{x}_k(t) \quad (13)$$

$$\mathbf{X}_k(t+1) = (1 - \gamma_X)\mathbf{X}_k(t) + \gamma_X (\mathbf{x}_k(t) - \bar{\mathbf{x}}_k(t))(\mathbf{x}_k(t) - \bar{\mathbf{x}}_k(t))^t \quad (14)$$

位置:

$$\bar{\mathbf{p}}_k(t+1) = (1 - \gamma_p)\bar{\mathbf{p}}_k(t) + \gamma_p \mathbf{p}_k(t) \quad (15)$$

$$\mathbf{P}_k(t+1) = (1 - \gamma_P)\mathbf{P}_k(t) + \gamma_P (\mathbf{p}_k(t) - \bar{\mathbf{p}}_k(t))(\mathbf{p}_k(t) - \bar{\mathbf{p}}_k(t))^t \quad (16)$$

動き:

$$\mathbf{G}_k(t+1) = (1 - \gamma_G - \gamma_{m_s})\mathbf{G}_k(t) + \gamma_G \mathbf{g}_k(t)\mathbf{g}_k(t)^t + \gamma_{m_s} \mathbf{m}_s \mathbf{m}_s^t \quad \mathbf{m}_s = (1, 1, 0)^t, \gamma_{m_s} \ll \gamma_G \quad (17)$$

なるデルタ規則が得られる。ここで、 $0 < \gamma_\theta < 1, \theta = \{\bar{x}, X, \bar{p}, P, G\}$  は利得係数である。 $i \neq k$  なる全ての  $\theta_i$  については、デルタ規則は  $\theta_i(t+1) = \theta_i(t)$  となる。

式 17 に関して、動きの先験的情報が既知であり、事前確率  $p(\omega_i)$  が

$$p(\omega_i) \sim \exp[(\mathbf{m}_s^t \mathbf{m}_i)^2], \mathbf{m}_s = (1, 1, 0)^t \quad (18)$$

の形式で与えている。 $(\mathbf{m}_s^t \mathbf{m}_i)^2$  の項は、大きな動きベクトルに対するペナルティーである。式 17 の  $\gamma_{m_s} \mathbf{m}_s \mathbf{m}_s^t$  は式 18 を  $\mathbf{m}_i$  について勾配を求めることにより得られ、デルタ規則の安定化項となる。

### 3.3 ロバスト統計

領域断片の境界から得られる画像特徴は、はずれ値を持つ特徴 (outlier) となり、パラメータ推定に悪影響を与える可能性がある。そこで、デルタ規則をロバスト統計の考えに沿って、以下に変更する。変更点は、利得係数  $\gamma_\theta, \theta = \{\bar{x}, X, \bar{p}, P, G\}$  を  $\chi^2(\mathbf{x}, \mathbf{p}, \mathbf{g})$  の評価値の関数にすることである。ここで、 $\chi^2(\cdot)$  はマハラノビス距離関数で、 $(\mathbf{x} - \bar{\mathbf{x}}_i)^t \mathbf{X}_i^{-1} (\mathbf{x} - \bar{\mathbf{x}}_i) + (\mathbf{p} - \bar{\mathbf{p}}_i)^t \mathbf{P}_i^{-1} (\mathbf{p} - \bar{\mathbf{p}}_i) + (\mathbf{m}_i^t \mathbf{g}) / s_i^2$  として定義され、距離関数  $d_i(\cdot)$  に含まれている。変更は、 $\theta = \{\bar{x}, \bar{p}\}$  に対して、

$$\gamma_\theta(\chi_i^2(\mathbf{x}, \mathbf{p}, \mathbf{g})) = \begin{cases} \gamma_\theta & \chi_i^2(\cdot) < \chi_{th}^2 \text{ の場合} \\ \gamma_\theta \sqrt{\chi_{thr}^2 / \chi_i^2(\cdot)} & \text{上記以外} \end{cases} \quad (19)$$

また、 $\theta = \{X, P, G\}$  に対して、

$$\gamma_\theta(\chi_i^2(\mathbf{x}, \mathbf{p}, \mathbf{g})) = \begin{cases} \gamma_\theta & \chi_i^2(\cdot) < \chi_{th}^2 \text{ の場合} \\ \gamma_\theta \frac{\chi_{thr}^2}{\chi_i^2(\cdot)} & \text{上記以外} \end{cases} \quad (20)$$

となる。仮定より  $\chi_i^2(\cdot)$  は  $\chi^2$  乗分布に従うことから、しきい値  $\chi_{th}^2$  は自由度 (この場合は 6) と  $\chi^2$  乗検定の危険率から求めることができる。以上の利得係数に関するデルタ規則の変更は、Huber のミニマックス型  $\psi$  関数を用いた M-推定 [18] の近似になっている。

### 3.4 領域断片の分割と消去

競合学習の中で、デルタ規則に加えて 2 つの規則を加える。

分割規則:  $(\ln |\mathbf{X}_i| + \ln |\mathbf{P}_i| + \ln s_i^2)$  が所定のしきい値以上である断片は  $\mathbf{P}_i$  の 1 番目の固有値に相当する固有ベクトルに沿って分割する。

消去規則: 所定の期間、一度も標本の割当が得られなかった領域断片を取り消す。

以上の規則は、画像中の物体移動に伴う領域の出現と消失に対応するためである。

## 4 実験

全ての実験で、条件は:

- 入力画像は家庭用 VCR から得られた RGB カラー画像。サイズは  $256 \times 240$  画素で、偶数フィールドの画像をフレーム画像として使用。
- 全ての  $\theta = \{\bar{x}, X, \bar{p}, P, G\}$  に対して、 $\gamma_\theta = 0.03$ 。
- ランダムな標本化を最初のフレームについては 30,000 回、その後のフレームについては、10,000 回行なう。
- 領域断片数  $c$  は最大 100。

図 4(a) に領域記述の初期化の例を示す。競合学習の結果、図 4(b) が得られた。楕円は、クラスタリングされた画素の集合を表しており、その半径は  $\mathbf{P}_i$  の固有値平方根の  $\sqrt{3}$  倍である。楕円中心のドットとそのドットより延びる線分は、位置平均と動きベクトルを表している。得られた領域記述を用いて、各画素について、画素が属する断片の動きベクトルを求めると、図 2(d) が得られる。この競合学習には、SPARCstation2 で 2 分程度要した。以下の節では、提案手法の 3 つの応用例を示して、その能力を明らかにする。

### 4.1 領域の追跡

図 5 に、図 4 に示した人物像領域を 30 フレーム追跡した結果である。左上の画像で、人物領域に属する断片を予

め選んでいる。左上から右下にかけて、追跡結果の中で代表的なフレームを示した。この実験では、追跡能力を向上させるために、事前確率密度関数  $p(\omega_i)$  を:

$$p(\omega_i) \sim \exp[(m_0^t m_i)^2] \exp[(\bar{x}_i - \hat{x}_i)^2] \exp[(\hat{p}_i - \bar{p}_i)^2], \quad (21)$$

と定めた。ここで、 $\hat{x}$  と  $\hat{p}$  は前フレームでの現フレームの予測色と予測位置である。前フレーム最後の標準化の時刻を  $t-1$  とすると、 $\hat{x}$  と  $\hat{p}$  は:

$$\hat{x} = \bar{x}_i(t-1)$$

$$\hat{p}_i = \bar{p}_i(t-1) + (u_i(t-1), v_i(t-1))^t$$

として求められる。この事前確率密度に対応して、式13と式14に書かれたデルタ規則を、式17と同様に、 $\bar{x}_i$  と  $\bar{p}_i$  に関する勾配を加えることにより変更する。

## 4.2 物体／背景分離

図6は、物体／背景分離に、提案手法を適用した結果である。この画像では、物体毎に見かけの速度が異なるため、動き情報が物体／背景分離の大きな手がかりとなる。分離誤りは少しあるものの、全体として走行物体と背景が分離できている。この画像には、車、道路等に輝度勾配がほとんど観察されない画素が多くあるが、色の情報により動きベクトルが補間されている。

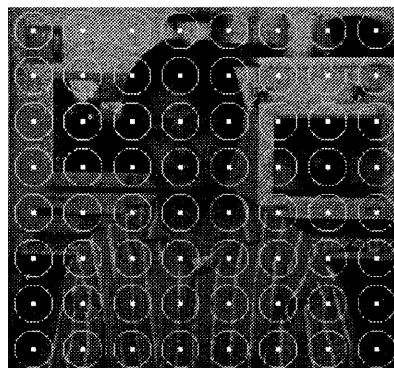
## 4.3 ステレオ

カメラと物体間の相対的な並行移動により、相対的な距離を求めることができる。提案手法が、動領域の境界付近でも安定に精度良く求めることができることを示すために、図7に、相対的な距離を求めて、距離に応じてレリーフとして表示した例を示す。手、頭、胴体の相対的な位置が微妙に求まっている。

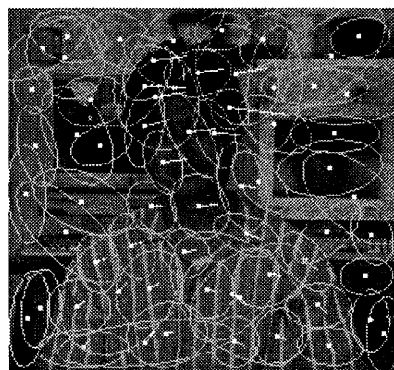
## 5 議論

従来の輝度勾配を用いる方法に比べて、提案手法では、動きベクトルの粒度が画素単位ではないものの、安定かつ正確に動きベクトルを求めることができる。特に、従来では輝度勾配の値が十分でなく大きな推定誤りとなっていた領域や、物体境界付近での動き推定に効果がある点が特長である。競合学習と距離関数の性質を考えてみると、 $(\ln |X_i| + \ln s_i^2)$  が小さな領域(輝度変化の小さな領域)では、 $\ln |P_i|$ (領域の大きさを表現)が大きくなる。逆に  $(\ln |X_i| + \ln s_i^2)$  が大きな領域(輝度変化の大きな領域)では  $\ln |P_i|$  が小さくなる。これは、ちょうど、「適応サイズ/形状窓」を用いて動きベクトル推定を行なっていることに等しい。

1章で述べた窓問題に関して、共分散行列  $G$  の固有値から得られた動きベクトルの信頼度を評価することができる。第2固有値  $\lambda_{2(i)}$  が第3固有値に比べて大きい場合は、動きベクトルは安定であるが、 $\lambda_{2(i)} \approx \lambda_{3(i)}$  であるような場合は、動きベクトル推定に必要な独立した方向を持つエッジ成分がなく、不安定であることが分かる。



(a) Initialization of a fragmental description.



(b) A resulting description.

図4 領域記述獲得例

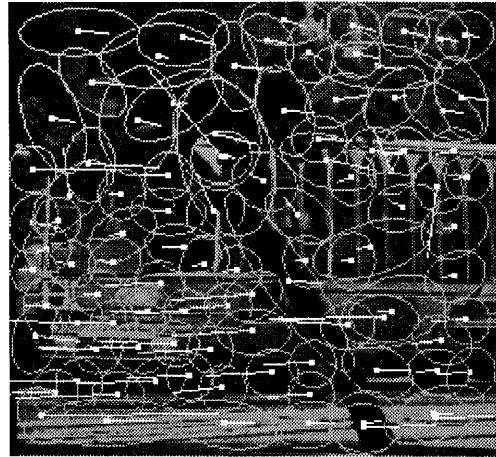
初期化の段階で、中央にドットを持つ円は64個、等間隔に配置され、 $\bar{x}_i$  と  $\bar{p}_i$  には中央の画素値を、 $X_i, M_i$  には単位行列を、そして  $P_i$  には、円の半径に相当する対角行列を用いた。



図5 人物領域30フレームの追跡結果  
サイズ  $3 \times 3$  のブロック単位で、追跡対象の断片に属しているブロックを選び、ブロック内画素の輝度平均値を表示している。



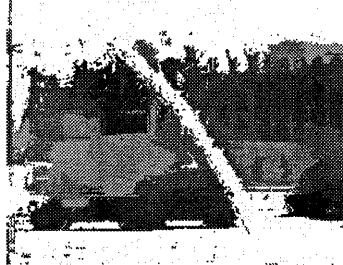
(a) A road scene.



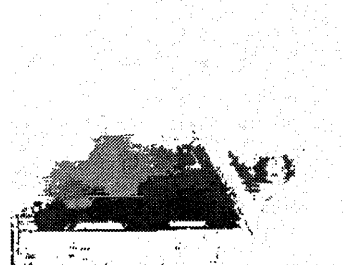
(b) A resulting description.



(c) Ground Segments.



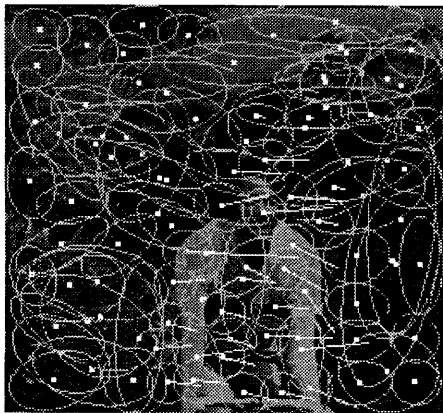
(d) Tracked Segments.



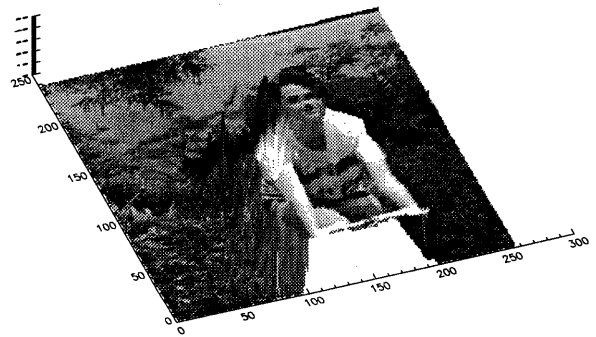
(e) A fast car segments.

図6 物体／背景分離の道路シーンへの適用例

観測者のカメラは中央のバスを追っており(バスの動きベクトルは左へ約1画素/フレーム)、背景は左から右へ動いている。手前の車(fast car)が左へ5画素/フレームの速度でバスを追い抜いている。(b)は得られた領域記述。同じ動き成分をもつ断片をグループ化して、背景領域(c)、車の領域(d)、追い越し車領域(e)を各々得た。各領域の画素値は、属する領域断片の平均色 $\bar{x}_i$ で置き換えている。



(a) A resulting description.



(b) A relief produced from the motion.

図7 ステレオへの応用例

自転車に乗る人物がカメラの前を左から右に横切っている。動きベクトルは約4画素/フレームである。(a)は得られた領域記述で、(b)は水平方向の動き成分より得られた深さ情報を用いてレリーフとして入力画像を表示している。

## 6 まとめ

本報告では、色と輝度勾配の領域分割による2次元動き推定法提案し、実験により効果を示した。提案手法は単なる2次元動き推定手法ではなく、領域断片毎に時系列画像中で連続的に色、位置、動きのパラメータを求めていくという、動画像記述の意味を持っている。

今後の課題として、シーン切替時の領域記述の更新手法、デルタ規則の収束性の検討などがある。

**謝辞** 本研究の機会を与えて頂いた松下電器 中央研究所 新田恒治 取締役 所長ならびに 電子機器基礎研究所 山下貞彦 所長に深謝します。また日頃より御助言頂く、同研究所 山田隆弘室長、大阪大学浅田 稔助教授に感謝します。

## 参考文献

- [1] D. Marr and S. Ullman: "Directional selectivity and its use in early visual processing" Proc. Royal Soc. London, B211, pp.151-180 (1981).
- [2] J. Bergen, P. Anandan, K. Hanna, and R. Higorani: "Hierarchical Model-Based Motion Estimation", Proc. of ECCV-92, pp.237-252, (1992)
- [3] B. Horn and B. Schunk: "Determining Optical Flow", Artificial Intelligence, 17, pp.185-203, (1981).
- [4] B. Lucas and T. Kanade: "An iterative image registration technique with an application to stereo vision", Proc. DARPA Image Understanding Workshop pp.121-130, (1981).
- [5] M. Srinivasan: "Generalized Gradient Schemes for the Measurement of Two-Dimensional Image Motion", Biol. Cybern. 63, pp.421-431 (1990).
- [6] J. Barron, D. Fleet, S. Beauchemin and T.A.Burkitt: "Performance of Optical Flow Techniques": Proc. CVPR'92, pp.236-242, (1992).
- [7] G. Adiv: "Determining Three-Dimensional Motion and Structure from Optical Flow, Generated by Several Moving Objects", IEEE Trans. on Patt. Anal. & Mach. Intell., Vol.7, No.4, pp.384-401, 1985.
- [8] 柴藤 稔, 白井良明, 浅田 稔: "クラスタリングから得られる領域の記述に基づく動的な輪郭抽出", 信学論 (D-II), J75-D-II, 7, pp.1111-1119, (1992-7).
- [9] J. Harris, C. Koch, E. Staats, and K. Luo: "Analog Hardware for Detecting Discontinuities in Early Vision", IJCV, 4, 211-223, (1990).
- [10] C. Fennema and W. Thompson: "Velocity Determination in Scenes Containing Several Moving Objects", Comp. Graphics & Image Process., 9, pp.301-315 (1979).
- [11] B.G.Schunck: "Image Flow Segmentation and Estimation by Constraint Line Clustering" IEEE, Trans. on Patt. Anal. & Mach. Intell., Vol.11, No.10, pp.1010-1027, (1989).
- [12] T. Darel and A. Pentland: "Robust Estimation of a Multi-Layered Motion Estimation", IEEE Workshop on Visual Motion, pp.173-178 (1991).

- [13] T. Kohonen: "The Self-Organizing Map", Proceedings of the IEEE, Vol. 78, no.9 (Sept. 1990).
- [14] G. Klinker, S.Shafer, and T.Kanade: "A physical approach to color image understanding", Int. J. Computer Vision, Vol.4, no.1, pp.7-38 (1990).
- [15] S. Tominaga: "Surface Identification Using the Dichromatic Reflection Model", IEEE, Trans. on Patt. Anal. & Mach. Intell., Vol.13, No.7, pp.658-670, (1991).
- [16] G. Healey: "Segmenting Images Using Normalized Color", IEEE Trans. on Sys., Man, and Cyber., vol.22, no.1, pp.64-73 (1992).
- [17] R. Duda and P. Hart: *Pattern Classification and Scene Analysis*. John Wiley and Sons (1973)
- [18] P.J. Huber: *Robust statistics*, Wiley, (1981).

## 付録: デルタ規則の導出

色に関するパラメータ  $\bar{x}_i, \mathbf{X}_i$  に関して、式7  $d_i(\cdot)$  の勾配を考える。他のパラメータについても導出は同じである。簡単のため、 $pp(\omega_i)$  が色に関して独立であるとする。ここで、 $x_p$  を  $\mathbf{x}$  の  $p$  番目の要素、 $\bar{x}_{p(i)}$  を  $\bar{x}_i$  の  $p$  番目の要素、 $\sigma_{pq(i)}$  を  $\mathbf{X}_i$  の  $p$  行  $q$  列の要素、 $\sigma'_{pq(i)}$  を  $\mathbf{X}_i^{-1}$  の  $p$  行  $q$  列の要素とする。 $d_i(\cdot)$  の  $\bar{x}_i$  と  $\sigma'_{pq(i)}$  に関する勾配は式22, 式23として得られる ([17] 参照)。

$$\nabla_{\bar{x}_i} d_i(\mathbf{x}, \mathbf{p}, \mathbf{g}) = -\mathbf{X}_i^{-1}(\mathbf{x} - \bar{x}_i) \quad (22)$$

$$\frac{\partial d_i(\mathbf{x}, \mathbf{p}, \mathbf{g})}{\partial \sigma'_{pq(i)}} = -(1 - \frac{\delta_{pq}}{2}) \{ \sigma_{pq(i)} - (\mathbf{x}_p - \bar{x}_{p(i)})(\mathbf{x}_q - \bar{x}_{q(i)})^t \} \quad (23)$$

クロネッカーのデルタ  $\delta_{pq}$  の項は共分散行列とその逆行列が対称行列であることによる。一方、 $\frac{\partial \sigma'_{pq(i)}}{\partial \sigma_{pq(i)}}$  は、

$$\frac{\partial \sigma'_{pq(i)}}{\partial \sigma_{pq(i)}} = \frac{(1 - \delta_{pq}) \partial \mathbf{X}_{pq(i)}}{|\mathbf{X}_i|} - (2 - \delta_{pq})(\sigma'_{pq(i)})^2 \quad (24)$$

と表せる。 $\mathbf{X}_{pq(i)}$  は  $\mathbf{X}_i$ 、 $pq$  成分の余因子行列である。従って、 $t$  を時間変数にとり、 $\sum_R \sum_{\omega_i} d_i(\mathbf{x}(t), \mathbf{p}(t), \mathbf{g}(t))$  を減少させる最急降下ステップを用いたデルタ規則は式25, 26で表せる。

$$\bar{x}_i(t+1) = \bar{x}_i(t) + \gamma_1 \mathbf{X}_i^{-1}(t)(\mathbf{x}(t) - \bar{x}_i(t)) \quad (25)$$

$$\sigma_{pq(i)}(t+1) = \sigma_{pq(i)}(t) + \gamma_2 \left[ \left( 1 - \frac{\delta_{pq}}{2} \right) \{ \sigma_{pq(i)}(t) - (\mathbf{x}_p(t) - \bar{x}_{p(i)}(t))(\mathbf{x}_q(t) - \bar{x}_{q(i)}(t))^t \} \right] \left[ \frac{1 - \delta_{pq}}{|\mathbf{X}_i(t)|} \frac{\partial \mathbf{X}_{pq(i)}(t)}{\partial \sigma_{pq(i)}(t)} - (2 - \delta_{pq})(\sigma'_{pq(i)}(t))^2 \right] \quad (26)$$

$0 < \gamma_1, \gamma_2 < 1$  は利得係数である。以上の最急降下ステップによるデルタ規則は実際の実験では不安定であった。理由は、降下ステップが、まだ得られていない(妥当な値に収束していない)共分散行列に依存していることによる。そこで、 $\mathbf{X}_i^{-1}$  が半正値、Eq.24 が半負値であることに注目してこれらは各々、正と負の定数値に簡単化し、さらに  $\delta_{pq}$  を無視すると3章に示したデルタ規則が得られる。