

# 2次元ワープを用いた顔画像処理

内田誠一，松本直樹，迫江博昭 (九州大学)

Two-Dimensional Warping for Face Image Matching

Seiichi Uchida, Naoki Matsumoto, Hiroaki Sakoe (Kyushu University)

## Abstract

Two-dimensional warping, or elastic image matching is a deformation-tolerant image matching technique and has been applied to various problems in pattern recognition, computer vision, and image processing areas. In this paper, a two-dimensional warping algorithm for face images is investigated. The algorithm is developed using dynamic programming and therefore can provide globally optimal matching between a pair of face images. Several complexity reduction techniques for practical use are also investigated.

キーワード: 顔画像, 2次元ワープ法, 弾性マッチング, 動的計画法  
(Face image, elastic matching, two-dimensional warping, dynamic programming)

## 1 まえがき

画像の2次元ワープ法(弾性マッチング法, 弾性テンプレート法) [2, 3, 6, 8] は, 画像パターン認識問題の基本的な道具の1つである. その実体は2画像間の非線形2次元-2次元写像(2次元ワープ)の最適化アルゴリズムであり, 直感的には一方の画像をゴム膜的に変形させてもう一方の画像との最大一致を図る方法として説明される. 2次元ワープによる画素対応の下での2画像の類似度(もしくは相違度)は, 理想的には変形不変量となるので, それを識別関数値として利用することで非線形変形に対して頑強な認識系を実現できる. 一方, 2次元ワープ法には構造解析手法としての側面もある. すなわち, 一方の画像をモデル, もう一方の画像を構造未知の入力画像として考えれば, 2次元ワープによる画素対応関係を介して入力画像の構造を把握できる.

2次元ワープ法のこれらの特徴は, 顔画像処理において好都合である. まず非線形変形に対する頑強性は, 表情変化や向き変化といった複雑な変形が発生する顔画像に適している. 実際, 顔認識手法のもう1つの主流である固有顔法よりも高い認識精度を発揮しうることが報告されている [15]. 次に構造解析手法であることは, 入力顔画像中の造作(目, 鼻, 口など)の位置関係や変形状況の同定に有効である. この原理を利用した顔画像解析手法は既に1970年代にFischlerら [1]により提案されており, 最近でも表情認識に応用されている [5].

2次元ワープ法を顔画像のマッチングに利用する際の

第一の検討課題は, 2次元ワープ最適化問題の定式化である. 具体的には, ワープの制御点, 制御点間の位置関係制約, 画素特徴量, および類似度/相違度を評価する目的関数などの設定が検討事項となる. ここで制御点とは, その対応位置が最適化変数となる画素のことである. 従来法の多くは, Dynamic Link Architecture [4]に代表されるように, 一方の顔画像上から目尻や眼球中心など幾つかの特徴点を選定しておき, それらを制御点としていた. すなわち, すべての画素ではなく, 選ばれた幾つかの画素(およびその近傍)の特徴のみを利用して, いわば粗な点群と点群のマッチングを行うものであった.

第二の検討課題は, こうして定式化された最適化問題の解を探索するアルゴリズムの開発である. 従来, 顔画像マッチングで用いられていたアルゴリズムの多くは局所的かつ逐次的な探索処理を採用していた. 例えば, 各制御点毎に独立にその対応先を反復更新していた. こうした局所最適解を求めるアルゴリズムは, 2次元的最適マッチング問題の計算量的困難性もしくは複雑さを回避したいという消極的な理由で採用されたものであり, 理想的には大局的最適解を与えるものが望ましい.

本稿では2次元ワープ法として, 単調連続2次元ワープ法 [10, 11]を出発点として採用する. この手法は, (i)画像のすべての画素を制御点とすることで, 完全に面的なマッチングを実現, (ii)動的計画法(DP)に基づくアルゴリズムにより大局的最適ワープを求めることが可能, という従来の顔画像マッチング法には無かった特徴を有する. 前者については, 特徴点を予め検出する必要がなく

なる，画像中のすべての画素特徴を利用できる，位相保存性が自然にワーブモデルに組み込める，という利点が生ずる．後者については，理論的興味だけではなく，得られるワーブの精度を向上できるという点で実用的にも大きな意味がある．しかし，大局的最適ワーブを求める場合は，計算量的な困難性に直面する．実際，少くとも  $64 \times 64 \sim 100 \times 100$  程度のサイズが必要な顔画像について，それらに最適な単調連続2次元ワーブを求めるのは，事実上不可能となる．そこで本稿では，この単調連続2次元ワーブ法を出発点とし，それにワーブの区分線形化を適用することで計算量の低減を図る．さらにその際に起る変形追従能力（ワーブの自由度）の低下を回避するために，幾つかの工夫を提案・導入する．以上をまとめると，顔の面的な弾性マッチング問題に対し，実行可能な程度の計算量と十分な自由度を持つワーブ法を実現することが本稿の目的である．

顔画像マッチングにおいては，精度向上のために画素特徴量の工夫や顔独特の事前知識の活用が為されるのが一般である．特徴量として代表的なものは Zernike Moment [7, 9] と Gabor jet [4, 14] である．一方，顔画像特有の事前知識としては，各造作毎の局所的な特徴および相互位置制約がある．前述のように本稿では， $100 \times 100$  程度の顔画像について計算時間および自由度の点で実用的に動作するワーブ法の実装を第一目的とする．従って，これら精度向上のための工夫については次ステップの課題として考えることとする．

## 2 単調連続2次元ワーブ法

本節では，本稿で提案する顔画像マッチングの基本となる単調連続2次元ワーブ法 [10] について，その定式化とアルゴリズムについて述べる．また，その計算量が画像サイズについて指数オーダーで増加し，そのままでは顔画像マッチングに用いることが困難なことを指摘する．

### 2.1 定式化

マッチングの対象とする2画像パターンをそれぞれ  $A = \{a_{i,j} \mid i, j = 1, 2, \dots, N\}$ ， $B = \{b_{x,y} \mid x, y = 1, 2, \dots, N\}$  と表記する．ここで  $a_{i,j}$ ， $b_{x,y}$  はそれぞれ  $A$  と  $B$  の画素  $(i, j)$ ， $(x, y)$  における特徴ベクトルである．このとき  $A$  から  $B$  への2次元-2次元写像，すなわち2次元ワーブは， $2N^2$  次元の整数値ベクトル  $F = ((x_{1,1}, y_{1,1}), \dots, (x_{i,j}, y_{i,j}), \dots, (x_{N,N}, y_{N,N}))$  とし

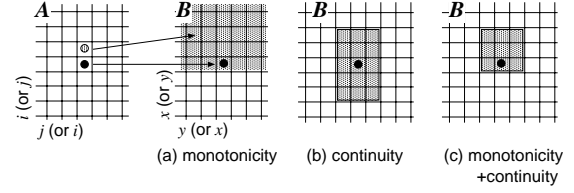


図 1: 単調連続性条件．

Fig. 1 Monotonicity and continuity constraints.

```

/* Initialization */
1 for all  $f_1 \in \Omega_1$  do  $g_1(f_1) = d(f_1)$ 
/* DP-recursion */
2 for  $i := 2$  to  $N$  do
3   for all  $f_i \in \Omega_i$  do
4      $g_i(f_i) = d_i(f_i) + \min_{f_{i-1} \in \Phi(f_i)} g_{i-1}(f_{i-1})$ 
/* Termination */
5  $D(A, B) = \min_{f_N \in \Omega_N} g_N(f_N)$ 

```

図 2: DP アルゴリズム．

Fig. 2 DP algorithm.

て表現できる．ここで  $(x_{i,j}, y_{i,j})$  は画素  $(i, j)$  のワーブ関数  $F$  による像を表している．

2次元ワーブ法の目的は，変数ベクトル  $F$  を制御して， $B_F = \{b_{x_{i,j}, y_{i,j}} \mid i, j = 1, 2, \dots, N\}$  を  $A$  に最も近づけることである．これは目的関数

$$J_{A,B}(F) = \sum_{i=1}^N \sum_{j=1}^N \delta_{i,j}(x_{i,j}, y_{i,j}) \quad (1)$$

の最小化問題として定式化できる．ここで  $\delta_{i,j}(x, y)$  は2つの画素特徴ベクトル  $a_{i,j}$ ， $b_{x,y}$  の間の距離を与える関数である．以下，特に断りの無い限り，その距離として  $L_1$  距離を用いる．

目的関数 (1) の最小値を  $D(A, B)$  とする．すなわち，

$$D(A, B) = J_{A,B}(F_{\text{opt}}) = \min_F J_{A,B}(F). \quad (2)$$

ここで  $F_{\text{opt}}$  は (1) を最小化するワーブ関数  $F$  である．最小値  $D(A, B)$  はワーブにより歪みを除去したあとの2画像間  $A, B$  間の距離である．また  $A$  と  $B_{F_{\text{opt}}}$  との単純重ね合わせ距離として見ることもできる．

極端でない変形の多くは，パターン有位相を保存した変形である．そこで，補償可能な変形を位相保存性を満たすものに限定する．具体的には目的関数 (1) の最小化においてワーブ関数に図 1(c) の単調連続性制約 [10] を課

する．この制約条件はすべての4隣接画素間のワーブに課されるもので，上下左右関係の保存制約(単調性，同図(a))および近傍関係の保存制約(連続性，同図(b))を合成したものである．なお本稿では， $A$ の周囲4边上の画素は $B$ の対応する边上に写像されるという境界条件も併せて課する．以下ではこの単調連続性制約下での2次元ワーブ法を単調連続2次元ワーブと呼ぶ．

変位がある範囲内であることが既知の場合，以下の最大許容変位制約も有効である．

$$|i - x_{i,j}| \leq w, \quad |j - y_{i,j}| \leq w \quad (3)$$

ここで $w$ は許容最大変位を表す正定数である．

## 2.2 動的計画法による解法

本節では前節までに定式化された単調連続2次元ワーブの最適化アルゴリズムについて述べる．まず，画像 $A$ を各行(もしくは列)毎に分解し，それに応じてワーブ $F$ も分解する．ベクトル $f_i$ を $A$ の第 $i$ 行のワーブ $((x_{i,1}, y_{i,1}), \dots, (x_{i,j}, y_{i,j}), \dots, (x_{i,N}, y_{i,N}))$ とする．この部分ワーブの系列 $f_1, \dots, f_i, \dots, f_N$ によりワーブ $F$ が定まる．

単調連続性条件は，4隣接画素間の制約であるから，部分ワーブ $f_i$ そのものに対する制約(行内制約)と， $f_i$ と $f_{i-1}$ の関係に対する制約(行間制約)に分解できる．部分ワーブ $f_i$ のうち行内制約を満たすものの集合を $\Omega_i$ と表す．一方，ある $f_i$ について，行間条件を満たすものの集合を $f_{i-1}$ の集合を $\Phi(f_i)$ と表す．なお最大許容変位制約(3)は行内制約に含まれる．

こうしたワーブ関数と制約条件の分解の結果，単調連続2次元ワーブ決定問題は， $f_i \in \Omega_i$ かつ $f_{i-1} \in \Phi(f_i)$ の系列 $f_1, \dots, f_i, \dots, f_N$ の最適化問題として考えることができる．その解は図2の動的計画(DP)アルゴリズムにより効率的に求めることができる．同図ステップ4がDP漸化式である．また， $d_i(f_i) = \sum_{j=1}^N \delta_{i,j}(x_{i,j}, y_{i,j})$ である．なお，構造解析処理や画像 $B_{F_{opt}}$ のために具体的に最適ワーブ $F_{opt}$ が必要となる場合は，各DP漸化式計算時に最小値を与えた $f_{i-1}$ をテーブルに記憶しておき，上記のアルゴリズム終了後にそのテーブル上でバックトラック処理を行う．

## 2.3 計算量

可能な部分ワーブの集合のサイズ $|\Omega_i|$ は $N$ に対して指数オーダーになるため，図2のDPアルゴリズムの計算量も指数オーダーとなる．このため，現在の計算機能力で

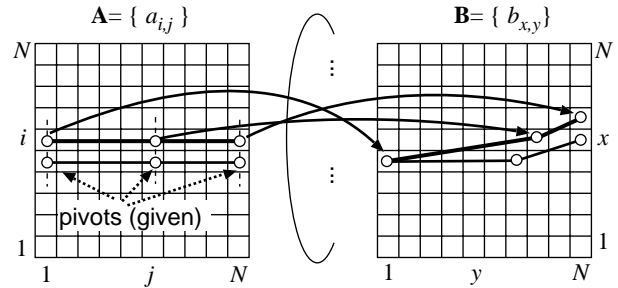


図 3: 区線線形 2次元ワーブ ( $K = 3$ ) .

Fig. 3 Piecewise linear 2D warping ( $K = 3$ ).

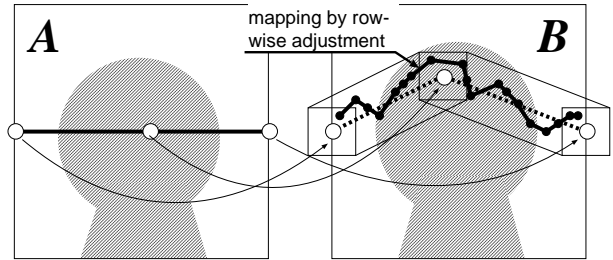


図 4: 行単位補正処理 .

Fig. 4 Row-wise adjustment.

は $N = 8 \sim 10$ 程度が限界であり， $N = 64$ 程度以上のサイズを要する顔画像を対象とするのは不可能である．

計算量低減のために，これまでもアルゴリズムの再構成[11]，ビームサーチの導入[10, 11]が検討されている．アルゴリズムの再構成とは，ワーブの分解形式を変えることで，図2のDPアルゴリズムと同じ最適解を与えながらも計算量が少なく済むDPアルゴリズムを構成する方法である．しかし，問題の本質的な困難性から， $N$ に関する多項式アルゴリズムは得られていない．もう一方のビームサーチとは，DP計算の各段階において， $|\Omega_i|$ 個の部分ワーブ $f_i$ のうちで最適解が得られる可能性の高いものだけを累積コスト値 $g_i(f_i)$ を基準に選出し，それ以外は以後の探索から除外するものである．選出基準を厳しくすれば，相当の計算量低減効果が見込めるが，同時に探索途中で最適ワーブを構成する $f_i$ が除外される可能性も高くなり，結果として最適解より大きく逸脱した局所解が求まり易くなる．顔画像のように $N \geq 64$ となると，元の探索幅 $|\Omega_i|$ は非常に大きくなり，実用上は選出基準を相当厳しくしなければならない．よって，局所解の危険性を考えると，ビームサーチ単独での効率化には限界があると言える．

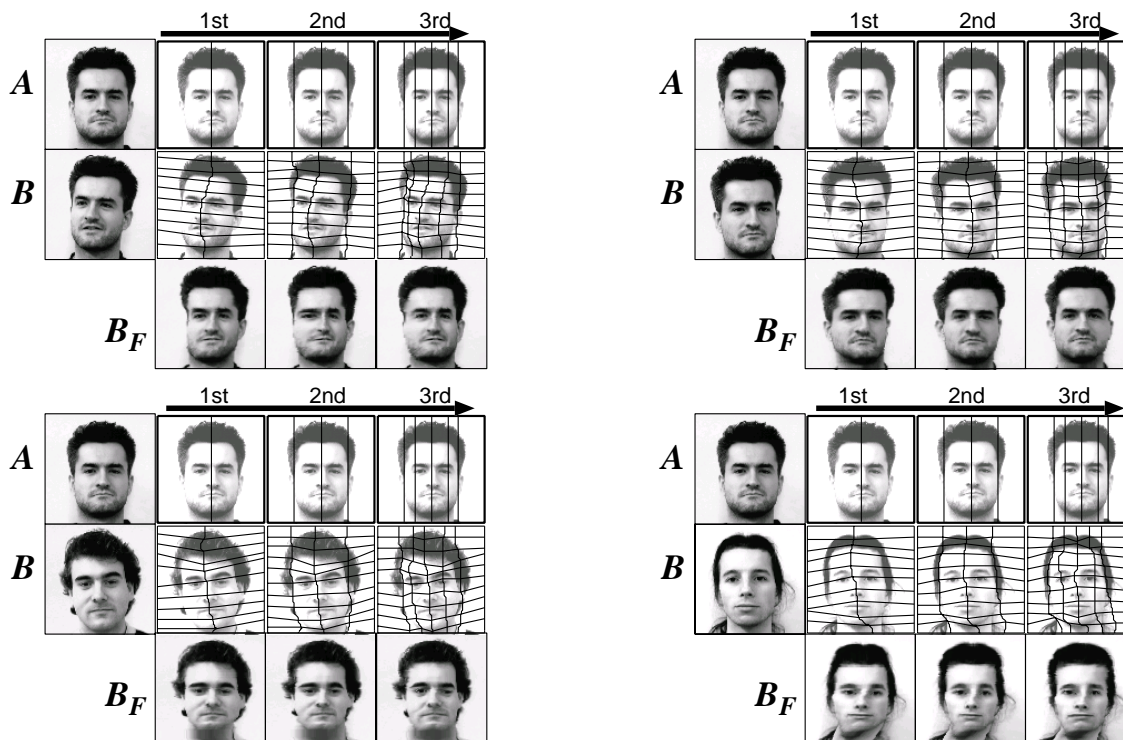


図 5: 粗密探索によるマッチングの結果．見易さのため，垂直方向についてはワーブメッシュを粗く描画した．

Fig. 5 Results of face image matching by coarse-to-fine piecewise linear 2D warping.

### 3 区分線形化による計算量の低減

#### 3.1 2次元ワーブの区分線形化 [12]

本節では，ワーブの区分線形化について述べる．その基本となる考え方は，制御変数  $(x_{i,j}, y_{i,j})$  を持つ画素  $(i, j)$  を少数に限定し，他の画素についてはその対応先を補間により決定することにある．こうして制御変数を全ての行で一様に減らすことで，計算量を指数的に減らすことができ，大域的最適解を得ることが可能となる．

図3に区分線形化された2次元ワーブを示す．画像  $A$  の各行には，ピボットと呼ばれる制御点を両端を含めて  $K$  個設定し，その  $B$  上での像を制御変数としてワーブの最適化を行う．ピボット以外の画素の像は，ピボットの像の線形補間により定める．各行の制御変数数は  $O(K)$  であるので，時間計算量は  $K$  に関する指数オーダーであり， $N$  については多項式オーダーとなる．具体的には  $O((9K + N)NW^{2K})$  となる．よって  $K$  が小さければ  $N$  が大きくても十分実行可能となる．

区分線形化によって，ワーブの自由度は水平方向において大きく低下する．顔はある程度の構造を保ったまま変

形するので，この自由度低下の影響は劇的なものではないが，それでも  $K \leq 5$  程度では不自然な変形として  $B_F$  に表出する場合がある．そこで以下では，計算量を現実的な範囲に抑えたまま自由度の低下に対処するための，2つの手法について述べる．

#### 3.2 粗密探索

本節では粗密探索(階層的探索)を用いて段階的にワーブを求めることで，計算量の増加を抑えることを考える．粗密探索には，何の密度を変化させるかにより様々な実現形態がある [6]．ここでは，ワーブ関数の密度を変化させる方法を採用する．

具体的には，ピボット数を  $K_1, \dots, K_t, \dots, K_T$  と増加させながら区分線形2次元ワーブを繰り返し適用することを考える．第  $t$  段のワーブを  $x_{i,j}^t, y_{i,j}^t$  のように表すと，反復過程においては次のような制約を課する．

$$|x_{i,j}^t - x_{i,j}^{t-1}| \leq w_t, \quad |y_{i,j}^t - y_{i,j}^{t-1}| \leq w_t \quad (4)$$

ここで整合窓幅  $w_t$  は  $t$  に従って徐々に小さくするものとする．全  $T$  回の反復により，最大許容変位は  $w_{\text{tot}} = \sum_{t=1}^T w_t$

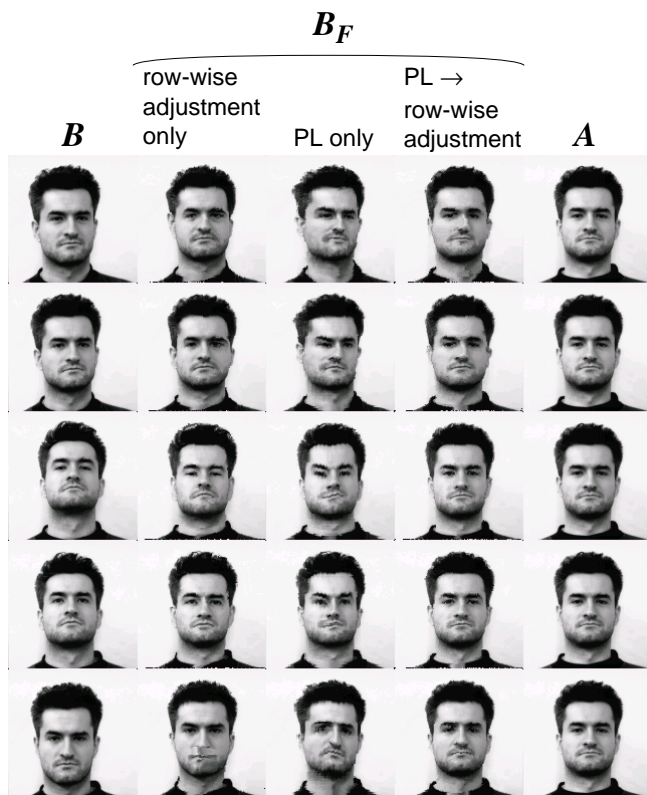


図 6: 行単位補正処理の効果。

Fig. 6 Effect of row-wise adjustment.

となる。区分線形 2 次元ワープの計算量から明らかなように、粗密探索法による計算量は、ピボット数を  $K_T$ 、最大許容変位量を  $w_{\text{tot}}$  として一度に区分線形 2 次元ワープを行う場合の計算量に比べて、大幅に少なくて済む。

### 3.3 行単位補正処理

計算量を大きく増加させることなくワープの自由度を高めるもう 1 つの方法として、各行独立にワープ関数  $F$  を事後補正することを考える。具体的には、まず区分線形 2 次元ワープを行い、そうして得られた第  $i$  行のワープ結果を中心として、今度は線形補間無しで第  $i$  行のワープを再び最適化する (図 4)。自由度の高いワープを事後処理として施すという考え方は粗密探索と同じであるが、行単位補正ではワープのモデルそのものも変える点に違いがある。行単位補正処理にも DP が使え、全  $N$  行分の再最適化処理を行っても、区分線形 2 次元ワープに比べて十分少い計算量で済む。

行単位補正は効果的であるが、あまり大きな補正を許すと、隣接行のワープ間に交差が発生するという問題が

ある。この交差は顔パターンの構造を部分的に破壊していることになるので、それを避けるためにも、この行単位補正単独で利用するよりも、あくまで区分線形 2 次元ワープ結果の詳細な補正という形で利用すべきと言える。

## 4 実験

本節では、前節で述べた区分線形 2 次元ワープを基本とした粗密探索および行単位補正を、実際に顔画像マッチングに用いた結果について述べる。実験で用いた顔画像は、Bern 大学によって提供されたデータベース [16] 中のもので、切出しと縮小処理によって  $N = 128$  とした。いずれも、顔中心と画像中心はおよそ一致している。なお、以下の実験結果の図中では濃淡画像のみを示すが、ワープ最適化時にはこの濃淡画像を処理して得たエッジ画像も併用した。すなわち、画素特徴  $a_{i,j}, b_{x,y}$  は濃淡特徴とエッジ特徴よりなる多次元ベクトルとした。

### 4.1 粗密探索の効果

図 5 に、区分線形 2 次元ワープを用いた粗密探索によるマッチングの 4 つの結果を示す。反復回数は 3 回で、それぞれピボット数  $K_t$  を 3, 5, 7、窓幅  $w_t$  を 10, 3, 2 と変化させた。画像  $A$  上でのピボットの位置はワープの精度を左右する重要な要素であるが [13]、本実験では単純にすべての行  $i$  において同じとした。その結果、同図の各結果の最上段に示したように、ピボットは画像  $A$  上で (両端を含め)  $K$  本の垂線を為す。ただし、ピボットは顔中心 ( $t = 1, 2, 3$ )、輪郭付近 ( $t = 2, 3$ ) および眼球中心 ( $t = 3$ ) にあるようにした。本実験では、このピボットの大きかな設定に関してのみ、対象顔パターンに関する事前知識を利用した。

実験では、各反復において、ピボット数を  $K_t$  として区分線形 2 次元ワープを求める際に、ビームサーチを利用し、計算時間の更なる高速化を図った。その結果、全実行時間は PC 上で約 90 秒であった。なお、ビームサーチを用いることを許せば、粗密探索を用いるまでもなく、同程度の計算時間で同様の自由度を持つ区分線形 2 次元ワープ ( $K = K_T = 7, w = w_{\text{tot}} = 15$ ) を 1 回で実行することもできる。しかしその場合、探索空間の絞り込みは粗密探索の場合に比べて  $10^9$  倍程度も厳しくする必要があり、従って局所解に落ち込む可能性は増加する。

図 5 の結果から、粗密探索の反復により、最終的には良好な結果 (すなわち  $A \sim B_F$ ) が得られていることがわか



る。造作の位置に着目すると、造作の相互位置や造作毎の形状といった事前知識を用いながらも関わらず、同一人物でも異なる人物でも、ほぼ同位置に整合できていることがわかる。このことは各結果中段に示したワープ関数からもわかる。ただし「見え」については一部に不自然な感じも残った。これは (i) 顔画像の場合、画像的には数画素程度の違いであっても「見え」においては非常に大きな違いとして感じられてしまうことと、(ii) 1つの顔画像だけから異なる表情・角度や別人物の顔画像を合成することの困難さ、が原因と考えられる。

## 4.2 行単位補正の効果

図6に行単位補正の結果を示す。本実験では、 $K = 3$  (ピボットは各行両端と中央に配置) という、非常に自由度の低い区分線形2次元ワープを採用した。よって、区分線形2次元ワープ単独の結果 (同図各行中央) には、A (右端) とのやや大きな差異が見られる。それに対し行単位補正を行った結果 (右から2番目) は、A に十分近づいており、その効果を示している。なお、補正に要した時間はPCで約2秒程度であった。

同図に、区分線形2次元ワープを介さずに行単位補正法のみを用いて直接2次元ワープを行った結果も示した (左から2番目)。十分な結果が得られているものもあるが、最下段の結果を見ると、鼻と口の間の部分が不自然に歪んでいるのがわかる。これは、3.3でも指摘した隣接行のワープ間の交差によるものであり、各行独立な最適化に基づく行単位補正を単独で用いることの限界を示している。

## 5 むすび

動的計画法 (DP) に基づき、画像上の全画素について対応関係を最適化する2次元ワープ (弾性マッチング) アルゴリズムについて、その顔画像処理における利用を検討した。具体的には、出発点となる単調連続性2次元ワープ法について述べ、その計算量が膨大となる点に触れ、解決策として区分線形化による効率化について述べた。さらに粗密探索および行単位補正処理によるワープ自由度の補強について述べた。実験の結果、現実的な時間で、高精度なマッチング結果が得られるとの見通しを得た。

謝辞 本研究においては、Bern大学提供の顔画像データベースを利用させて頂いた。ここに感謝する。なお、本研究の一部は文部科学省科学研究費補助金 (No.14780293) によった。

## 参考文献

- [1] M. A. Fischler and R. A. Elschlager, "The representation and matching of pictorial structures," IEEE Trans. Computers, Vol. C-22, No. 1, pp. 67-92, 1973.
- [2] C. A. Glasbey and K. V. Mardia, "A review of image-warping methods," J. Applied Statistics, vol. 25, no. 2, pp. 155-171, 1998.
- [3] A. K. Jain, Y. Zhong, and M.-P. Dubuisson-Jolly, "Deformable template models: A review," Signal Processing, vol. 71, no. 2, pp. 109-129, 1998.
- [4] M. Lades, J. C. Vorbruggen, J. Buhmann, J. Lange, C. Malsburg, P. P. Wurtz and W. Konen, "Distortion invariant object recognition in the dynamic link architecture", IEEE Trans. Computers, Vol. 42, No. 3, pp. 300-311, 1993.
- [5] A. Lanitis, C. J. Taylor, and T. F. Cootes, "Automatic interpretation and coding of face image using flexible modes," IEEE Trans. PAMI, Vol. 19, No. 7, pp. 743-756, 1997.
- [6] H. Lester and S. R. Arridge, "A survey of hierarchical non-linear medical image registration," Pattern Recog., vol. 32, no. 1, pp. 129-149, 1999.
- [7] A. Ono, "Face recognition with Zernike moments," Trans. IEICE, Vol. J85-D-II, No. 7, pp. 1149-1156, 2002. (in Japanese)
- [8] A. Redert, E. Hendriks and J. Biemond, "Correspondence estimation in image pairs," IEEE SP Mag., vol. 16, no. 3, pp. 29-46, 1999.
- [9] M. Sugimura, Y. Iiguni, and N. Adachi, "A 2-dimensional dynamic programming for image matching with Zernike moments," Trans. IEICE, Vol. J80-D-II, No. 1, pp. 101-108, 1997. (in Japanese)
- [10] S. Uchida and H. Sakoe, "Monotonic and continuous two-dimensional warping based on dynamic programming," Trans. IEICE, Vol. J81-D-II, No. 6, pp. 1251-1258, 1998. (in Japanese)
- [11] —, "An efficient two-dimensional warping algorithm," IEICE Trans. Info & Syst., Vol. E82-D, No. 3, pp. 693-700, 1999.
- [12] —, "Piecewise linear two-dimensional warping," Trans. IEICE, Vol. J83-D-II, No. 12, pp. 2622-2629, 2000. (in Japanese)
- [13] —, "A priori knowledge free piecewise linear two-dimensional warping," Journal of Institute of Image Info. & Television Eng., Vol. 55, No. 12, pp. 1643-1649, 2001. (in Japanese)
- [14] L. Wiskott, J. -M. Fellous, N. Krüger, and C. Malsburg, "Face recognition and gender determination," Proc. Int. Workshop on Automatic Face and Gesture Recog., pp.92-97, 1995.
- [15] J. Zhang, Y. Yan, and M. Lades, "Face recognition: eigenface, elastic matching, and neural nets," Proc. IEEE, Vol.85, No.9, pp.1423-1435, 1997.
- [16] <http://www.iam.unibe.ch/~fkiwww/staff/achermann.html>