

動き情報に基づくエッジセグメントの最適分割

顧海松 浅田稔 白井良明

大阪大学工学部電子制御機械工学科

あらまし 本報告では、シーン全体に対するフレーム間の動き記述を最小化することに基づいて、人体のような複数運動部位を持つ物体のエッジセグメントを分割し各部位のフレーム間の動きを推定する方法について述べる。従来、オプティカルフローに基づく分割法が多く提案されているが、本手法は時空間勾配情報を利用し、MDL原理に基づいて分割を行う。まず、信頼性の高い空間的に安定なエッジセグメントを抽出し、これらの上の時空間勾配情報によって拘束された運動モデルによる運動推定量の良さ、モデルの複雑さ、および分割自身の記述、の三つの成分に基づく評価式の最小化によって、各エッジセグメントの動きを最適に記述する。さらに、シーン全体の動き記述長を最小化するための併合規則を用いて、エッジセグメント同士を併合させる。各運動部位をエッジセグメント集合として抽出すると同時にその動きパラメータの最尤推定値を求める。得られたシーン全体のフレーム間の最適運動記述は、結果として、各エッジセグメント集合をそれぞれ、単一剛体運動モデルとして記述している。本手法を人体時系列画像に適用し、実験でその有効性を示す。

和文キーワード 動画画像の運動記述, 分割, 最尤推定, 時空間勾配, MDL原理

The optimal partition of moving edge segments

Haisong Gu, Minoru Asada and Yoshiaki Shirai

Mech. Eng. of Computer-Controlled Machinery, Osaka University

Email:ko@ccm.osaka-u.ac.jp

Abstract In this paper, we describe a method to segment moving parts and estimate their motions by the minimization of motion description of the whole scene between two consecutive images. Unlike existing approaches which assume that a correct optical flow is given, the proposed approach relies on MDL (Minimum Description Length) principle and directly use the motion constraints of spatiotemporal gradients for segmentation. With a criterion consisting of the modeling precision, the model complexity, and the segmentation description, the motion on an edge segment is optimally modeled. On the basis of a combining rule for finding the minimum motion description for the whole scene, we divide all moving edge segments into several sets, each of which corresponds to one moving part, and simultaneously to obtain the maximum likelihood motion estimation for each part. We apply our method to both synthetic and real images, as a result, each set of segments corresponds to a rigid part of moving objects in a dynamic scene.

英文 key words Dynamic scene description, Segmentation, Maximum likelihood estimation, Spatiotemporal gradient, Minimum description length (MDL) principle

まえかき

分割はコンピュータビジョンにおける極めて重要な問題の一つである。初期処理としての分割はある画像特徴の類似性を持つ画素集合を抽出することと同等である。これに対して、明るさ、色、距離、動きなど様々な手がかりが利用可能である。中でも動き情報は、有力な手段であることが以前から指摘されている。オプティカルフローが得られたならば、移動ベクトルの類似性により対象の検出と対象の分割が可能である [1, 2, 3]。しかしながら、オプティカルフローを正確に抽出することが容易ではない。そこで、本研究では、フローを抽出せずに、エッジセグメント上の時空間勾配情報を直接に利用し、MDL 原理に基づいて、分割を行う一つのアプローチを提案する。

従来、モデルの推定問題に対して、MAP(Maximum-a-Posteriori) 原理はよく使われている。MAP では、モデルの構造的な知識を持つ上に、処理を行なう。しかし、使われているモデル自身の適切性が問題になっている。そこで、Rissanen ら [4, 5, 6, 7] は最適モデル化について、MDL 原理を提案した。彼らは情報と符号化の理論からモデルの複雑さとのモデルを使ってデータを表す正しさを評価する一つの記述長式を作った。この記述長を最小にすることによって、最適モデルが得られる。即ち、データに当てはめるべき記述は精度の良さを満たす最短の記述である。しかし、この統計学の理論を用いる時、応用面では、記述最小化をどのように行なうべきかを検討しなければならない。最近、コンピュータビジョンの分野では、Leclerc[8] がはじめに物体表面のモデル化問題に MDL 原理を応用した。その後、様々な問題に対して、それぞれ有効なアプローチが提案された [9, 10, 11, 12, 13]。本研究では、MDL 原理に基づいて、画像から最適な運動記述を得ることを試みた。まず、エッジセグメントのフレーム間の動きを記述するために、複数運動モデルの中から最適なモデルを選択する。次に、MDL 分割規範に基づく二進探索法により、最適な記述を得る。そして、同じ運動モデルパラメータを持つエッジセグメント同士を併合することによって、複数動物体を含むシーン全体の動き記述長を減少させる。

時空間勾配法の解法は、クラスタリング法と最尤推定法（或は最小二乗推定法）に分けられる [14]。クラスタリング法 [15, 16] は複数対象の扱いが容易である反面、高次元の運動を精密に推定する場合、ヒストグラム空間が巨大化する欠点がある。最尤推定法はノイズに強く、推定精度も良いが、推定対象が異なる動きを含む場合、誤った結果が得られる。そこで、本研究では、分割と最尤推定の繰り返しから、シーン全体の最適な記述を得る。最尤推定した残差は分割の正しさを反映し、最適な分割までの最尤推定値を運動の最適なモデルのパラメータとして求める。この繰り返し推定により、推定精度が良い、かつ複数対象の扱いが容易である。

以下、まず 2 章で、空間的に安定なエッジセグメントの役目と抽出法を述べる。3 章では、各エッジセグメントの最適な運動記述の決定法を述べる。4 章では、シーン全体の動き記述長を最小化することに基づいてエッジセグメントにおける物体の分割と動き推定を説明する。最後に、実験でこのアプローチの有効性を示す。

1 空間的に安定なエッジセグメント

1.1 空間的に安定なエッジセグメントの役目

複数個の動物体の分割を行なうにはオプティカルフローを利用する方法がある。しかし、我々が普段目している対象では、その表面には必ずしもオプティカルフローが得易い模様が記されているとは限らない。一方、エッジセグメントは画像の大まかな構造が表現できる。しかも、零交差点のような二階微分値が零となるエッジ点からなるエッジセグメント上の時空間勾配は物体のフレーム間の動きに関する情報を与える。しかし、複数運動部位を含むシーンの中で、異なる運動によって、エッジセグメントが互いに隠れたり、背景のものと融合したりすることが多い。この時、勾配情報に大きな誤りが生じる。そこで、このような問題が生じない空間的に安定なエッジセグメント (Spatially-Stable-Edge-Segments, or SSES) を尺度空間フィルタリングを用いて、取り出す。これらのエッジセグメントの近傍は、空間周波数が低く、勾配情報の信頼性が高い。以下、このようなエッジセグメントに基づいて、処理を行う。

1.2 空間的に安定なエッジセグメントの検出

波形を階層的に表現するために、波形をさまざまな尺度でとらえた波形集合に拡張するフィルタリングを尺度空間フィルタリング (Scale-space filtering) とよぶ [18]。

二次元変数 (x, y) 上の波形 $f(x, y)$ に対して

$$F(x, y, \sigma) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(\eta, \xi) \phi(x - \eta, y - \xi, \sigma) d\eta d\xi \quad (1)$$

を $f(x, y)$ の一般化波形とよぶ。ただし、関数 $\phi(x, y, \sigma)$ はパラメータ σ を持つガウス関数である。

尺度空間の零交差点 (z_c) は

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) F(x, y, \sigma) = 0 \quad (2)$$

この尺度空間における安定点と空間的に安定なエッジセグメントは次のように定義される。

定義：

(a) 安定点：尺度空間で方向が σ によって変化しないかつコントラストの強い零交差点である。

(b) 空間的に安定なエッジセグメント：連結している幾つかの安定点を含む安定点列である。

空間的に安定なエッジセグメント検出アルゴリズム：

・ $\sigma = 1, 2, 4$ に対して、それぞれ原画像を $\nabla^2 f$ でたたみ込む； ($\nabla^2 = (\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2})\phi$)

・ 得られた尺度空間から、 z_c 点を検出する；

・ 尺度空間では、 z_c が若干の位置変動を持つので、大きい尺度での z_c 点に基づいて、近傍により、安定点を抽出する；

これらの安定点の接続性を探しながら、空間的に安定なエッジセグメントを検出する。

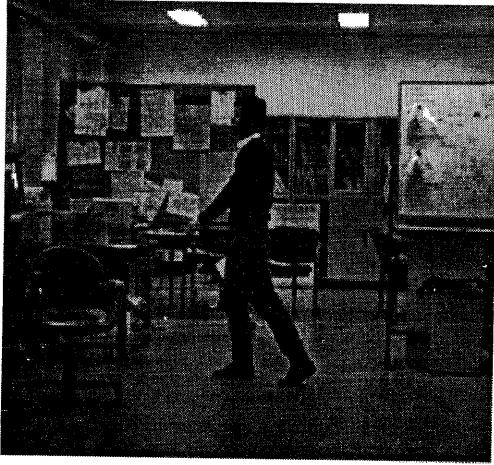


図 1: 入力動画像の例 (一枚目)

図 1 に示すのは入力動画像の例 (1 枚目) である。この画像に対して、検出された空間的に安定エッジセグメントを図 2 に表す。図中の数字はエッジセグメントの強さ (コントラストの総和と長さの積) 順の番号である。この図に示す様に、尺度空間フィルタリングを用いると、空間的に安定なエッジセグメントのみを対象としているので、オクルージョンの発生確率が低い。それゆえ、このようなエッジセグメントの近傍は勾配情報の信頼性の高いところと対応している。

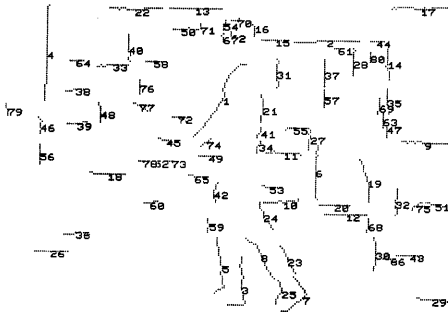


図 2: 検出した安定エッジセグメント

2 エッジセグメントの最適な動き記述

2.1 フレーム間の運動モデル

明るさの時間変化率 (以下、時間勾配と呼ぶ) は空間勾配と共に見かけの動きを束縛する。この拘束を利用して移動ベクトルを求めるのが勾配法である。フレーム間の運動を剛体運動と仮定すると、エッジセグメントのフレーム間の運動は回転と並進の二種類に分けられる。回転中心と回転角速度からなるベクトル $(X_0, Y_0, Z_0, \omega_x, \omega_y, \omega_z)$ で回転運動を表す。並進運動を並進速度ベクトル $T = (T_x, T_y, T_z)$ で表す。三次元空間に直交座標系 (X, Y, Z) を設定した時、対象物体の像は xy 平面 ($Z=1$) に中心投影されるものとする。物体上の点 $P = (X, Y, Z)$ の速度ベクトル P_t は、次式で表される。

- 回転運動で

$$P_t = \Omega \times (P - P_0) \quad (3)$$

ただし、 $\Omega = (\omega_x, \omega_y, \omega_z)$ は回転角速度ベクトルで、 $P_0 = (X_0, Y_0, Z_0)$ は回転中心ベクトルである。

- 並進運動で

$$P_t = T \quad (4)$$

一方、画面上の点 (x, y) での時刻 t における明るさを $E(x, y, t)$ 、空間勾配を (E_x, E_y) 、時間勾配を E_t とすると、画像上の移動ベクトル (u, v) が近似的に次の式に拘束されることが知られている。

$$E_x \cdot u + E_y \cdot v + E_t = 0 \quad (5)$$

中心投影モデルの下に、Horn ら [19] と同じ導き方で、回転運動と並進運動に対して、運動パラメータと時空間勾配の拘束をそれぞれ式 (6) と式 (7) のように表す。

$$\begin{aligned} & -\omega_x[E_y + y(xE_x + yE_y)] \\ & + \omega_y[E_x + x(xE_x + yE_y)] \\ & + \omega_z(xE_x - yE_y) + E_x \frac{(Y_0\omega_z - Z_0\omega_y)}{Z} \\ & + E_y \frac{(Z_0\omega_x - X_0\omega_z)}{Z} + \frac{(Y_0\omega_x - X_0\omega_y)}{Z} = -E_t \end{aligned} \quad (6)$$

$$\frac{E_x T_x}{Z} + \frac{E_y T_y}{Z} - \frac{(xE_x + yE_y)T_z}{Z} = -E_t \quad (7)$$

フレーム間の動きを推定する際に、 $\omega_x = \omega_y = T_z = 0$ 、かつ Z が定数であると仮定すると、上の式はそれぞれ次のように簡略化される。

$$\frac{E_x Y_0 \omega_z}{Z} - \frac{E_y X_0 \omega_z}{Z} + \omega_z(xE_y - yE_x) = -E_t \quad (8)$$

$$\frac{E_x T_x}{Z} + \frac{E_y T_y}{Z} = -E_t \quad (9)$$

もし、背景のモデル (10) を含めると、

$$E_t = 0 \quad (10)$$

全てのエッジセグメントの運動が複雑さの違うモデルを含む確率的なモデルクラス Υ で表現できる。モデルクラス Υ では、画像からの時間勾配をガウス確率密度関数 $p(E_t|E_x, E_y, x, y, \theta, k)$ で表す。このモデルクラス中で、 $k=0, 2, 3$ はそれぞれ、静止、並進および回転モデルと対応する。 θ は k に依存するパラメータベクトルである。

$$\Upsilon = \{p(E_t|E_x, E_y, x, y, \theta, k) : k \in [0, 2, 3], \theta \in \Theta(k)\} \quad (11)$$

ただし、 $\Theta(k)$ は固定した k の基に運動パラメータベクトル θ がとり得る範囲である。

2.2 MDL による動きの最適モデル化

モデルクラス Υ に基づいて、エッジセグメントの時空間勾配データを利用し、エッジセグメントの運動を最適

にモデル化できる。まず、エッジセグメント上の各点に関するデータ $(x_i, y_i, E_{xi}, E_{yi}, E_{ti})$ を、この三つのモデルにそれぞれ、代入する。統一的に次式のような線形回帰方程式が得られる。

$$A_{n \times k} U_{k \times 1} = B_{n \times 1} \quad (12)$$

ただし、

k	$A_{n \times k}$	$U_{k \times 1}$	$B_{n \times 1}$
3	$\begin{pmatrix} E_{x1} & E_{y1} & (x_1 E_{y1} - y_1 E_{x1}) \\ \vdots & \vdots & \vdots \\ E_{xn} & E_{yn} & (x_n E_{yn} - y_n E_{xn}) \end{pmatrix}$	$\begin{pmatrix} \frac{Y_0 \omega_x}{\omega_z} \\ -\frac{X_0 \omega_x}{\omega_z} \\ \frac{Y_0 \omega_y}{\omega_z} \\ -\frac{X_0 \omega_y}{\omega_z} \end{pmatrix}$	$\begin{pmatrix} E_{t1} \\ \vdots \\ E_{tn} \end{pmatrix}$
2	$\begin{pmatrix} E_{x1} & E_{y1} \\ \vdots & \vdots \\ E_{xn} & E_{yn} \end{pmatrix}$	$\begin{pmatrix} \frac{T_x}{T} \\ \frac{T_y}{T} \end{pmatrix}$	$\begin{pmatrix} E_{t1} \\ \vdots \\ E_{tn} \end{pmatrix}$
0	$\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$	$\begin{pmatrix} E_{t1} \\ \vdots \\ E_{tn} \end{pmatrix}$

式 (12) の最尤推定量は

$$U^* = [A^T A]^{-1} A^T B \quad (13)$$

この推定量による残差平方和は

$$\begin{aligned} S_k^* &= \|B_{n \times 1} - A_{n \times k} U^*\|_2^2 \\ &= B_{n \times 1}^T B_{n \times 1} - (A U^*)^T A U^* \end{aligned} \quad (14)$$

もし、この n 個の残差が独立かつ平均値が零の正規分布に従うと、それぞれのモデル M_k に基づいて、エッジセグメント B_t の動きを表す記述長が式 (15) のように二つ部分からなる。 $I(E_t|M_k)$ はモデルのパラメータと推定誤差を表す記述長（これを確率的な複雑度と呼ぶ）、 $L(M_k)$ はモデル同士を区別するための記述長である。

$$DL_{k,t} = I(E_t|M_k) + L(M_k) \quad (15)$$

ここで使われているモデルは全て線形回帰モデルなので、回帰変数の個数 k を記述すれば、モデル同士が区別

できる。自然数 k に対する記述長を語頭符号化に従って、式 (16) で計算する [7]。

$$\begin{aligned} L^*(k) &= \log_2 2.856 + \log_2(k) + \\ &\log_2 \log_2(k) + \log_2 \log_2 \log_2(k) + \dots \end{aligned} \quad (16)$$

ただし、 $\log_2 \log_2 \dots \log_2 k > 0$ の項まで、加計算する。

Rissanen[5] らによる確率的な複雑度の式を (15) に代入すると、動き記述長 $DL_{k,t}$ を式 (17) のように表す。

$$\begin{aligned} DL_{k,t} &= (n \log_2(2\pi S_k^*)) / 2 \\ &+ (\log_2 |A_{n \times k}^T A_{n \times k}|) / 2 + L^*(k) \end{aligned} \quad (17)$$

この式の第一項は推定されたデータと実際のデータのずれを表す記述長である。この部分でモデルの推定良さを表す。第二項は、モデル自身の運動パラメータを表す記述長である。MDL 原理に従って、 $MDL_t = \min_k \{DL_{k,t}\}$ により、運動のモデルが決定できる。

2.3 エッジセグメントの最適な記述

自然界の滑らかな境界を持つ物体のシーンにおける、一つのエッジセグメントの中には異なる物体のエッジを含む場合がある。この時、上の一つの剛体モデルに基づく最小化によっても、最適な運動記述が得られない。例えば、図 4(a) のように重なっている三つの同じ輝度の円盤の場合、尺度空間フィルタリングにより、得られたエッジセグメントには三つ物体の境界線を含んでいる（図 4c）。このエッジセグメントの運動を一つの剛体モデルとして記述すると、大きな誤りが生じる。分割を行わなければならない。問題となるのは、どのように分割するかと、どこまで分割すればよいかということである。誤差の観点から見れば、細かく分割するほど、誤差が小さくなるが、そうすると、直観的な三分割の結果が得られない。そこで、分割の複雑度を含む評価式 (18) を用いて、二進決定木探索による最適な分割方法を提案する。

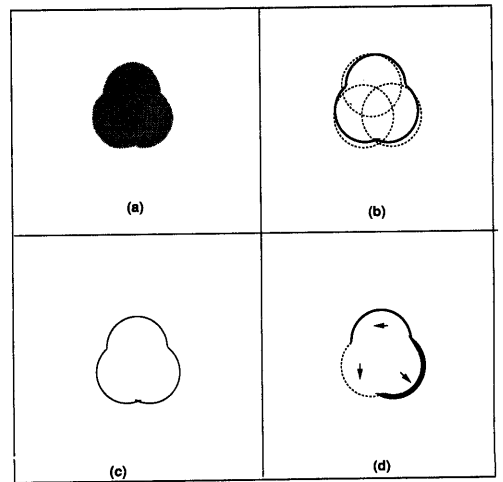


図 4: 動きの違いによる分割

$$\min_b \{MDL_t^{l+1} + SDL\} < MDL_t^l \quad (18)$$

$$MDL_t^{l+1} = \min_k [DL_{k,t1}^{l+1}(b)] + \min_k [DL_{k,t2}^{l+1}(n-b)]$$

$$SDL = L^*(\max(b, n-b)) + L^*(2)$$

MDL_t^{l+1} と MDL_t^l はそれぞれ、図 3 の長さ n のエッジセグメント B_i に対して、第 l レベルでの分割した後と分割する前の記述長である。 SDL は分割を行なった時、分割と分割の位置を表す記述長である。

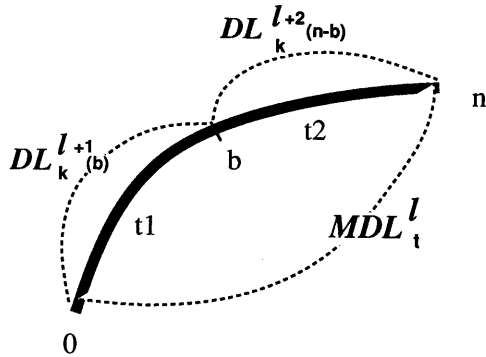


図 3: 評価式の説明図

もし、この評価式を満たせば、二分割を行なう；満たさなければ、エッジセグメント B_i を一つの部分として記述する。

図 4(b) ではフレーム間の運動を表す。

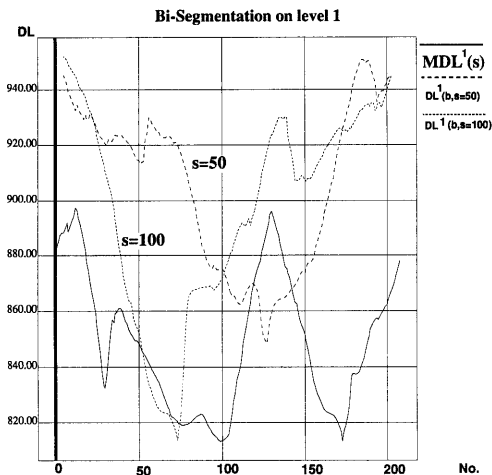


図 5: 記述長の分割点による変化

図 5では図 4のエッジセグメントに対して、レベル 1 で分割を行なった際に、分割点による運動記述長の変化

を示す、最初の分割（レベル 1）では閉じた輪郭線なので、二つの分割点を発見しなければならない。それぞれ始点（ s ）と分割点（ b ）とする。ある固定された s 点に対して b 点による分割の記述長 $MDL^1(b)$ を点線で示す（二本の点線はそれぞれ $s = 50, 150$ の時の b 点の位置変化（横軸）による記述長の変化に対応する）。最適な b 点に対する最小記述長を s の関数として表したのが実線である。ここで、二つの谷 $s = 99, 173$ では、 s 点と b 点を入れ換えた場合を示し、それぞれ同じ最小記述長に対応する。

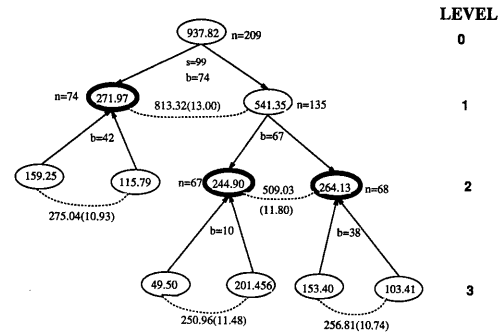


図 6: 三分割の二進決定木

図 6では、式 (18) の評価式を用いて、二進決定木探索により、エッジセグメントの分割を行なったことを示す。図の中の矢印は分割を進める方向を表すもので、括

弧の中の数字は分割自身を表すための記述長（ SDL ）である。太い線の節点の所で、求められた最適な記述を示す。この探索図に基づいて、図 4のような三つの異なる動きの物体を、レベル 1 でのパート 1 ($length = 74$) とパート 2 に分割して、レベル 2 ではパート 1 を分割せず、パート 2 をパート 21 ($length = 67$) とパート 22 ($length = 68$) に分けた。それぞれのパートに対応するエッジセグメントと運動を図 4(d) に表す。

3 物体の分割と運動推定

二進決定木探索により、各エッジセグメントの最適な動き記述が求められる。これらのエッジセグメントを動きの類似性と距離により、物体ごとに併合し、シーン全体の最適な記述を求める。

3.1 シーン全体の動き記述評価

シーン全体に対するフレーム間の最適な動き記述は、各運動部位を正確に分割した上で、各部位の最適な運動モデルに基づいて、運動パラメータを得ることである。フレーム間の最適な動き記述を求める際に、分割と併合の繰り返す処理をさせながら、以下の評価関数を最小化させる。

$$\Gamma = \sum_{B_i \in B^{(n)}} [\min_{k \in \{3,2,0\}} (DL_{k,B_i})] \quad (19)$$

ただし、 B_i は処理段階 n でのエッジセグメント集合で、 $B^{(n)}$ はシーンにおける全体 B_i の集合である。

3.2 物体分割と運動推定アルゴリズム

エッジセグメントを、エッジセグメント集合に集める際に、以下の併合規則と評価式を用いる。

併合規則：もし、二つの同種の動くエッジセグメントの併合により、新しい記述長がそれぞれ独立の記述長の和より短ければ（即ち、評価式 (20) を満たせば）、二つのエッジセグメントが同じ運動を行うと判断する。しかも、互いに近ければ、同じの運動部位のエッジセグメントとして併合する。

評価式：

$$MDL_i < MDL_{i1} + MDL_{i2} \\ -L^*(\max(n1, n2)) - L^*(2) \quad (20)$$

アルゴリズム：

step1: 画像の中での全ての空間的に安定なエッジセグメントをモデルの違いにより回転、並進および静止の三つのグループに大別する。

step2: グループ毎にエッジセグメントを大きさの順に並べる、これをエッジセグメント列とよぶ。

step3: エッジセグメントをエッジセグメント列の上から取り出す。それを B_i とおく。

step4: B_i との最小距離が、ある一定の距離 D 以下なるエッジセグメントをエッジセグメント列の中から順に一つ取り出す。それを b_{ij} とおく。

step5: B_i と b_{ij} が併合規則を満たせば、 $B_i \leftarrow B_i \cup b_{ij}$ 、満たさなければ、 b_{ij} をエッジセグメント列に戻す。

step6: 併合規則を満たすエッジセグメントがエッジセグメント列の中に見つからなくなるまで、step4-5 を繰り返す。

step7: 残ったエッジセグメント列に対し、Step3-6 をエッジセグメント列が空になるまで、繰り返す。それによって、いくつかのエッジセグメント集合ができる。

step8: 併合された各々のエッジセグメント集合における、二進決定木探索により、最適な分割を求める。

step9: 得られたエッジセグメント集合の全体に対して、シーン全体の動き記述長 (19) が最小になるまで、step1-step8 を繰り返す。

こうして各物体に属するエッジセグメントの集まりが得られる。しかも、各物体のフレーム間の動きの最尤推定も同時に求められる。

4 実験結果

4.1 歩行者の部位分割と動き推定

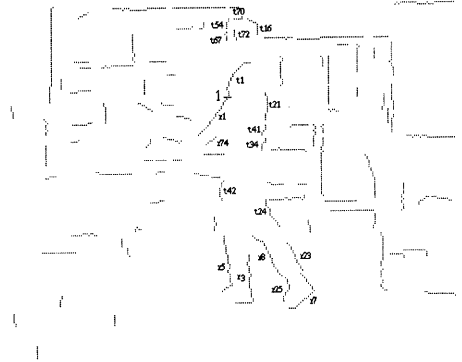


図 7: 各エッジセグメントの最適な動き記述

図 1 のような歩行者の動画像に対して、空間的に安定エッジセグメントの最適な動き記述を図 7 と表 1 に表す。図中の番号の前の t と r はそれぞれ、並進運動と回転運動を示す。このシーンで、番号 1 のエッジセグメントには、胴体と右手の境界が含まれているので、二進決定木探索により、正確に分離されたことがわかった。

表 1: エッジセグメントの動き記述長

番号	運動パラメータ	記述長 (bit)
1'	(-1.68,0.11)	148.23 (263.20)
1''	(116.08,112.77,2.83°)	100.95
3	(124.71,193.20,1.29°)	180.73
5	(115.38,188.32,2.71°)	153.58
7	(146.15,204.70,4.33°)	142.13
8	(136.69,229.31,2.02°)	135.37
16	(-2.20,-0.05)	87.06
21	(-2.36,0.03)	104.80
23	(155.89,221.37,2.01°)	91.44
24	(-2.46,-0.27)	77.75
25	(147.08,200.94,5.32°)	78.26
34	(-2.20,0.34)	61.43
41	(-1.44,-1.03)	49.05
42	(-2.07,0.88)	84.92
54	(-1.28,0.63)	44.48
62	(-1.85,0.82)	39.86
67	(-1.44,0.24)	27.42
70	(-0.50,-0.33)	45.72
74	(111.21,100.96,2.87°)	44.25
シーン全体の動き記述長 (bit) =		1697.43

表 1 では、エッジセグメントに基づく最適な動き記述長と最尤推定された運動パラメータを示す。括弧の中の数字は分離される前の記述長である。

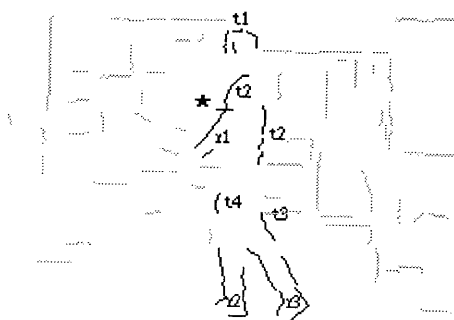


図 8: シーン全体の最適な動き記述

表 2: シーン全体に対する最適な動き記述

番号	運動パラメータ	記述長(分離時)
r1	(114.17,109.81,3.20°)	144.20(145.20)
r2	(122.64,192.24,1.76°)	328.05(334.31)
r3	(143.73,212.73,2.65°)	416.04(447.20)
t1	(-1.80,0.39)	230.78(244.54)
t2	(-2.09,0.31)	336.18(363.51)
t3	(-2.46,-0.27)	77.75(77.75)
t4	(-2.07,-0.88)	84.92(84.92)
シーン全体の動き記述長(bit)= 1617.92		

併合処理による、得られたシーン全体の最適な記述を図 8 (* は分割を行なった所を表す) と表 2 に表す。これらの結果により、シーン全体の動き記述長を減少させながら、同じ部位に属するエッジセグメントが併合できたことが分る。

同じ時系列画像の中で、第四枚目の画像(図 9)に対して、本手法を用いて、得られた最適な分割は図 10 に表される。

人体運動による、人体部位を隠れたりすることが多いです。見えるところに対して、本手法により、正しく分割できる。本手法を人体運動解釈の初期処理として、この有効性を示した。

5 まとめ

本研究では、シーン全体に対するフレーム間の最適動き記述を求めると共に、複数動物体の分割とフレーム間

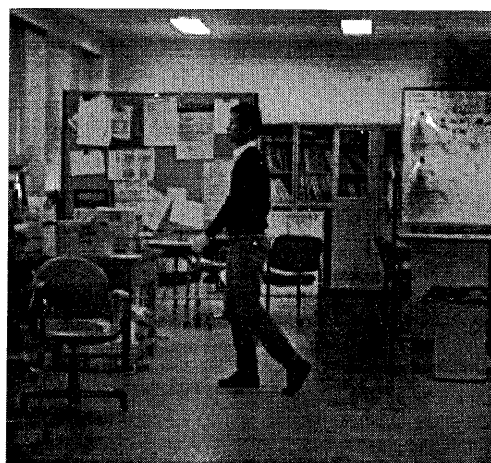


図 9: 入力動画画像の例(第四枚目)

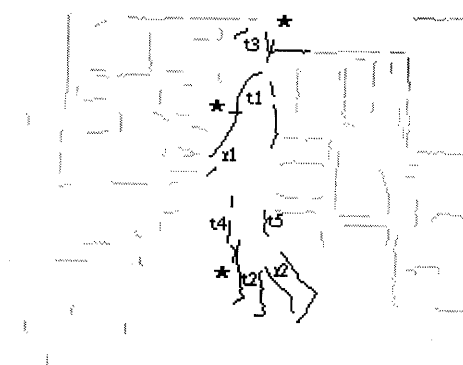


図 10: 第四枚目の最適な動き記述

の運動推定を行なう手法を提案し、実験で、その有効性を示した。すべての評価は画像から有効な時空間勾配情報により運動モデル化の適切性という観点から行われ、閾値の問題が避けられた。併合規則により、最尤推定と最適な分割を同時に求めることができた。

異なる対象でも、動きに差がない場合は分割が難しい。その際、長い時系列追跡情報の利用すれば、さらに精密の分割が可能であろう。

本手法に基づいて、人体のような複数動部位を持ち物体の追跡と運動解釈への発展が期待される。

参考文献

- [1] E.C.Hildreth, "The Measurement of Visual Motion" The MIT press,(1984)
- [2] W.B.Thompson,M.Mutch and V.A.Berzins, "Dynamic Occlusion Analysis in Optical Flow Fields" IEEE,PAMI-7,4,pp.374-383 (1985)
- [3] S.Tsuji,M.Osada and M.Yachida, "Tracking and Segmentation of Moving Objects in Dynamic Line Images" IEEE, PAMI-2,6,pp.516-549(1980)
- [4] J.Rissanen, "Modeling by shortest data description" Automatica,Vol.14,pp.465-471(1978).
- [5] J.Rissanen, "Stochastic Complexity in Statistical Inquiry" Vol.15 World Scientific Series in Computer Science, (1989)
- [6] J.Rissanen, "Stochastic complexity" Journal of the Royal Statist. Society B, Vol.49, no. 3, pp.223-239 and 252-265(1987).
- [7] J.Rissanen, "A universal prior for intergers and estimation by minmum description length" Annals. of Statistics, Vol.11, pp.416-431(1983)
- [8] Y.G.Leclerc, "Constructing Simple Stable Descriptions for Image Partitioning" Intl. J.Computer Vision, Vol. 3, pp.73-102(1989)
- [9] T.Darrell, S.Sclaroff, and A.Pentland, "Segmentation by minimal description" Proc. ICCV'90,Osaka, Japan, pp. 112-116,(1990)
- [10] J.Dengler, "Estimation of Discontinuous Displacement Vector Fields with the Minimum Description Length Criterion" Proc. CVPR'91,Lahaina, Maui,Hawaii, pp.276-282(1991)
- [11] Y.Sato, J.Ohya and K.Ishii, "Recovery of Hierarchical Part Structure of 3D Shape from Range Image" Proc. ICCV'92,pp.699-702,(1992)
- [12] K.Keeler, "Map Representations and Coding-Based Priors for Segmentation" Proc. CVPR'91,Lahaina, Maui,Hawaii, pp.420-425(1991)
- [13] J.Sheinvald, B.Dom, W.Niblack and D.Steel, "Unsupervised Image Segmentation using the Minmum Description length Principle" Proc. IAPR'92, pp.709-712
- [14] W.B.Thompson and S.T.Barnard, "Lower-level Estimation and Interpretation of Visual Motion" IEEE, Computer, 14, 8, pp. 20-28(1981)
- [15] C.Cafforio and F.Rocca, "Methods for Measuring Small Displacements of Television Images" IEEE, IT-22,5,pp.573-579(1976)
- [16] C.L.Fennema and W.B.Thompson, "Velocity Determination in Scenes Containing Several Moving Objects" C.G.I.P.,9,pp.301-315(1979)
- [17] D.Marr, "Vision" W.H.Freeman, San Fransisco,(1982).
- [18] A.P.Witkin, "Scale-Space Filtering" Proc. of 8th International Joint Conference on Artifical Intelligence, pp. 1019-1022, Karlsruhe,(1983).
- [19] B.K.P.Horn and E.J.Weldon JR., "Direct Methods for Recovering Motion" International Journal of Computer Vision,2, pp. 51-76(1988).