

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[†]

石川 繁樹^{**} 宮内 ミナミ^{**} 小沢 慎治^{**}

動画像から、移動物体を検出し追跡するシステムにおいて、対象物体の検出結果は、追跡能力を大きく左右する。しかしながら、実際の場面では一時的に対象物体を検出するのに十分な画質を得られないことや、対象物体が他の物体と見かけ上重なり合っさえぎられること、対象物体を視野内に保持できないことなどが起こりうる。本論文では、このような状況に対応するため、システムが状態を推定して、対象物体を正しく検出することができないと判定した場合には、対象物体を正しく検出することができると判定された隣接する過去のフレームの対象物体の形状に関する情報を用いて、より正しい対象物体の位置を求める方法を提案する。本手法の有効性を確認するため、本手法を動画像追跡システムに適用して実験を行った。形状の一部が正しく検出されており、誤り領域との重なりがあまり大きくない画像では、本手法は有効であり、正しく対象物体の位置を求めることができた。形状の情報を用いることにより、動画像追跡システムの汎用性・適応性を高めることができることが確認された。

1. ま え が き

近年、人間の視覚機能の代行という立場からコンピュータ・ビジョン、ロボット・ビジョンに関する研究・開発がさかんに行われるようになってきている。これらのうち、工場・空港・街頭などに、画像入力をもつ計算機システムを備え、監視・観察・分類などの作業を代行させることを目的として、動画像から移動物体を検出し、追跡するシステムの研究・開発が行われている^{1),2)}。これらのシステムは一般に、人間の視覚機能に比べて、適応性・汎用性に乏しく、制約も多い。そのため、対象物体の形状・大きさ・位置などに関する制約を緩和するための研究も多く行われている³⁾⁻⁷⁾。

実際の場面では一時的に対象物体を検出するのに十分な画質が得られないことや、対象物体が他の物体と見かけ上重なり合っさえぎられること、物体を視野内に保持できないことなどが起こりうる。

このような状況に対応するため、ビデオ記録された魚の運動を追跡するシステム⁸⁾では、ある時刻までのフレーム解析から個々の魚をモデルとして、予測を用いたトップダウンの手法を用いている。また、動画像から、物体像の運動軌跡が描かれるような画像を合成し、合成画像から運動軌跡を抽出して物体の追跡を行

う手法⁹⁾や、過去の運動軌跡の情報からカルマン・フィルタを用いて対象物体の位置を予測し追跡する手法⁷⁾などが提案されている。

一般に、動画像における移動物体追跡の手法は、二つのプロセスから成り立っている。まず、各フレームにおける対象物体を検出する。次に、各フレームで検出された対象物体を対応づけることにより追跡を行う。したがって、対象物体の検出を正確に行うことが、追跡能力を左右する。しかし、実際には、明度情報のみでは対象物体を正しく検出できない状況が起こりうる。本論文では、このような状況に対処するために、形状情報を用いて、より正しく対象物体の位置を求める手法を提案する。

筆者らは、動画像追跡システムにおいて入力された画像を評価し、システムの機能状態をモニターすることができれば、その状態に適した処理方法を選択することができると考え、検討を行ってきた⁸⁾。本論文で提案する手法は、明度情報によって対象物体の検出を行った結果(二値画像)に対して、対象物体を正しく検出できないと判定された場合に、形状特徴量、周辺分布、距離変換画像等を用いてモデルとの比較を行い、対象物体のより正しい位置を求める。モデルは、対象物体が正しく検出されると判定された過去のフレームの対象物体の形状から作成する。これを形状モデルと呼ぶ。本手法を動画像追跡システムに適用して実験を行い、本手法の有効性を確認した。

2. システム構成

本論文では、図1に示すような動画像追跡システム

[†] A Method of Estimating Target Positions Using Shape Information for an Image Tracking System by SHIGEKI ISHIKAWA, MINAMI MIYAUCHI and SHINJI OZAWA (Department of Electrical Engineering, Faculty of Science and Technology, Keio University).

^{**} 慶応義塾大学理工学部電気工学科

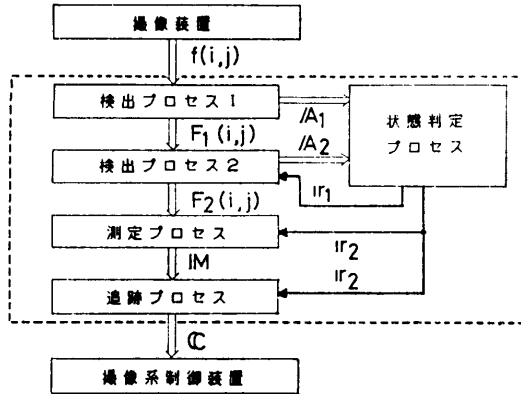


図1 システム構成
Fig. 1 System block diagram.

の構成を想定している。このうち、本論文で提案する手法は検出プロセス2に適用される。以下、システムの構成を述べる。

撮像装置から画像 $f(i, j)$ を入力し、この $f(i, j)$ に対し、検出プロセス1で明度情報に基づいて対象物体を検出する。動画像では扱うデータ量が膨大であり、高速かつ正確に動画像から対象物体を検出するために、全画面のうち、対象物体が含まれると思われる小領域（ウィンドウ）を設定し、その小領域内に対して処理を行う。この検出プロセス1では、画像中から指示されたウィンドウ画像において、対象物体は周囲の背景と異なる明度分布をもつという前提に基づき、各画素の明度データを情報として、フレームごとに対象物体が本来もっている明度分布を推定して、背景から物体を分離する⁸⁾。この結果、二値画像 $F_1(i, j)$ を得る。

検出プロセス1で対象物体を背景から分離する際に得られる特徴量 A_1 を用いて、状態判定プロセスで“検出の確からしさと追跡の状態”を判定する。この判定結果は、図1に示すとおり各プロセスに送られる⁸⁾。

検出プロセス2では、明度情報から検出された結果 $F_1(i, j)$ を入力として、形状の情報を用いてより正しい検出結果である二値画像 $F_2(i, j)$ を得る処理を行う。

このようにして得られた対象物体の検出結果の画像 $F_2(i, j)$ から、測定プロセスにおいて対象物体の位置（重心位置）および速度が測定される。これらの結果から、追跡プロセスにおいて次のフレームにおける対象物体の位置を線形予測し、ウィンドウの位置の設定および撮像系の制御を行う。

3. 形状情報を用いた対象物体の位置推定

本手法の適用条件を以下に述べる。

- (1) フレーム間では対象物体の形状の変化は急激でない。
- (2) フレーム間では対象物体の運動の変化は急激でない。
- (3) 最初のフレームでは対象物体の位置は既知であり、正しい検出結果が得られる。
- (4) 対象物体の形状の変化は進行方向に沿った二次元平面上の回転のみとする。

本手法の処理手順を図2に示す。

3.1 孤立点除去

検出プロセス1で、明度情報に基づいて対象物体と背景の分離を行った結果 $F_1(i, j)$ は二値画像であり、対象物体は正しく検出された場合、連結領域を形成する。しかしながら、このときに背景が誤って対象物体として検出されることがあり、このような画素は空間的孤立点あるいは小領域として現れることが多い。一方、対象物体が誤って背景となることも考えられ、この場合には連結領域の中に欠損（穴）として現れる（図3(a)）。

こうしたことを考慮して、以下の手順で孤立点除去を行う。

- (1) 連結領域内における一点欠損を充填する。すなわち、注目している画素 (i, j) が背景で、その周囲4近傍が対象物体の画素である場合、その画素を対象物体の画素に変更する（図3(b)）。
- (2) 各画素について8-隣接画素に関する縮退と伝搬を行う方法⁹⁾を用い、孤立点および 3×3 未満の小領域を除去する（図3(c), (d)）。

これらの操作により、孤立点などの雑音成分は除去されるが、対象物体の形状・大きさなどの特徴はほぼ保たれる。

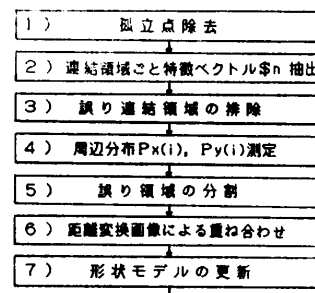


図2 検出プロセス2の処理手順
Fig. 2 Procedures of detecting process.

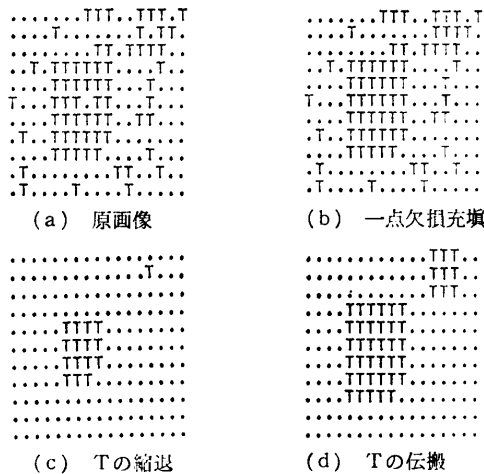


図3 孤立点除去. T: 対象物体, .: 背景
Fig. 3 Noise clearing operation.

3.2 連結領域の分類

各連結領域について、幾何学的特徴量を求める。特徴量としては、面積、周囲長、円形度、サイズ、輪郭の偏角差分ヒストグラムを用いる。各領域ごとに、(1)式のような形状特徴ベクトル S_i (i は連結領域の番号) を求める。

$$S_i = (a, l, k, c, H) \tag{1}$$

- i : 連結領域の番号
- a : 面積
- l : 周囲長
- c : 円形度 $= 4a/l^2$
- k : サイズ $= 2a/l$
- H : 輪郭の偏角差分ヒストグラム

(1)式の形状特徴ベクトル $S_i (i=1 \sim n)$ をモデルの形状特徴ベクトル S_m と比較し、最も類似した連結領域を対象物体である領域として、それ以外の領域を排除する。

3.3 周辺分布を用いた対象物体の範囲推定

X, Y 軸方向への周辺分布 $P_x(i), P_y(i)$ をそれぞれ

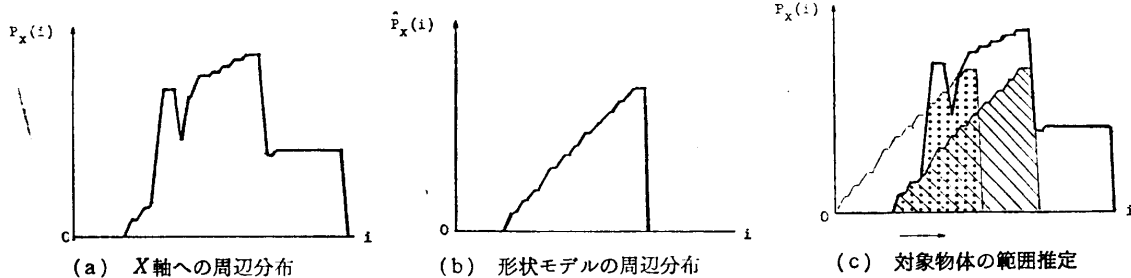


図4 対象物体の存在する範囲の推定
Fig. 4 Estimation of the object extent.

求め、形状モデルの周辺分布 $\hat{P}_x(i), \hat{P}_y(i)$ と比較して、対象物体のある範囲を推定する。図4に示すように、形状モデルの周辺分布をテンプレートとして X, Y 軸方向にそれぞれ移動し、(2)式が最大となる範囲、すなわち、測定された周辺分布とテンプレートの重なりが最も大きくなるような範囲を求める。

$$P \cap \hat{P}(k) = \sum_i \min [P(i), \hat{P}(i+k)] \tag{2}$$

この処理により、対象物体が存在すると推定される範囲が限定される。範囲外の部分は誤まり部分と考えられ、5画素の余裕を考えて、推定された範囲の外周5画素の外側の対象物体画素を排除する。

3.4 距離変換を用いた画像重ね合せによる対象物体の位置推定

前節までで得られた画像に対して、距離変換を施す。距離変換は、連結成分に属する各画素の値を背景画素からの最小距離値に置き換える操作をいう⁹⁾。この操作により、連結成分を背景画素からの最小距離値で重みを付けた画像(距離変換画像)が得られる。距離変換画像は図5に示すような三次元的な高さをもったものと考えられる。

この距離変換画像を用いて対象物体の位置推定を行う。形状モデルについて求めた距離変換画像をテンプレートとして、前節で推定された対象物体の存在すると考えられる範囲内でテンプレートを移動し、SSDA法¹⁰⁾による画像重ね合せを行う。距離変換画像の重ね合せを行うことによって、対象物体の形状のマッチングを取ることができる。また、モデルと最も形状の一致する位置が対象物体の存在する位置と考えられる。したがって、残差が最小の位置を対象物体が存在する位置とする。SSDA法は画像間の相関を残差の総和で評価するものであり、閾値を工夫することにより、高速な画像重ね合せを行う。

3.5 形状モデルの更新および予測

各ステップにおいて用いた形状モデルの形状特徴ベ

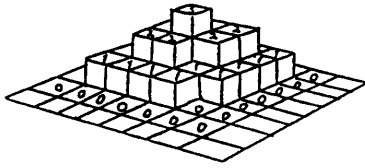


図 5 距離変換画像

Fig. 5 Distance transformation.

クトル, 周辺分布, 距離変換画像は, 状態判定プロセスにおいて, 検出・追跡の状態が正しいと判定されたフレームにおいて更新される. モデルはそのフレームの対象物体像から作成される. 検出・追跡の状態が悪いと判定されたときは, システムはモデルを更新せず, 検出の状態がよくなるまで過去のモデルを用いて予測モデルを作成する.

実際の場面では対象物体は三次元の広がりを持ち, 画像として得られる対象物体は二次元平面への投影であるため, モデルを予測することはむずかしい. そこで, 本論文では前述の仮定のように, 対象物体の形の変化は, 二次元平面上の回転のみを考える. 対象物体はつねに運動方向と同じ方向を向き, 運動方向の変化に伴って, 見かけ上二次元の回転をするとする(図 6).

図 6 のように, 時刻 T において, 時刻 $T-1$ からの対象物体と位置の変位 \vec{A}_T と, モデルが最終的に更新された時刻 T' と $T'-1$ との位置の変位 $\vec{A}_{T'}$ の方向変化を $\delta\theta$ とする. このとき, (3)式に示すアフィン変換を使って形状の変化を予測する.

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} u \\ v \end{bmatrix} \quad (3)$$

ただし,

$$a = d = \cos \delta\theta$$

$$b = \sin \delta\theta$$

$$c = -\sin \delta\theta$$

$$u = X_0(1 - \cos \delta\theta) - Y_0 \sin \delta\theta$$

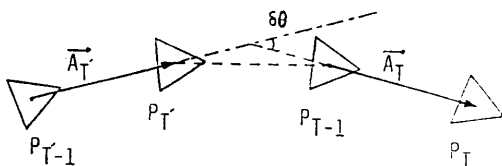


図 6 対象物体の運動方向の変化

Fig. 6 Change of the object orientation.

P_T : 時刻 T における対象物体の位置, $P_{T'}$: 最後にモデルが更新された時刻 T' における対象物体の位置, $\delta\theta$: 運動方向の変化

$$v = -X_0(-\sin \delta\theta) + Y_0(1 - \cos \delta\theta)$$

(X_0, Y_0) は回転の中心

4. 実験

本システムの有効性を確認するため, 実験を行った. 標準画像データベース SIDBA¹¹⁾より, HOME等の画像を背景の画像として用いた. 図 7 に示すように, 背景画像に対象物体像を合成し, 本システムの入力とした. 対象物体の各画素における明度は, 正規分布乱数で与えている.

4.1 検出プロセス 2 の実験

明度情報に基づいて対象物体を検出する検出プロセス 1 の処理結果の二値画像を入力として, 検出プロセス 2 の実験を行った. 図 8~10 に検出プロセス 2 における各ステップの処理結果の例を示す.

図 8, 9(a)では, 対象物体(三角形)と類似した明度分布をもつ領域が背景の一部に存在したため, 検出プロセス 1 では背景の一部を誤って対象物体として

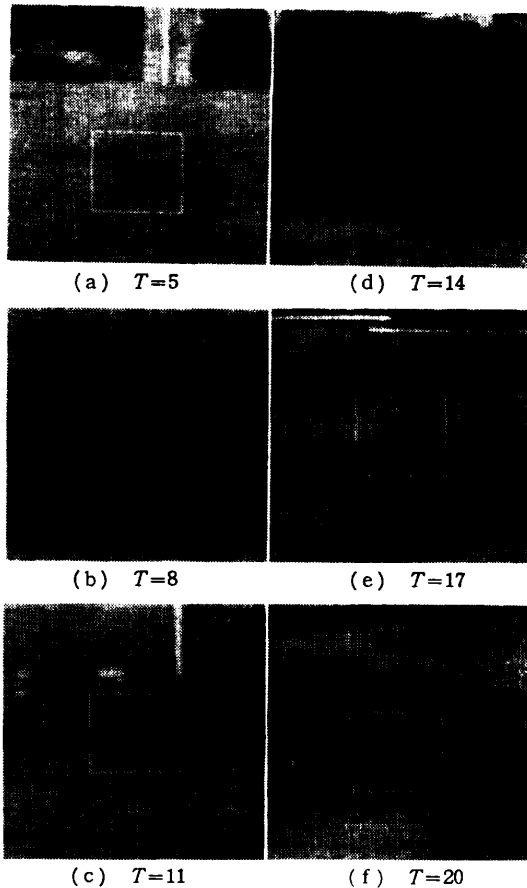
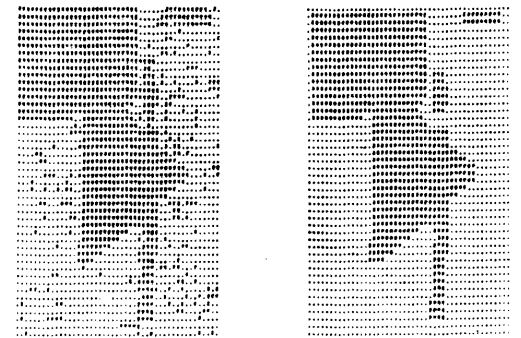
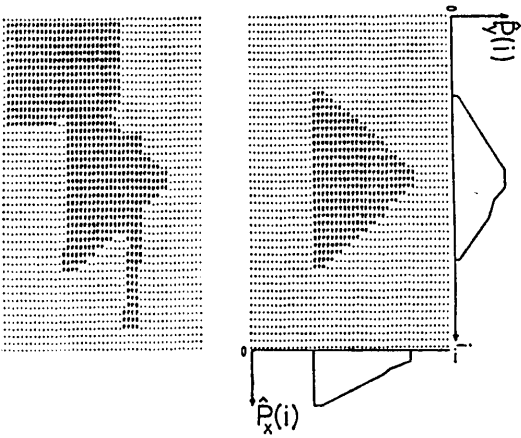


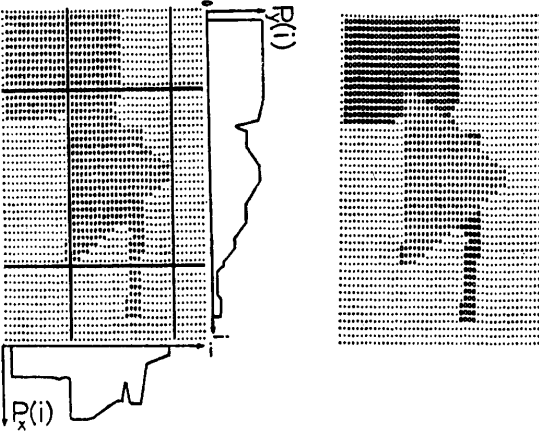
図 7 対象物体像と背景画像の合成画像の例
Fig. 7 Examples of the superimposed object image.



(a) 検出プロセス1の出力画像 $F_1(i, j)$ (b) 孤立点除去



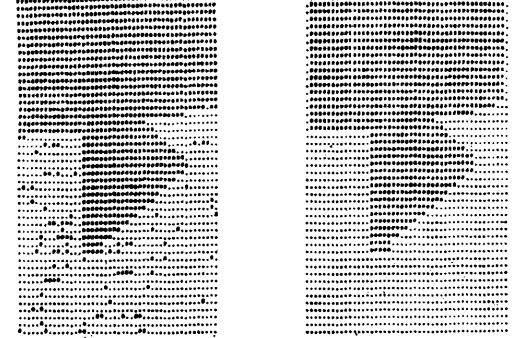
(c) 誤り連結領域排除 (d) 形状モデルと周辺分布



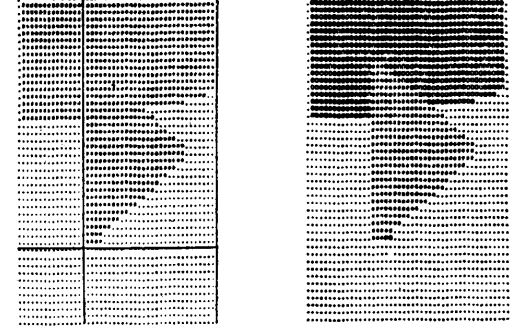
(e) 周辺分布と対象物体の範囲推定 (f) 距離変換画像による画像重ね合せ

図8 検出プロセス2の処理結果(1)
Fig. 8 Results (1) of detecting process-2.

検出している。また、図10(a)では背景と対象物体のコントラストが低く明度分布に偏りがあったため、背景領域を誤って対象物体として検出している。いずれの場合も、状態判定プロセスにおいて、“対象物体



(a) 検出プロセス1の出力画像 $F_1(i, j)$ (b) 孤立点除去



(c) 対象物体の範囲推定 (d) 距離変換画像による画像重ね合せ

図9 検出プロセス2の処理結果(2)
Fig. 9 Results (2) of detecting process-2.



(a) 検出プロセス1の出力画像 $F_1(i, j)$ (b) 距離変換画像による画像重ね合せ

図10 検出プロセス2の処理結果(3)
Fig. 10 Results (3) of detecting process-2.

を正しく検出できない”と判定されている。

図8(a)の検出結果 $F_1(i, j)$ に対して、孤立点除去を施した結果を同図(b)に示す。雑音成分は除かれ、また、形状はほぼ保たれている。対象物体の含まれる連結領域以外にも誤って対象物体であるとされた領域が存在している。各領域の形状特徴ベクトルを、形状モデル(d)の特徴ベクトルと比較することによ

り、誤って検出した連結領域を除き (同図(c)), X, Y軸方向に測定した周辺分布と, 形状モデルの周辺分布 (同図(d)) から対象物体の存在する範囲を推定した結果を同図(e)に示す. また, (f)は推定された範囲内で距離変換画像を施した画像に, 形状モデルの距離変換画像を重ね合せ, 対象物体の位置を求めた結果である. ●で示す部分が求められた対象物体であり, ○で示す部分が検出プロセス2で排除された領域である. この例では, 周辺分布による対象物体の範囲の推定において, 正しい位置がほぼ求まっている.

図9に示す例では, 周辺分布による対象物体の範囲の推定結果は広範となっているが, 距離変換画像を用いた画像重ね合せの結果, 対象物体の正しい位置が求まっている.

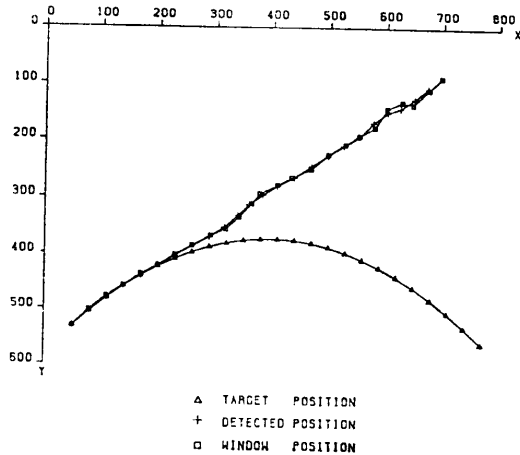
図10に示す例では, 検出プロセス1の結果において, 対象物体の形状情報がほとんど失われている. このような場合には, 対象物体の位置を正しく求めることができない.

本手法を用いることにより, 対象物体を明度情報に基づいて検出した結果において, 対象物体の一部が正しく検出されている場合には対象物体の位置を正しく求めることができる. 本手法は, 形状の情報を用いているので, 形状の特徴がほとんど失われている場合には有効でない.

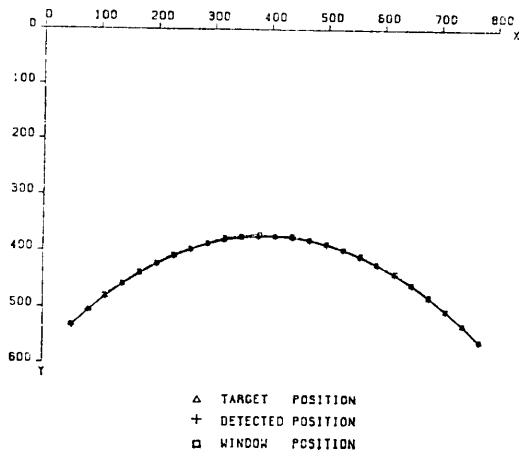
4.2 追跡の実験

本手法を動画像追跡システムに適用した場合の有効性を確認するための追跡実験として, 本手法 (検出プロセス2) を除いたシステム (孤立点除去の処理のみ行う) と, 本手法 (検出プロセス2) を用いたシステムについて追跡能力の比較を行った. 背景画像, 対象物体の運動軌跡・形状を変化させ, 100通りの画像系列を作成し, それぞれ25フレームを処理して対象物体の真の運動軌跡と追跡結果を比較した. 対象物体の運動について実験した.

追跡実験の一例を図11に示す. 同図(a)は本手法を適用しないシステムの結果である. この例では, 第7フレームにおいて明度情報に基づく対象物体の検出が正しく行われなかったため, 運動の予測が正しく行われず, 追跡を持続できなかった. 一方, 同図(b)に示すように本手法を用いたシステムでは検出の状態が悪化したフレーム (第7フレーム) においても, 対象物体の正しい位置を求めることができるので追跡を正しく行うことができた. この例では図7に示した合成画像を用いている. この例の第7フレームにおける



(a) 検出プロセス2を用いない場合の追跡結果



(b) 検出プロセス2を用いた場合の追跡結果

図11 追跡実験結果の例

Fig. 11 An example of tracking results.

ウィンドウ内の画像と検出結果を, 本手法を用いない場合と用いる場合について図12, 13に示す.

また, 第7フレームから第21フレームまでは, 対象物体が類似した明度をもつ背景中を移動する. この間, 状態判定プロセスでは, “対象物体を正しく検出できない”と判定しており, 形状モデルは更新されなく予測モデルによって対象物体の位置を求めている.

図14に示す画像は, 検出の状態が悪くなる直前に更新された形状モデル (第6フレーム) と, 第12フレームにおける予測モデルである.

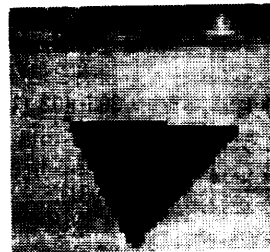
次に, 追跡の実験結果をまとめたものを図15に示す. 図15は正しく追跡されなくなったフレーム番号を横軸に, 画像系列の数を縦軸に示すヒストグラムである. フレーム番号25は最後まで正しく追跡が行われたことを示す.

同図より、本手法（検出プロセス2）を適用しないシステムと適用したシステムの比較を行うと、本手法を用いたシステムでは最後のフレームまで追跡が正しく行われることが、本手法を除いたシステムに比べ圧倒的に多いことがわかる。

このように、明度情報による対象物体の検出の状態が悪化した場合においても、本手法を用いることにより、正しく追跡が行えることが確認できた。しかしながら、形状の情報がまったくあるいは大部分失われている場合には、本手法では正しい位置を求めることがむずかしい。

5. まとめ

動画から、移動物体を検出し追跡するシステムにおいて、入力された画像の状態やシステムの機能状態をシステム自身が知り、その情報を使って処理を行うことが必要である。本論文では、動画追跡システムのうち、“物体を正しく検出することができない”と判定された画像に対して、“物体を正しく検出できる”と推定された過去の画像における形状の情報を用いることで、正しい対象物体の位置を求める手法について述べ、また、本手法の有効性を確認した。



(a) 入力ウィンドウ画像 (T=7)



(b) 検出プロセス1の出力画像 $F(i, j)$



(c) 孤立点除去

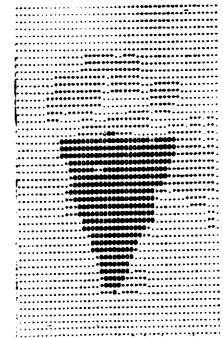
図 12 図 11 に示した追跡実験の処理結果例（検出プロセス2を用いない場合）
Fig. 12 Examples of the tracking experiment shown in Fig. 11 (without detecting process-2).



(a) 入力ウィンドウ画像 (T=7)

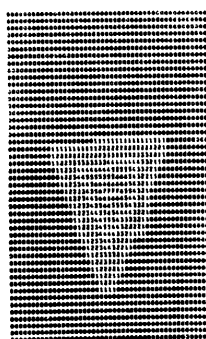


(b) 検出プロセス1の出力画像 $F(i, j)$

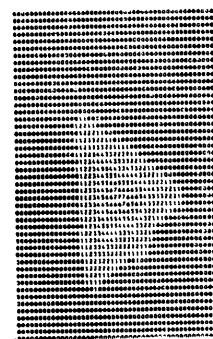


(c) 求められた対象物体の位置

図 13 図 11 に示した追跡実験の処理結果例（検出プロセス2を用いた場合）
Fig. 13 Examples of the tracking experiment shown in Fig. 11 (with detecting process-2).



(a) 最後に更新されたモデル (T=6)



(b) (a)より予測されたモデル (T=12)

図 14 形状モデルと予測モデル
Fig. 14 A shape model and a predicted model.

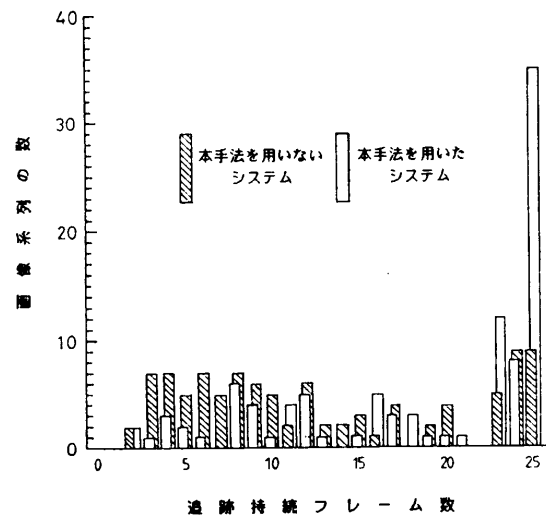


図 15 追跡実験の集計結果
Fig. 15 Results of tracking experiments.

形状の一部が正しく検出されており、誤り領域との重なりがあまり大きくない画像では、正しく対象物体の位置を求めることができた。また、本手法を適用することにより、動画像追跡システムの汎用性・適応性を向上させることができる。

しかしながら、形状がまったく、あるいは大部分失われている場合には、本手法は有効でなかった。こうした場合には、運動軌跡の情報を利用する手法などと組み合わせることが必要であると考えられる。

また、本論文において、形状モデル作成のための条件として簡単な場合を扱ったが、今後の課題として三次元的な広がりを考慮することにより、より汎用性・適応性を向上させることができると考えられる。

参 考 文 献

- 1) Aggarwal, J. K. et al.: Special Issue on Motion and Time-Varying Imagery, *IEEE Trans.*, PAMI-2, No. 6, pp. 493-516 (1980).
- 2) Martin, W. N. and Aggarwal, J. K.: Survey; Dynamic Scene Analysis, *Comput. Gr. Image Process.*, Vol. 7, pp. 356-374 (1978).
- 3) Roach, J. W. and Aggarwal, J. K.: Computer Tracking of Objects Moving in Space, *IEEE Trans.*, PAMI-1, No. 2, pp. 127-135 (1979).
- 4) Gilbert, A. L. et al.: A Real Time Video Tracking System, *IEEE Trans.*, PAMI-2, No. 1, pp. 47-56 (1980).
- 5) 浅田 稔, 谷内田正彦, 辻 三郎: 運動物体の検出と追跡, *信学論*, Vol. J 62-D, No. 6, pp. 395-402 (1979).
- 6) 山本正信: 画像化された運動軌跡による動画像処理, *情報処理*, Vol. 22, No. 5, pp. 442-449 (1981).
- 7) Legters, G. R., Jr. and Young, T. Y.: A Mathematical Model for Computer Image Tracking, *IEEE Trans.*, PAMI-4, No. 6, pp. 583-594 (1982).
- 8) 山田ミナミ, 小沢慎治: 動画像追跡システムにおける検出の確からしさと追跡の状態の判定, *電気学会論文誌 C*, Vol. 104, No. 7, pp. 165-172 (1984).
- 9) Rosenfeld, A. and Kak, A. C.: *Digital Picture Processing*, Academic Press, New York (1976).
- 10) 尾上, 前田, 斎藤: 残差逐次検定法による画像の重ね合わせ, *情報処理*, Vol. 17, No. 7, pp. 634-640 (1976).
- 11) SIDBA 79, ver 1, from MCIDS IMAGE SAVE TAPE (1979).

(昭和 59 年 5 月 2 日受付)

(昭和 59 年 10 月 18 日採録)