

解説



動画像処理†

辻 三 郎††

1. ま え が き

われわれ人間の視覚は、テレビジョンや映画のように一定時間ごとにサンプルされた時系列の画像信号から、興味のある対象、それらの動きのパターン、さらに人物の行動やストーリーといった高次の情報を捉えている。イメージプロセッシングも、従来の一枚ずつの静止画像から一歩進んで、一連の時間的に変化する画像、すなわち動画像を研究の対象とするようになった^{1),2)}。

コンピュータを用いた動画像の処理は、画像伝送のための帯域圧縮の研究から始まったが、最近ではビデオテープや映画フィルムなどに記録された各種の実験の動画像をコンピュータが解析し、運動物体を検出したり、運動パターンや形状変化を調べる研究が盛んになった。例えば、衛星写真の雲の動きから風速を推定したり³⁾、生物の個体や組織の運動解析⁴⁾⁻⁷⁾、心臓の運動パターンの計測⁸⁾⁻¹⁰⁾、交通量調査¹¹⁾などがあげられる。これらの動画像はデータ量が膨大で、従来行われてきた人手による解析では多くの時間と労力を必要とし、コンピュータによる解析が強く要望されている。

このような応用指向のアプローチとともに、人工知能の色彩の濃い基礎的研究も行われている。そこでは、コンピュータが短いアニメーション映画入力を時間とともに変動する世界として解釈し、さらにそこで起っている行動を推論するシステム^{12),13)}などが発表されている。

最近盛んになった視覚研究の一つの流れは、視覚情報処理では何が本質で、それは何故かを明らかにしようという立場である。すなわち、種々の心理実験で人間の視覚の持つ機能を明確にし、人工知能の立場からそれを再構成するものである¹⁴⁾。運動視覚でも、多くの興味ある事実とその仕組みが明らかにされ、動画像

理解の有力な手がかりを与えている¹⁵⁾。

2. 動画像解析

動画像は、普通一定時間ごとに撮像される時系列の画像で、各サンプル画像をフレームと呼ぶ。動画像を解析するには、各フレームに一枚ずつ独立に静止画像処理のやり方を適用しても解決できない。その問題点としては、

(1) データ量が膨大：動画像のデータ量は膨大である。TVカメラから送られる各画面を256×256点、各点8ビットのデジタル画像に変換すると、1分間の動画像は約120Mバイトの情報量がある。したがって、能率良く処理することが不可欠である。

(2) フレーム間の変化の検出と対応づけ：単に各フレーム内の対象を検出し、その構造を明らかにするだけでなく、異なるフレーム間でそれらがどう変化し、各々の部分がどう対応するかを知る必要がある。

(3) 物体の重なり：物体が移動したり回転したりすると、他の物体に重なって見えない部分ができたり、あるいは物体自身の一部が現れたり隠れたりする。このような重なりによって起る画面内のイメージの構造変化を、コンピュータが解釈する必要がある。以上の他に、抽出された動パターンを解析し、それを意味づけるといった高次の問題がある。

さて、このような動画像を解析するため、次のようなアイデアが提案されている。

[A] 変化情報に基づく解析画像の選択：動画像は膨大なデータ量を持っているが、動画像解析ではそのすべての部分を詳しく調べる必要はない。画面内の変化のある部分だけを解析したり⁶⁾、前に解析した画面に比べてほとんど変化のないフレームはスキップする¹⁰⁾方法が計算量を節約するため有効である。

[B] 運動情報を利用した画面の分割：明るさ、色、テクスチャといった静止画像の解析に利用した情報の他に、運動ベクトルの一様性に着目した領域分割ができる。

† Analysis of Dynamic Images by Saburo TSUJI (Osaka University, Faculty of Engineering Science).

†† 大阪大学基礎工学部

【C】前画面の処理結果の利用：以前の画面の処理結果から、画面内に存在する物体の性質とその運動のモデルを作成し、このモデルを用いて以降の画面の処理の効率化をはかる。

【D】前後画面を利用した解析：一枚のフレームだけでは、いくら詳しく調べても不確かな部分や解釈し難い部分も、前後の画面の処理結果から推論し解釈が可能となる。

動画像処理には、次のレベルが考えられる。

- (1) 運動領域の検出
- (2) 運動ベクトルの解析
- (3) 運動対象の追跡
- (4) 変化するシーンとしての理解
- (5) 行動の理解

(1)と(2)は低レベルの処理で、対象についての知識はほとんど利用しないで、フレーム間の変化成分を利用して運動領域を検出し、各部の運動ベクトルを算出してその類似性から対象を手、足といった意味のある部分に分割する。(3)は、発見された動領域を以降の画面で追跡し、その運動を測定する。(1)(2)が対象についての知識を利用しないボトム・アップの処理であるのに対し、追跡は(1)で得られたモデルを利用するトップ・ダウンの処理を用いている。(4)と(5)は、動画理解のレベルで、画像内の各部の動きをカメラで写されている三次元世界の変化やカメラの動きとして解釈し、さらにそれから“走る”“投げる”といった行動を推論する。現在の所、(1)~(3)のレベルは比較的良好に研究されているが、動画理解のレベルの研究はあまり進んでいない。

3. 運動領域の検出

動画像解析の第一段階は、動物体の検出である。すなわち、一連の画像間の変化分を解析し、定常領域と動領域に分割する。

差分画像の利用

動物体を検出する最も単純な方法は、二つの画面の差の利用であろう。対象が二つのフレームの間に、図-1の点線から実線の位置に移ったとする。前後のフレームの差分画像には、図-1のように負の領域と正の領域ができる。これから直接には動物体の一部しか検出できないが、動物体が背景より暗い(または明るい)ことが判っていると、その移動ベクトルを推定できる。この情報を第一近似とし、相関法の探索領域を狭くして能率良く対象の移動を測定できる⁷⁾。

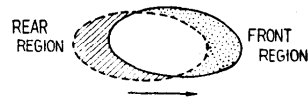


図-1 差分画像

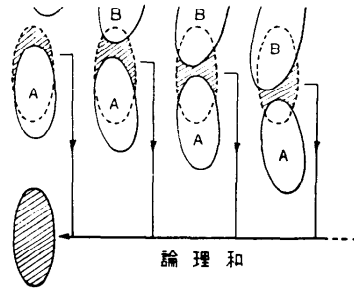


図-2 連続差分画像⁹⁾

連続差分画像

差分画像から動物体の輪郭を検出するには、図-2に示すように差分画像のうち負(または正)の成分のみを利用し、連続したフレームで論理和をとる。これは動物体が複数個あり、前フレームで物体 A のあった場所に後続フレームで物体 B が入ってくる影響を除くためである⁹⁾。

類似性検定を利用する差分画像

実世界の動画像に対して上記の方法を適用すると、雑音に弱い欠点が問題となる。そこで画素間のグレイレベルの差でなく、小領域(例えば 4×6)間の類似度を調べる方法が提案されている^{16),17)}。すなわち、

$$\frac{[2(S_c + S_R) + (m_c - m_R)^2]^2}{S_c S_R} > t$$

の時ミスマッチとする。ただし、 m_c と m_R は現フレームと比較フレームの小領域の濃淡値の平均値、 S_c 、 S_R はそれぞれの分散、 t はしきい値である。連続したフレームに対して類似度を調べ、ミスマッチの回数を累積して FODP (一次差分画像) を作成し、その性質から動画像の解析を行っている。またフレーム間の FODP の差を SODP (二次差分画像) として求めると、対象検出に有効である。図-3 は、暗い対象が(a)から(b)(d)と動いた時(速度は1画素1フレーム)の FODP, SODP を示す。

4. 運動ベクトルの解析

差分画像を利用した動物体の検出では、物体が移動したかどうかの情報を用いるだけで、その移動ベクトル(移動量と方向)は利用していない。したがって、

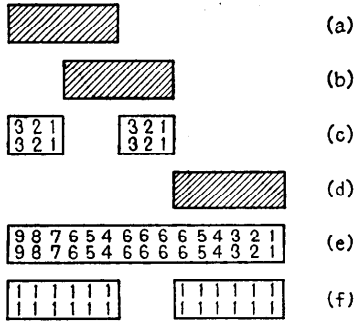


図-3 FODP (c) (e) と SODP (f)

重なって見えるが速度の異なる物体、人の手や足のよう異なる動きのパターンを示す対象を、一つの動領域として検出する。

このような異なる動きをする対象を分離するには、動領域の各部分の移動ベクトルを求め、その類似性により領域分割をすればよい。しかし、対象の画面内の移動は二次元であるから、図-4 に示すような小さい局所領域の観測では移動方向は決定できない。そこで近傍のパターンの対応関係を調べて運動ベクトルを算出する相関法¹⁷⁻¹⁹⁾やプレート照合法¹⁸⁾、明度の勾配とフレーム間の明度変化から可能な移動方向を累積して求める方法¹⁹⁾などが研究されている。ここでは、最近発表された運動ベクトルによる対象分割の手法を紹介する。

可変プレート照合法²⁰⁾

図-5 (a) に示す線画から運動ベクトルを求め、手足などの部分に分割する問題を考えてみよう。一方の線画 F_1 を短いセグメントに分割し、それぞれに対応するセグメントを他方の線画 F_2 で発見できれば、移動ベクトルが算出できる。まず F_1 、 F_2 から閉じた領域を発見し、その対応づけを形状パラメータなどを用いて行う。 F_1 の各領域の境界を分岐点、端点でアークに分割し、対応するアークを F_2 内で探す。しかし図-5 に示すように物体は三次元的に動くので、フレーム間で領域の大きさや形が変わり、普通のプレート照合法ではマッチングがうまくとれない。そこで図-5(b) に示すように、アークをセグメントに分割し、対応づけを行う。セグメントは F_1 では互いに接続しているが、対応づけられた F_2 では順序は同じであるが互いに接続するとは限らず、大きさや形の変化に応じて重なったり離れたりする。図-6 は、こうして求められたセグメントの中点の移動ベクトルである。画

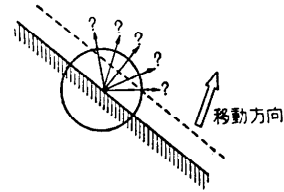
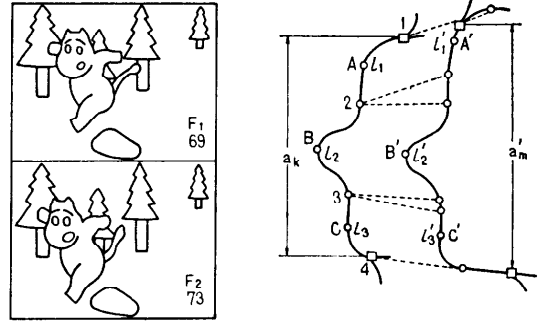


図-4 局所領域の観測だけでは、移動方向は分らない。



(a) 入力画像 (b) 可変プレート照合法²⁰⁾

図-5

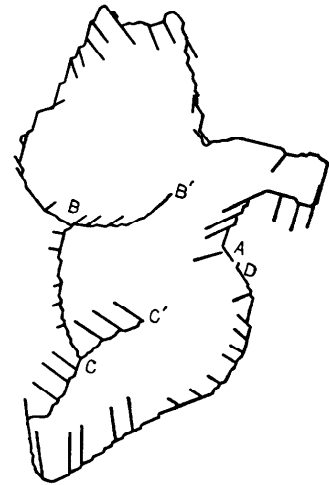
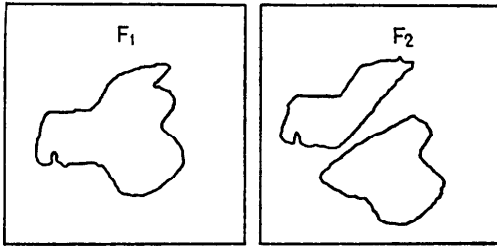


図-6 移動ベクトル¹⁹⁾

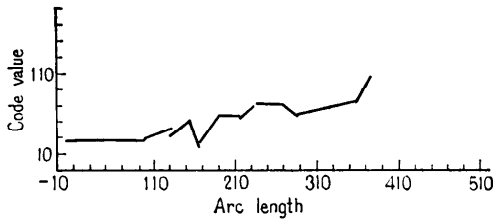
面内の移動ベクトルの類似性から、動領域と背景の分離、さらに動領域を手や足などの部分に分割できる。

$\phi-S$ 関数による対応づけ²¹⁾

チェーン符号で記述された曲線は、容易に $\phi-S$ 関数で表現できる。ここで ϕ は出発点からのチェーン符号 (すなわち角度) の差を加算した値、 S は出発点からの長さである。曲線を平面上で回転させても $\phi-S$ 曲線は変わらないから、出発点の対応がとれると二



(a) 入力画像



(b) $\phi-S$ 曲線の直線近似
図-7

つの曲線の形の類似度は $\phi-S$ 関数の差を計算するだけで測定できる。また $\phi-S$ グラフ上の水平な部分は、もとの曲線では直線に相当し、グラフ上の直線は原図形の円弧に相当する。

図-7 (a) のフレーム F_1 と F_2 から異なる速度で動く物体を検出するため $\phi-S$ 関数を利用する。各フレームの $\phi-S$ 曲線を探し、それを線分近似する。次に二つの $\phi-S$ グラフで長さ、傾きなどの性質が最も似ている線分を探し、それから出発して隣接する線分の類似度を $\phi-S$ 関数の差で調べ、しきい値以下のものをグループ化する。この操作をくり返し、両フレームでグループ化した線分を対応づける。図-8 は、その結果である。次に対応するグループの移動ベクトルの類似度を求め、個々の物体を発見した。図-8 で、グループ A と B は移動ベクトルがほぼ等しいので結合され、最終的に画面で二つの物体が発見された。

5. 動物体の追跡

動物体が発見され、その性質が判明すると、それから各動物体のモデルを作り、モデルを利用して以降の画面内の動物体を能率良く追跡し、その時間的変化を求める。

モデルを用いた個体の追跡

個体の運動の追跡には、各個体のモデルの利用が有効である。ビデオ記録されたサカナの運動を追跡する

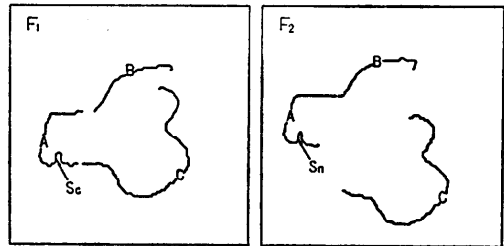
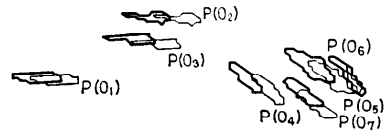


図-8 $\phi-S$ 関数による対応づけ²¹⁾



太線：モデル 細線：予測
図-9 モデルからの予測²¹⁾

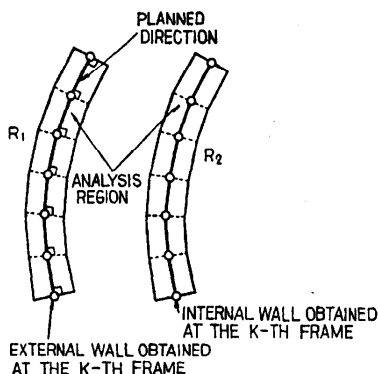
システム⁵⁾では、今までのフレームの解析結果から個々のサカナのモデルとして、面積、周囲長、重心位置、明度のしきい値、境界線、速度を持つ、これから図-9に示すように次のフレームの状態、すなわち各個体の位置、明度、他の物体と重なるかどうかを予測し、予測を用いたトップ・ダウンの手法で能率良く次画面から物体を検出する。モデルの利用は、特に複数個の物体が重なって見える場合の解釈に有効である^{22), 23)}。

計画を用いた動パターン解析

入力画面をまず縮小し、それについて簡単な処理を行って解析の計画を立て、計画に従ってトップ・ダウンの処理をする考えは静止画像の解析で提案され、計算量の節約や結果の信頼性向上という利点があることはよく知られている。しかし、このアイデアは膨大な情報量を持ち、しかもフレーム間の相関の大きい動画像を解析する時、その真価を発揮する。

心臓の X 線写真の動画像(シネ・アンジオグラム)から、心臓壁厚の動パターンを測定するシステム¹⁰⁾では、まず計画を利用する静止画像解析法により第1フレームの心臓の外壁と内壁を検出する。この結果に基づいて、以降のフレームを解析する計画を作成する。計画は、(1)解析するフレームの選択、(2)そのフレーム内で調べる領域、(3)そこで用いる特徴抽出のオペレータについて指示する。

まず発見された外壁と内壁のまわりに図-10に示すような解析領域を設け、その中だけを調べることにより処理の能率を向上させる。心臓壁の動きは一定でなく、動きの早い期間と遅い期間があるので、遅い部分

図-10 解析領域¹⁰⁾

ではサンプル間隔を広げて差支えない。新しい入力フレームの解析領域のグレイ・レベルと解析済みフレームのそれを比べ、差が少ないとそのフレームはスキップする。差がしきい値以上であると、解析領域の中だけを計画が指示したオペレータで特徴抽出し、壁を探索する。この方法で入力フレームの約 1/3 が解析されるが、心臓の拡張・収縮の末期では約 1/6 が解析されるに過ぎない。解析時間も、計画を用いないとミニコンピュータで 1 フレームに約 1 時間を要するが、この方式では約 14 秒と著しい能率の向上が達成された。なおスキップするかどうか判断する時間は、1 秒以下と非常に短い。

6. む す び

動画の解析の現状を、主として低レベルの運動情報の抽出と利用、すなわち動物体の検出、運動ベクトルの計測と利用、モデルを活用する効率的な動物体の追跡に重点をおいて述べた。しかし、使用したモデルは簡単なもので、今後はより複雑な構造モデル⁶⁾を利用した解析が必要となろう。

動画解析システムの実用化には、その処理に適した画像プロセッサの開発が必要である。特に実時間処理、大量の動画データデータの圧縮、蓄積、アクセスには専用ハードウェアが不可欠である。

基礎研究の面から考えると、本稿で述べてきたシステムは、画面内の変化に着目して解析をしているが、写された世界での変化としては理解していない。静止画像を三次元世界として理解するコンピュータ・ビジョンの対象を、動画に拡張するのは今後の重要課題であろう。画面内の変化の原因としては、物体の運動の他にカメラの動き、照明条件の変化がある。motion

stereo や photometric stereo などの知識を包含し、時空間世界を理解できるシステムに発展することを期待する。

参 考 文 献

- 1) Martin, W. N. and Aggarwal, J.K.: Dynamic Scene Analysis, Computer Graphics and Image Processing, Vol. 7, No. 3, pp. 356-374 (1978).
- 2) Nagel, H. H.: Analysis for Image Sequences, Proc. 4th Int. Joint Conf. Pattern Recognition, pp. 181-211 (1978).
- 3) Hall, D. J. et al.: Objective Methods for Registering Land Marks and Determining Cloud Motions from Satellite Data, IEEE Trans. Comput., Vol. C-21, pp. 768-776 (1972).
- 4) Greaves, J. O.: The Software Structure for Reduction of Quantized Video Data of Moving Organisms, Proc. of IEEE, Vol. 63, pp. 1415-1425 (1975).
- 5) Yachida, M., Asada, M. and Tsuji, S.: Automatic Motion Analysis System of Moving Objects from Records of Natural Processes, Proc. 4th Int. Joint Conf. Pattern Recognition, pp. 726-730 (1978).
- 6) Ariki, Y., Kanade, T. and Sakai, T.: An Interactive Image Modeling and Tracing System for Moving Pictures, *ibid.*, pp. 681-685 (1978).
- 7) Takagi, M. and Sakaue, K.: The Analysis of Moving Granules in a Pancreatic Cell by Digital Moving Image Processing, *ibid.*, pp. 735-739 (1978).
- 8) Kaneko, T. and Macini, P.: Straight-Line Approximation for the Boundary of the Left Ventricular Chamber from a Cardiac Cineangiogram, IEEE Trans. Biomed., Vol. BME-20, pp. 413-416 (1973).
- 9) Jong, L. P. and Slager, C. J.: Automatic Detection of the Left Ventricular Outline in Angiographs Using Television Signal Processing Technique, IEEE Trans. Biomed., Vol. BME-22, pp. 230-237 (1975).
- 10) Yachida, M., Ikeda, M. and Tsuji, S.: Plan-Guided Analysis of Noisy Dynamic Images, Proc. 6th Int. Joint Conf. Artificial Intelligence, pp. 978-983 (1979).
- 11) Onoe, M. and Ohba, K.: Digital Image Analysis of Traffic Flow. Proc. 3rd. Int. Joint Conf. Pattern Recognition, pp. 803-808 (1976).
- 12) Badler, N.I.: Temporal Scene Analysis: Conceptual Description of Object Movement, Univ. of Toronto, TR-80 (1975).
- 13) Tsuji, S., Morizono, A. and Kuroda, S.: Understanding a Simple Cartoon Film by a Com-

- puter Vision System, Proc. 5th Int. Joint Conf. Artificial Intelligence, pp. 601-610 (1977).
- 14) Marr, D.: Representing and Computing Visual Information, in (Winston, P. H. and Brown, R. H. Eds.) Artificial Intelligence: An MIT Perspective, Vol. 2, pp. 19-80, MIT Press (1979).
 - 15) Ullman, S.: The Interpretation of Visual Motion, MIT Press (1979).
 - 16) Jain, R., Miltzer, D. and Nagel, H. H.: Separating Non-Stationary from Stationary Scene Components in a Sequence of Real World TV Images, Proc. 5th Int. Joint Conf. Artificial Intelligence, pp. 612-618 (1977).
 - 17) Jain, R. and Nagel, H. H.: On the Analysis of Accumulative Difference Pictures from Image Sequence of Real World Scenes, IEEE Trans. Pattern Analysis & Machine Intelligence, Vol. PAMI-1, pp. 206-214 (1979).
 - 18) Potter, J.L.: Scene Segmentation by Velocity Measurement Obtained with a Cross-Shaped Template, Proc. 4th Int. Joint Conf. Artificial Intelligence, pp. 803-808 (1975).
 - 19) Fenema, C.L. and Thompson, W.B.: Velocity Determination in Scenes Containing Several Moving Objects, Computer Graphics and Image Processing, Vol. 9, pp. 301-315 (1979).
 - 20) Tsuji, S., Osada, M. and Yachida, M.: Three-Dimensional Movement Analysis of Dynamic Line Images, Proc. 6th Int. Joint Conf. Artificial Intelligence, pp. 896-912 (1979).
 - 21) Martin, W. N. and Aggarwal, J. K.: Computer Analysis of Dynamic Scenes Containig Curvilinear Figures, Pattern Recognition, Vol. 11, No. 3, pp. 169-178 (1979).
 - 22) Roach, J. W. and Aggarwal, J. K.: Computer Tracking of Objects Moving in Space, IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. PAMI-1, pp. 127-135 (1979).
 - 23) Chow, W. K. and Aggarwal, J. K.: Computer Analysis of Planar Curvilinear Moving Objects, IEEE Trans. Comput., Vol. C-26, pp. 179-185 (1977).

(昭和 55 年 2 月 12 日 受付)
