

対象の動きを利用した画像分割法

今泉 竜一 本谷 秀堅 出口 光一郎 (東大工学部)

本稿では対象の動きを利用し、画像を異なる動きを持つ複数の領域に分割する一手法を示す。画像中の一点はその点における画像の時空間微分で与えられる Optical Flow Constraint をみだす。このことから、動きからの領域分割は時空間微分で張られる空間における複数の平面の当てはめ問題となる。これを正規分布の Mixture Models でモデル化し、EM アルゴリズムによって解を求める。更に、あるモデルにおける隣接する 2 点間の距離を点が同時に同じモデルに属す事後確率から定義する。この距離を用いて各点からそれぞれのモデルまでの距離を測ることができ、複数の平面の交線上の点についても適切に割り当てをすることが可能となる。

Image segmentation from motion

Ryoichi IMAIZUMI, Hidekata HONTANI and Koichiro DEGUCHI
(University of Tokyo)

We present a method to segment an image from its motion. A pixel in an image satisfy an Optical Flow Constraint, whose coefficients are given by the spatio-temporal derivatives. The segmentation problem becomes the plan fitting problem in this space. We solve this problem with Mixture Models and EM algorithm. Then we define a distance of two points on a model from the posterior probability that the two points belong to the same model. Measuring distance of a point to a model with this distance, we can assign points on the intersection line of multiple planes properly, which cannot be assigned with Mixture Models.

1 はじめに

画像中に動きの異なる複数の物体が写っているとき、動きを元に画像を分割する問題を考える。対象の画像上での動きを表すベクトル (u, v) はオプティカルフローと呼ばれる。同一の対象は一定のオプティカルフローを持つと仮定する。

時刻 t ，画像上の点 $p = (x, y)$ におけるオプティカルフローが $(u(p, t), v(p, t))$ であるとしたとき

$$I(x, y, t) = I(x - u\Delta t, y - v\Delta t, t - \Delta t) \quad (1)$$

が成立するという仮定 (brightness constancy assumption) をおく。 $(u\Delta t, v\Delta t)$ が微小であると仮定すると、式 (1) を Taylor 展開し一次近似をすることで、optical flow constraint と呼ばれる次式が求まる。

$$I_x(p, t)u(p, t) + I_y(p, t)v(p, t) + I_t(p, t) = 0 \quad (2)$$

式 (2) は自由度を 1 つしか拘束しないので、オプティカルフロー (u, v) を画像中の 1 点から求めることはできない。 (u, v) を求めるためには同じ対象の異なる時空間微分値が必要となる。 (u, v) を求めるために利用するこの画像中の領域を「窓」と呼ぶ。しかし、どの点が同一の対象に属するかをあらかじめ知ることはできない。正確にフローを求めるためには、窓の内部に複数の対象が含まれないように窓の大きさ及び形を適切に定める必要がある [2]。しかし、形には膨大な組み合わせが存在し、ある判定基準を満たすような窓の形を探索することは容易ではない。

本稿では全領域から (I_x, I_y, I_t) で作られる時空間微分空間の点を求め、それらの点に対して複数の平面を当てはめることで、各点の所属とそれぞれの対象の動きの推定値を求める (Fig. 2)。複数平面の当てはめ問題は Mixture Models として定式化し、EM アルゴリズムによって解を求める。

しかし、平面同士の交線上の点についてはどちらの平面に属しているかを決定することができない。これらいずれのモデルに属するかを決定できない点を不確定点と呼ぶことにする。不確定点をいずれかのモデルにあてはめるために、同一の対象は画像上で連結しているという仮定をおく。この仮定から、不確定点周り

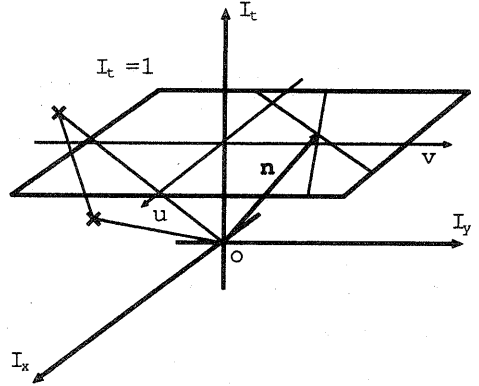


Fig.1: 時空間微分とオプティカルフロー。独立な 2 点が張る平面の原点を通る法線と $I_t = 1$ の交点がオプティカルフローを与える。

に連結した窓を定め、その窓内の点の所属を調べることによって不確定点の割り当てを定める。しかし、点の所属は確率的な変数であり完全に定まっているわけではないということが問題となる。従って、所属の確率を考慮に入れた窓を定める必要がある。

1 点があるモデルに属す確率は、Mixture Models によって各点の平面への所属の事後確率として得られる。離れた 2 点を結ぶ経路は複数存在するが、その中で経路上の点すべてがあるモデルに属す確率を最大にするような経路を求める。そして、その最大確率を 2 点間の距離とする。更にあるモデルへの距離を、不確定点とそのモデルに確定している点との間の距離とする。このモデルへの距離は不確定点とあるモデルが連結している確率を与えるので、各モデルに対してそれぞれモデルへの距離を求め、比較することで不確定点がどのモデルに属するかを決定できる。

また、 (u, v) を運動パラメータ a の線形和で表すことで対象の運動の自由度を増すことができる [1]。これは

$$u(p; a) = Ma \quad (3)$$

と表せる。具体的には、回転による変位が小さく対象

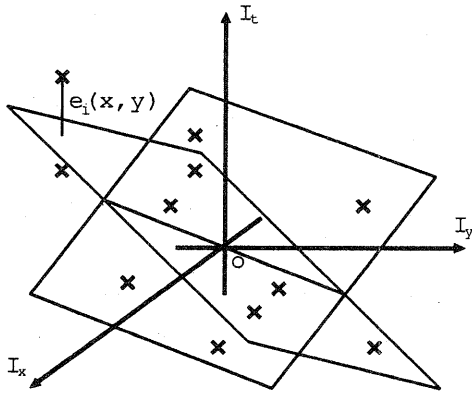


Fig.2: 平面当てはめによる領域分割

までの距離に比べて深さ方向の変化が小さいときには

$$M = \begin{pmatrix} 1 & x & y & 0 & 0 & 0 & x^2 & xy \\ 0 & 0 & 0 & 1 & x & y & xy & y^2 \end{pmatrix} \quad (4)$$

また、直交投影下での平面の動きや、カメラの pan, tilt, zoom, 微小移動を仮定すると

$$M = \begin{pmatrix} 1 & x & y & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & x & y \end{pmatrix} \quad (5)$$

となる。

式(3)を式(2)に代入すれば

$$(I_x, I_y)M\mathbf{a} + I_t = 0 \quad (6)$$

となり、 $(I_x, I_y)M$ は式(4)式(5)のときそれぞれ $(I_x, I_y, xI_x, yI_y, xI_x, yI_y, x^2I_x + xyI_y, xyI_x + y^2I_y)$, $(I_x, I_y, xI_x, yI_y, xI_y, yI_x)$ となる。よって、上で述べた (I_x, I_y, I_t) の3次元空間がそれぞれ9次元空間、7次元空間となるだけで、基本的に同じ平面の当てはめ問題として扱える。

以下では簡単のため2次元の運動モデルを扱う。

2 Mixture Models による分離

複数平面の当てはめを行うために Mixture Models を用いる [3].

Mixture Models とはある確率密度分布が複数のモデルの確率密度分布の重み付きの和で表されるとする

ものである。すなわち、観測値 $D = \{d_1, \dots, d_n\}$ が得られたとき d_j の確率密度分布 f を、モデル i の確率密度分布 f_i の重み付きの和として

$$f(d_j; \phi) = \sum_{i=1}^g \pi_i f_i(d_j; \theta) \quad (7)$$

のように表わす。ここで π_i はモデル i の重みを表し

$$\sum_{i=1}^g \pi_i = 1 \quad (8)$$

である。 g はモデル数、 θ は確率密度分布 f のパラメータベクトル、 $\phi = (\pi', \theta')$ はすべての未知パラメータを表すベクトルである。

本手法では点 p_j をモデル i にあてはめたときの残差を

$$e_{ij} = e_i(p_j) = I_x(p_j, t)u_i + I_y(p_j, t)v_i + I_t(p_j, t) \quad (9)$$

と置き、 e_{ij} の確率密度分布を平均0分散 σ_i の正規分布と仮定する。このとき、 e_{ij} は点 p_j についての I_t 方向へ測ったモデル i の平面への距離である (Fig.2)。このとき、式(7)は

$$f(p_j) = \sum_{i=1}^g \pi_i \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{e_{ij}^2}{2\sigma_i^2}\right) \quad (10)$$

となる。

上で述べたようなモデルに対する最尤推定を行うアルゴリズムに EM アルゴリズムが知られている。まず点 p_j がモデル $i(M_i)$ に属することを示す変数 z_{ij} を次のように定義する。

$$z_{ij} = \begin{cases} 1 & \text{if } p_j \in M_i \\ 0 & \text{if } p_j \notin M_i \end{cases} \quad (11)$$

$\{z_{ij}\}$ が与えられた時の全体の対数尤度は

$$L_C(\phi) = \sum_{i=1}^g \sum_{j=1}^n z_{ij} (\log \pi_i + \log f_i(d_j; \theta)) \quad (12)$$

となる。また、 d_j が得られたときの点 p_j がモデル i に属す事後確率を τ_{ij} とすると、

$$\tau_{ij} = f(z_{ij} | d_j; \phi) \quad (13)$$

$$= \frac{\pi_i f_i(d_j | u_i, v_i, \sigma_i)}{\sum_{i=1}^g \pi_i f_i(d_j | u_i, v_i, \sigma_i)} \quad (14)$$

である。

EM アルゴリズムは E step と M step の繰り返しからなる。E step は現在のパラメータの推定値が与えられたときの L_C の条件付き期待値 Q を求め、M step では Q を最大にするようなパラメータを求める。この E step と M step を繰り返すことで ϕ を求める。

1. E step

$$\begin{aligned} Q(\phi, \phi^{(k)}) &= E[L_C(\phi)|X; \phi^{(k)}] \\ &= \sum_{i=1}^g \sum_{j=1}^n \tau_{ij} (\log \pi_i + \log f_i(x_j; \theta)) \end{aligned} \quad (15)$$

2. M step

$$\phi^{(k+1)} = \arg \max_{\phi} Q(\phi, \phi^{(k)}) \quad (17)$$

式(17)より、 ϕ の要素である π と θ について

$$\hat{\pi}_i = \sum_{j=1}^n \hat{\tau}_{ij} / n \quad (18)$$

$$\sum_{i=1}^g \sum_{j=1}^n \hat{\tau}_{ij} \frac{\partial \log f_i(x_j; \theta)}{\partial \theta} = 0 \quad (19)$$

となる。式(10)の場合さらに式(19)も陽に解ける。すなわち、分散 σ_i については式(19)を解けば

$$\hat{\sigma}_i^2 = \frac{1}{n\hat{\pi}_i} \sum_j \tau_{ij} e_j^2 \quad (20)$$

となる。また、オプティカルフロー (u_i, v_i) についても式(19)より

$$\sum_{j=1}^n \hat{\tau}_{ij} I_x(\mathbf{p}_j) (I_x(\mathbf{p}_j) u_i + I_y(\mathbf{p}_j) v_i - I_t(\mathbf{p}_j)) = 0 \quad (21)$$

$$\sum_{j=1}^n \hat{\tau}_{ij} I_y(\mathbf{p}_j) (I_x(\mathbf{p}_j) u_i + I_y(\mathbf{p}_j) v_i - I_t(\mathbf{p}_j)) = 0 \quad (22)$$

となる。これは、 $\hat{\tau}_{ij}$ を重みとした重み付き最小2乗法を解くことになっている。

結局 EM アルゴリズムは以下のステップの繰り返しとなる。

1. z_{ij} の事後確率 τ_{ij} を式(14)により計算

2. 重み確率 π_i を式(18)により更新

3. 分散 σ_i を式(20)により更新

4. オプティカルフロー (u_i, v_i) を式(21)(22)により更新

3 再割り当て法

時空間微分空間における平面の割り当てによりオプティカルフローが求まり領域分割がなされるが、平面の交線上の点については領域の割り当ては不確定である。このような不確定点に対する割り当てを行うために、同一の対象は連結して、それら連結している点が時空間微分空間中で平面をなすという仮定をおく。ここで、Mixture Models から得られたモデル i に属す事後確率である τ_{ij} を使って画素間の距離を定義する。連結した点がすべて同じモデルに属すもってもらしさを、この τ_{ij} による距離によって測ることができる。

また、画素間の距離の形で表現したことで他の画像情報(輝度勾配や色等)を同時に用いることが容易になる。

3.1 画素間の距離

隣接する2点 $\mathbf{p}_j, \mathbf{p}_k$ 間のモデル i における距離を

$$d(\mathbf{p}_j, \mathbf{p}_k; i) = -\frac{1}{2} \log \tau_{ij} \tau_{ik} \quad (23)$$

と定義する。

すると、隣接していない2点 $\mathbf{p}_a, \mathbf{p}_b$ のモデル i における距離は次のようになる。点 \mathbf{p}_a から点 \mathbf{p}_b へのある経路を \mathcal{P} とおいたとき $\mathcal{P} = \{\mathbf{p}_{p1}, \mathbf{p}_{p2}, \dots, \mathbf{p}_{pN}\}$ とすれば

$$D_{\mathcal{P}}(\mathbf{p}_a, \mathbf{p}_b; i) = \min_{\mathcal{P}} \sum_{j=1}^{N-1} d(\mathbf{p}_j, \mathbf{p}_{j+1}; i) \quad (24)$$

となる。よって

$$D(\mathbf{p}_a, \mathbf{p}_b; i) = D_{\mathcal{P}}(\mathbf{p}_a, \mathbf{p}_b; i) - \frac{1}{2} (\log \tau_{i1} + \log \tau_{iN}) \quad (25)$$

により \mathcal{P} 上の点 \mathbf{p}_a がすべてモデル i に属す最大の確率が分かる。よって、 $D(\mathbf{p}_a, \mathbf{p}_b; i)$ は点 \mathbf{p}_a と点 \mathbf{p}_b がモデ

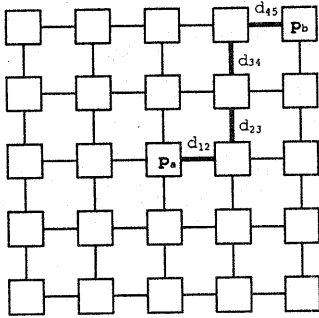


Fig. 3: 一点からの距離の定義

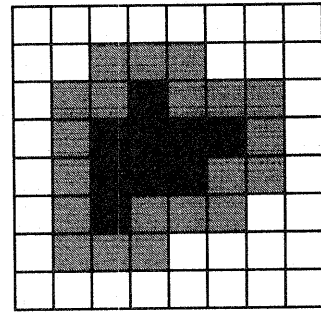


Fig. 4: 距離の求め方: 黒は確定領域, グレーは候補領域, 白は未確定領域

ル i 上でつながっているという条件の確からしさの基準となる。

距離 $D(p_a, p_b; i)$ は次のようにして求めることができる (Fig. 4)。ある領域までの距離が求められているとする。この時、距離が確定した領域 (A_D) と接している未確定領域 (A_U) を候補領域 (A_C) と呼ぶ。 A_C の一点 p_c の暫定的な値は p_c に隣接する A_D 内点 p_d までの距離と p_a, p_d 間の距離の和の最小値として求まる。このとき、 A_C の中でこの暫定値が最小となる点 (p_m) は A_D に加えることができる。なぜなら、 p_m へ至る他の経路は必ず p_m 以外の A_C 中の点を含むからである。結局、ある点 p_a からの距離を求める手順は以下ようになる。

1. 点 p_a に隣接する各点を A_U に加える。
2. A_U の暫定値を求める
3. A_U の中で最小の点を A_D に加える
4. 3で加えた点に隣接する点を A_U に加え、2へ戻る

この手順により p_a から順に A_C を広げていくことで、ある点からの等距離線が求まる (Fig. 8)。これにより、近傍の情報を用いて一点の運動パラメータを求めるときの窓のサイズおよび形状の決定が一次元の探索により可能となる。

3.2 再割り当ての手順

上で定義した距離 D を用いて各点の再割り当てを行う。画像上で不確定点とモデル内部の点とを結ぶ道筋を考える。その道筋上の各点と同じモデルに高い確率で属しているならば、注目している不確定点とそのモデルとの距離を近いと判断する。

具体的には、まず点 p_a から距離の小さい順に窓を広げていく。窓内の点により時空間微分空間中で平面が張られたとき、 p_a から平面を張った点までの距離を点 p_a のモデル i への距離 $D_M(p_a, i)$ と定める。ここで、 D_M が小さい順に $D_M(i_1), D_M(i_2), \dots, D_M(i_g)$ であるとする。このとき $D_M(p_a, i_2)/D_M(p_a, i_1)$ がある閾値 $R_D (R_D \geq 1)$ より大きいとき、点 p_a はモデル i_1 に属すと決定する。 R_D より小さい点については動きだけでは決定できない点と判定する。

ある点 p_j について、残差 e_{ij} の2乗を小さい順に $E_{ij}^{(1)}, E_{ij}^{(2)}, \dots, E_{ij}^{(g)}$ とおく。このとき $E_{ij}^{(1)}/E_{ij}^{(2)}$ の比が r 以上となる点 p_j を未確定点と判定し、それらの点についてのみ再割り当てを行う。 r は適当な定数。未確定点以外の点は $E_{ij}^{(1)}$ を与えたモデルに属すと判定する。また、平面が張られたかの判定は単純に不確定点以外の連結した2点が窓の内部に含まれた時とした。

3.3 勾配情報の統合

モデルへの距離の比が閾値 R_D 以下となる点については上記の手法を用いても特定のモデルに確定できな

い。これは不確定な領域が候補となる複数の領域の境界に存在する場合に生じる。このような場合には、動きだけから割り当てを行うのは困難である。そこで、その他の情報、特に輝度の勾配の情報を統合する。

物体の境界には輝度のコントラストが多くの場合存在する。よって、画素間の距離として輝度の勾配を利用することができる。前節で定義した動きに基づいた画素間の距離に、勾配の距離として画素間の輝度差を加える。すなわち新たな隣接する2点 p_j, p_k 間の距離は

$$d_C(p_j, p_k; i) = d(p_j, p_k; i) + \alpha \|I(p_j) - I(p_k)\| \quad (26)$$

となる。この距離 d_C を未確定領域に用いる。モデルへの距離は前節と同じ方法により求める。このとき、 α を0から徐々に増してゆき比が R_D を越えた時点での最も近いモデルに割り当てをおこなう。

4 実験結果

合成画像を用いて実験を行った。実験に用いたパラメータは $R_D = 10, r = 0.8$ である。距離を求めるときの近傍は8近傍とした。また、モデル数は既知とした。

まず、動きだけから分割を行った結果を示す。テキスチャのある背景が静止しており正方形の対象が右下へ x 方向、 y 方向それぞれ2画素、1画素の速さで移動している。ただし、正方形の対象は輪郭付近にのみ勾配があり中心付近では輝度一定である (Fig.5)。

EM アルゴリズムを適用した結果は Fig.6 の通り、異なる平面を張る点については正しく分割されているが、正方形の中央の部分が背景に割り当てられている。運動パラメータ (オプティカルフロー) は Fig.7 にあるように、正しい値に収束している。

点 (13, 21) からの距離を窓の大きさ400画素で表示した結果が Fig.8。正方形の対象を回り込むように適切に窓が求められているのが分かる。

第3節で述べたように、まず再割り当てが必要な領域を抽出する (Fig.9)。さらに再割り当てを行った結果が Fig.10。正方形中央のテキスチャのない部分や時間微分が0となる領域について適切に割り当てがなされていることが分かる。

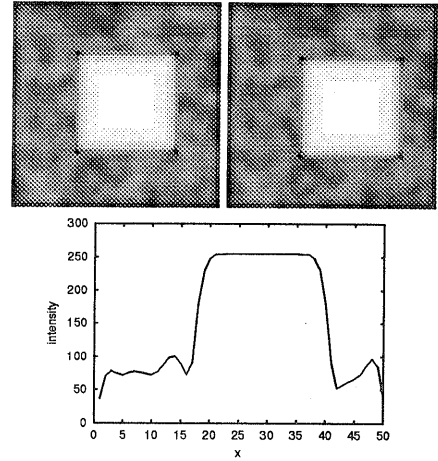


Fig.5: 実験画像とその断面

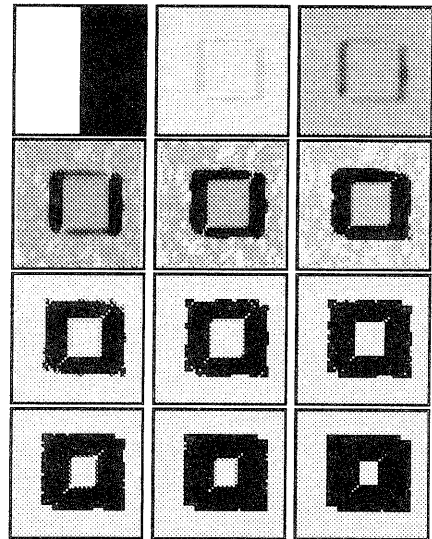


Fig.6: EM アルゴリズムの結果。左上から右下へ

次に輝度の勾配情報を利用した場合の結果を示す。実験画像は Fig.11 の通りで、テキスチャのある背景が静止しており円形の対象が右下へ x 方向、 y 方向共に1画素の速さで移動している。ただし、円形の対象は輪郭付近にのみテキスチャがあり中心付近では輝度一定である。

EM アルゴリズムを適用した結果は Fig.12 の通り。

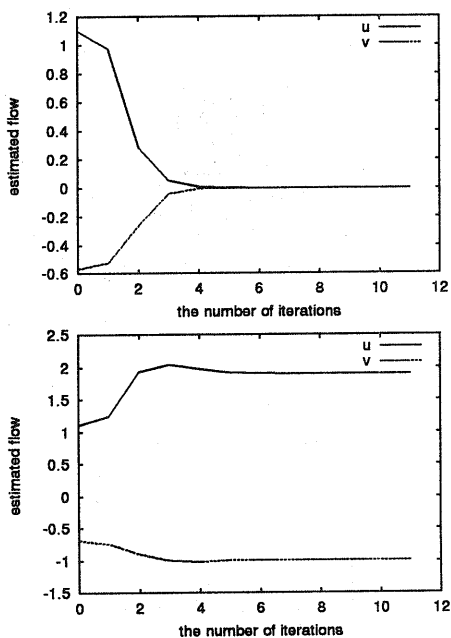


Fig.7: オプティカルフローの推定値. 上: 背景, 下: 前景

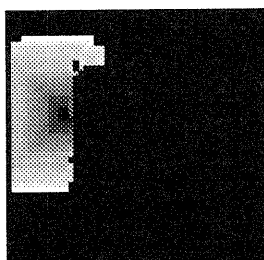


Fig.8: 点 (13,21) からの距離. 窓の大きさ 400

上と同様に円の中央及び時間微分が0となる点については割り当てができていない (Fig.13).

輝度情報を用いず動きだけで再割り当てを行った結果が Fig.14. 円の右上の部分に2つの領域に接した未確定領域が存在するために, 円中央部分が未確定と判定されている. この中央部分についても輝度情報を用いた再割り当てにより適切に割り当てられた (Fig.15).

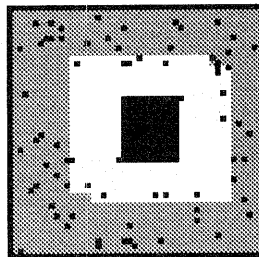


Fig.9: EM アルゴリズムによる分割結果: 黒い領域は不確定点

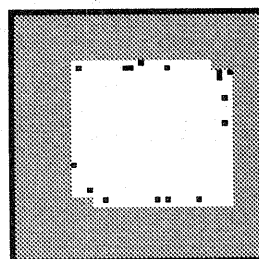


Fig.10: 動きだけからの再割り当て結果. 黒い領域は不確定点

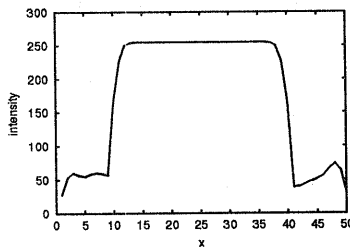
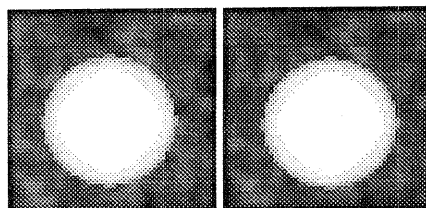


Fig.11: 実験画像とその断面

5 おわりに

画像の動きを利用して領域分割を行う手法を示した. 画像の時空間微分で与えられる Optical Flow

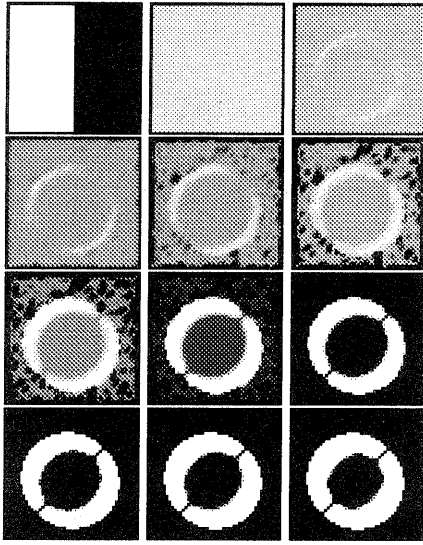


Fig.12: EM アルゴリズムの結果. 左上から右下へ

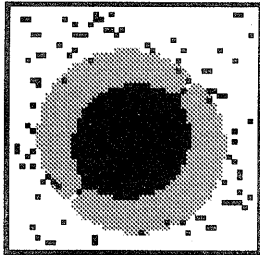


Fig.13: EM アルゴリズムによる分割結果: 黒い領域は不確定点

Constraint の残差が正規分布をなすとモデル化した。それを Mixture Models によりモデル化し、EM アルゴリズムにより割り当てと運動パラメータの推定を行った。Mixture Models によっては適切に割り当てができない複数のモデルに同時に属するような点についても、各モデルへの所属確率を考慮した画素間の距離を用いて点からモデルへの距離を測ることで適切な割り当てが可能となった。

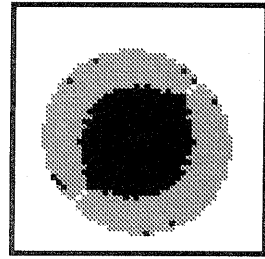


Fig.14: 動きだけからの再割り当て結果. 黒い領域は不確定点

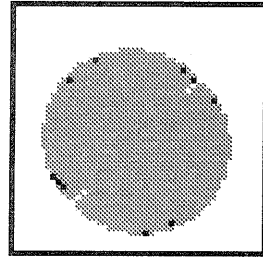


Fig.15: 輝度情報も利用した再割り当て結果. 黒い領域は不確定点

参考文献

- [1] Gilad Adiv. Determining three-dimensional motion and structure from optical flow generated by several moving objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-7, No. 4, pp. 384-401, 1985.
- [2] Micheal J. Black and P. Anandan. The robust estimation of multiple motions: Parametric and piecewise-smooth flow fields. *Computer Vision and Image Understanding*, Vol. 63, No. 1, pp. 75-104, 1996.
- [3] Geoffrey. J. McLachlan and Kaye E. Basford. *Mixture Models Inference and Applications to Clustering*. New York and Basel: Marcel Dekker Inc., 1988.