

大局解抽出と局所解追跡の協調による心臓動画像認識

Recognition of Echocardiogram by Cooperation of Global Extraction and Local Tracking

山田博三
H. YAMADA

In model-based dynamic image tracking, in order to stabilize the results, global extraction for initial frame or key frames and local tracking for successive frames are combined. Elastic Edge Sequence Matching method is proposed where the constraint by a window for local tracking and rough-to-fine mechanism for high speed computation are taken into account in a unified manner by introducing each coordinate system for each evaluation point into a conventional dynamic programming matching method.

For the problem of time difference between local high-speed tracking and time-consuming global search, we introduce time delay for the input image of tracking in order not to make the control of the system complex. It can be called delayed real time processing. Some experimental results for echocardiograms are given.

We extend the model to have multiple contours where the evaluation between contours is performed hierarchically, and applied to inner- and outer-contour of the heart in order to obtain more stable results.

§1 はじめに

動画像解析において、追跡の対象を見失うことが不安定性の要因となる。本論文では、モデルマッチングを用いた追跡において、動的計画法を用いることにより、局所解による追跡の他に、初期画像・キー画像に対する大局解抽出により、安定化を図る。

ここで、従来提案されてきた手法について考察してみる。まず、手法は剛体モデルと弾性体モデルに分けられる。剛体モデルとしては、テンプレートマッチングの他、Hough 変換²⁾があげられるが、これらは、マッチングが剛体であることの限界を持つ。一方、弾性体モデルとしては、Snakes^{1),4),5),10)}や弛緩法^{6),17)}のように局所最適解を探索する局所弾性体モデルがある。これらは追跡とか局所解の探索には適しているが、大局解の探索は不可能である。一方、大局解を抽出する弾性モデルマッチング手法としては、動的計画法を用いた RS マッチングがあげられる。オリジナルな RS マッチング^{7),9)}は 2 値画像を対象としていたが、方向性の導入、モデル線

分の方向の任意化、巡回性の評価等の改良^{11),12)}により、幅広い適用が可能となった。しかし、問題点としては、計算量が多いことが、ひとつのネックとなっている。MAP 整合法^{13),14)}では、入力表現に複数の方向面表現を用いることにより角度差の計算を単なる画像間の加算に置き換えると共に、モデル線分に沿った計算を評価点上の計算にすることにより、画素単位の並列計算を可能にした。しかし、画像全体に関する並列演算の導入は、ある限られた窓に関する計算や、粗密構造導入による計算をむしろ困難にしている。

本論文では、これらの方法の持つ大局的最適性を基本としつつ、各モデル評価点毎の位置拘束ウインドウ座標系の導入により、局所解の抽出、大局解の粗密化による高速化を統一的に扱う手法を提案する。この方法は、RS マッチングに比較して、計算量の削減、制御構造の簡素化、MAP 整合法に比較して、データ量の削減、粗密化・局所化の容易化、という特長を持つ。

次に、限られた画像に対する大局的な探索処理と、継続的な画像に対する局所的な追跡処理を同じ時間軸上に

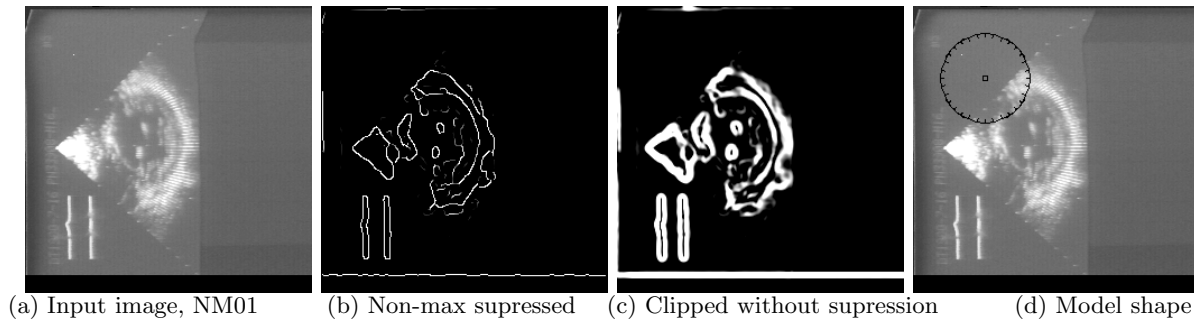


Fig. 1 Input Image, Edge Feature, and Shape Model.

並べてみると、一般に前者の方が処理時間がかかり、大局解が求められた時には既に時間は先に進んでいるという問題が生じる。この問題に対して、処理・制御形態を複雑にしないために、ここでは、追跡処理に入力画像の遅延を与える。即ち処理時間のかかる大局解抽出は入力画像に対して直ちに処理を開始するが、局所解追跡は常にある時間だけ遅延した画像に対して処理を行ない、両者の終了時刻を揃える。いわば、遅延型リアルタイム処理である。このような形態では、真の即時応答は不可能であるが、計測など遅延が許される場合は、キャッチアップ処理などの複雑な制御が不要となる。なお、本論文の実験は、このようなリアルタイム処理でもなく、その考え方を模擬したものである。

§ 2 弾性エッジ列整合法

輝度または濃度入力画像 (Fig.(a)) を b とする。

$$b(z) = b(i, j), \quad 0 \leq i < I, \quad 0 \leq j < J. \quad (1)$$

ここで、 I, J は画像サイズ、 x - y 座標系 $z = (i, j)$ はそれぞれ右方向、下方向を正とする。方向は右 (x 軸) 方向から時計回りで定義する。

次に、標準偏差 σ のガウシアンオペレータ G_σ をかけた後の勾配画像 (ベクトル表現) $g(z)$ を得る。

$$\begin{aligned} g(z) &= (g_X(z), g_Y(z)) = -\nabla G_\sigma b(z) \\ &= \left(-\frac{\partial G_\sigma}{\partial x} b(z), -\frac{\partial G_\sigma}{\partial y} b(z) \right). \end{aligned} \quad (2)$$

式に示すように、実際には、ガウス分布を x 方向に偏微分したフィルタ $\frac{\partial G_\sigma}{\partial x}$ と、 y 方向に偏微分したフィルタ $\frac{\partial G_\sigma}{\partial y}$ を原画像にかけることにより抽出する。ここで、勾配の向きが逆なのは、値の大きい方から小さい方へ (通常は物体の内部から外部へ) 方向付けるためである。

この勾配画像から、エッジ強度 $|u|^{(0)}$ とエッジ方向 $\angle u^{(0)}$ を要素として持つエッジ画像 $u^{(0)}$ を得る。

$$\begin{aligned} u^{(0)}(z) &= (|u|^{(0)}(z), \angle u^{(0)}(z)), \\ |u|^{(0)}(z) &= \begin{cases} 0, & \text{if } |g(z)| < \left| g\left(z - \frac{g(z)}{|g(z)|}\right) \right| \\ & \text{or } |g(z)| < \left| g\left(z + \frac{g(z)}{|g(z)|}\right) \right|, \\ 0, & \text{if } |g(z)| < \Theta_L \text{ or } \Theta_L = \Theta_U = |g(z)| = 0, \\ 1, & \text{if } |g(z)| > \Theta_U \text{ or } \Theta_L = \Theta_U = |g(z)| \neq 0, \\ \frac{|g(z)| - \Theta_L}{\Theta_U - \Theta_L}, & \text{otherwise,} \end{cases} \quad (3) \\ \angle u^{(0)}(z) &= \arctan\left(\frac{g_Y(z)}{g_X(z)}\right), \quad 0 \leq \angle u^{(0)} < 2\pi. \end{aligned}$$

式 (3) の強度 $|u|^{(0)}$ の計算において、右辺第 1 項は、法線方向に沿った非最大抑制³⁾ (Fig.(b)) であり、 $\frac{g(z)}{|g(z)|}$ は法線方向の隣接ピクセルである。右辺第 2~4 項は下限 $\Theta_L = EX(|g(z)|)$ 、上限 $\Theta_U = EX(|g(z)|) + SD(|g(z)|)$ 、によるクリッピングである。ここで EX および SD はそれぞれ平均と標準偏差である。Fig.(c) に非最大抑制なしのクリッピングを図示する。エッジ画像 $u^{(0)}$ は、Fig.(b) の中心線の上に Fig.(c) の濃度を持つ画像である。

エッジ画像 $u^{(0)}$ は反復処理によりぼかされる。

$$u^{(\tau')}(z) = u^{(\tau'-1)}\left(z + \arg \max_{z' \in \{(0 \pm 1, 0 \pm 1)\}} \left\{ |u|^{(\tau'-1)}(z+z') \right\}\right). \quad (4)$$

ここで “arg max” は以降のカッコ内が最大値をとる z' (は 3×3 近傍内) の値である。即ち 3×3 近傍内で強度が最大であるエッジのベクトル値をその点に持ってくる。この処理を τ 回行なうことにより以降のマッチングで用いる画像 U を得る。

$$U = u^{(\tau)}(z) = (|u|^{(\tau)}(z), \angle u^{(\tau)}(z)). \quad (5)$$

次に形のモデル R (Fig.(d)) は、輪郭上の評価点の系列として表され、各々の評価点は 5 つの属性を持つ。

$$\begin{aligned} R &= \{(a_m, x_m, y_m, |p|_m, d_m) \\ & \quad | m = 1 (=m_S), \dots, M (=m_E)\}. \end{aligned} \quad (6)$$

点列は対象物を進行左に見るように順序付けられている。 (x_m, y_m) は参照点 (X_R, Y_R) からの相対位置、 d_m はエッジ方向、 $|p|_m$ はエッジ強度 (実験では用いない) である。 $a_m < 0$ は、その評価点がネガであることを示す。これは、例えば、入力 “E” に対するモデル “F” のように、その場所に特徴があることによって評価を下げる時に用いられる。 $a_m > 0$ はポジであることを示し、 a_m は許容変形範囲を指定する。即ち、 $m-1$ から m への相対位置が、 A_{a_m} で指定される範囲内で変形可能であることを意味する。ここで A_{a_m} は次のように定義される。

$$\begin{aligned} A_1 &= A_{-1} = \{(0, 0)\}, \\ A_5 &= \{(0, 0), (\pm 1, 0), (0, \pm 1)\}, \\ A_9 &= \{(0, 0), (\pm 1, 0), (0, \pm 1), (\pm 1, \pm 1)\}, \\ A_{04} &= \{(0, 0), (1, 0), (-1, 0)\}, \\ A_{15} &= \{(0, 0), (1, 1), (-1, -1)\}, \\ A_{26} &= \{(0, 0), (0, 1), (0, -1)\}, \\ A_{37} &= \{(0, 0), (-1, 1), (1, -1)\}. \end{aligned} \quad (7)$$

前半の 3 個は等方的近傍であり、後半の 4 個は方向性を持つ近傍である。本実験では a_m は常に 9 である。 κ は近傍までの長さを変える係数であり、 $\kappa \cdot A_9$ は、 $\{(0, 0), (\pm \kappa, 0), (0, \pm \kappa), (\pm \kappa, \pm \kappa)\}$ である。

マッチングの目的は、モデル R と入力 U の間に定義された評価を許容されるモデル形状の変形の範囲内で最

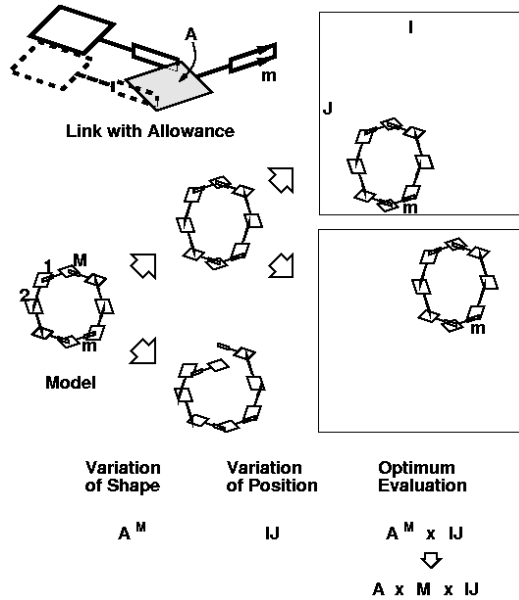


Fig. 2 Conceptually, a shape model is a chain of pieces each of which has rings at both ends by which the pieces link each other. The shape of the whole chain changes within the allowance of the rings while the direction of each piece is fixed. The generated shape is put onto an arbitrary location of an input image and an evaluation of that case is measured.

適化することである (Fig.2)。モデルと入力との全体の評価は次のように表される

$$E(R, U) = \max_{\substack{\{(i_m, j_m) \mid m=1, 2, \dots, M\} \\ (i_m - i_{m-1}, j_m - j_{m-1}) \in \kappa A_{a_m} \\ -(x_m - x_{m-1}, y_m - y_{m-1}) \in \kappa A_{a_m}}} \left\{ \sum_{m=1}^M \frac{e(m; i_m, j_m)}{M_N} \right\}. \quad (8)$$

ここで (i_m, j_m) は、 m 番目の評価点に対応する入力の位置である。 M_N は正規化係数でありポジ評価点の数である。 m 番目の評価点 (i, j) に対応する時の評価量 $e(m; i, j)$ は次のように定義される。

$$e(m; i, j) = \text{sign}(a_m) \cdot |u|^{(\tau)}(i, j) \times \max \left\{ 0, \left(1 - \frac{|\angle u^{(\tau)}(i, j) \ominus d_m|}{(\pi/2)} \right) \right\}. \quad (9)$$

ここで “sign” は a_m の符号、 \ominus は方向の周期性を考慮した角度の差であり $-\pi$ と π の間の値をとる。即ち引き続き 2 点間が許容範囲で変形するあらゆる組合せの中で e による評価が全体として最適な形状を抽出するのがこの目的である。

各評価点は許容変形によって相対的に拘束されるばかりでなく、窓 W により絶対位置の拘束を受ける。窓の指定には 2 種類ある。

窓内サーチの場合：

$$\begin{aligned} (i_m, j_m) &= (X_R, Y_R) + (x_m, y_m) + \kappa \cdot (k, l), \\ W &= \{(k, l) \mid \max(|k|, |l|) \leq W_R\}. \end{aligned} \quad (10)$$

全面サーチの場合：

$$\begin{aligned} (i_m, j_m) &= (x_m, y_m) + \kappa \cdot (k, l), \\ W &= \{(k, l) \mid 0 \leq \kappa \cdot k < I, 0 \leq \kappa \cdot l < J\}. \end{aligned} \quad (11)$$

上の関係式は、位置の拘束を指定することの他に、それぞれの m によってそれぞれ異なる座標系 (k, l) を導入している。ここで式 (10) の (k, l) は窓の中央からの変位を示し、式 (11) の (k, l) は (i, j) の粗な座標系を示す。標準化間隔 κ を大きくすると、計算する点がまばらになり計算量が下がる。また同時に、ある点から次の点への変形の許容範囲が増大するため大きな形の変形が許容されるようになる。

ここで m に関して 1 から m までの中間 (漸化) 的な評価関数を導入する。

$$f(m, i, j) = \max_{\substack{\{(i_{m'}, j_{m'}) \mid m'=1, 2, \dots, m\} \\ (i_{m'} - i_{m'-1}, j_{m'} - j_{m'-1}) \in \kappa A_{a_{m'}} \\ -(x_{m'} - x_{m'-1}, y_{m'} - y_{m'-1}) \in \kappa A_{a_{m'}} \\ (i_m, j_m) = (i, j)}} \left\{ \sum_{m'=1}^m e(m'; i_{m'}, j_{m'}) \right\}. \quad (12)$$

式 (8) の全体の評価関数との相違は、評価が m までに制限されていることと、最後の m の位置が (i, j) に拘束されていることである。分母 M_N は本質的ではない。

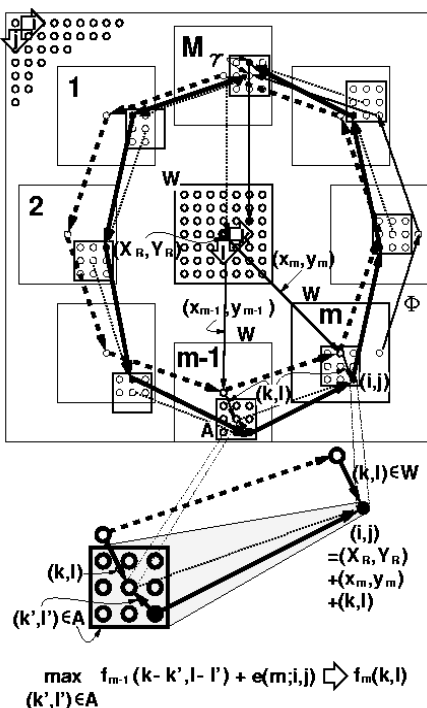


Fig. 3 EES Matching Method. The large-sized square is an input image. The m -th evaluation point is constrained by the window W (middle-sized square). The m -th stage of recurrence calculation is enlarged at the bottom. At the last stage of forward recurrence calculation, the preliminary evaluation Φ and the penalty for the gap γ are subtracted. Bold solid arrows show the extracted correspondence by the trace-back after the forward recurrence calculation.

ここで動的計画法⁸⁾の最適性の原理の適用により、問題は次のような漸化関係に分解される。

$$f(m, i, j) = \max_{\substack{(i', j') - (i', j') - (x_{m-1}, y_{m-1}) \\ + (x_{m-1}, y_{m-1}) \in A_{am}}} \{f(m-1, i', j') + e(m; i, j)\}. \quad (13)$$

ここで新しい関数 $f_m(k, l) = f(m, i, j)$ を導入すると、式 (13) は次のようになる。

$$f_m(k, l) = \max_{(k', l') \in A_{am}} \{f_{m-1}(k-k', l-l') + e(m; i_m, j_m)\}. \quad (14)$$

ここで (i, j) の代わりに (i_m, j_m) を用いたのは、 m 番目の評価点に関わる座標値であることをはっきりさせるためである。これが第 m 段階における基本的な漸化関係式である。式 (8) を満足する最終結果は、 $m = 1, 2, \dots, M$ に対する漸化式計算によって得られる。

しかし、この漸化式では、モデルの形状が閉じているかどうかの評価は行っていない。その評価のために、1 周以上 ($1.0 < \Phi \leq 2.0$ 周) の評価を行なう。この過程を表現するため、 m に上つき記号 (1) または (2) をつけて表す。 $m^{(1)}$ が予備的な 1 回目のサイクル、 $m^{(2)}$ がメイ

ンの 2 回目のサイクルである。従って $m^{(1)}$ の値自身は m と違いはない。

巡回性を考慮した漸化式は次のようになる (Fig.3)。

INITIAL SET: For $m = m_P^{(1)} = (2 - \Phi) \cdot M$,

$$f_m(k, l) = 0. \quad (15)$$

PRE-CYCLE: For $m = m_P^{(1)} + 1, m_P^{(1)} + 2, \dots, M^{(1)}$,

$$f_m(k, l) = \max_{(k', l') \in A_{am}} \{f_{m-1}(k-k', l-l') + e(m; i_m, j_m)\}. \quad (16)$$

END OF PRE-CYCLE: For $m = M^{(1)}$,

$$\begin{aligned} f_0(k, l) &= f_m(k, l), \\ s_0(k, l) &= (k, l). \end{aligned} \quad (17)$$

MAIN CYCLE: For $m = 1^{(2)}, 2^{(2)}, \dots, M^{(2)} - 1$,

$$\begin{aligned} f_m(k, l) &= \max_{(k', l') \in A_{am}} \{f_{m-1}(k-k', l-l') + e(m; i_m, j_m)\}, \\ c(k, l, m) &= \arg \max_{(k', l') \in A_{am}} \{f_{m-1}(k-k', l-l') + e(m; i_m, j_m)\}, \\ s_m(k, l) &= s_{m-1}((k, l) - c(k, l, m)). \end{aligned} \quad (18)$$

関数 $c(k, l, m)$ には決定 (k', l') が保持され、 m 番目の点 (k, l) に対応する時、 $m-1$ 番目の点が入力のどの点に対応していたかの情報を持つ。関数 $s_m(k, l)$ は、第 2 サイクルの最初の時点で、対応付けられた点の情報を m 段階で受け継いで次の段階へ受け渡す役割を持つ。漸化式計算の過程でネガ評価点に関しては減算が行なわれることになる。

END OF MAIN CYCLE: For $m = M^{(2)}$,

$$\begin{aligned} f_m(k, l) &= \max_{(k', l') \in A_{am}} \{f_{m-1}(k-k', l-l') \\ &\quad + e(m; i_m, j_m) - f_0(s_{m-1}(k-k', l-l')) \\ &\quad - \gamma \cdot \kappa \cdot |(k, l) - s_{m-1}(k-k', l-l')|\}, \\ c(k, l, m) &= \arg \max_{(k', l') \in A_{am}} \{f_{m-1}(k-k', l-l') + \dots\}. \end{aligned} \quad (19)$$

この最終段階で第 1 サイクルの最後と第 2 サイクルの最後の対応点のずれが係数 γ で評価されると共に、第 1 サイクルの分の評価量 f_0 が差し引かれる。

ここで、 $f_M(k, l)$ の最適値をとることにより、式 (8) で定義された値を得る。

BEST RESULTS: For $m = M$ and $\rho = 1, 2, \dots$,

$$\begin{aligned} E^{(\rho)}(R, U) &= \begin{cases} \max_{(k, l) \in W} \{f_M(k, l)/M_N\}, & \text{if } \rho = 1, \\ \max_{(k, l) \in W} \{f_M(k, l)/M_N\}, & \text{if } \rho \geq 2, \end{cases} \\ &\quad \forall \rho' < \rho \max_{(k, l) \in W} \{f_{M'}(k, l)/M_N\} > \Delta \\ (k_M^{(\rho)}, l_M^{(\rho)}) &= \arg \max_{(k, l)} \{f_M(k, l)/M_N\}. \end{aligned} \quad (20)$$

ここで対応する入力的位置 $(i_M^{(\rho)}, j_M^{(\rho)})$ は、式 (10) または (11) の $(k_M^{(\rho)}, l_M^{(\rho)})$ から、また、参照点 (X_R, Y_R) の対応点 $(i_R^{(\rho)}, j_R^{(\rho)})$ も $(k_M^{(\rho)}, l_M^{(\rho)})$ から得られる。ランク ρ が

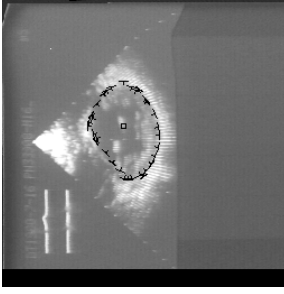


Fig. 4 Result for NM01.

1 より大、すなわち最良対応以外の結果を引続き求める時、半径 Ω の不応領域を導入し、上位の結果から離れた領域から次の結果をサーチする。

最終点以外の評価点に関する対応点は、マトリックス $c(k, l, m)$ をトレースバックすることにより得られる (Fig.4)。

TRACE-BACK: For $m = M, M-1, \dots, 3, 2$, and $\rho = 1, 2, \dots$,

$$\begin{aligned} & (k_{m-1}^{(\rho)}, l_{m-1}^{(\rho)}) \\ &= \begin{cases} (k_m^{(\rho)}, l_m^{(\rho)}) - c(k_m^{(\rho)}, l_m^{(\rho)}, m), & \text{if } a_m > 1, \\ (k_m^{(\rho)}, l_m^{(\rho)}), & \text{if } a_m \leq 1. \end{cases} \quad (21) \end{aligned}$$

この計算で最大の記憶容量を必要とするのは $c(k, l, m)$ であり、 $\|W\| \times M_F$ となる。 $\|W\|$ は窓 W の点の数、 $M_F = \|\{(a_m, \dots, \dots) \mid a_m \geq 2\}\|$ は弾性評価点の数である。個々の要素は9値をとる。 $f_0(k, l)$ を除く関数 $f_m(k, l)$ と $s_m(k, l)$ は、 m と $m-1$ でダブルバッファ化できる。 $e(\dots)$ の計算数でカウントする計算量は、 $EX(\|A_{a_m}\|) \times M \times \|W\| \times \Phi$ 、 $EX(\|A_{a_m}\|)$ は $\|A_{a_m}\|$ の平均、である。

§3 従来法との比較

3.1 動的計画法

RS マッチング^{11),12)} における漸化関係において、許容変形は、ベクトル $(x_m - x_{m-1}, y_m - y_{m-1})$ に沿った画素上に、例えば $\lambda = 2$ のときは長さ $1/2$ から 2 の範囲で、とられる。

$$\begin{aligned} f(m, i, j) &= \max_{(i', j') \in A(m, i, j)} \{f(m-1, i', j') \\ &\quad + e(m; i', j', i, j)\}, \quad (22) \\ A(m, i, j) &= \{(i', j') = (i, j) - \lambda' \cdot (x_m \\ &\quad - x_{m-1}, y_m - y_{m-1}) \mid 1/\lambda \leq \lambda' \leq \lambda\}. \end{aligned}$$

また、評価関数 e は、 (i', j') から (i, j) までの線上の画素から計算される。

$$\begin{aligned} & e(m; i_{m-1}, j_{m-1}, i_m, j_m) \\ &= \sum_{\substack{(i, j) = \lambda'' \cdot (i_{m-1}, j_{m-1}) \\ + (1 - \lambda'') \cdot (i_m, j_m) \mid 0 \leq \lambda'' < 1}} |u(i, j) \cdot \max\{0, (1 - \frac{|\angle u(i, j) \ominus d_m|}{(\pi/2)})\}|. \quad (23) \end{aligned}$$

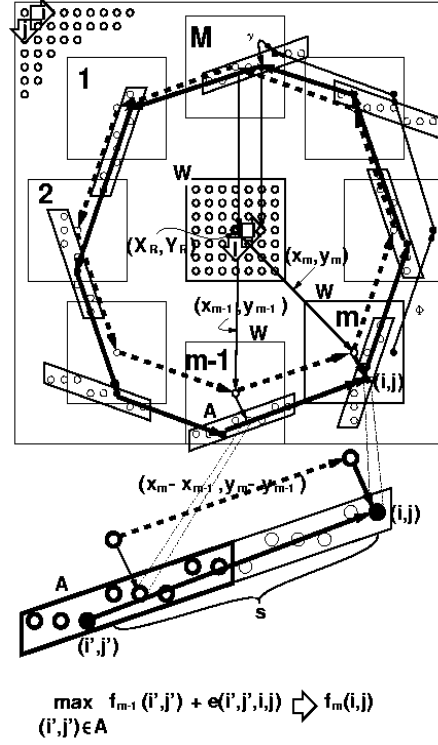


Fig. 5 RS Matching. Main difference from EES method is: (k, l) is not introduced, and A is set along a line segment.

これに対して、式 (13) または (14) においては、 (x_m, y_m) の周りの近傍 (即ち $\lambda' = 1$) がとられる。また、式 (9) の評価量 e は、1 点 (i_m, j_m) のみで計算される。

RS マッチングにおける接線に沿った変形は、精密ではあるが計算量が大きくなる。また、入力座標系による漸化関数計算をおこなっているため、粗密計算や位置拘束窓による計算の際の制御が複雑になる。

3.2 MAP 整合法

MAP 整合法^{13),14)} においては、入力特徴は、各々がその方向成分の強さを持つ複数の方向面で表現される。

$$\begin{aligned} U &= \{u_d(z) \mid d = 0, \pi/4, 2\pi/4, \dots, 7\pi/4\}, \quad (24) \\ u_d(z) &= \max\{0, |u(z) \cdot \cos(\angle u(z) - d)|\}. \end{aligned}$$

また、評価量 e は、いずれかの方向成分で近似される。

$$e(m; i, j) = u_{d_m}(i, j). \quad (25)$$

窓の指定が、 $\kappa = 1$ の全面マッチングの時、画像 $f_m(k, l)$ 全体を記号 F_m で表し、 $u_d(i, j)$ を U_d で表現すると、式 (14) は次のように書き直すことができる。

$$F_m = \max_{(k', l') \in A_{a_m}} \left\{ F_{m-1} [x_m - x_{m-1} + k', y_m - y_{m-1} + l'] + U_{d_m} \right\}. \quad (26)$$

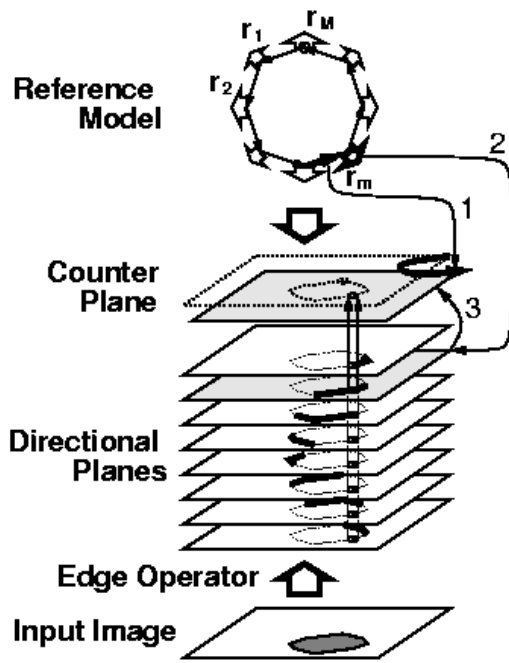


Fig. 6 MAP Matching Method. Recurrence calculation is executed in parallel by preparing multiple directional images for input feature.

ここで、 $(x_0, y_0) = (x_M, y_M)$ 、記法 $[z']$ は、画像全体を z' だけ平行移動(シフト)することを示す。即ち処理全体が、画像の平行移動、最大値、加算という画素単位の並列演算に適した演算のみで実現できる。

3.3 本手法の特長

以上、2つの手法に比べて、次のような特長を持つ。

(1) 座標系 (k, l) の考え方を整理することにより、大局解抽出と局所解追跡を統一的に実現できる。さらに、局所解追跡時に、計算量、トレースバック用メモリの節約が図られる。

(2) (i, j) 座標系と (k, l) 座標系の分離、および κ の導入により、入力画像を作り直す事なく、間引いた時の計算が可能になり、粗密計算が容易になった。従来、特にMAP 整合法では上の2点に不向きであった。

(3) MAP 整合法に比べて角度差計算が精密になった。MAP 整合法では、入力画像を方向別表現する事により、距離演算を和演算のみで行っていたが、本手法では、元に戻して角度差 \times 強度の計算を用いている。表現面を増やして演算を単純にすると、表現面を減らして演算でカバーするのがどちらが良いかは、ハードウェア環境による。本手法では、各評価点の存在範囲をウィンドウ W で制限した際に、トレースバック用メモリも節約される。

(4) 次節に示すように弾性近傍の制御をパラメータ化

することにより、適用に対する柔軟性が増した。ここで、本手法では近傍を大きくすることと、標準化ピッチ κ は連動していることに注意する必要がある。なお、全体の变形を小さくするには、数点おきに变形を許容するなどで制御できる。

§4 超音波心臓動画像による実験

4.1 実験

提案する手法の有効性を検証するため、超音波心臓像の追跡に適用した。実験には、標準画像データベース SIDBA、Vol.701 の3連の動画像の内、最初の1連(レコード番号2~37、データ名 NM01~NM36、僧帽弁先端付近短軸断層像)を用いた。これらの動画像から輪郭として顕著な左室心外膜付近を抽出を試みる。入力画像は、 256×256 各点8ビットの濃淡画像である。これに対して、 $\sigma = 2$ のガウシアンオペレータによるエッジ抽出の後、 $\tau = 1$ 回のぼかしを行なう。

処理の基本的な考え方は、次の3つのプロセスである。以下、モデルの巡回性 $\phi = 1.5$ 周、始点と終点の差の重み $\gamma = 1$ である。

(1) 36枚の6枚毎(1, 7, 13, 19, 25, 31)をキーとする大局的探索：大局解抽出のためのモデルは、 $M = 32$ 個の評価点による半径40の円である。大局解抽出は全面サーチとなり時間がかかるため、 $\kappa = 4 \rightarrow 2 \rightarrow 1$ の順に粗密制御を行ない、サーチの高速化を図る。

(2) 全てのフレームでの局所解による追跡：モデルは前フレームの抽出結果に対し、評価点を等間隔に並べ直したものである。絶対位置拘束窓 $W_r = 2$ 、マッチングの解像度 $\kappa = 4$ とした。

(3) 2つの結果の検定：キーフレームについて局所解と大局解の評価値の比較を行ない、その値の良好な方から以降の追跡を行なう。ここで、遅延型リアルタイム処理の考え方では、大局解を採用した場合でも、局所解の追跡を次のフレームに対して行なえばよい、という仮定がおかれているため、切替の制御が単純である。

4.2 結果

Fig.7(c) が抽出結果である。比較のため、全フレームに対する大局解抽出 (Fig.7(a))、初期フレーム以外の全フレームに対する局所解追跡 (Fig.7(b)) の結果を示す。

Fig.7(a) の大局解の抽出結果のうち、第1フレームと第19フレームの結果を Fig.8 に示す。第1フレームでは、ほぼ左室心外膜付近を正しく抽出しているが、第19フレームでは、一部が左室内膜輪郭に引っかかっている。この他に第28フレーム付近でも前後とは不連続な解を抽出しているが、これ以外では、前後と連続した

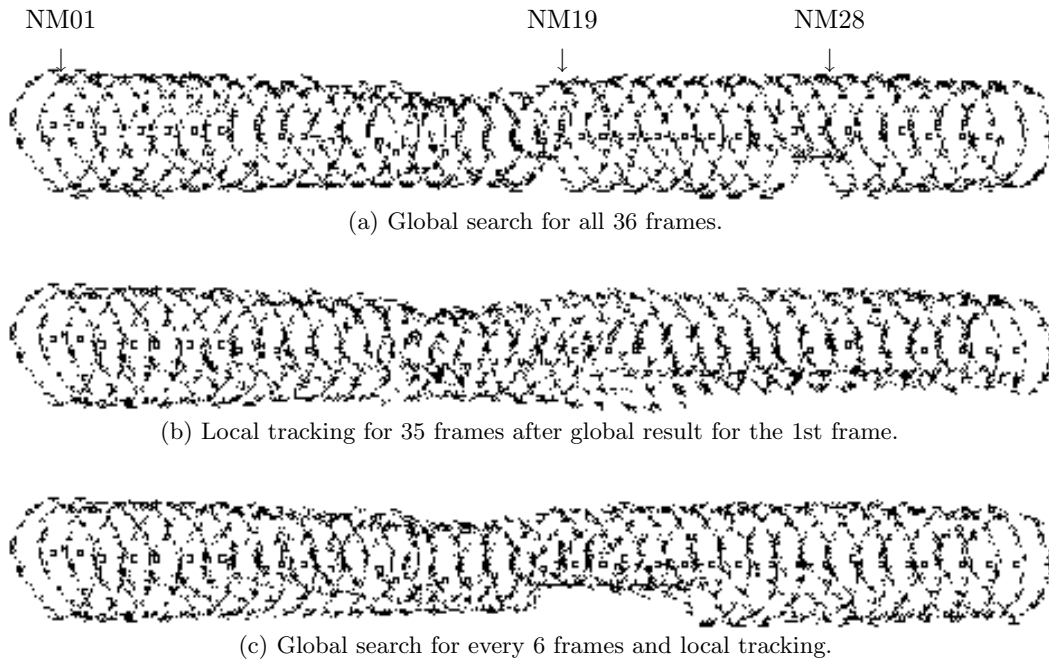
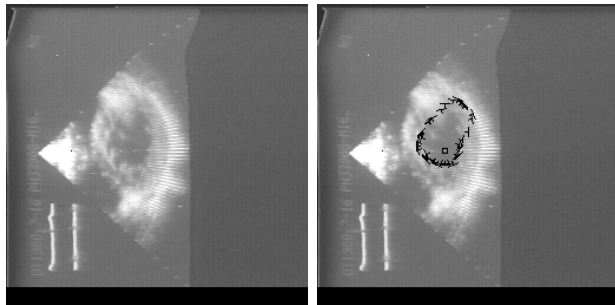


Fig. 7 Experimental Results.



(a) Original Image NM19 (b) Result for NM19

Fig. 8 Result by Global Search.

結果を適正な位置に抽出している。大局解抽出は前後とは独立に抽出するにも拘らずこのような結果が得られたことは、本手法が安定であることを示している。

一方、Fig.7(b)の局所解による追跡において、サーチの窓を比較的大きくとしたことと、対象の動きがある範囲に限られるため、局所解追跡のみでも一度外れかけた結果も元に戻る現象も起こっているようであるが、対象によっては、元に戻らない危険性がある。

Fig.7(c)の局所解と大局解の併用において、第(7, 13, 19, 25, 31)フレームでの検定のうち、局所解から大局解への変更が起こったのは第(7, 13, 19)フレーム、局所解がそのまま選ばれたのが第(25, 31)フレームであった。動きの激しい前半部では大局解への変更により安定性が保持され、動きの少ない後半部では局所解が選ばれ連続性が保持されていることが分かる。この中で第19フレームについては、前のフレームとは不連続的な大局解

が選ばれたために結果も不連続になっている。ただし、Fig.7(b)の局所解の追跡結果から分かるように、この付近では局所解の追跡のみでも収縮期の形状抽出に失敗しておりむしろ大局解に戻したことにより全体としての安定性は保たれていると言えよう。

4.3 大局解と局所解の時間差に関する考察

本実験におけるSS10(/M2)による処理時間は、1フレームあたり、エッジ抽出部8秒、その後の大局解抽出10秒、局所解追跡0.2秒であった。このように、限られた画像に対する大局的な探索処理と、継続的な画像に対する局所的な追跡処理を同じ時間軸上に並べてみると、どうしても前者の方が処理時間がかかる。

現状のハードウェア技術では、局所的追跡処理に対してリアルタイム性を仮定するのはリーズナブルになるつつあるが、大局解抽出は、当面かなり困難であろう。従って、リアルタイム処理において、大局解が求められた時には既に時間は先に進んでいるという状況が生じる。ここで、実時間性を重視すると、この遅れた時間をとり戻すためのキャッチアップ処理を施すのが基本であろうが、処理・制御形態が複雑になるのは避けられない。

この問題に対して、ここでは、追跡処理に入力画像の遅延を与える。即ち処理時間のかかる大局解抽出は入力画像に対して直ちに処理を開始するが、局所解追跡は常にある時間だけ遅延した画像に対して処理を行ない、両者の終了時刻を揃える。いわば、遅延型リアルタイム処理である。このような形態では、真の即時応答は不可

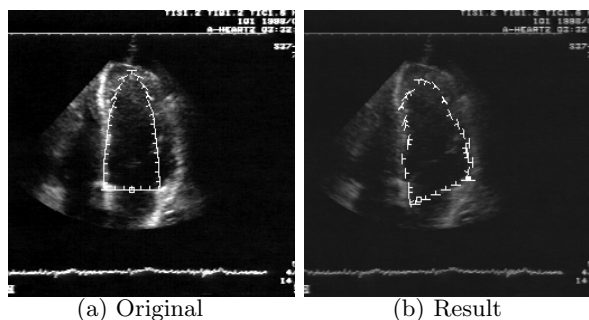


Fig. 9 Result for Vertical Cut

能であるが、計測など遅延が許される場合は、キャッチアップ処理などの複雑な制御が不要となる。このような前提を、本実験の処理時間に当てはめると、今後の(準)リアルタイム化のための高速化において、エッジ抽出部のハードウェア化による高速化と、特に大局解抽出に重点の置いた高速化が必要であることが分かる。

4.4 縦断面に対する結果

前節の SIDBA データベースに対する実験では、心臓に対して横方向の断面であるのに対し、ここでは、Fig.9 の様な心臓下部からの縦方向断面に対する実験を行なった。この場合、前節と同様の円形モデルでもある程度の抽出は可能であったが、特に収縮期には追従が困難であったため、Fig.9(a) の様な楕円の上半分を切り出した形のモデルを用いた。これにより、ほぼ満足な結果が得られた。なお、Fig.9 の (a) と (b) の背景は同一画像であるが、(a) は濃度階調を見易いように変換してある。

§5 手法の拡張に関する検討

5.1 許容域 A の変形に関する考察

モデルの変形は、一つ前の評価点からその点までのベクトル(遷移ベクトル)を棒とし、ベクトルの始点の周りに許容域 A の輪をもつ鎖を、棒の方向は変えないで、棒の終点を次の輪に入れながら次々につないでゆくことに相当する。この時、全体としてのあらゆる変形の可能性の中で、置かれた棒の先毎に定義された評価値の全体としての最適な場所と形状を探索する。

本実験のようにモデルの許容域 A を正方形(3×3 近傍)にすると、元の形状からの全体としての変形は(次に述べる 2 つの場合に比べて)最も大きくなる(MAP 整合法の場合)。許容域の形を遷移ベクトルの方向に偏平にすると、棒の接線方向にしか変形が許容されず、全体として、元の形の相似形に近い形のみの変形になる。もとの形状が滑らかな場合、結果の形状も滑らかとなる(DP 整合法の場合)。輪の形を遷移ベクトルに直角な方向に偏平に小さくすると、長さ方向の変形は許容せず、

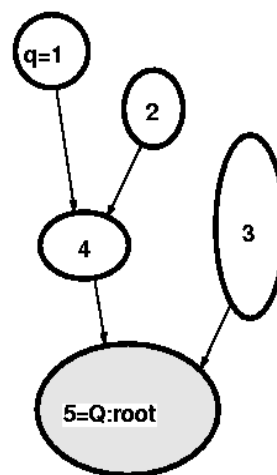


Fig. 10 Tree Model with Multiple Contours.

横方向に変形が許容される。このような変形は、回転する対象に対して角度の違う複数のモデルによるマッチングで対応する場合に、生成するモデルの数を少なく押さえる効果を持つ。もちろん、これらの 2 つのいずれについても、最大の変形をもつ正方形の許容域の場合の一部となるが、変形を不必要に大きくとると、不自然な形状にマッチングしてしまう危険性が生ずるため、目的に応じて適切に選択できることは重要である。

Snakes では 2 次関係の漸化式を計算することにより、輪郭の平滑性評価を実現しているが、ここでは、接線方向のみに変形を許容することにより、そこまで計算量を上げないでその一部を実現することができる。また、関連して、(0, 0) 以外の A の要素に対してペナルティを加えることによりモデルからの変形の程度を評価することも可能であるが、評価式を簡略化するために許容範囲内における画像との差異だけで評価関数を構成している。

5.2 弾性エッジ列整合法の多輪郭化

前章まででは、一繋がりエッジ点の系列で表現された単一の輪郭の弾性モデルを扱ってきた。ここでは、複数輪郭で構成されるモデルへの拡張について述べる。

モデルは Fig.10 の様に Q 個の輪郭で構成され、それらは、向きつけられた木構造を持つグラフとして表される。結合される輪郭(ノード)を親輪郭と呼び、結合する方を子輪郭と呼ぶ。2 つ以上の子輪郭から同時にひとつの親輪郭に結合できるが、子輪郭 q' が結合する親輪郭 q は一つである。この時、 $q' < q$ なるように番号付けられている。一番最後の親輪郭 $q=Q$ をルートと呼ぶ。

$$R = \{(a_m^q, x_m^q, y_m^q, |p|_m^q, d_m^q) \mid m = m_S^q, \dots, m_E^q, q = 1, \dots, Q, m_E^q < m_S^{q+1}\}. \quad (27)$$

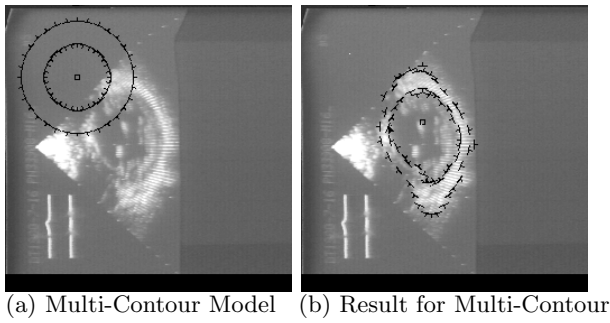


Fig. 11 Multiple Contour Model.

それぞれの輪郭 $q (= 1, 2, 3, 4, \dots, Q)$ に対するマッチングは、単一輪郭モデルの場合の漸化式 (14) を $m = m_S^q$ から $m = m_E^q$ について計算し、更に最後の評価点 $m = m_E^q$ において、モデルの形状が閉じているかどうかの評価を加えて $f^{[q]}(k, l) = f_{m_E^q}(k, l)$ として得る。この値は、各評価点が局所窓による絶対位置の拘束を受けると同時に、引き続き 2 評価点が許容変形によって相対的に拘束され、かつ輪郭 q の参照点 (代表点) が q 毎に定義された座標系上の (k, l) に位置する時の各評価点での局所的評価の和が最適となるマッチングの度合を示す。

次に、多輪郭の場合は、新たに結合している輪郭間で漸化式を計算する。輪郭間漸化式として、親輪郭 q とそれに結合する子輪郭 q' 群の間で、次の漸化式計算を tree の深い (q の小さい) ノードから順に行なう。

$$F^{[q]}(k, l) = f^{[q]}(k, l) + \sum_{q'} \max_{\max(|k'-k|, |l'-l|) \leq W_C} \left\{ F^{[q']}(k', l') - \gamma_C \cdot \kappa \cdot (|k' - k| + |l' - l|) \right\}. \quad (28)$$

ここで、 W_C は、輪郭間サーチ領域許容幅、 γ_C は、輪郭間相対位置ズレに対する重みである。 γ_C が大きくなるほど、モデルの輪郭間相対位置関係が強く保たれる。

更にこの時 $q - q'$ 輪郭間の変動量を対応点 (k, l) を要素とするマトリックスとして記憶しておく。

$$C(k, l, q') = \arg \max_{(k, l)} \{ F^{[q]}(k, l) \}. \quad (29)$$

この漸化式の計算終了後、全体に関する最適解を求める。

$$E(R, U) = \max_{(k, l)} \{ F^{[Q]}(k, l) \}, \quad (30)$$

更に、輪郭間対応点マトリックスと各輪郭内対応点マトリックスに対するルート側からのトレースバックにより、対応点を求める。

Fig.11(a) の様な内側の内向き輪郭と、外側の外向き輪郭が同心円状になっている多輪郭モデルを背景の画像 NM01 に適用した結果を Fig.11(b) に示す。心臓左心室像の場合、外側の輪郭の方がエッジ画が明確で、かつ形状も滑らかであるという傾向があるため、最終的な目的

が内壁の抽出にある場合でも、外側輪郭の援用が、抽出の高速化に有効であろう。

§ 6 おわりに

一般に動画像は情報量が大きく、処理の任意性が多すぎるため、様々な処理内容の拘束が必要になる。そこでは、既存手法の組み合わせや対象の知識・拘束条件による手法の精錬も重要であるが、更なるブレークスルーの為に、頑強な要素技術、統合コンセプトの確立が必要である。従来から我々は、要素技術として、動的計画法を用いたマッチング手法およびその並列計算版である MAP 整合法を提案してきた。本論文では、各評価点毎に座標系を導入することにより、従来から持つ大局的最適解抽出という特長に加えて、局所解の位置拘束窓設定と粗密化による高速化を統一的に実行する弾性エッジ列整合法を提案した。モデルを用いた動画像追跡において、局所解による追跡と、初期画像・キー画像に対する大局解抽出にこれを適用し、解の安定化を図った。

今後の課題として、解のスムーズネスの導入、エッジの情報だけでなく、領域内外部の特徴の取り込み、大局解と局所解の判定基準の強化などが残されている。

謝 辞

研究を手伝って頂いた日本板硝子 (株) 筒井啓輔氏、および日頃ご指導ご討論頂く坂上勝彦ラボリーダはじめ画像理解ラボメンバの皆様に感謝します。

参 考 文 献

- 1) A.A.Amini, T.E.Weymouth, and R.Jain: "Using dynamic programming for solving variational problems in vision", IEEE Trans. Pat. Anal. Mach. Intel., Vol.12, No.9 (1990) pp.855-867.
- 2) D.H.Ballard: "Generalizing the Hough transform to detect arbitrary shapes", Pattern Recognition, Vol.13, No.2 (1981) pp.111-122.
- 3) J.Canny: "A computational approach to edge detection", IEEE Trans. Pat. Anal. Mach. Intel., Vol.8, No.6 (1986) pp.678-698.
- 4) 藤村恒太, 横矢直和, 山本和彦: "多重スケール画像を用いた動的輪郭モデルによる非剛体物体の輪郭追跡と動きの解析", 信学論, Vol.J76-D-II (1993) pp.382-390.
- 5) M.Kass, A.Witkin, and D.Terzopoulos: "Snakes: active contour models", Int. Jour.

- Comput. Vision (1988) pp.321-331.
- 6) L.Kitchen and A.Rosenfeld: “Discrete relaxation for matching relational structures”, IEEE Trans., Vol.SMC-9, No.2 (1979) pp.869-874.
 - 7) V.A.Kovalevsky: “Sequential optimization in pattern recognition”, Proc.IFIP Congr.68 (1968) pp.1603-1607.
 - 8) 大田友一, 山田博三: “動的計画法によるパターンマッチング”, 情報処理, Vol.30, No.9 (1989) pp.1058-1066.
 - 9) 迫江博昭: “Rubber String Matching 法による手書き文字認識”, 信学技報, PRL74-20 (1974).
 - 10) 上田修功, 間瀬健二, 末永康仁: “弾性輪郭モデルとエネルギー最小化原理による輪郭追跡手法”, 信学論, Vol.J75-D-II (1992) pp.111-120.
 - 11) H.Yamada, C.Merritt, and T.Kasvand: “Recognition of kidney glomerulus by two-dimensional dynamic programming matching method”, IEEE Trans. Pat. Anal. Mach. Intel., Vol.10, No.5 (Sep.1988) pp.731-737.
 - 12) 山田博三, 山本和彦: “DP 整合法による超音波心臓動画像の認識”, 信学論, Vol.J71-D, No.4 (1988) pp.678-684.
 - 13) 山田博三: “並列演算による方向パターン弾性整合法の提案—MAP 整合法”, 信学論, Vol.J73-D-II, No.4 (1990) pp.553-561.
 - 14) H.Yamada, K.Yamamoto, T.Saito, and K.Hosokawa: “Directional mathematical morphology and reformalized Hough transformation for the analysis of topographic maps”, IEEE Trans. Pattern Anal. Machine Intell., Vol.15, No.4 (1993) pp.380-387.
 - 15) 山田博三, 筒井啓輔: “弾性エッジ列整合法による動画像の大局解抽出と局所解追跡”, 信学論, Vol.J80-D-II, No.1 (1997) pp.198-205.
 - 16) 山田博三, 筒井啓輔: “弾性エッジ列整合法の多輪郭化による顔の位置決め”, 信学論, Vol.J80-D-II, No.8 (1997) pp.2239-2243.
 - 17) 山本和彦: “弛緩整合法による手書き教育漢字認識”, 信学論, Vol.J65-D, No.9 (1982) pp.1167-1174.

(1998.12.21 受付)

著者紹介



山田博三

Hiromitsu YAMADA

知能情報部画像理解ラボ

E-mail: yamada@etl.go.jp

パターン認識、特に手書き文字認識、並列演算による特徴抽出、画像認識、動的計画法による弾性マッチング法の研究に従事。