

指示者からの情報と領域変化のダイナミクスを併用した対象追跡法

菊地 明美[†] 高橋 弘太[†]

[†] 電気通信大学 情報通信工学科 〒182-8585 東京都調布市調布ヶ丘 1-5-1

E-mail: †{kikuchi,kota}@ice.uec.ac.jp

あらまし 信号分離問題では、音源の位置を知ることで分離性能の向上が可能となる。そこで音響信号分離に画像信号を併用し、画像から得られる目的音源の位置情報を利用した信号分離を行ない、それを補聴システムへ応用することを考えた。提案する補聴システムの画像処理部には、使用者の指示に基づいて目的対象を決定し、その対象を追跡して音源位置情報をリアルタイムで出力することが要求される。本報告ではこの要求を満たす新しい対象追跡法を提案する。提案手法のシステムは、まず使用者の指示によって音源位置やテンプレート情報を決定する。そしてその情報を、観測される画像を用いて数理統計的な根拠に基づき更新し、対象追跡を実現する。この手法を用いて拡大縮小・回転などを含む変形をしながら複雑な運動をする対象の追跡シミュレーションをおこない、従来のテンプレートマッチングよりもロバストな対象追跡性能が得られることを示した。

キーワード 動画像, 対象追跡, テンプレートマッチング, 領域分割, ベイズ推定

A method for object tracking using user intention and dynamics of object area

Akemi KIKUCHI[†] and Kota TAKAHASHI[†]

[†] Department of Information and Communication Engineering, University of Electro-communications,
1-5-1 Chofugaoka, Chofu-shi, Tokyo, 182-8585 JAPAN

E-mail: †{kikuchi,kota}@ice.uec.ac.jp

Abstract In the source separation problem, knowledge of source locations improve the performance of separation. This is a motivation to develop our new hearing system which uses not only acoustical signals but also image sequences for source localization. The functions of the image processing sub-system of the hearing system are decision of the object region according to the intention of user and the real-time object tracking. In this paper, we propose a new method for object tracking which has the ability for these functions. In this method, at first, a location of a sound source and a shape of the template for image matching are estimated according to a pointing light located by user. Then, using image sequence, the location and the template are updated based on a statistical model; the updating induces the object tracking. Numerical simulations using synthesized image sequences including a object, changing its shape and location randomly, demonstrate robustness and efficiency of the proposed method by comparison with conventional template-matching method.

Key words image sequence, object tracking, template matching, region segmentation, Bayesian estimation

1. はじめに

カクテルパーティのような状況で目的音を聞き分けるとき、人間は聴覚だけでなく視覚による情報も利用していると言われている。このような生体の機能を見本とし、視覚情報を併用する補聴システムを提案する。

補聴システムでは、混合された音響信号から使用者が聞きたい信号を分離しなくてはならない。しかし音響信号のみで信号

源を特定し、分離をおこなうことは困難である。ところが画像信号を併用すれば、画像情報から音源の位置を知ることができ、この位置情報を利用することで分離処理を簡単にすることができる。また画像情報があれば、使用者が目的音源を補聴システムに指定することも容易となる。これは補聴システムにおいて不可欠な重要な機能である。さらに無音状態でも音源位置を指定したり追跡させたりすることが可能になる。

ところで、使用者が常に音源を指示し続けなければならない

とすると、補聴システムでは、実用性に乏しいと言わざるを得ない。そこでこの補聴システムの画像処理部には、使用者の指示に基づいて決定した目的音源を追跡する機能が必要となる。この機能を実現する従来法として、テンプレートマッチングがある。テンプレートマッチングでは指示情報から対象の位置とテンプレートを決定した後、対象を追跡する。しかし人間の顔を追跡することを考えると、表情や顔の向きの変化で画像が変化するため、補聴システムに必要な追跡性能が得られない。

そこで、視覚情報を併用した補聴システムの画像処理部に適した対象追跡法を提案する。この手法は、使用者の指示に基づいて初期の対象位置を決定した後、対象の位置情報、対象領域であるか否かの度合、テンプレート、テンプレートの信頼度情報を内部変数として保持する。そして画像の観測によって数理統計的な手法でそれら内部変数を逐次変化させ、更新させていくことで対象の追跡をおこなうものである。本報告ではその手法について説明し、さらに提案法の有効性を示すために計算機シミュレーションをおこなった。

本報告の構成は次の通りである。

まず2章でテンプレートマッチングによる対象追跡について説明し、その問題点を指摘する。次に3章で提案する対象追跡法の概要を説明する。提案法は大きく分けて予測と推定の過程からなる。そこで4章では予測過程について説明する。そして5章では推定過程の処理について説明し、特に提案法における内部変数の更新法について詳細に説明する。6章では計算機シミュレーションによる実験について示し、最後に7章でまとめる。

2. テンプレートマッチングによる対象追跡

2つの画像が一致するかどうか、一致するならばその相対的位置はどこかを調べるための手法をマッチングと呼び、特に画像そのものをを用いて画素値でマッチングをおこなうものをテンプレートマッチングという。

時刻 t における観測画像を F_t (サイズ $M \times N$)、テンプレートを T_t (サイズ $W \times W$) とする。位置 (i, j) におけるそれぞれの値を $F_t(i, j)$, $T_t(i, j)$ とあらわすとき、

$$S_t(i, j) = \sum_{i'=0}^{W-1} \sum_{j'=0}^{W-1} |F_t(i+i', j+j') - T_t(i', j')| \quad (1)$$

を全画素について求め、 $S_t(i, j)$ の値がしきい値 Z_t を越えない場合、位置 (i, j) にテンプレートと一致する画像部分があると考えられる。ここでしきい値 Z_t は過去の $S_t(i, j)$ の値で最小のものを採用することで自動的に与えることができる。このテンプレートマッチングの処理を時刻 $t=1, 2, \dots$ における観測画像に対しておこなうことで、対象の追跡が可能となる。

従来のテンプレートマッチングでは、対象の位置とテンプレートを決定し、テンプレートを用いてマッチングをおこなっていた。ここで、テンプレートの抽出に失敗し、観測画像中の対象画像とテンプレート画像の相関が低くなった場合を考える。テンプレートの信頼性に関する指標があればそれを反映させてマッチング計算をおこなえるので、ある程度の追跡性能を得る

ことができる。しかし従来のテンプレートマッチングではテンプレートの信頼性に関する指標がないため、両者の相関が低くなったなら追跡に失敗してしまうという問題があった。

また、テンプレートのサイズが大きいと追跡性能がよいことが知られているが、従来のテンプレートマッチング法ではテンプレートの大きさは常に一定であり、テンプレートの大きさや形を逐次変える機能がなかった。そのため追跡性能に限界があった。

このような欠点があるため、従来のテンプレートマッチングは拡大縮小・回転するような対象の追跡性能が悪い。[1]ではこの欠点を解決するため、テンプレートマッチングと他の手法との融合を試みている。本報告では、この欠点を克服し、ロバストな対象追跡を実現する手法を提案する。

3. 指示と領域の変化を併用した対象追跡

3.1 提案する手法で用いるパラメータ

2章で述べたような従来のテンプレートマッチングには、テンプレートの信頼度を考慮する機能と、テンプレートの大きさや形を変化させていく機能がない。この機能を持たせるため、使用者の指示に基づいて時刻 t における画像 F_t (サイズ $M \times N$) から3つの情報を抽出する (図1)。

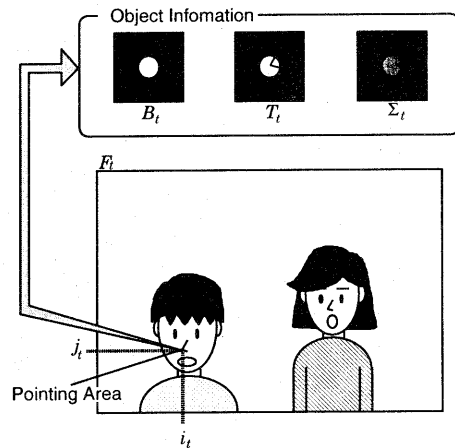


図1 使用者の指示に基づいた情報の抽出のイメージ

1つめの情報は、対象領域の境界に関する情報 B_t である。以下 B_t を境界画像と呼ぶ。 B_t は対象の領域であるかどうか、領域内だとしたら対象である可能性はどれくらいか、という情報を持っており、0~1の値をとる。初期の境界画像は大きさや値を適当に与えて生成される。そして画像を観測しながら領域を更新させていく。

2つめはテンプレート T_t である。 T_t を B_t で切出したものが観測画像 F_t と比較すべきものであるとする。 T_t と F_t を比較することで、 T_t を更新していく。

そして3つめはテンプレート T_t の各画素の濃度分布の分散 Σ_t である。以下 Σ_t を分散画像と呼ぶ。この分散画像は T_t の各画素の濃淡値の信頼度を表している。対応する座標の Σ_t の

値が大きければ、その座標における T_t の濃淡値の信頼性は低いと考えられる。 Σ_t もまた、 F_t を観測することによって更新していく。

さて、観測画像上で使用者が座標 $P_t = (i_t, j_t)$ を指示したとしよう。指示直後の T_t は、 P_t を中心としたある一定の領域と設定する。 B_t はその領域をあらわすように設定する。 Σ_t は適当な定数としておく。

この3つのパラメータを用いることで、テンプレートの信頼性を考慮しながら、テンプレートの大きさや形を変化させることができる。

3.2 提案手法の概要

処理の流れは予測過程と推定過程からなる(図2)。ここでは時刻 $k-1$ における各パラメータを $B_{k-1|k-1}$, $T_{k-1|k-1}$, $\Sigma_{k-1|k-1}$ と表記する。

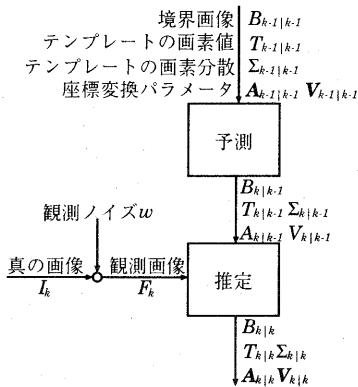


図2 提案手法の処理の流れ

図中の $A_{k-1|k-1}$ は対象の拡大縮小・回転などの変形、 $V_{k-1|k-1}$ は対象の並進を幾何学的にあらわす行列で、それぞれ

$$A_{k-1|k-1} = \begin{pmatrix} a_{k-1|k-1} & b_{k-1|k-1} \\ c_{k-1|k-1} & d_{k-1|k-1} \end{pmatrix}, V_{k-1|k-1} = \begin{pmatrix} u_{k-1|k-1} \\ v_{k-1|k-1} \end{pmatrix}$$

である。座標変換前の座標を (i, j) とすると、座標変換後の座標 (i', j') は

$$\begin{pmatrix} i' \\ j' \end{pmatrix} = \begin{pmatrix} a_{k-1|k-1} & b_{k-1|k-1} \\ c_{k-1|k-1} & d_{k-1|k-1} \end{pmatrix} \begin{pmatrix} i \\ j \end{pmatrix} + \begin{pmatrix} u_{k-1|k-1} \\ v_{k-1|k-1} \end{pmatrix} \quad (2)$$

となる。この $a_{k-1|k-1}, b_{k-1|k-1}, c_{k-1|k-1}, d_{k-1|k-1}$ および $u_{k-1|k-1}, v_{k-1|k-1}$ を求めることで、対象の変形や運動を知ることができる。対象の追跡は $A_{k-1|k-1}, V_{k-1|k-1}$ から、図2の予測と推定の過程により、 $A_{k|k}, V_{k|k}$ を求めることで可能となる。

「予測」の処理は、時間がひとつ進んだ各パラメータを求める処理である。時間による更新であるので、「時間更新」とも呼ぶことにする。時間更新により、各パラメータの状態は $k-1|k-1$ から $k|k-1$ となる。

「推定」の処理は、時刻 k で観測された観測画像 F_k を用い

て各パラメータを更新する処理である。観測画像による更新であるので、「画像更新」とも呼ぶことにする。 F_k は真の画像 I_k に平均0で分散 σ_w^2 のガウス性白色ノイズ w が加わったものとしてモデル化する。これにより各パラメータを数理統計的根拠に基づいて更新することができる。画像更新により、パラメータの状態は $k|k-1$ から $k|k$ となる。すなわち時刻 k における各パラメータが求められることになる。

予測と推定の過程によって $B_{k-1|k-1}, T_{k-1|k-1}, \Sigma_{k-1|k-1}$ から、 $B_{k|k}, T_{k|k}, \Sigma_{k|k}$ を求める過程で $A_{k|k}, V_{k|k}$ もまた求められる。 $A_{k|k}, V_{k|k}$ から対象の動きを知ることができるので、対象追跡が実現される。

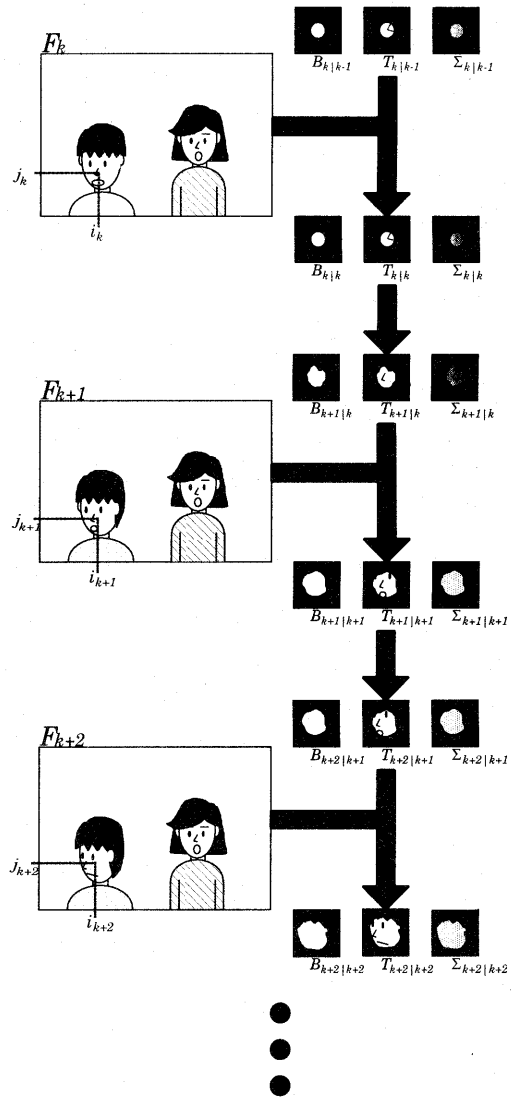


図3 領域変化のイメージ

時刻 k における境界画像、テンプレート、分散画像が変化していくイメージは図3の通りである。時刻 $k-1$ における状態から時刻 k における状態への更新処理の詳細を以下に説明す

る。4章は予測過程(時間更新)の説明であり、5章は推定過程(画像更新)の説明である。

4. 予測過程

$A_{k-1|k-1}$ および $V_{k-1|k-1}$ を用いて $B_{k-1|k-1}$, $T_{k-1|k-1}$, $\Sigma_{k-1|k-1}$ を座標変換することで $B_{k|k-1}$, $T_{k|k-1}$, $\Sigma_{k|k-1}$ が予測される。位置 (i, j) での $B_{k-1|k-1}$, $T_{k-1|k-1}$, $\Sigma_{k-1|k-1}$ の値は、位置 (i', j') における $B_{k|k-1}$, $T_{k|k-1}$, $\Sigma_{k|k-1}$ の値となる。

$$B_{k|k-1}(i', j') = B_{k-1|k-1}(i, j) \quad (3)$$

$$T_{k|k-1}(i', j') = T_{k-1|k-1}(i, j) \quad (4)$$

$$\Sigma_{k|k-1}(i', j') = \Sigma_{k-1|k-1}(i, j) \quad (5)$$

これを $0 \leq i' < W, 0 \leq j' < W$ についてすべて求めることで、各画像を予測できる。

ここで (i, j) が整数ではないときは、4つの近隣画素から線形に内挿して位置 (i, j) の画素値を計算する。

また、対象は慣性運動をすると仮定しているため、 $A_{k|k-1}$, $V_{k|k-1}$ の予測式は $A_{k|k-1} = A_{k-1|k-1}$, $V_{k|k-1} = V_{k-1|k-1}$ となる。

5. 推定過程

予測過程では、対象は慣性運動をし、 $B_{k-1|k-1}$, $T_{k-1|k-1}$, $\Sigma_{k-1|k-1}$ は $A_{k-1|k-1}$, $V_{k-1|k-1}$ によりそのまま変換されると仮定していたが、常にこの仮定が成り立つわけではない。そこで観測画像 F_k を用いて各パラメータの修正をおこなう。

5.1 テンプレートと分散画像の推定

テンプレート $T_{k|k-1}$ と分散画像 $\Sigma_{k|k-1}$ の更新は、ベイズ的手法でおこなう。ここで簡単のため、位置 (i, j) における真の画像 I_k の値を i_k , 観測画像 F_k の値を f_k , $B_{k|k-1}$ の値を $b_{k|k-1}$, $T_{k|k-1}$ の値を $t_{k|k-1}$, $\Sigma_{k|k-1}$ の値を $\sigma_{k|k-1}^2$ とする。

テンプレートの画素分布を平均 $t_{k|k-1}$, 分散 $\sigma_{k|k-1}^2$ のガウス分布と仮定する。さらに画像の観測ノイズ w は平均0, 分散 σ_w^2 のガウス性白色雑音であると仮定することで、 i_k の分布は平均 f_k , 分散 σ_w^2 のガウス性とみなせる。

テンプレートの画素分布 $p(x|t_{k|k-1})$ は、ガウス性の仮定から

$$p(x|t_{k|k-1}) = \frac{1}{\sqrt{2\pi\sigma_{k|k-1}^2}} \exp\left[-\frac{(x-t_{k|k-1})^2}{2\sigma_{k|k-1}^2}\right] \quad (6)$$

となる。さらに観測画像に関する確率密度 $p(f_k|x, t_{k|k-1})$ は、同じくガウス性から

$$p(f_k|x, t_{k|k-1}) = \frac{1}{\sqrt{2\pi\sigma_w^2}} \exp\left[-\frac{(x-f_k)^2}{2\sigma_w^2}\right] \quad (7)$$

となる。ここでベイズの定理を適用することで、 f_k が観測されたときの事後確率密度 $p(x|f_k, t_{k|k-1})$ は

$$p(x|f_k, t_{k|k-1}) = \frac{p(x|t_{k|k-1})p(f_k|x, t_{k|k-1})}{p(f_k|t_{k|k-1})} \quad (8)$$

$$= \frac{1}{\sqrt{2\pi\sigma_{k|k}^2}} \exp\left[-\frac{(x-t_{k|k})^2}{2\sigma_{k|k}^2}\right] \quad (9)$$

となる。よって $p(x|f_k, t_{k|k-1})$ は平均 $t_{k|k}$, 分散 $\sigma_{k|k}^2$ のガウス分布であり、それぞれ

$$t_{k|k} = \frac{\sigma_w^2 t_{k|k-1} + \sigma_{k|k-1}^2 f_k}{\sigma_w^2 + \sigma_{k|k-1}^2} \quad (10)$$

$$\sigma_{k|k}^2 = \frac{\sigma_w^2 \sigma_{k|k-1}^2}{\sigma_w^2 + \sigma_{k|k-1}^2} \quad (11)$$

である[2]。以上より、式(10)、(11)の値を位置 (i, j) における $T_{k|k}(i, j)$, $\Sigma_{k|k}(i, j)$ の推定値とすればよいことがわかる。

5.2 座標変換に関する行列と境界画像の推定

座標変換に関する行列 $A_{k|k}$ および $V_{k|k}$ は、 $T_{k|k-1}$ を座標変換したものを、観測画像 F_k にできるだけ近くなるように決めるべきである。そこで $T_{k|k}$ と F_k の類似度に関する評価式を考える。

$T_{k|k-1}(i, j)$ の画素値を $t_{k|k-1}^{(i, j)}$ とする。これを $A_{k|k}$, $V_{k|k}$ によって座標変換したものを $T_{k|k-1}(\tilde{i}, \tilde{j})$ とし、画素値を $t_{k|k-1}^{(\tilde{i}, \tilde{j})}$ とあわらすとしよう。同様に $\Sigma_{k|k-1}(i, j)$ と $\Sigma_{k|k-1}(\tilde{i}, \tilde{j})$ を $\sigma_{k|k-1}^{2(i, j)}$, $\sigma_{k|k-1}^{2(\tilde{i}, \tilde{j})}$ とする。式(9)の x に $t_{k|k-1}^{(\tilde{i}, \tilde{j})}$ を代入し、 $\delta_{i, j} = f_k - t_{k|k-1}^{(\tilde{i}, \tilde{j})}$ とおくと、

$$p(f_k|t_{k|k-1}^{(\tilde{i}, \tilde{j})}|f_k, t_{k|k-1}) = \frac{1}{\sqrt{2\pi\sigma_{k|k-1}^{2(\tilde{i}, \tilde{j})}}} \exp\left[-\frac{\delta_{i, j}^2}{2\sigma_{k|k-1}^{2(\tilde{i}, \tilde{j})}}\right] \quad (12)$$

となる。この式(12)を最大化すればよい。ここで $1/\sqrt{2\pi\sigma_{k|k-1}^{2(\tilde{i}, \tilde{j})}}$ は f_k と $t_{k|k-1}^{(\tilde{i}, \tilde{j})}$ の関数ではないので無視できる。さらに T_k の全画素に関して

$$L = \prod_{i=0}^{W-1} \prod_{j=0}^{W-1} \exp\left[-\frac{\delta_{i, j}^2}{2\Sigma_{k|k}(\tilde{i}, \tilde{j})}\right] \\ = \exp\left[-\frac{1}{2} \sum_{i=0}^{W-1} \sum_{j=0}^{W-1} \frac{\delta_{i, j}^2}{\Sigma_{k|k}(\tilde{i}, \tilde{j})}\right] \quad (13)$$

を求める。この L が最大になるとき、つまり

$$l = \sum_{i=0}^{W-1} \sum_{j=0}^{W-1} \frac{\delta_{i, j}^2}{\Sigma_{k|k}(\tilde{i}, \tilde{j})} \quad (14)$$

が最小になるとき、 $T_{k|k}$ と $F_{k|k}$ が最も類似していると考えられる。

l に境界画像 $B_{k|k}$ に関する条件と、 $A_{k|k}$, $V_{k|k}$ の変化幅に関する項

$$A = \frac{(a_{k|k} - a_{k|k-1})^2}{\sigma_a^2} + \frac{(b_{k|k} - b_{k|k-1})^2}{\sigma_b^2} \\ + \frac{(c_{k|k} - c_{k|k-1})^2}{\sigma_c^2} + \frac{(d_{k|k} - d_{k|k-1})^2}{\sigma_d^2}$$

$$V = \frac{(u_{k|k} - u_{k|k-1})^2}{\sigma_u^2} + \frac{(v_{k|k} - v_{k|k-1})^2}{\sigma_v^2}$$

を加えることで次の評価式を得る。

$$S = \sum_{i=0}^{N-1} \sum_{M=0}^{W-1} \frac{\delta_{i, j}^2}{\Sigma_{k|k}(\tilde{i}, \tilde{j})} \times B_{k|k}(i, j) + A + V \quad (15)$$

この式が最小になるように $A_{k|k}$, $V_{k|k}$ を選ぶことで、

$B_{k|k}, T_{k|k}, \Sigma_{k|k}, A_{k|k}, V_{k|k}$ が推定できる。ここで、 $\sigma_a^2, \dots, \sigma_d^2$ および σ_u^2, σ_v^2 は座標変換に関するパラメータの変動の大きさを表現する量である。

この処理に次に説明する境界画像を拡大するための工夫をすることで、境界領域を変化させていくことができる。

5.3 境界画像を拡大する処理

領域拡大の処理は多くの課題を抱えており、そのため難しい問題となっている。今回用いた領域拡大の手法は暫定的なものであり、まだ改良の余地がある。

境界画像は対象領域であるかどうかを示す度合に関する情報である。境界画像を変化させるためには、観測画像の対象と背景の領域分割をする必要がある。

対象は動いていると考える。それにより、対象領域内である条件として以下の2つを用いることができる。

- 対象の動きに関する条件
対象領域であれば連続したフレーム間の同じ位置の濃淡値は異なる
- 対象の濃淡値に関する条件
対象領域内では隣接する画素間の濃淡値の変化はゆるやかである

これらの条件を元に対象の領域と背景の領域を切り分ける。状態 $k|k-1$ の境界画像から $k|k$ の境界画像へと更新させていく。以下、表記の簡略のために状態 $k|k-1$ を $k-1$ とし、状態 $k|k$ を k とあらわすことにする。

ここでは画素の位置をベクトル $p = (m, n)^T$ ($0 \leq m < W, 0 \leq n < W$) であらわすことにする。そして方向ベクトル r_i ($i=0, \dots, 8$) を定義する。 $r_0=(0,0)^T$, $r_1=(1,1)^T$, $r_2=(1,0)^T$, $r_3=(1,-1)^T$, $r_4=(0,-1)^T$, $r_5=(-1,-1)^T$, $r_6=(-1,0)^T$, $r_7=(-1,1)^T$, $r_8=(0,1)^T$ である。これらを用いて p とその8近傍は $p+r_i$ ($i=0, 1, \dots, 8$) としてあらわされ、位置関係は図4のようになる。ここで $p=p+r_0$ である。

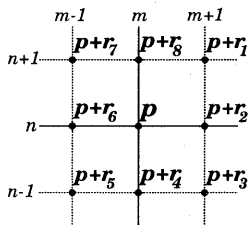


図4 p と r_i の位置関係

観測画像のサイズが $M \times N$ であるのに対し、境界画像のサイズは $W \times W$ である。定式の簡略化のため、図5のように観測画像の座標を変換する。これにより境界画像の位置 p に対応する観測画像の位置を同じベクトル p を使って表現することができる。

5.3.1 対象の動きに関する処理

境界画像 B_k を求めるため、時刻 $k-1$ における観測画像 F_{k-1}

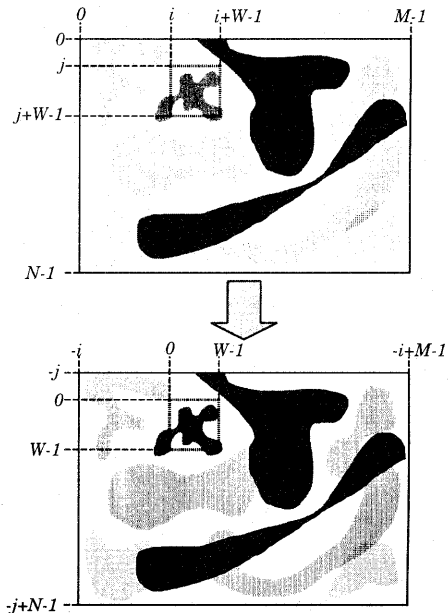


図5 座標変換

と時刻 k における観測画像 F_k を用いる。 $p+r_i$ のうち、次の条件を満たすものの集合を P_p とする。

$$P_p = \{p+r_i | \forall i=0, \dots, 8; |F_{k-1}(p+r_i) - F_k(p+r_i)| > C_V\} \quad (16)$$

ここで C_V は動きの量に関するしきい値である。 P_p の要素数を $n(P_p)$ とあらわすとしよう。 $n(P_p)$ は $p+r_i$ のうち、差がある位置の数であり、最大値は9、最小値は0となる。この値が大きければ p が対象領域である確率が高くなり、反対に小さければ確率が低くなる。そこで

$$V_p = \frac{n(P_p)}{9} \quad (17)$$

を B_k の全画素について求めることで、対象領域候補を決定することができる。

5.3.2 濃淡値の変化に関する処理

B_k の値に関するしきい値 C_B を考え、 $B_k > C_B$ の領域内の p について $F_k(p) - F_k(p+r_i)$ ($i=1, \dots, 8$) の値が小さいとき、同じ対象の領域であると考えられる。そこで次のような集合 Q_p, R_p を以下のように定義する。

$$Q_p = \{p+r_i | \forall i=1, \dots, 8; B_{k-1}(p+r_i) > C_B\} \quad (18)$$

$$R_p = \left\{ p+r_i | \forall i=1, \dots, 8; |F_k(p) - F_k(p+r_i)| < \frac{C_C}{8} \right\} \quad (19)$$

ここで C_C は画像の階調に関する定数である。時刻 $k-1$ で対象領域の近傍であり、時刻 k で濃淡値の変化がなめらかであれば、位置 p は時刻 k で対象領域候補となる。 p における濃淡評価式を

$$G_p = \sum_{p+r_i \in Q_p \cap R_p} |F_k(p) - F_k(p+r_i)| + C_n(Q) + 8C_n(R_p) \quad (20)$$

とし、この値が $8C_c$ 以下のときに濃淡値の変化がなめらかと
考え、対象領域候補とする。

5.3.3 境界画像の変化

式(17)および(20)を合わせることで対象領域を求めること
ができる。よって時刻 $t+1$ の境界画像の位置 p の値を

$$B_{t+1}(p) = \begin{cases} B_t(p) & (n(Pp)=0) \\ 0 & (n(Pp) \neq 0 \text{ and } Gp \geq 8C) \\ \frac{8C-Gp}{8C} \times \frac{n(Pp)}{9} & (\text{otherwise}) \end{cases} \quad (21)$$

によって求める。

6. 計算機シミュレーション

6.1 入力画像

静止した背景の中を球が回転しながら並進運動をするような
全 50 フレームの連続画像を生成するプログラムを組み、それ
から得られる画像を入力画像とした。サイズは 192×128 [pixel]、
球の半径は 30 である。1 フレームめと 25 フレームめを図 6 に
示す。

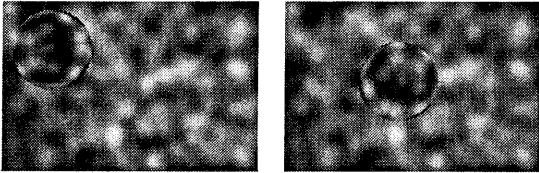


図 6 入力画像 (左:1 フレームめ 右:25 フレームめ)

6.2 領域変化の様子

領域の変化の様子を確認するため、 $A_{k|k}$ 、 $V_{k|k}$ を既知で実
験をおこなった。境界画像の初期領域は、座標 (35, 35) を中
心とした半径 15 の領域とした。また境界画像の初期値は座標
(35, 35) からの距離を r としたとき、位置 (i, j) における値は

$$B_{0|1}(i, j) = \begin{cases} 0.5 + 0.5 \cos\left(\frac{r}{15}\pi\right) & (r \leq 15) \\ 0 & (\text{otherwise}) \end{cases}$$

とした。1 フレームめから 10 フレームめまでの結果を図 7 に
示す。

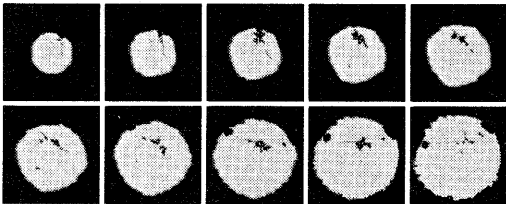


図 7 境界画像の変化 (左上が 1 フレームめ 右下が 10 フレームめ)

最初は小さな円であったものが次第に大きくなっていき、10
フレームめくらいで対象とほぼ同等の大きさとなった。領域の
拡大は 10 フレームめくらいで停止した。図 7 によって切出さ
れたテンプレートは図 8 の通りである。これを見ると、対象の
動きに合わせてテンプレートが変化していることがわかる。

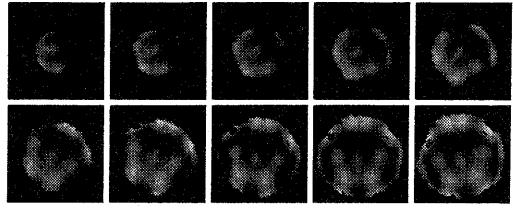


図 8 切出されたテンプレート (左上が 1 フレームめ 右下が 10 フレームめ)

7. まとめ

今回の報告では、対象領域であるか否かの度合に関する情報
 $B_{k|k}$ 、テンプレート $T_{k|k}$ 、テンプレートの信頼度に関する情報
 $\Sigma_{k|k}$ を内部変数として保持し、予測と推定の過程によって更新
していく対象追跡法を提案した。これにより従来のテンプレ
ートマッチングよりもロバストな対象追跡が実現できると期待で
きる。

現時点では $B_{k-1|k}$ の更新法に改良の余地がある。さらに今
後の課題として、リアルタイム処理に向けてこのアルゴリズム
を DSP (Digital Signal Processor) に実装し、補聴システムを慣
性させることが挙げられる。

文 献

- [1] 林 健太郎, 橋本 学, 鷲見 和彦, “頑強性と精緻性を備えた顔
特頂点追跡による顔方向推定,” 信学論 D-II Vol. J84-D-II No. 8
pp. 1762-1771 Aug. 2001.
- [2] 有本 卓, “カルマン・フィルター,” 産業図書, 1977.