

単調連続 2 次元 DP アルゴリズムの階層化

勝股 充^{†*} 鈴木 尚^{†**} 徳田 恵一[†] 北村 正[†]

A Hierarchical Algorithm for Monotonic and Continuous Two-Dimensional Dynamic Programming

Mitsuru KATSUMATA^{†*}, Hisashi SUZUKI^{†**}, Keiichi TOKUDA[†],
and Tadashi KITAMURA[†]

あらまし 画像マッチングの手法である単調連続 2 次元 DP アルゴリズムの改良手法について述べる。従来の 2 次元 DP による 2 次元画像マッチングでは、直接的な探索に比べれば大きく計算量を削減できる。しかし、現実的な時間で計算を終えるためには、ビームサーチを行う必要があり、現実的な計算時間のビーム径では、最適解からのずれが大きくなるという問題があった。本論文では、まず、ワーブ時に起こる画素の位置関係の反転を完全に回避することのできる手法を提案した上で、既存のアルゴリズムを階層化することにより、現実的なビーム径で性能改善を図る。階層化は、計算量の増加につながるが、ビーム径を従来法に比べてかなり小さくでき、結果的に計算量の削減を実現できることが実験により示された。また、反転問題では従来法では完全に反転を防止できていなかったが、本手法により完全に反転を防止することができるようになった。

キーワード 2 次元ワーブ, パターンマッチング, 動的計画法, 階層化, 反転防止

1. ま え が き

画像認識は、画像処理や、パターン認識といった研究分野における一つの大きなテーマとなっており、その応用範囲は、人間と計算機間のインタラクションの促進、ビデオ監視システム、文字認識システム、顔や目の画像認識による個人識別、個人認証など多岐にわたっている。近年、様々な方法で画像認識は研究されてきたが、多くの手法に共通の問題点として、認識すべき画像の、幾何学的な非線形伸縮による認識率低下の問題がある。このような問題を解決する手法は、2 次元ワーブ決定問題といわれ、2 画像の最大一致を与える画素間のノンパラメトリックなマッチングを求め、この最適化問題は爆発的にその数を増大させる組合せの問題を抱えているため、アルゴリズムの検討にあたっては計算量の考察が重要となってくる。

このような、2 次元ワーブ問題を解決するため、動

的計画法 (DP: Dynamic Programming) に基づくアルゴリズム [1] ~ [3] 隠れマルコフモデル (HMM: Hidden Markov Model) に基づくアルゴリズム [4] ~ [7]、が検討されている。音声認識の分野では発話者の発音の癖の違いによって生じる時間軸上の非線形の伸縮を吸収するために、DP あるいは HMM が利用されており [8]、これらのアルゴリズムを 2 次元に拡張することにより、2 次元ワーブ問題に適用することができる。しかし、現在提案されている手法では、ある特定の方向への変形は吸収可能であるが、傾きや回転には対応できず、完全な 2 次元自由度をもっていないものが多い。これらの中で [1] は、2 次元的自由度をもっているが、計算量の大きさや制約条件が単調性だけであるため不適切なワーピングが起こり得るという問題がある。それに対して [3] では制約条件に単調連続性を導入している、ビームサーチを用いて計算量の削減をしているなどの工夫をしていることから注目されるが、実現可能なビーム径では、以下のような問題があった。

[問題 1] 画素位置の反転が起きる。

[問題 2] ビームサーチに起因する不適切なワーピングが発生する。

問題 1 は、単調連続性の条件が、垂直、水平方向に

[†] 名古屋工業大学知能情報システム学科, 名古屋市

Department of Computer Science, Nagoya Institute of Technology, Nagoya-shi, 466-8555 Japan

* 現在, ソニー株式会社

** 現在, 日本 IBM 株式会社

対してのみ課せられているため、全体としての単調連続性が成立してないためである。平面ワーピングのための DP アルゴリズムは、パターンの位相を保存するゴム膜変換を実現することが本質であり、この条件を厳密に課すことは重要な課題と考える。問題 2 は当該画素までの局所的なコストに基づいて、ビームサーチが行われるため、ビームサーチの早い段階で、適切な解がすべて失われてしまうためである。文献 [3] においては、ヒューリスティックなペナルティを目的関数に算入することにより、これらの問題を回避しようとしている。それに対して、本論文は、問題 1 の画素反転を完全に回避する手法として、直接的な制約を与える方法を、そして問題 2 のより良い回避法として、階層化した 2 次元 DP を提案する。階層化 2 次元 DP においては、画像全体の大局的な構造を考慮したビームサーチが行われるため、不適切な変形が起こりにくいだけでなく、ビーム径を小さくできることから、計算量の削減にもつながる。

以下、2. ではもともとなる単調連続 2 次元 DP アルゴリズム、3. では画素反転防止処理、4. では階層化アルゴリズム、5. では手書き文字マッチング実験、そして、6. で結果を述べる。

2. 単調連続 2 次元 DP アルゴリズム [3]

2.1 2 次元ワーブ

2.1.1 2 次元ワーブ決定問題

入力画像 (Input Image) を

$$A = \{a(i, j) | i = 1, 2, \dots, I, j = 1, 2, \dots, J\} \quad (1)$$

モデル画像 (Model Image) を

$$B = \{b(x, y) | x = 1, 2, \dots, X, y = 1, 2, \dots, Y\} \quad (2)$$

と表す。ここで $a(i, j)$, $b(x, y)$ はそれぞれ入力画像、モデル画像の特徴量である。ここで、画像 A から画像 B へのワーブ関数を $x(i, j)$, $y(i, j)$ とする。

このとき、ワーブ決定問題は、ある制約条件下で目的関数

$$D = \sum_{i=1}^I \sum_{j=1}^J |a(i, j) - b(x(i, j), y(i, j))| \quad (3)$$

の最小値を与えるワーブ関数の決定問題と定式化される。このワーブ関数で変形された画像 $\{b(x(i, j), y(i, j))\}$ を \tilde{B} (Output Image) と表す。

目的関数 D の最小値を $D(A, B)$ と表す。

$$D(A, B) = \min_{\substack{x(i, j) \\ y(i, j)}} \sum_{i=1}^I \sum_{j=1}^J |a(i, j) - b(x(i, j), y(i, j))| \quad (4)$$

この $D(A, B)$ は、変形画像 \tilde{B} と入力画像 A に対して、テンプレートマッチングを行った場合の距離に相当する。

2.1.2 制約条件

単調性及び連続性を同時に満たすようなワーブを構成するためにワーブ関数に次の条件を与える。

$$0 \leq x(i, j) - x(i-1, j) \leq 2 \quad (5)$$

$$0 \leq y(i, j) - y(i, j-1) \leq 2 \quad (6)$$

$$|x(i, j) - x(i, j-1)| \leq 1 \quad (7)$$

$$|y(i, j) - y(i-1, j)| \leq 1 \quad (8)$$

これらは、 A 上で隣接の関係にある 2 ピクセルのワーブ後の位置関係を制限していることになる。この連続単調性により、パターンの位相を近似的に保存可能な非線形変形が実現される^(注1)。

更に、境界条件

$$x(1, j) = 1, \quad x(I, j) = X \quad (9)$$

$$y(i, 1) = 1, \quad y(i, J) = Y \quad (10)$$

を用いる。

2.2 カラムワイズビームサーチ DP アルゴリズム

この節では、3. 以降で述べる手法のもともとなるカラムワイズ DP アルゴリズム [3] を説明する。このアルゴリズムは、2 次元 DP マッチングを列単位に分解して各列における 1 次元 DP マッチングととらえている。

2.2.1 縦方向、横方向への分解

画像の A の第 i 行を

$$\alpha(i) = (a(i, 1), \dots, a(i, j), \dots, a(i, J)) \quad (11)$$

と表す。ここで $i = 1$ から I までの I 段の決定過程を考え、その第 i 段では $\alpha(i)$ のマッピングを決定する。また、 B 上における $\alpha(i)$ の像を

(注1): 厳密に言えば、1 画素分の不連続を許している。また、これにより単調性も部分的に満たされない場合が生じる。この問題に対する処理は 3. で述べる。

$$\begin{aligned} \mathbf{xy}(i) = & ((x(i, 1), y(i, 1)), \dots \\ & \dots, (x(i, j), y(i, j)), \dots, (x(i, J), y(i, J))) \end{aligned} \quad (12)$$

と表し局所ひずみパターンと呼ぶ。 $\mathbf{xy}(i)$ は、単調連続性 (6), (7), 境界条件 (9), (10), を満たす範囲で任意に形状を選ぶことができる。その集合を $XY(i) = \{\mathbf{xy}(i)\}$ と表す。局所ひずみパターンの導入により、式 (4) の 2 次元的な組合せ最適化問題は、次のように局所ひずみパターンの 1 次元系列 $\mathbf{xy}(1) \dots \mathbf{xy}(I)$ の最適化問題に帰着される。

$$D_C(\mathbf{A}, \mathbf{B}) = \min_{\mathbf{xy}(1) \dots \mathbf{xy}(I)} \sum_{i=1}^I d(\mathbf{a}(i), \mathbf{xy}(i)) \quad (13)$$

ここで、 $d(\mathbf{a}(i), \mathbf{xy}(i))$ は $\mathbf{a}(i)$ を $\mathbf{xy}(i-1)$ を用いて B 上へマッピングした場合の局所距離で次式で定義される。

$$\begin{aligned} d(\mathbf{a}(i), \mathbf{xy}(i)) \\ = \sum_{j=1}^J |a(i, j) - b(x(i, j), y(i, j))| \end{aligned} \quad (14)$$

更に、 $\mathbf{xy}(i)$ と単調連続性条件 (5), (8) で決定される $\mathbf{xy}(i)$ に隣接可能な $\mathbf{xy}(i-1)$ の集合を、 $XY'(\mathbf{xy}(i))$ と表すことにする。図 1 にカラムワイズ DP アルゴリズムを示す。

2.2.2 ビームサーチ

カラムワイズ DP アルゴリズムは時間空間量及び空間計算量はともに画像サイズに関して指数オーダーで増

加するため [3]、最適解探索のアルゴリズムでは、実用的な画像サイズで実行することは事実上不可能である。したがって、ビームサーチによる準最適解探索アルゴリズムを適用する。

$i-1$ 行目の局所ひずみパターン $\mathbf{xy}(i-1)$ (画像 A の $(i-1)$ 行目の B 上での像) の集合 $XY(i-1)$ が定まっているとする。 $\mathbf{xy}(i-1)$ と $\mathbf{xy}(i)$ との差分 $(x(i, j) - x(i-1, j), y(i, j) - y(i-1, j))$ は $(\{2, 1, 0\}, \{-1, 0, 1\})$ で表せる 9 種類となる。したがって $\mathbf{xy}(i-1)$ に隣接する $\mathbf{xy}(i)$ のすべては $\mathbf{xy}(i-1)$ と各ノードがそれぞれ 9 種類の差分の一つに対応する子ノードをもつような深さ J の 9 分木で表される。すなわち、この 9 分木の根から葉までの 9^J 本の道により、 $\mathbf{xy}(i-1)$ に隣接可能な $\mathbf{xy}(i)$ の全部の道が表現される。この 9 分木を各 $g(i-1, \mathbf{xy}(i-1))$ の根とし、そのそれぞれの根から並列的に枝を伸ばしていく。そこで木の深さ j を一つ増すごとに、累積距離が最小の R 個のノードだけを残す枝刈処理を行う^(注2)。この処理を $j = J$ まで繰り返した後、同じ $\mathbf{xy}(i)$ の表す葉のうち、最小の累積距離をもつものを選択し第 i 行目の処理が終了する。

3. 画素反転防止処理

制約条件 (5) ~ (8) は近似的には単調性、連続性を満たしている。しかし図 2 のような部分的な反転を許容してしまう。この部分的な反転は、単調性のための制約が、本来、すべての方向について満たされるべきであるが、それが局所的かつ離散的近似になっていることに起因する。反転は、正確にゴム膜変換の 2 次元ワーブを期待する場合は望ましくない。ペナルティを導入することにより反転を防止しようとする手法に関しては [3] で検討されているが、それは完全に反転を回避するものではなかった。

3.1 反転判別

反転を避ける方法として、4 点 $(i, j), (i-1, j), (i, j-1), (i-1, j-1)$ が、ワーブ後に形成する四角形の面の向きを判定する (図 3)。

4 点 $(i, j), (i-1, j), (i, j-1), (i-1, j-1)$ が、ワーブ後に形成する四辺形の面の向きを判定し、反転と判断された場合 (図 3 の下段の場合) は、以後の探索範囲から除外する。反転の判別は以下のように行う。

(i, j) のワーブ先 $(x(i, j), y(i, j))$ (図 3 の B に

```

1. 初期状態
for all  $\mathbf{xy} \in XY(1)$ 
 $g(1, \mathbf{xy}) = d(1, \mathbf{xy})$ 
2. DP 漸化式
for  $i := 2$  to  $I$ 
for all  $\mathbf{xy} \in XY(i)$ 
 $g(i, \mathbf{xy}) = d(i, \mathbf{xy}) + \min_{\mathbf{xy}' \in XY'(\mathbf{xy})} g(i-1, \mathbf{xy}')$ 
3. 終了
 $D(\mathbf{A}, \mathbf{B}) = \min_{\mathbf{xy} \in XY(I)} g(I, \mathbf{xy})$ 

```

図 1 カラムワイズ DP アルゴリズム

Fig. 1 DP algorithm for the monotonic and continuous 2D warping.

(注2): R はビーム径 (Beam Size) と呼ばれる。

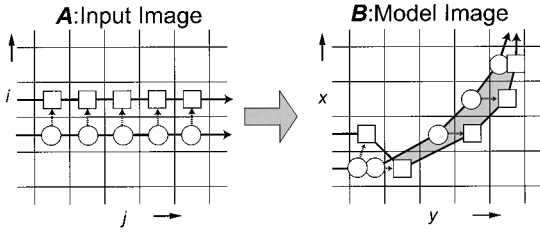


図 2 制約条件を満たす反転の例
Fig. 2 Violation of monotonicity.

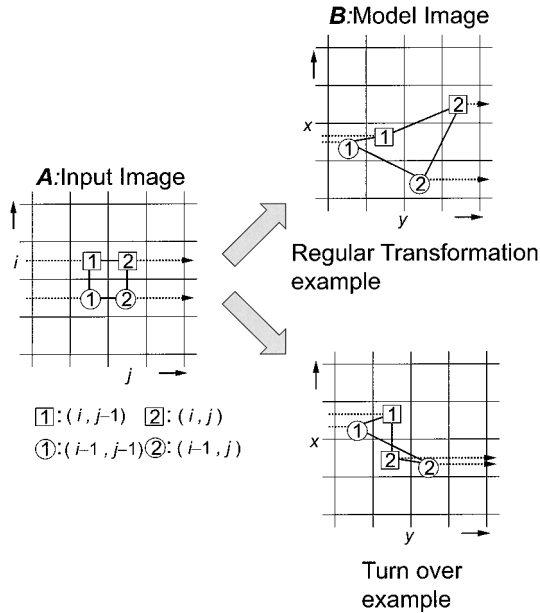


図 3 反転判別
Fig. 3 Detection of turn over.

おける 2) を考える . $(x(i, j), y(i, j))$ のワーブ可能範囲は , $(x(i-1, j), y(i-1, j))$ (図 4 の B における ②) と $(x(i, j-1), y(i, j-1))$ (図 4 の B における 1) と制約条件 (5) ~ (8) によりいくつかに定まる (図 4 の網かけの部分) . そのワーブ可能範囲の中で 1 点 $(x(i, j), y(i, j))$ を定め , 出来上がる四辺形の形を判別し , 反転とみなされる場所 (図 4 の x) は探索範囲から除外する .

$(x(i, j), y(i, j))$ 以外の 3 点の相対的な位置は制約条件 (5) ~ (8) により , たかだか 81 通りである . その中から反転場所を列挙し制約条件に加える .

3.2 ビームサーチによる探索経路消失回避処理

ビームサーチ DP (2.2) に画素反転防止処理 (3.1) を用いた場合 , 探索経路が消失してしまう

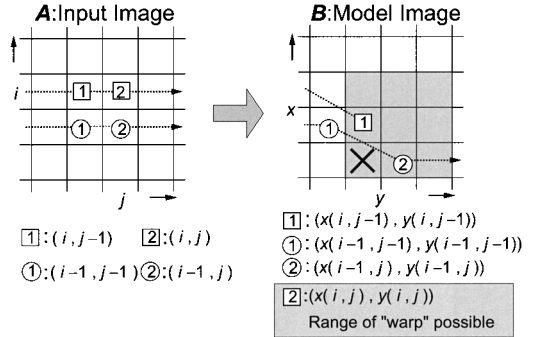


図 4 $(x(i, j), y(i, j))$ のワーブ可能範囲
Fig. 4 Range of "warp" possible.

問題がある . それは , 反転を判別する四辺形の 3 点 $(x(i-1, j), y(i-1, j))$, $(x(i, j-1), y(i, j-1))$, $(x(i-1, j-1), y(i-1, j-1))$ (図 4 の 1 , ① , ②) の位置関係によっては , 制約条件 (5) ~ (8) によって定まる探索場所すべてが反転とみなされるものがあり (図 5) , もしビームサーチで残ったコストの小さい R 個の探索経路すべてがこの場合に当てはまる場合には , 次の探索経路 $(x(i, j), y(i, j))$ はすべて反転とみなされ探索経路は一つも残らなくなる .

この問題を解決するため , $(x(i, j), y(i, j))$ の位置をワーブ可能範囲の中から決める際に , その後の探索経路が一つ以上存在するか先読みをすることにより調べる . 探索経路が一つもなくなってしまうパターンは図 5 の 5 通りである . 図 5 を詳しく見ると , どの場合も探索経路が存在しなくなる $(x(i-1, j-1), y(i-1, j-1))$ (図 5 の ①) と $(x(i, j-1), y(i, j-1))$ (図 5 の 1) の y の値の差は 1 である . よって ,

$$y(i, j) - y(i-1, j) = 1 \tag{15}$$

のときにのみ先読みを行えばよい . 画素反転防止処理を行った場合でも $xy(i-1)$ に対して $xy(i)$ は必ず存在するので , 先読みは $j = J$ まで行えばよい .

式 (15) と先読みは $j = J$ まで行えばよいことより , 先読みアルゴリズムは以下ようになる .

(i, j) のワーピング先 $(x(i, j), y(i, j))$ を決定する場合 (ただし $j > 2$) . $y(i, j) - y(i-1, j) = 1$ ならば先読みを行う .

(1) $k = 1$

(2) $(i, j+k)$ のワーピング先が存在しなければ他の点を先読みする . もし , すべての先読み点に対してワーピング先がない場合は探索範囲から除外する .

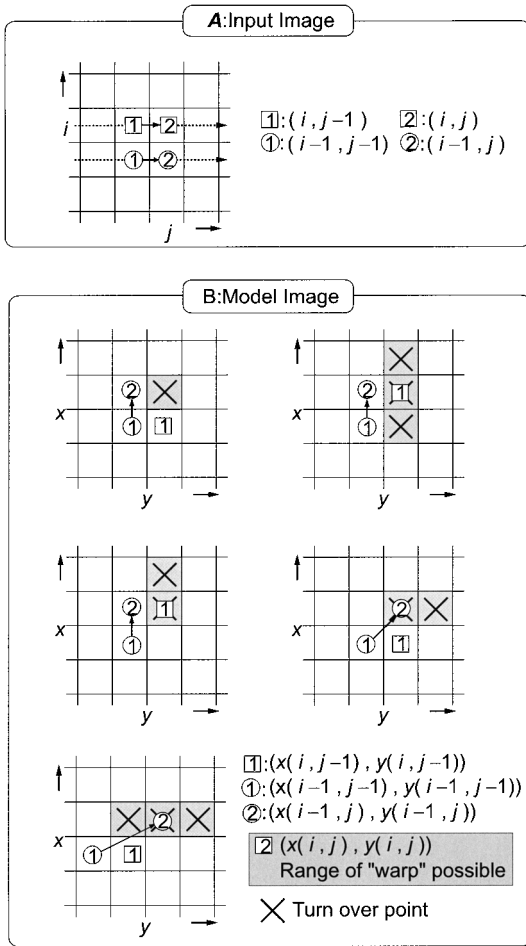


図5 探索経路がなくなるパターン
 Fig. 5 Cases where no search path will be found.

(3) $y(i, j+k) - y(i-1, j+k) = 1$ 若しくは $j+k \neq J$ ならば (4) へ。そうでなければ先読み終了 (探索経路が存在する保証のあるパス)。

(4) $k = k+1$ として (2) へ
 このような手順で先読みは行われる。J になるまで先読みをした場合に計算時間は最悪になるが、実際には先読み自体行われる場合は少ない。また、 $(i, 1)$ の場合は境界条件式 (10) より $y(i, 1) - y(i-1, 1) = 0$ となり探索経路が存在する保証がある。よって、再帰的に考えれば探索経路の存在の保証ができる。

4. 階層化アルゴリズム

これまでに述べたアルゴリズムをそのまま適用した場合、ビーム径を大きくとっても、現実的な計算時間

ではマッピング先の集中や、全体としての滑らかさの欠如など、不自然な2次元ワーブが生じる場合がある。

この不自然なワーブの原因として考えられるのは

- ビームサーチによる最適解からのずれ
- 単調連続性制約条件の局所性

が考えられる。この問題について文献 [3] では一様性ペナルティ関数という変形の量や質を定量的に評価する関数を本来の目的関数に加え、ビームサーチの判断基準、そして最適解の判断基準とし、解決を図っている。しかし、通常、最適解への変移の大きさは画像によって差がある。一様性ペナルティを使用する場合は、最適解への変移量が小さい場合は、ペナルティの効果を制御する重みを大きくし、変移量が大き場合は、逆にその重みを小さくしなければならない。よって、最適解への変移の大きさをあらかじめ知らなければならないという意味で汎用性に欠ける。また、通常、最適解の入力画像に対する変移量は一様ではない。例えば、モデル画像を回転したものを入力画像とした場合、画像の中心部はあまり変移しないが画像の端の部分の変移量は大きくなる。このような場合、ペナルティに対する重みを定めるのは容易ではないと考えられる。

4.1 画像サイズとビーム径

画像サイズが大きい場合には、現実的な計算時間となるビーム径が全探索範囲に対して小さすぎ、最適解からの差が大きくなりやすくなる。

画像サイズが小さい場合は、小さなビーム径でも品質はかなり良いものとなることが実験的にわかっている。そこで、対象とする画像が大きい場合、そのモデル画像と入力画像をサブサンプリングし画像サイズを下げ、そこで得られた準最適解をもとに、画像サイズの大きな画像のワーブを行うことを考える。

4.2 アルゴリズムの階層化

階層化アルゴリズムを示す (図6)。まず、もとの入力画像、モデル画像を 2×2 (pixel) のサイズまで、サブサンプリングする (本節以降の入力画像、モデル画像のサイズは2のべき乗のサイズ $2^M \times 2^M$ とする)。得られた画像サイズ 2×2 の入力画像を用い、3. で述べた DP アルゴリズムによりマッチングを行い、コストが最小のワーピングを求める。この階層を Layer 1 とする。Layer 1 でマッピングされる画像は、式 (9)、(10) からただ一つ求まる。

次にもとの画像を 4×4 のサイズにサブサンプリングしたもので、マッチングを行う (Layer 2)。一つ前の Layer の構造を反映させるため目的関数を以下のよ

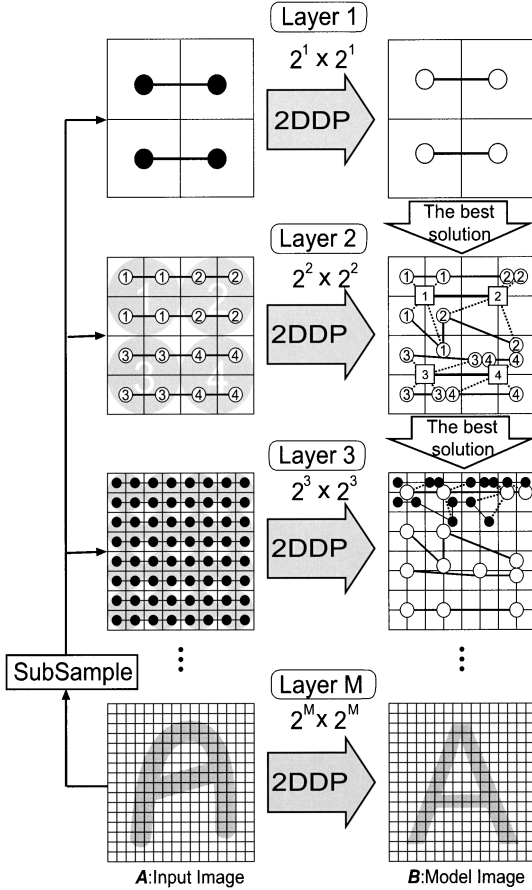


図 6 アルゴリズムの階層化
Fig. 6 Hierarchical algorithm.

うに変更する。

$$D_{new} = D_C + \alpha P \quad (16)$$

ここで、 P は一つ前の Layer の最適解からの距離、 $\alpha (\geq 0)$ は一つ前の Layer の結果の影響を制御する重みであり、 α が大きければ大きいほど上位階層における最適解に沿ったワーピングが最適解として選ばれやすくなる。ただし、 α は、画像ごとの変移量、あるいは各画素ごとの変移量とは独立に決まる値であり、一度、値を定めれば、汎用的に用いることができる。 P はユークリッド距離を用い、次式のように定義される。

$$P = \sum_{i=1}^I \sum_{j=1}^J U(i, j) \quad (17)$$

$$U(i, j) = \left\{ \left(x_p \left(\frac{i}{2}, \frac{j}{2} \right) - x(i, j) \right)^2 \right.$$

$$\left. + \left(y_p \left(\frac{i}{2}, \frac{j}{2} \right) - y(i, j) \right)^2 \right\}^{1/2} \quad (18)$$

$x_p(i, j)$, $y_p(i, j)$ は、一つ前のクラスにおける (i, j) のワープ先の x 座標及び y 座標である。上位階層における 1 画素は次の階層では 4 画素に対応する。したがって、図 6 の Layer 2 の P は、それぞれの番号の白丸の位置と同じ番号が付いている黒丸の位置とのユークリッド距離（同図中の点線の長さ）の総和となる。このアルゴリズムで Layer 2 の最適解を得る。Layer 3 以降も同様に、上位階層における画像サイズの縦横 2 倍の大きさの入力画像、モデル画像を使用し、最適解を求める。これを Layer M まで繰り返すことにより、全体構造を考慮した解を得ることができる。

また、入力サイズとモデルサイズが違う場合や、画像サイズが $2^M \times 2^M$ でない場合においても、式 (18) を変更することにより、同様の方法を用いることができる。

5. 手書き文字マッチング実験

5.1 実験条件

提案法の有効性を検討するために、8 種類の手書き文字のマッチング実験を行う。モデル画像にはゴシック体のアルファベット、及び記号を用い、入力画像には各文字につき 5 種類の癖をつけた手書き文字を用い、マッチングを行い、評価実験を行う。実験に用いた入力画像、モデル画像を図 7 に示す。

[実験 1] 単調性ペナルティを用いた場合、反転防止を用いた場合、両方用いない場合について比較する (表 1)。

[実験 2] 画素反転防止処理を用いた上で階層化を用いる場合と用いない場合について比較する (表 2)。

5.2 実験結果

5.2.1 実験 1

単調ペナルティと画素反転防止処理なし (手法 1)、単調性ペナルティを用いた手法 (手法 2)、画素反転防止処理を用いた手法 (手法 3) とする。実験結果は表 3 に示す。反転を禁止することにより、可能なワーピングに対する制約は強くなり、結果として目的関数の値は増加することが予想されるが、提案手法 (手法 3) の手法 1 に対する目的関数の増加はわずかであることが表より示されている。また、提案手法が完全に反転を回避しているのに対し、手法 2 では、反転箇所を残していることがわかる。なお、手法 2 において、

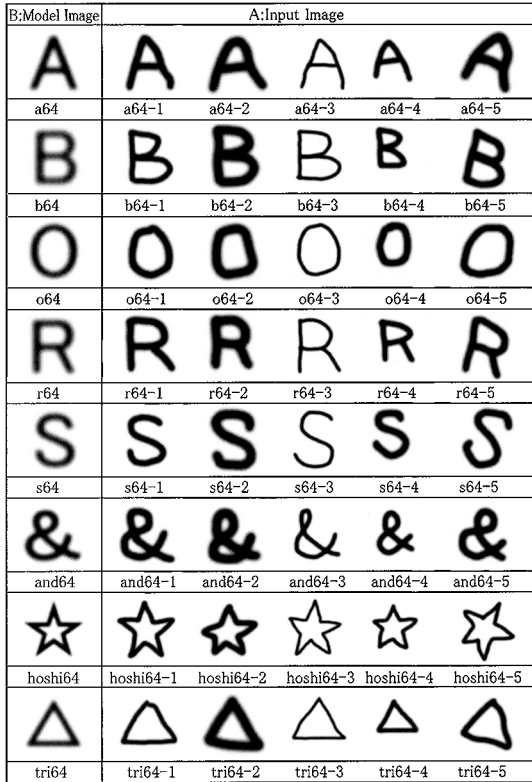


図 7 入力画像とモデル画像
Fig. 7 Input image and model image.

表 1 実験条件 1
Table 1 Experimental condition 1.

入力画像	64 × 64 (pixel) 256 階調 グレイスケール
モデル画像	
ビーム径 R	1000
単調性ペナルティ β (単調性ペナルティ使用時)	100
特徴量	輝度値 (1 次元)

使用計算機 : COMPAQ Professional Workstation XP1000
(SPECfp95:52, SPECint95:27)

目的関数が若干減少しているのは、ビームサーチと関連した影響によると考えられる。

5.2.2 実験 2

図 8~図 10 に変形画像 \tilde{B} の例を示す。各変形画像の下の数字は目的関数値であり、ここに前 Layer とのユークリッド距離 P (式 (17)) は含まれていない。図 8 が階層化なし、ビーム径 2000、図 9 が階層化あり ($\alpha = 20$)、ビーム径 2000、図 10 が階層化あり、ビーム径 100 である。なお、階層化ありの場合、すべての階層で同じビーム径を用いた。図 9 より、階層化

表 2 実験条件 2

Table 2 Experimental condition 2.

入力画像	64 × 64 (pixel) 256 階調 グレイスケール
モデル画像	
ビーム径 R	10,50,100,500,1000,2000
反転防止処理	あり
階層化の重み α (階層化ありの場合)	1,10,20
特徴量	輝度値 (1 次元) 上下左右斜め方向差分 (4 次元)

使用計算機 : COMPAQ Professional Workstation XP1000
(SPECfp95:52, SPECint95:27)

表 3 実験結果

Table 3 Experimental results.

	手法 1	手法 2	手法 3
平均計算時間 (s)	129.00	133.00	126.81
目的関数の平均	95769	93180	97788
平均反転箇所	152.7	28.1	0

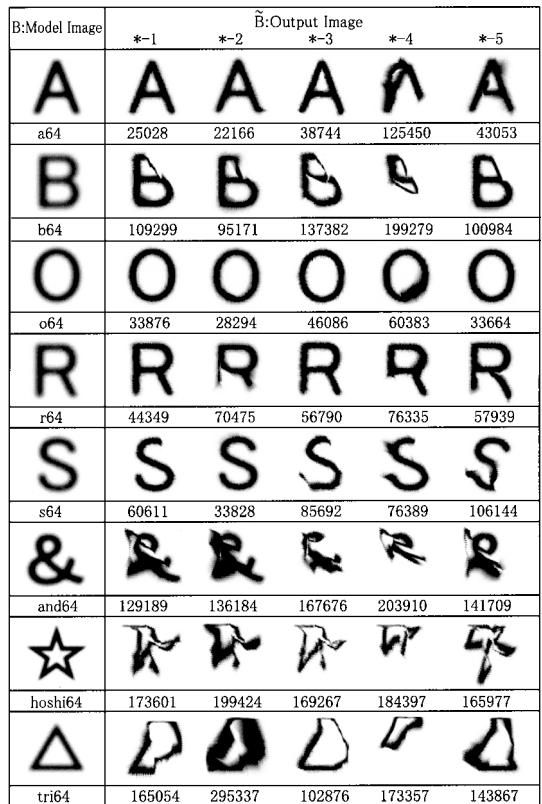


図 8 マッチング結果 ($R = 2000$, 階層化なし)
Fig. 8 Experimental results not using hierarchical algorithm. ($R = 2000$)

B:Model Image	\tilde{B} :Output Image				
	*-1	*-2	*-3	*-4	*-5
A	A	A	A	A	A
a64	32722	34749	52263	30181	34130
B	B	B	B	B	B
b64	66262	81080	57810	49504	66876
O	O	O	O	O	O
o64	38572	37690	35503	32747	44089
R	R	R	R	R	R
r64	55159	49969	40735	39362	52642
S	S	S	S	S	S
s64	46551	58145	38126	51204	54386
&	&	&	&	&	&
and64	35103	33219	42339	26393	40175
☆	☆	☆	☆	☆	☆
hoshi64	32400	24095	51895	42981	87648
△	△	△	△	△	△
tri64	40032	46904	31870	33765	39261

図 9 マッチング結果 ($R = 2000$, 階層化あり, $\alpha = 20$)

Fig. 9 Experimental results using hierarchial algorithm. ($R = 2000$, $\alpha = 20$)

B:Model Image	\tilde{B} :Output Image				
	*-1	*-2	*-3	*-4	*-5
A	A	A	A	A	A
a64	32436	34883	53913	35492	34088
B	B	B	B	B	B
b64	66309	81031	58456	55300	66724
O	O	O	O	O	O
o64	39016	37711	35896	33351	44729
R	R	R	R	R	R
r64	55120	49664	40987	39176	52385
S	S	S	S	S	S
s64	46378	58709	39410	55862	56028
&	&	&	&	&	&
and64	35026	36149	42901	26332	40636
☆	☆	☆	☆	☆	☆
hoshi64	58761	24355	55552	49126	87235
△	△	△	△	△	△
tri64	40358	48121	33343	34351	39145

図 10 マッチング結果 ($R = 100$, 階層化あり, $\alpha = 20$)

Fig. 10 Experimental results using hierarchial algorithm. ($R = 100$, $\alpha = 20$)

を行わない場合は特に、“b64”、“and64”、“hoshi64”、“tri64”では不自然なマッチングが生じてしまっている。どの場合もビームサーチの前半での間違いにより、ワーブ先が画像の左上に集中してしまい、後半でその間違いにつつまを合わせるようなワーブが生じている。

次に図 9 では、図 8 に比べ、どの文字においてもほぼモデル画像に近い出力画像となっているといえる。変形量の少ない入力画像 (*-1) から、変形量の多い入力画像 (*-4,5) まで、変形量によらず、準最適解を求めることに成功しているといえる。これは、階層化することにより、全体構造を考慮したワーブが求められているためと考えられる。更に、図 10 より、階層化を行った場合ではビーム径を 100 とした場合でも 2000 の場合とほぼ同じ品質を保っているといえることがわかる。

図 11 に、すべての文字画像についての階層化なし、

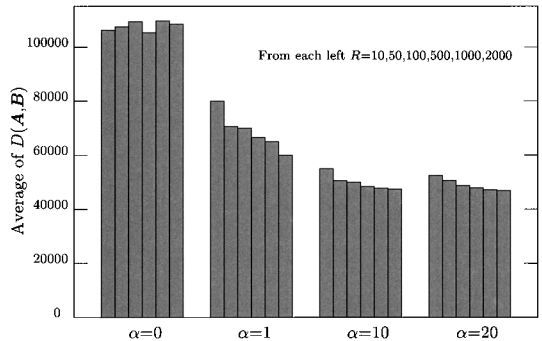


図 11 α の値と目的関数 $D(A, B)$ の平均値

Fig. 11 Average distance.

階層化あり ($\alpha = 1, 10, 20$) の場合のビーム径を 10, 50, 100, 500, 1000, 2000 としたときの目的関数の平均を示す。この値が最適解未満になることはないの、小さいほど最適解に近い解を求めていることにな

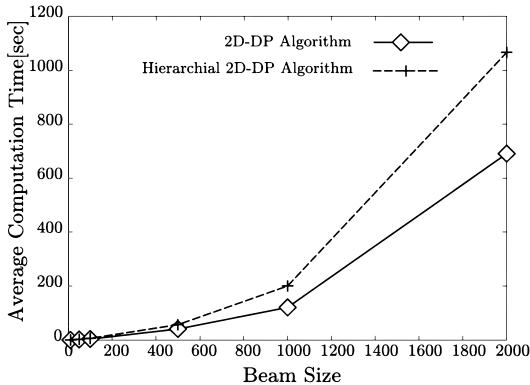


図 12 平均計算時間
Fig. 12 Average computation time.

る。しかし、 α をこれ以上大きくしても目的関数の値はあまり変わらない。また、 α の値が大きいと、目的関数の平均値が減少している。これにより、階層化し、解像度の低い画像での最適解を利用することの有効性が確認できる。階層化あり ($\alpha = 20$) と、なしの場合の各ビーム径における計算時間の平均を図 12 に示す。図 11, 図 12 より、例えば、階層化なしでビーム径が 2000 の場合で目的関数の平均が約 107980, 計算時間の平均が約 691 秒。階層化ありでビーム径が 100 の場合で目的関数の平均が 47993, 計算時間が約 6 秒であった。したがって、階層化を用いることにより約 100 分の 1 の時間で約半分の距離の解を求めることに成功していることがわかる。これらから、階層化により、ビーム径を小さくすることができ、結果として計算量が大きく削減されたことになる。

6. む す び

本研究では、既存の 2 次元 DP アルゴリズムをもとに性能を向上させるためのアルゴリズムを提案した。まず、完全な単調連続性を満たすよう反転防止処理を行った。その後、ワープの集中といった、制約条件の局所性や、ビームサーチによる最適解からのずれといった問題を解決するため、アルゴリズムを階層化し、解像度の低い画像のマッチング結果を利用する手法を提案した。そして、実験により、画素反転防止処理の効果と、大局的構造を考慮したアルゴリズムの性能の評価を行った。

手書き文字マッチング実験 1 では、画素反転防止処理を用いることにより、計算量をほとんど増加させることなく、完全に画素反転を防止することができた。

実験 2 では、階層化手法を用いることにより、不自然な変形や、ワープ先の集中といった問題を回避し、全体構造を考慮した最適解に近い解が得られることが確認できた。更に、ビーム径を下げてその品質を保つことができ、従来法に比べ、約 100 分の 1 の計算時間で、より高品質なマッチングを実現した。

また、文献 [9] における、ピクセルワイズ DP アルゴリズムにおいても、反転防止、階層化は利用可能と考えられ、更なる高性能化が期待される。

今後の課題として、2 次元隠れマルコフモデルへの拡張、文字認識、顔画像認識への応用などが挙げられる。

文 献

- [1] E. Levin and R. Pieraccini, "Dynamic planar warping for optical character recognition," Proc. IEEE International Conference on Acoustic, Speech & Signal Processing, vol.3, pp.149-152, March 1992.
- [2] R.K. Moore, "A dynamic programming algorithm for the distance between two finite areas," IEEE Trans. Pattern Anal. & Mach. Intell., vol.PAMI-1, no.1, pp.86-88, 1979.
- [3] 内田誠一, 迫江博昭, "動的計画法に基づく単調連続 2 次元ワープ法の検討," 信学論 (D-II), vol.J81-D-II, no.6, pp.1251-1258, June 1998.
- [4] S. Kuo and O.E. Agazzi, "Keyword spotting in poorly printed documents using pseudo 2-D hidden Markov model," IEEE Trans. Pattern Anal. & Mach. Intell., vol.16, no.8, pp.842-848, Aug. 1994.
- [5] S. Eickeler, S. Muller, and G. Rigoll, "High quality face recognition in JPEG compressed images," Proc. IEEE International Conference on Image Processing, 25PP6B.5, 1999.
- [6] F. Samaria and S. Young, "HMM-based architecture for face identification," Image and Computer Vision, vol.12, pp.537-583, Oct. 1994.
- [7] A.V. Nefian and M.H. Hanes III, "Hidden Markov models for face recognition," Proc. IEEE International Conference on Acoustic., Speech & Signal Processing, vol.5, pp.2721-2725, May 1998.
- [8] H. Sakoe and S. Chiba, "Dynamic programming algorithm optimization for spoken word recognition," IEEE Trans. Acoust. Speech & Signal Process., vol.ASSP-26, no.1, pp.43-49, 1978.
- [9] S. Uchida and H. Sakoe, "An efficient two-dimensional warping algorithm," IEICE Trans. Inf. & Syst., vol.E82-D, no.3, pp.693-700, March 1999.
(平成 13 年 8 月 22 日受付, 14 年 1 月 28 日再受付)



勝股 充 (正員)

平 12 名工大・工・知能情報システム卒 . 同大学院博士前期課程了 (電気情報工学専攻). 在学中, 画像処理の研究に従事. 現在 (株) ソニー勤務 .



鈴木 尚

平 10 名工大・工・電気情報卒 . 平 12 同大学院博士前期課程了 (電気情報工学専攻). 在学中, 画像処理の研究に従事. 現在 (株) 日本 IBM 勤務 .



徳田 恵一 (正員)

昭 59 名工大・工・電子卒 . 平 1 東工大大学院博士課程了 . 同年東工大電気電子工学科助手 . 平 8 名工大知能情報システム学科助教授 . 工博 . 音声分析・合成・符号化・認識, マルチモーダルインタフェースの研究に従事 . 平 13 電気通信普及財団賞, 平 13 本会論文賞, 猪瀬賞各受賞 . 日本音響学会, 人工知能学会, 情報処理学会, IEEE, ISCA 各会員 .



北村 正 (正員)

昭 48 名工大・工・電子卒 . 昭 53 東工大大学院博士課程了 . 同年東工大精密工学研究所助手 . 昭 58 名工大・工・電子工学科講師 . 昭 59 同助教授 . 平 7 名工大知能情報システム学科教授 . 工博 . 音声情報処理, マルチメディア情報処理の研究に従事 . 日本音響学会, 情報処理学会, IEEE, ISCA 各会員 .