

第4章 連続画像の処理

やまもと まさ のぶ
山本正信

電子技術総合研究所制御部

1. はじめに

時間的に連続する複数枚の画像から、対象の動きや構造を理解しようとする研究は、動画画像処理、あるいはステレオ立体視を含めて多重画像処理⁽¹⁾などと呼ばれている。動画画像処理は、テレビ信号の帯域圧縮などの実用面からの要請や、運動を知覚する生物の視覚メカニズムの解明のような科学上の動機に端を発し、画像処理のなかでも比較的新しい分野である。この分野は、ここ数年、特に基礎理論面を中心に活発に研究が繰り広げられている。最近では、この理論成果をロボットビジョンなどへ応用することも検討され始めた⁽²⁾。ここでは、このような分野の代表的な課題の幾つかを選んで紹介することにする。

動画画像処理では、次の三つの課題がよく取り上げられている

- (1) 動きの推定：移動ベクトルや運動パラメータの推定
 - (2) 対象の追跡
 - (3) 三次元構造の復元：動いている剛体の投影像からその三次元構造を推定する⁽³⁾
- このうち、動画画像処理特有の問題である(1)と(2)を中心に解説する。

2. 動きの推定

動画画像から動きを推定するとき、画像間の動きを推定する場合と、画面上に投影されたシーンの動きを推定する場合とに分けられる。

2.1 画像間の動きの推定

連続する画像間の動きは、画像間の対応づけ、あるいは移動ベクトルによって表わすことができる。移動ベクトルは比較的小きな動きを表わす場合が多く、その画面上の分布は、オプティカルフローと呼ばれている。

小さな動きの推定には、以下に示すように画像間差分画像の利用が便利である。画像上の点 (x, y) の移動ベクトル (u, v) は、画像の明るさの空間こう配 (E_x, E_y) と、画像間の差分値 E_t により、近似的にはあ

るが次の簡単な方程式に拘束されることが知られている⁽⁴⁾⁽⁵⁾。

$$E_x \cdot u + E_y \cdot v + E_t = 0 \dots\dots\dots (1)$$

E_x, E_y, E_t は動画画像から差分演算などにより容易に得られる。従って、 u, v を未知数とするこの方程式を利用して移動ベクトルを推定するためには、もう一つ別に移動ベクトルの拘束が必要である。

第二の拘束として、動きを平行移動とか回転運動など、特定の運動に限定する方法が考えられる。最も簡単な例として、平行に移動しているパターンの動きを求めてみよう。この場合、パターン上の各点では同じ移動ベクトルを未知数とする方程式(1)が成立する。従って、移動ベクトルはこれらの線形連立方程式の最小二乗解により得られる⁽⁶⁾。別解⁽⁶⁾として、(1)式をパラメータ空間 (u, v) へ Hough 変換したとき、二次元ヒストグラムのパーク点を探索することによっても移動ベクトルは得られる。この方法は複数個の動きの解析に便利である。

パターンの回転や拡大・縮小を表わす運動は、次のアフィン変換式でモデル化される。

$$\begin{cases} u = ax + by + u_0 \\ v = cx + dy + v_0 \end{cases} \dots\dots\dots (2)$$

(2)式を(1)式に代入すれば、6つの運動パラメータ a, b, c, d, u_0, v_0 を未知数とする次の線形推定式が導かれる。

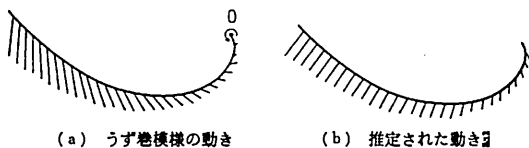
$$\begin{aligned} x \cdot E_x \cdot a + y \cdot E_x \cdot b + x \cdot E_y \cdot c + y \cdot E_y \cdot d \\ + E_x \cdot u_0 + E_y \cdot v_0 + E_t = 0 \dots\dots\dots (3) \end{aligned}$$

この推定式から、最小二乗解により運動パラメータを求める⁽⁶⁾。次いで、(2)式を経て移動ベクトルが得られる。

更に複雑な画像間の動きは、(2)式のような簡単なモデルでは表現し切れない。そこで、Horn 氏⁽⁷⁾は“移動ベクトルは滑らかに分布する”といった大域的な拘束を(1)式の拘束に付け加えた。すなわち、

$$\begin{aligned} \Sigma \{ (E_x \cdot u + E_y \cdot v + E_t)^2 + \alpha^2 (u_x^2 + u_y^2 \\ + v_x^2 + v_y^2) \} \longrightarrow \text{最小} \end{aligned}$$

となるように分布 (u, v) を定めるのである。ここで、



(a) うず巻模様の動き (b) 推定された動き
図1 滑らか拘束による動きの推定⁽⁸⁾

上式の第1項は(1)式, u と v の空間微分からなる第2項は滑らかさを表わす拘束に対応している。また, α は両拘束間のバランスを取る重みである。Hildreth氏⁽⁸⁾も同様な拘束を利用して, ゼロクロッシング上でオプティカルフローを求めた。その際, “この推定法は真の動きと異なる動きをもたらす場合があり, その動きは人間が知覚した結果に近い” ことを指摘している。例えば, 中心点まわりに回転しているうず巻模様の動き(図1(a))を推定した結果(図1(b))は, うずの中心に吸い込まれるような動きを示している。

動きが大きな場合には, パターンの特徴を利用した対応付けが行なわれる。対応付け問題の詳細はほかの解説⁽¹⁾⁽⁹⁾に委ね, ここでは第3節で具体例をあげるにとどめる。

2.2 シーンの動きの推定

シーンの動きの推定では, シーン中の剛体の三次元運動パラメータを推定することが多い。また, シーンは中心投影法(図2), あるいは平行投影法(図3)によって画面上に投影されている。この投影像から運動パラメータを推定する方法は次の二つに大別される⁽¹⁴⁾。

(1) 二段階推定法 最初に画像間の動きを, 移動ベクトルあるいは対応付けとして求め, 次にこれらの変動情報から運動パラメータを推定する方法。

(2) 直接的推定法 移動ベクトルなどを介せず, 幾つかの方程式を連立させて, 直接動きを推定す

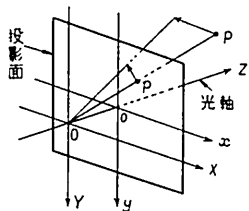


図2 中心投影法

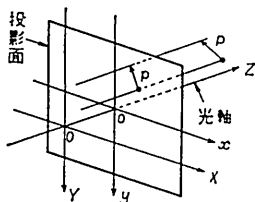


図3 平行投影法

る方法。

以下, それぞれ例をあげて説明する。

(1) 二段階推定法 二段階推定法により剛体の三次元運動パラメータを求める方法が幾つか提案されている⁽¹⁰⁾。ここでは, Higgins氏⁽¹¹⁾と Tsai氏⁽¹²⁾により, それぞれ独立に提案されたエレガントな手法を, 少し詳しく紹介しよう。

剛体上の点 $X=(X, Y, Z)$ が $X'=(X', Y', Z')$ に移動したとき, その運動は, 原点を通る軸まわりの回転を表わす直交行列 R と, 並進ベクトル $T=(T_x, T_y, T_z)$ を使って次式で表わされる。

$$X' = R \cdot X + T \dots \dots \dots (4)$$

点 X, X' はそれぞれ画面上の点 $x=(x, y), x'=(x', y')$ に中心投影されるとする。

$$\left. \begin{aligned} (x, y) &= (X/Z, Y/Z) \\ (x', y') &= (X'/Z', Y'/Z') \end{aligned} \right\} \dots \dots \dots (5)$$

x と x' が対応付けられているとき, 運動パラメータ R と T を求めるのが問題である。なお, 物体面までの距離が未知であるため, 並進ベクトルは要素の比でしか定められない。ここで簡単のため T の大きさを1としておく。

さて, (4), (5)式から Z と Z' を消去すれば, 運動パラメータだけを未知数とする次の推定式が導かれる。

$$(x', y', 1)Q(x, y, 1)^T = 0 \dots \dots \dots (6)$$

但し,

$$Q = R \cdot S = R \cdot \begin{bmatrix} 0 & T_z & -T_y \\ -T_z & 0 & T_x \\ T_y & -T_x & 0 \end{bmatrix} \dots (7)$$

なお, 添字の T は行列の転置を示す。

最初に, R と T からなる Q の要素を定めておく。(6)式は, Q の9つの要素に関して斉一次方程式であるので, 8組の対応付けから得られる8個の一次方程式から, Q の要素の比を定めることができる。次に, (7)式から,

$$\begin{aligned} Q^T Q &= S^T \cdot R^T R \cdot S = S^T \cdot S \\ &= \begin{bmatrix} 1 - T_x^2 & -T_x \cdot T_y & -T_x \cdot T_z \\ -T_y \cdot T_x & 1 - T_y^2 & -T_y \cdot T_z \\ -T_z \cdot T_x & -T_z \cdot T_y & 1 - T_z^2 \end{bmatrix} \dots \dots \dots (8) \end{aligned}$$

を得る。最右辺の行列のトレースは2に等しい。従って, 両辺の行列のトレースが等しくなるように, Q の要素を確定した上で両辺を比較して T を定める。 R は S が正則ではないので(7)式から直接得られないが, Q の行ベクトルと T のベクトル積⁽¹¹⁾, あるいは

Q の特異値分解によって得られる二つの直交行列⁽¹²⁾を使って、一意に定め得ることが証明されている。

(2) 直接推定法 2・1 項でみてきたように、平面運動のあるクラスは、オプティカルフローや対応付けを必要とせずに直接運動パラメータを推定することができた。三次元運動でも同様に、その運動パラメータを直接推定することができないであろうか。

剛体上の点 $X=(X, Y, Z)$ の移動ベクトル $V=(U, V, W)$ は移動が小さな場合、原点を通る軸まわりの角速度ベクトル $\Omega=(\omega_x, \omega_y, \omega_z)$ と、並進ベクトル $V_0=(u_0, v_0, w_0)$ により次式で表わされる。

$$V = \Omega \times X + V_0 \dots\dots\dots (9)$$

シーンが、平行投影

$$x = X, y = Y \dots\dots\dots (10)$$

されるならば、 (U, V) がオプティカルフローとなる。この U, V を(1)式の u, v にそれぞれ代入して整理すれば次の推定式が導かれる。

$$E_x \cdot u_0 + E_y \cdot v_0 - z \cdot E_y \cdot \omega_x + z \cdot E_x \cdot \omega_y + (x \cdot E_y - y \cdot E_x) \omega_z + E_t = 0 \dots\dots (11)$$

今、物体までの距離 z が既知であるとすれば、運動パラメータ Ω と V_0 はこの推定式の最小二乗解として得られる⁽¹³⁾。実際に、この方法で動きを推定した例を図4に示す。但し、結果はオプティカルフローと回転軸により表わされている。この手法は、ステレオ立体視など、適当なレンジファインダと組み合わせることにより、実用的な運動解析法として期待できる。

物体までの距離 z が未知の場合は、常に未知数の個数が方程式の個数を上回る。そのため、直接推定は不可能である。そこで、Tsai 氏と Huang 氏⁽¹⁴⁾は、物体を平面と仮定し、面の傾きと運動パラメータが同時に得られる非線形の連立方程式を導いた。解析解を得るためには⁽¹⁵⁾、更に、非線形方程式の線形近似が必要である。

3. 対象の追跡

対象を追跡するには、まず、画像ごとに対象を検出し、次に対象の特徴を利用して、画像間の対応付けにより追跡を行なうのが、オーソドックスな方法⁽¹⁶⁾である。具体例として、廊下のシーンについて築山氏ら⁽¹⁷⁾が行なった歩行者の追跡を取り上げてみよう。

図5に示される4枚の連続した画像から、各々つま先に相当する特徴点を検出する。この特徴点と床面に対するカメラの設置情報から実際の床面上の歩行者の位置が算出される。この位置情報を利用して、画像間で歩行者を対応付けるのであるが、歩行者が複数存在



図4 三次元運動の直接的推定

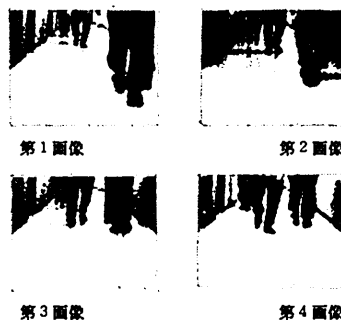


図5 人の動きの連続画像⁽¹⁷⁾

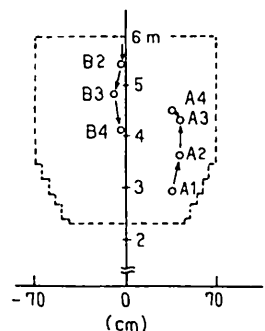
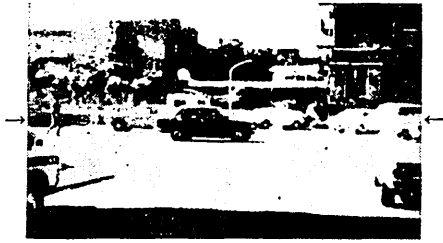
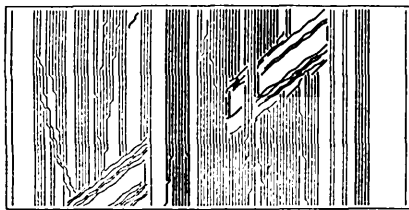


図6 検出された人の動き⁽¹⁷⁾

するので、真の対応付けを見つけなければならない。それには、対応付けを評価する機構を設け、最適な対応付けを選択する。この例では、歩行者達の移動距離が、全体として最小になるような対応付けを選ぶ。その際、歩行者の最大移動速度や壁に沿って歩くという歩行者の性質が対応付けの最適化に利用されている。このようにして得られた追跡結果を図6に示す。

ところで、三次元シーンでは物体同志が重なり合ってみえる場合が多い。このオクルージョン現象は、対象の特徴を利用する追跡法にとってやっかいな問題である。オクルージョンが起る対象の追跡には、対象の特徴利用にはよらず、直接運動軌跡を抽出する方法⁽¹⁸⁾が有効である。街頭で行き交う人や車の追跡を例に、この手法を説明しよう。

街頭シーンの1コマを図7に示す。この動画像を時

図7 街頭シーンの1コマ⁽¹⁸⁾図8 水平断面の合成画像⁽¹⁸⁾図9 人と車の運動軌跡⁽¹⁸⁾

間順に重ねて、一冊の本のような三次元画像を構成する。人や車が左右にしか移動しないとすれば、この三次元画像の水平断面には対象の見掛の運動軌跡が画像化される。実際、図7の矢印の位置での水平断面画像は図8のようになる。図の真ん中付近の黒っぽい方形の領域や、その左右の白い斜めの条痕は、それぞれ停車中の車や通行人に対応している。また、右上から左下に向けた黒い条痕を含む白い領域は、右から左に移動している車に対応したものである。この合成画像から、移動する車や通行人などの運動軌跡を抽出した結果が図9に示されている。この軌跡をたどることにより、車はオクルージョンによる消滅や再登場の場合でも、また通行人は車が背後を通過することによる背影変動にもかかわらず、追跡可能である。

4. あとがき

連続画像の処理について、動きの理解を中心に現状を概説した。動画像処理の適用分野は、医療や気象、交通など広く考えられるが、データ量が膨大なため処理の高速化が必要である。特に視覚による移動体の誘導などには専用ハードウェアによる高速処理が不可欠

である。これに対し、イメージセンサと演算機能が一体化された超々 LSI デバイス⁽¹⁹⁾が開発されつつある。このようなハードウェアの進歩は、実時間処理が重要な意味をもつ動画像プロセッサの実現に大きな希望を与えるものである。

(昭和59年12月25日受付)

文 献

- (1) 谷内田:「多重画像を利用したシーンの理解」情報処理 24, 12, 1429 (昭 58-12)
- (2) J. Fang & T.S. Huang: "Some experiments on estimating the 3-D motion Parameters of a rigid body from two consecutive image frames" *IEEE Trans. Pattern Analysis Machine Intell.* PAMI-6, 5, 545 (1984)
- (3) S. Ullman: *The interpretation of visual motion* (1979) MIT Press
- (4) J.O. Limb & J.A. Murphy: "Estimation of velocity of moving images in television signals" *C.G.I.P.* 4, 311 (1975)
- (5) C. Cafforio & F. Rocca: "Methods for measuring small displacements of television images" *IEEE Trans. Information Theory* IT-22, 5, 573 (1976)
- (6) R.J. Schalkoff & E.S. McVey: "A model and tracking algorithm for a class of video targets" *IEEE Trans. Pattern Analysis Machine Intell.* PAMI-4, 1, 2 (1982)
- (7) B.K.P. Horn & B.G. Schunk: "Determining optical flow" *Artificial Intelligence* 17, 185 (1981)
- (8) E.C. Hildreth: "Computations underlying the measurement of visual motion" *ibid.* 23, 309 (1984)
- (9) J.K. Aggarwal, L.S. Davis & W.N. Martin: "Correspondence processes in dynamics scene analysis" *Proc. IEEE* 69, 5, 562 (1981)
- (10) H.C. Longuet-Higgins & K. Prazdny: "The interpretation of a moving retinal image" *Proc. R. Soc. Lond.* B208, 385 (1980)
- (11) H.C. Longuet-Higgins: "A computer algorithm for reconstructing a scene from two projections" *Nature* 293, 10, 133 (1981)
- (12) R.Y. Tsai & T.H. Huang: "Uniqueness and estimation of three-dimensional motion parameters of rigid objects with curved surfaces" *IEEE Trans. Pattern Analysis Machine Intell.* PAMI-6, 1, 13 (1984)
- (13) 山本:「動画像と距離情報を併用した3次元運動パラメータの直接的推定法」信学論誌 J-68 D 画像特集 (昭 60-4)
- (14) T.S. Huang & R.Y. Tsai: "Image sequence analysis; Motion estimation" *Image sequence analysis*. (1981) N.Y. Springer-Verlag
- (15) R.Y. Tsai & T.S. Huang: "Estimating three-dimensional motion parameters of arigid planar patch" *IEEE Trans. Acoust. Speech, Signal Process.* ASSP-29, 6, 1147 (1981)
- (16) J.W. Roach & J.K. Aggarwal: "Computer tracking of objects moving in Space" *IEEE Trans. Pattern Analysis Machine Intell.* PAMI-1, 2, 127 (1979)
- (17) 築山・白井:「連続する画像からの人の動きの検出」SICE 論文集 19, 1, 909 (昭 58-11)
- (18) 山本:「画像化された運動軌跡による動画像処理」情報処理学会論文誌 22, 5, 442 (昭 56-9)
- (19) 片岡, 他:「特集 "新機能素子"」日本の科学と技術 25, 225号 (昭 59-1)