

## 勾配法による動画像と距離情報からの複数動物体の 3次元運動解析

正 員 山本 正信<sup>†</sup>

Three-Dimensional Motion Analysis of Scene Containing  
Multiple Moving Objects from Image Sequence and Depth

Masanobu YAMAMOTO<sup>†</sup>, Member

あらまし 動きは画像から対象を抽出する有力な手がかりである。オプティカルフローに基づく解析法は、対象の抽出ばかりでなくその3次元構造も復元可能である。しかしながら、実際の世界にはオプティカルフローが得易い対象ばかりではなく、動きが一意に解釈できない対象も存在する。本論文は勾配法に基づき、動きが一意に解釈できない対象をも含む複数動物体の3次元運動解析法を提案する。勾配法によれば、対象の3次元構造が既知ならば、その3次元運動パラメータは線形連立方程式の最小2乗解としてオプティカルフローを介さずに推定することができる。推定時の残差平方和を基準に、画面上の領域に異なる動きが含まれているならば分割し、同じ動きと認識できれば領域を融合する。画面の分割と融合を繰返しながら対象を抽出する。動きの一意性は、連立方程式の係数行列の条件数により評価される。さらに、抽出された対象の動きが一意に解釈可能ならば、対象の動きを繰り返し推定することにより、移動ベクトルを精密に求める。

### 1. ま え が き

画像から対象を抽出するには、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。なかでも動きは、有力な手段であることが古くから指摘されている。

オプティカルフロー（画面上の移動ベクトル分布）が得られたならば、移動ベクトルの空間的变化から対象の境界が検出される<sup>(1),(2)</sup>。また、移動ベクトルの類似性から対象の抽出が可能である<sup>(3)~(6)</sup>。

さらに精密なセグメンテーションには、基準となる動きのモデルが必要となる。Adiv<sup>(7)</sup>は剛体の3次元運動モデルを使ってオプティカルフローから運動パラメータを推定した。得られたパラメータから計算されるオプティカルフローの推定誤差を評価基準に、複数個の動く剛体の抽出を行った。オプティカルフローさえ得ることが出来れば、対象の抽出ばかりでなく、その3次元構造も復元可能である。しかし、我々が普段目にしていない対象は人工物が多く、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。そればかりか、動きが一意に解釈できない対象も現実中存在する。

一方、明るさの時間変化率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動きを知るのが勾配法である。勾配法を使って単一剛体の3次元運動を解析するには対象の3次元構造についての何らかの知識が必要である<sup>(8)</sup>。対象を平面と仮定すれば、その面の傾きと運動パラメータをオプティカルフローを介することなく得ることができるが<sup>(9),(10)</sup>、その手続には6次方程式の解法が含まれている。対象の3次元形状が既知ならば、運動パラメータは線形連立方程式から直接推定できる<sup>(11)</sup>。また、動きの解釈の一意性はその方程式の独立性を調べることにより知ることが可能である。

本論文では、対象の3次元情報を既知とし、勾配法を使って、動きが一意に解釈できないような対象をも含む複数動物体のシーンの3次元運動を解析する。

勾配法の解法は、クラスタリング法と最小二乗推定法に分けられる<sup>(12)</sup>。クラスタリング法<sup>(13),(14)</sup>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。従って、3次元運動の解析には最小二乗推定法が有利である。最小二乗推定法はノイズに強く、推定精度も

<sup>†</sup>電子技術総合研究所制御部, 茨城県  
Automatic Control Division, Electrotechnical Laboratory,  
Ibaraki-ken, 305 Japan

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。これに対して、Schalkoff・McVey<sup>19</sup>は疑似3次元運動の場合に、画像をあらかじめ格子状に分割し、それぞれの区画上で得られた運動パラメータを比較することにより対象領域の抽出を行った。しかし、区画を小さくしてもやはり異なる動きを含む可能性は残る。また、動きが一意に解釈できない区画がある場合は、パラメータ自身の比較は無意味である。

本論文では、最小二乗推定時の推定残差を評価基準に、異なる動きを含む領域の分割と等質な動きをする領域の融合を繰り返すことにより、対象領域の抽出を行う。同時に、動きが不確定な対象の識別方法も示す。さらに、対象の動きが一意に解釈可能ならば、対象領域上で勾配法を反復適用することにより移動ベクトルの高精度推定を行う。

## 2. 3次元運動パラメータの推定

3次元運動パラメータの直接的推定法を簡単にまとめておこう<sup>11</sup>。

3次元空間中に直交座標系 $(x, y, z)$ を設定したとき、対象物体の像は $xy$ 平面に正射影されるものとする。対象が剛体運動しているならば、剛体上の点 $r = (x, y, z)$ の微小移動ベクトル $V_p = (u, v, w)$ は並進ベクトル $V_0 = (u_0, v_0, w_0)$ と角速度ベクトル $\Omega = (\omega_x, \omega_y, \omega_z)$ を運動パラメータとして、次式で表わされる。

$$V_p = \Omega \times r + V_0 \quad (1)$$

正射影を仮定しているので $(u, v)$ の分布がオペイカルフローである。以下、 $(u, v)$ を単に移動ベクトルとよぶ。

一方、画面上の点 $(x, y)$ の時刻 $t$ における明るさを $E(x, y, t)$ 、空間勾配を $(E_x, E_y)$ 、時間変化率を $E_t$ 、移動ベクトルを $(u, v)$ とする。そのとき移動ベクトルは近似的に次の式に拘束されることが知られている<sup>13</sup>。

$$E_x \cdot u + E_y \cdot v + E_t = 0 \quad (2)$$

式(2)が成立するためには、(a)  $(x, y)$ の近傍で明るさ分布が $(E_x, E_y)$ を勾配に持つ平面で近似できること、すなわち、明るさの2次以上の空間微分が0であること、(b)明るさの分布が移動先でも変化しないこと、が必要である。

さて、式(1)を式(2)に代入すれば次式を得る。

$$E_x \cdot u_0 + E_y \cdot v_0 - z \cdot E_y \cdot \omega_x + z \cdot E_x \cdot \omega_y + (x \cdot E_y - y \cdot E_x) \omega_z + E_t = 0 \quad (3)$$

$E_x, E_y, E_t$ は動画像から差分演算などで直接算出される。例えば、サイズが5の3次元エッジ検出オペ

レーターを5枚の連続画像に適用すれば信頼度の高い勾配値が得られる。本論文ではシーンの3次元情報は与えられているものとする。従って、対象までの距離、すなわち深さ量 $z$ が既知であるので、運動パラメータは、同じ剛体上の点から導かれる方程式(3)を連立させることにより、その最小二乗解として推定される。すなわち、連立方程式

$$Ax = b \quad (4)$$

の最小二乗解は

$$x_0 = (A^T A)^{-1} A^T b \quad (5)$$

となる。ただし、添字の $T$ は行列の転置を意味する。また、 $A$ と $b$ の行ベクトルは、各々 $(E_x, E_y, -z \cdot E_y, z \cdot E_x, x \cdot E_y - y \cdot E_x), (-E_t)$ であり、未知数は $x = (u_0, v_0, \omega_x, \omega_y, \omega_z)^T$ である。

動きを平面運動に限定すれば、 $\omega_x = \omega_y = 0$ であるから、式(3)は次式に簡略化される。

$$E_x \cdot u_0 + E_y \cdot v_0 + (x \cdot E_y - y \cdot E_x) \omega_z + E_t = 0 \quad (3)'$$

また、並進運動に限れば、さらに $\omega_z = 0$ となり、

$$E_x \cdot u_0 + E_y \cdot v_0 + E_t = 0 \quad (3)''$$

となる。式(3), (3)', (3)''でモデル化される動きを3D, 2D, 1D運動とよぶ。2D, 1D運動のパラメータも3D運動の場合と同様に得られる。

## 3. 動画像の入力および前処理

### 3.1 動画像の入力

CCDカメラ(XC-37, ソニー)で撮影した動画像を一旦ビデオディスク(VM-1000M, ビクター)に蓄える。ついで、TOSPIX-2を使って512×512, 8bit階調で量子化する。動画像はフィールド単位で入力できるが、さらにスムージングして128×128の大きさに縮小するためインターレースの影響は殆どない。

### 3.2 前処理

推定式(4)を使って動きを推定する場合、領域が単一の剛体に対応していなければならない。また精度の良い解を得るためには、式(2)が良く近似されている所を選ぶ。さらに、明るさが一様な所は式(2)が無意味なので省くことにする。

ゼロクロッシング上<sup>11)</sup>では、明るさの2次微分値が0となるため、式(2)は良く近似されている。厳密には明るさの3次以上の微分値も0となる必要があるが、実用上大差ない。ただし、ゼロクロッシングは異なる動きをする対象にまたがる可能性がある。5.で、異なる動きを含む対象の分割法を示すが、前処理の段階で、

出来るだけ単一剛体に対応するように分割しておくほうが、演算の効率上望ましい。ここでは、次のアルゴリズムを使用する。

まず、各画素点の勾配を求める。勾配の大きさがある閾値以上の点を残す。勾配の大きさが $x$ 方向あるいは $y$ 方向で極大となる点、すなわち明るさの2次微分値が0となる点を残す。8-隣接している残った画素点を互いに連結すれば、いくつかの線の集まりからなる線画が得られる。ついで、ひげ、4こ以下の画素点からなる線、3本以上の線が交差する点を除去する。さらに、深さが不連続な所で線を分割する。

得られた線画中の各々の線をセグメントとよび $S_i$ で表わす。 $S_i$ は画素点の集合である。また、いくつかのセグメント $S_i$ の和集合もセグメントとよび $U S_i$ と表わす。このアルゴリズムを図1に適用した結果を図2に示す。数字はセグメントの番号である。以下、動きの推定はセグメント上で行うことにする。

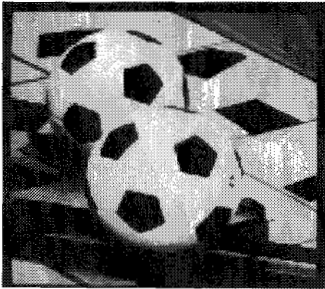


図1 二つの転がるボール  
Fig.1 Independently rolling balls scene.

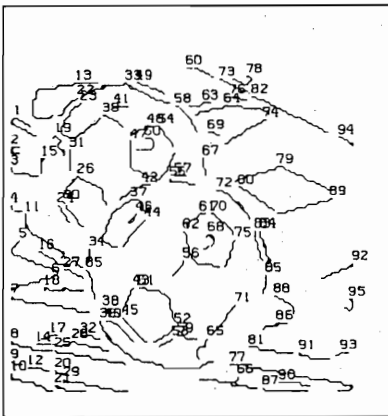


図2 図1のセグメント図  
Fig.2 Initial segments map of Fig.1.

#### 4. 動きの不確定な対象の識別

対象によっては、例えば、理髪屋の看板のように、その動きが一意に定まらないものもある。動きが一意に決定可能であるための必要十分条件は、理論的には与えられている<sup>10)</sup>。ここでは、実際に動きの決定可能性を判定する計算手続を示す。

推定式は線型連立方程式であるので、解の一意性すなわち動きの一意性はその係数行列 $A$ の独立性を調べることにより判定される。 $A$ の独立性はその条件数によって測ることができる<sup>20)</sup>。 $A$ の特異値分解を

$$A = U \Sigma V^T = U \begin{bmatrix} \lambda_1 & & & 0 \\ & \ddots & & \\ & & \lambda_n & \\ 0 & & & 0 \end{bmatrix} V^T \quad (6)$$

ただし、 $U, V$ は適当な直交行列、 $\lambda_i$ は $A$ の特異値で、 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$ 、 $n$ は運動パラメータの数としたとき、条件数 $\text{cond}(A)$ は $A$ の最大特異値と最小特異値の比

$$\text{cond}(A) = \lambda_1 / \lambda_n \geq 1 \quad (7)$$

で与えられる。 $A$ が1次従属の場合 $\lambda_n = 0$ となるので、 $\text{cond}(A) = \infty$ となる。従って、条件数が1に近いほど動きが一意に確定であり、大きくなるほど不確定になる。

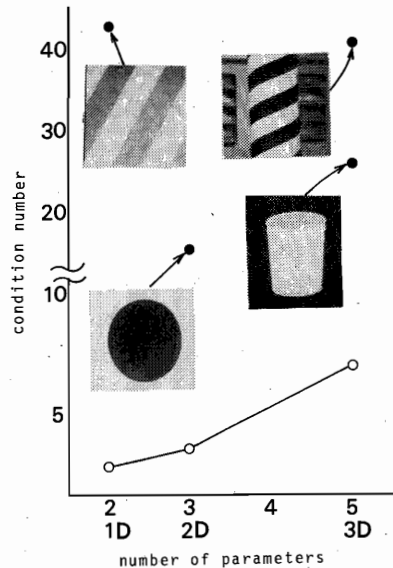


図3 運動モデルと対象の条件数  
Fig.3 Condition number versus the number of motion parameters.

各対象について条件数を比較するためには、 $A$ の行および列のノルムの大きさを、ほぼ同じオーダーにしておくことが望ましい。なぜならば、3D運動や2D運動の場合、係数行列 $A$ には位置情報 $(x, y)$ および深さ量 $z$ が含まれている。そのため相似な対象について、 $A$ の条件数が異なってくるからである。

それを避けるために、対象が単位球内に内接するように $x, y, z$ を相似変換するのである。こうすれば、 $|x|, |y|, |z| \leq 1$ であるので、 $A$ の各要素は $E_x, E_y$ のオーダーになり、行および列のノルムは均衡化され、相似な対象の条件数はほぼ等しくなる。以下、この条件数を対象の条件数とよぶことにする。

それでは、動きの決定可能性を判別する条件数の閾値はどのような値に定めればよいのであろうか。1D, 2D, 3D運動について、動きが一意に解釈可能な対象の条件数をそれぞれ20例ばかり測定した。その平均値を図3に○印で示す。また動きが一意に解釈不可能な対象についても、3D運動では理髪屋の看板とバケツ、2D運動では円模様、1D運動ではストライプ模様について測定した条件数を図3に●印で示す。この測定結果より判別の閾値を1D, 2D, 3D運動で、各々4, 8, 20とした。

5. 対象の抽出

動きに基づきセグメントを融合し対象を抽出する。

5.1 異なる動きと同じ動きの識別

セグメントの集合 $S = \cup S_i$ において、運動パラメータの最小二乗推定を行ったとき、残差平方和 $e^2$ は

$$e^2 = \|Ax_0 - b\|^2 = \sum_5 (E_x \cdot u + E_y \cdot v + E_z)^2 \quad (8)$$

となる。ノルムはユークリッドノルムとした。理想的には $S$ が単一の剛体に対応しているならば、 $e^2 = 0$ 、動きの異なる複数剛体に対応しているならば、 $e^2 \neq 0$ となる。しかし、実際の画像では、単一の剛体に対応している場合でも、 $e^2 = 0$ となることは、まずない。従って、残差平方和を用いて、異なる動きと同じ動きの識別を行うには、適当な閾値を定めねばならない。

閾値はどのように定めたら良いのであろうか。式(8)をながめると、各点の移動量が倍になれば、各点の明るさの時間差 $E_i$ も倍になる。従って、 $e$ は倍になる。その際、明るさの勾配に垂直な移動ベクトル成分はパターンの変化を起さないで、 $e$ に影響しない。 $e$ に影響するのは、勾配方向の移動成分だけである。また、各点の明るさの勾配の大きさが倍になれば、各点の明

さの時間差 $E_i$ も倍になり、従って、 $e$ も倍になる。すなわち、 $e^2$ は

$$\begin{aligned} & \sum_5 (E_x^2 + E_y^2)(u^2 + v^2) \cdot \cos^2 \theta \\ & = \sum_5 (E_x \cdot u + E_y \cdot v)^2 = \|Ax_0\|^2 \end{aligned}$$

(ここで、 $\theta$ は移動ベクトルと明るさの勾配のなす角度。)に比例する。このことを、実際に確かめてみよう。

静止背景中に単一剛体(ボール)が転がっている場合に、同じ剛体運動をするセグメントの集まりについて○印で、異なる動きを含むセグメントの集まりについて×印で、残差平方和と $\|Ax_0\|^2$ の関係を図4に示す。この図の第1グループが動いているボールに相当する。そこでは、残差平方和は、ほぼ $\|Ax_0\|^2$ に比例している。第2, 第3グループについては後述する。従って、 $S$ を単一の剛体運動とみなすには、 $k_1^2$ を比例定数として、

条件1

$$\|Ax_0 - b\|^2 \leq k_1^2 \|Ax_0\|^2$$

を満たせばよい。

静止あるいは静止に近い動きでは、上式の右辺は0あるいは0に近い値となる。ところが静止背景の残差平方和は図4の第2グループに対応している。これは、静止に近い状態では $E_x, E_y, E_z$ がラプラス分布に従う影響が強く出るためである。そこで、最大特異値の平方 $\|A\|^2$ と残差平方和の関係について調べてみると、

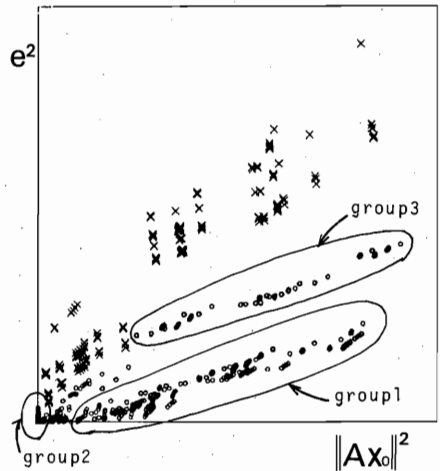


図4  $\|Ax_0\|^2$ と動きの最小二乗推定残差平方和  
Fig.4 The sum of squared error distribution for  $\|Ax_0\|^2$ .

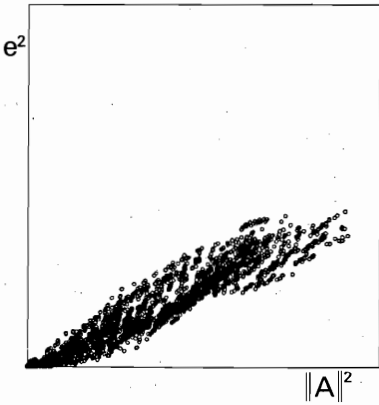


図5  $\|Ax\|^2$  と動きの最小二乗推定残差平方和  
Fig.5 The sum of squared error distribution for  $\|A\|^2$ .

図5に示すとおり  $e^2$  は  $\|A\|^2$  に比例している。従って、静止に近い対象とみなすためには、 $k_2^2$  を比例定数として

条件2

$$\|Ax_0 - b\|^2 \leq k_2^2 \|A\|^2$$

を満たせばよい。

さて図4の第3グループは第1グループを上方に平行移動しているように見える。実は、このグループはボールの輪郭部に対応する部分を含んだボールのセグメント集合に相当する。輪郭部では、異なる動きやオクルージョンの影響が残差平方和の増加となって表われる。少々異なる動きを含んでいる場合でも同じ剛体運動とみなすためには、 $k_1^2$  を大きくすることが考えられるが、 $k_1^2$  の増大は異なる動きの識別感度を鈍くするので注意しなければならない。ここでは、条件1にげたをはかせて、

条件1'

$$\|Ax_0 - b\|^2 \leq k_1^2 \|Ax_0\|^2 + k_3^2$$

を満たすものを単一の剛体運動とみなす。

条件1は条件1'に含まれるので、条件1'あるいは条件2が満たされるとき、セグメント集合  $S$  は1つの剛体に対応しているとみなす。これを、融合規則とよぶ。

$k_1^2, k_2^2$  は動画像の入力系に依存する。我々の入力系では、 $k_1^2$  は0.03~0.04であった。また、 $k_2^2$  は1フィールド間隔で約0.0001であり、画像の時間間隔が広がると大きくなる。 $k_3^2$  は異なる動きの大きさと広がりに関係しているのでその値を定めることはできないが、本論文では経験上~40程度の値とした。

以上は3D運動の場合であるが、2D, 1D運動の場合でも、 $k_1^2, k_2^2, k_3^2$  は3D運動の場合とほぼ同じであった。

### 5.2 動きの識別感度

融合規則はどの程度の異なる動きを許容しているのであろうか。また、条件1'と2の関係についても考察する。

$Ax_0 = b$  としたとき、異なる動きの最大許容限界が与えられる。条件1'の場合は、

$$\|Ax - Ax_0\|^2 \leq k_1^2 \|Ax_0\|^2 + k_3^2 \quad (9)$$

で  $x$  の限界が与えられる。式(6)より、

$$\left. \begin{aligned} z &= V^T x \\ d &= V^T x_0 \end{aligned} \right\} \quad (10)$$

とおけば、式(9)は、

$$\begin{aligned} &\lambda_1^2 (z_1 - d_1)^2 + \lambda_2^2 (z_2 - d_2)^2 + \dots + \lambda_n^2 (z_n - d_n)^2 \\ &\leq k_1^2 (\lambda_1^2 d_1^2 + \lambda_2^2 d_2^2 + \dots + \lambda_n^2 d_n^2) + k_3^2 \end{aligned} \quad (11)$$

となり、

$$r^2 = d_1^2 + d_2^2 / \left(\frac{\lambda_1}{\lambda_2}\right)^2 + \dots + d_n^2 / \left(\frac{\lambda_1}{\lambda_n}\right)^2 \quad (12)$$

とおけば、

$$\begin{aligned} &\frac{(z_1 - d_1)^2}{k_1^2 r^2 + \left(\frac{k_3}{\lambda_1}\right)^2} + \frac{(z_2 - d_2)^2}{k_1^2 r^2 \left(\frac{\lambda_1}{\lambda_2}\right)^2 + \left(\frac{k_3}{\lambda_2}\right)^2} + \dots \\ &+ \frac{(z_n - d_n)^2}{k_1^2 r^2 \left(\frac{\lambda_1}{\lambda_n}\right)^2 + \left(\frac{k_3}{\lambda_n}\right)^2} \leq 1 \end{aligned} \quad (13)$$

となる。同様に、条件2の許容限界は、

$$\|Ax - Ax_0\|^2 \leq k_2^2 \|A\|^2 \quad (14)$$

で表わされる。 $\|A\| = \lambda_1$  であることに注意すれば、

$$\begin{aligned} &\lambda_1^2 (z_1 - d_1)^2 + \lambda_2^2 (z_2 - d_2)^2 + \dots + \lambda_n^2 (z_n - d_n)^2 \\ &\leq k_2^2 \lambda_1^2 \end{aligned} \quad (15)$$

さらに、

$$\frac{(z_1 - d_1)^2}{k_2^2} + \frac{(z_2 - d_2)^2}{k_2^2 \left(\frac{\lambda_1}{\lambda_2}\right)^2} + \dots + \frac{(z_n - d_n)^2}{k_2^2 \left(\frac{\lambda_1}{\lambda_n}\right)^2} \leq 1 \quad (16)$$

となる。 $V$  が直交行列であることから、式(10)は座標軸の回転に過ぎず、式(13), (16)は、 $x$  の許容範囲が  $x_0$  を中心とした超楕円体の内部であることを示している。この2つの超楕円体は、その径に着目すれば、

$$k_1^2 r^2 + k_3^2 / \lambda_1^2 \leq k_2^2$$

のとき、すなわち、

$$\frac{d_1^2}{(k_2 \lambda_1)^2 - k_3^2} + \frac{d_2^2}{(k_2 \lambda_1)^2 - k_3^2} + \dots$$

$$\frac{d_1^2}{(k_1 \lambda_1)^2} + \frac{d_2^2}{(k_1 \lambda_2)^2} + \dots$$

$$+ \frac{d_n^2}{(k_2 \lambda_1)^2 - k_3^2} \leq 1 \quad (17)$$

のとき、条件1' は条件2に含まれる。式(17)も超楕円体であることから、運動パラメータが原点を中心とする超楕円体(17)の内側に存在する様な小さな動きのときは条件2に束縛され、超楕円体(17)の外側に存在するほど動きが大きくなると条件1' に束縛されることを示している。ただし、

$$\lambda_1 \leq k_3 / k_2$$

のとき1' は2を無条件に含むことを注意しておく。さらに、 $\lambda_1 / \lambda_n = \text{cond}(A)$ であることを思い起せば、2つの超楕円体(13), (16)の最大径はAの条件数とともに増大する。従って、条件数が多いほど解の許容空間が広がり、運動決定のあいまいさが増大することが解る。

### 5.3 対象抽出実験

3.で得られた各セグメントのなかには異なる動きを含み、融合規則を満たさないものもある。そのようなセグメントは、さらに分割しておく。

まず、そのようなセグメントの両端は他の動きの影響を受ける場合がある。そこで、端の2画素を削る。あるいは、セグメントを曲率の大きな点で分割する。それでも、融合規則を満たさなければ、そのセグメントを消去する。この様にして得られたセグメントの集まりを初期セグメント集合とよぶ。

初期セグメント集合から対象に対応するセグメントを抽出する。その前に、セグメント間の最小距離とセグメントの大きさを次の様に定義しておく。2つのセグメント $S_i, S_j$ 間の最小距離 $D_{ij}$ は、各セグメント

上の任意の点 $(x_i, y_i) \in S_i, (x_j, y_j) \in S_j$ とその深さ量 $z_i, z_j$ を使って、

$$D_{ij} = \min \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2}$$

と定義する。また、セグメントの大きさは含まれている画素点の数とする。

〈対象抽出アルゴリズム〉

Step1 セグメントを大きさの順に並べる。それをセグメント列とよぶ。

Step2 最も大きなセグメントをセグメント列から取り出す。それを $S_1$ と置く。

Step3  $S_1$ との最小距離が、ある閾値 $D_t$ 以下なるセグメントをセグメント列のなかから順に1つ取り出す。それを $S_i$ と置く。

Step4  $S_1$ と $S_i$ との和が融合規則を満たせば、 $S_1 \leftarrow S_1 \cup S_i$ 、満たさなければ、 $S_i$ をセグメント列に戻す。←印は左辺を右辺で置き換えることを意味する。

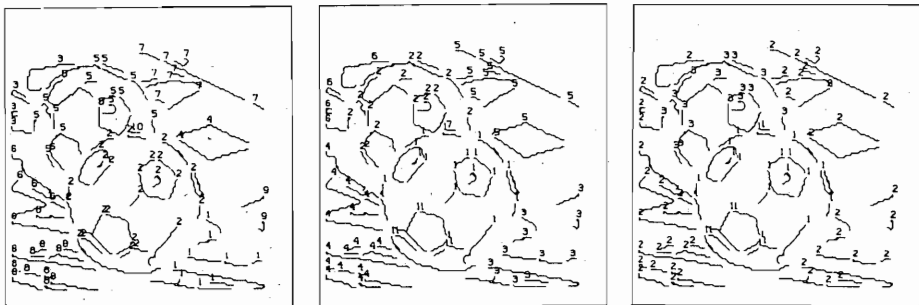
Step5 融合規則を満たすセグメントがセグメント列のなかに見つからなくなるまでStep3~4を繰り返す。

Step6 残ったセグメント列に対し、Step2~5をセグメント列が空になるまで、繰り返す。このようにして新たなセグメントの集まりが得られる。

Step7  $D_t$ を逐次大きくしてStep1~6を数回繰り返す。さらに、 $D_t$ を十分大きくして再びStep1~6を実行し、得られたセグメント集合を最終結果とする。

〈終り〉

上述の手続で最も大きなセグメントから融合を開始するのは信頼度の高い動きから始めるためである。ま



(a) 1st segmentation

(b) 2nd segmentation

(c) Final segmentation

図6 二つの転がるボールのシーンのセグメンテーション  
Fig.6 Segmentation process of independently moving balls scene.



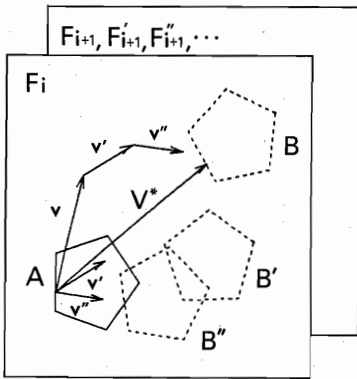


図8 対象の動きの反復推定

Fig.8 Object based iterative gradient method.

する。画像列  $F_i, F_{i+1}$  から  $A$  に対応する対象の3次元運動パラメータを推定する。式(1)を介して画面上の移動ベクトル分布を求める。 $B$ を $A$ に一致させるように、得られた移動ベクトル分だけ戻してセグメント  $B'$  を新しく画像  $F_{i+1}$  上に作る。画像列  $F_i, F_{i+1}$  から再び移動ベクトルを求め、同様に新しく画像  $F_{i+1}''$  上にセグメント  $B''$  を作る。この手続を、明るさの時間差の大ききの総和が減少しなくなるまで続ける。セグメント  $A$  上の点の移動ベクトル  $V^*$  は、各段階で得られた移動ベクトル  $v, v', v'' \dots$  の和として求められる。さらに、移動ベクトル  $V^*$  が対象像の歪みなどでセグメント  $B$  に到達しない場合には、移動ベクトルを移動方向に延長あるいは縮小させてセグメント  $B$  に一致させる。説明の簡単化のために、2枚の連続画像を使ったが、実際は5枚の連続画像を使って精度の向上を計る。

本推定法の精度を、転がるボールの追跡により評価した。なお、ここではボールの輪郭部などでのオクルージョン問題<sup>18,19</sup>や、対象検出の確からしさの問題<sup>20</sup>は扱わないことにする。使用した動画の枚数は150枚である。図9に10枚ごとの連続画像を表示する。図10に、反復推定による追跡結果を下側の曲線、反復推定しない場合の追跡結果を上側の曲線として示している。違いの理解のために、60番目と149番目の画像からボールだけを抜き出して、第3番目の画像に重ねた。60番目の画像上での追跡位置が、+印で示されている。反復推定しない場合は、追跡位置が五角形上の隣辺にずれている。図11にはいくつかの点の追跡結果が示されている。反復回数は、この実験例では、最高12回、最低3回、平均5回程度であった。

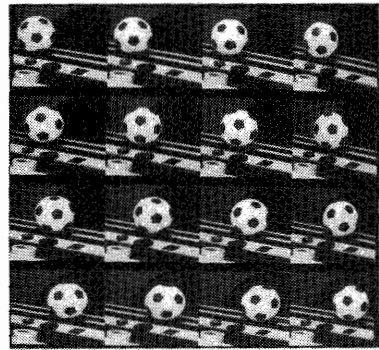


図9 転がるボールの連続画像(1番目から150番目まで10枚毎に表示)

Fig.9 Image sequence of rolling ball at intervals of 10 images up to 150th image.



図10 反復推定による追跡(下側)と反復しない場合の追跡(上側)

Fig.10 Upper/down curve is the estimated movement without/with the iterative gradient method.

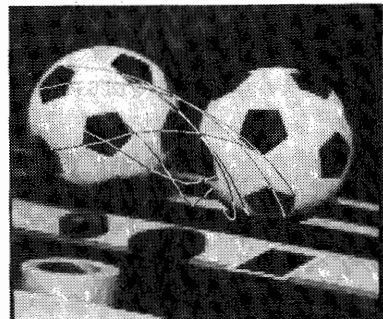


図11 複数点追跡結果

Fig.11 Movements of some points on the balls.

## 7. む す び

対象の3次元情報を既知として、勾配法により対象



の3次元運動を推定する手法を、複数動物体のシーンの解析に適用した。その際、シーンは動きが一意に解釈できないような対象を含んでいてもよい。その結果、個々の対象が抽出できた。同時に、異なる動きの本手法による識別限界を理論的に明らかにした。また、動きが一意に解釈できない対象を実際に判別した。さらに、抽出された対象の動きを反復推定することにより移動ベクトルを精密に推定した。

得られた移動ベクトルを用いて、簡単な追跡を行ったが、オクルージョンや対象検出の確からしさの問題を含んだ3次元追跡は別の機会に論じたい。

異なる対象でも、動きに差がない場合は抽出が難しい。その際、追跡情報の利用や他のセグメンテーション法を援用<sup>(6)</sup>すれば、さらに精密なセグメンテーションが可能であろう。

本手法は3次元情報を与えるステレオ立体視と併用することにより、3次元運動を解析するステレオ動画像処理システムとして発展が期待される。

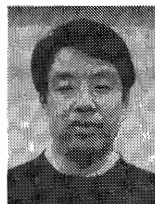
謝辞 日頃御指導頂く白井良明制御部長ならびに、貴重な御意見を頂く視覚システム研究室の皆様にご感謝致します。

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 (July 1985).
- (3) J. L. Potter: "Velocity as a Cue to Segmentation", IEEE, SMC-5, pp. 390-394 (May 1975).
- (4) J. L. Potter: "Scene Segmentation Using Motion Information", C. G. I. P., 6, pp. 558-581 (1977).
- (5) S. Tsuji, M. Osada and M. Yachida: "Tracking and Segmentation of Moving Objects in Dynamic Line Images", IEEE, PAMI-2, 6, pp. 516-522 (Nov. 1980).
- (6) W. B. Thompson: "Combining Motion and Contrast for Segmentation", IEEE, PAMI-2, 6, pp. 543-549 (Nov. 1980).
- (7) G. Adiv: "Determining Three-Dimensional Motion and Structure from Optical Flow Generated by Several Moving Objects", IEEE, PAMI-7, 4, pp. 384-401 (July 1985).
- (8) T. S. Huang and R. Y. Tsai: "Image Sequence Analysis: Motion Estimation", in Image Sequence Analysis, T. S. Huang ed., Springer-Verlag, New York (1981).
- (9) R. Y. Tsai and T. S. Huang: "Estimating Three-Dimensional Motion Parameters of a

- Rigid Planar Patch", IEEE, ASSP-29, 6, pp. 1147-1152 (Dec. 1981).
- (10) 丸山, 甘利: "物体の運動と立体構造の単眼視による認識の基礎", 信学論 (A), J67-A, 12, pp. 1107-1114 (昭59-12).
- (11) 山本正信: "動画像と距離情報の併用による3次元運動パラメータの直接的推定法", 信学論 (D), J68-D, 4, pp. 562-569 (昭60-04).
- (12) W. B. Thompson and S. T. Barnard: "Lower-Level Estimation and Interpretation of Visual Motion", IEEE, COMPUTER, 14, 8, pp. 20-28 (Aug. 1981).
- (13) C. Cafforio and F. Rocca: "Methods for Measuring Small Displacements of Television Images", IEEE, IT-22, 5, pp. 573-579 (Sep. 1976).
- (14) C. L. Fennema and W. B. Thompson: "Velocity Determination in Scenes Containing Several Moving Objects", C. G. I. P., 9, pp. 301-315 (1979).
- (15) R. J. Schalkoff and E. S. McVey: "A Model and Tracking Algorithm for a Class of Video Targets", IEEE, PAMI-4, 1, pp. 2-10 (Jan. 1982).
- (16) B. D. Lucas and T. Kanade: "An Iterative Image Registration Technique with an Application to Stereo Vision", 7th IJCAI, pp. 674-679 (1981).
- (17) 和田, 山口: "反復勾配法による動画像信号の動き量検出", 信学論 (D), J68-D, 4, pp. 663-670 (昭60-04).
- (18) 山本正信: "画像化された運動軌跡による動画像処理", 情処学論, 22, 5, pp. 442-449 (昭56-09).
- (19) K. M. Mutch and W. B. Thompson: "Analysis of Accretion and Deletion at Boundaries in Dynamic Scenes", IEEE, PAMI-7, 2, pp. 133-138 (March 1985).
- (20) 石川, 宮内, 小沢: "動画像追跡システムのための形状情報を用いた対象物体の位置推定法", 情処学論, 26, 3, pp. 399-406 (昭60-03).
- (21) 森, 名取, 鳥居: "数値計算, 岩波講座, 情報科学-18", 第3章, 岩波書店 (1982).

(昭和60年11月25日受付)



山本 正信

昭48九工大・工・制御卒。昭50東工大大学院修士課程了。同年通産省電子技術総合研究所入所。以来、動画像処理、知能ロボットなどの研究に従事。現在、同所制御部視覚システム研究室主任研究官。情報処理学会、ロボット学会、AVIRG等各会員。