区分線形2次元ワープによる手書き文字の変形吸収の試み

内田 誠一　　モハマッド アサッド ロニー　　石田 敏之　　迫江 博昭

九州大学大学院システム情報科学研究科
〒812-8581 福岡市東区箱崎 6-10-1
TEL:092-642-3888
E-mail: {uchida, ronee, tishida, sakoe}@human.is.kyushu-u.ac.jp

あらまし　本報告では，区分線形 2 次元ワープによる手書き文字歪みの吸収を検討する．区分線形 2 次元ワープは，画像の弾性マッチングを目的とした，区分的に線形化された 2 次元-2 次元写像の一種である．画像の各行は，その上に設定された節点で線形伸縮を伴いながら 2 次元的に屈曲され，もう一方の画像上に折れ線として重ね合わせられる．この仕組みにより文字形状の 2 次元的な手書き歪みを吸収する．文字画像では多くの場合，節点数は少なくて済むと考えられるので，ワープは現実的な計算量で実行できる．ETL6 の英大文字 26 文字種を対象とした認識実験を行い，1 次元弾性マッチングの組合せに基づく従来法よりも高い，97.5%の認識率を得た．また，誤認識の原因についても考察した．

キーワード　手書き文字認識　変形　弾性マッチング　2 次元ワープ　動的計画法

Experimental Study on Handwritten Character Recognition Using Piecewise Linear Two-Dimensional Warping

Seiichi UCHIDA, Mohammad Asad RONEE, Toshiyuki ISHIDA and Hiroaki SAKOE

Graduate School of Information Science and Electrical Engineering, Kyushu University
6-10-1 Hakozaki, Higashi-ku, Fukuoka-shi, 812-8581 Japan
TEL:+81-92-642-3888
E-mail: {uchida, ronee, tishida, sakoe}@human.is.kyushu-u.ac.jp

Abstract　In this paper, a character recognition experiment using piecewise linear two-dimensional warping, a dynamic programming-based elastic image matching technique, is reported. The present technique requires far less computations than the previous methods, such as monotonic and continuous two-dimensional warping, since line segment is used instead of pixel as the unit of mapping. This consideration does not decrease the recognition accuracy, as shown by the experimental result. Additional results also show the superiority of the present technique over template matching and two-dimensional warping techniques based on combinations of orthogonal one-dimensional warpings.

key words　off-line handwritten character recognition, elastic matching, two-dimensional warping, dynamic programming

1　まえがき

2 次元弾性マッチング [1]–[4] は，2 画像の最大一致を与える 2 次元-2 次元写像 (2 次元ワープ) を介して 2 つの画像パターンの類似度もしくは相違度を評価する方法であり，動画像処理や画像認識の分野において広く研究がなされている．文字認識の分野においても，文字の手書き歪みの吸収を目的として，その利用が検討されている．2 次元弾性マッチングは 2 次元ワープの最適化問題として定式化されるため，用いる最適化法によりマッチング法としての性質が左右される．従来用いられてきた最適化法としては，局所的な摂動法 (ずらし) [5]–[7] や反復解法 [8]–[13]，および動的計画法 (dynamic programming : DP) がある．このうち DP は，i) 解の最適性が保証される，ii) 目的関数が微分可能でなくてもよい，iii) 様々な制約条件を取り扱うことができる，iv) 演算誤差の集積による不安定性がない，など，他の方法には無い特長を持つ．

DP ベースの 2 次元ワープ法は 2 つの系統に大別できる．第一は，DP による 1 次元-1 次元弾性マッチング (1 次元ワープ) の組合わせに基づく手法である．中野ら [14], [15] は，1 次元ワープを水平，垂直方向の濃淡値の周辺分布に対し独立に行う手法を提案した．田中ら [16] および津雲 [17] は文字画像の行毎に独立に 1 次元ワープを行う手法を提案している．また，Levin ら [19] および Kuo ら [18] は水平方向の 1 次元ワープのループの内側で垂直方向の 1 次元ワープを行う手法を提案している．画像の一辺の長さを N とすると，これらの手法はいずれも$O(N^2) \sim O(N^4)$ 程度の計算量で済むので実用的であるが，回転など 2 次元的な文字変形を吸収できない点で検討の余地が残されている．

第二は，DP を真に 2 次元的に拡張し，1 次元ワープの組合せに基づく手法では対処できない変形も吸収できるようにした手法 [19]–[22] である．しかし，これらの手法においては，計算量が画像サイズに関して指数的に増加する DP アルゴリズムしか知られておらず，実行にはビームサーチなどの近似解法の導入が必要である．さらに，変形吸収能力，すなわちワープ自由度が相当に高いので，不正解カテゴリの標準パターンが過度に変形されて入力パターンに近づきすぎるという，いわゆる過変形の問題が顕著になる．例えば単調連続 2 次元ワープによる文字認識 [22] では，変形抑制ペナルティの導入や画素特徴量の高次元化により過変形の回避を図っても，依然として過変形が誤認識の原因の半数を占めたと報告されている．

本報告では，区分線形 2 次元ワープ [23] による手書き文字歪みの吸収を検討する．区分線形 2 次元ワープにおいてワープ関数は，画像の各行に設定された数個の節点 (ピボット) の像により制御される．ピボット以外の画素ついてはピボットの像の線形補間により定められる．直観的に言えば，一方の画像の各行はワープにより線形伸縮を伴いながら 2 次元的に折り曲げられ，もう一方の画像に重ね合わせられることになる．このため，すべての画素の像を制御変数とする単調連続 2 次元ワープよりは低い自由度となるが，部分的な回転や平行移動など 1 次元ワープの組合せに基づく手法では対応できない変形でも吸収できる．また，過変形の傾向は単調連続 2 次元ワープより緩和されると考えられる．計算量は，ピボットの個数に関して指数オーダーとなるが，画像サイズ N に関しては多項式オーダーとなる．文字画像では多くの場合，ピボット数は少なくて済むと考えられるので，近似解法を導入しなくても現実的な計算量でワープを実行できる．

本報告は次のように構成される．**2**では，文献 [23] で与えた区分線形 2 次元ワープの定式化と，その DP アルゴリズムを説明する．**3**では，区分線形 2 次元ワープにおいてピボット数を最少にして，計算量を 1 次元ワープの組合せに基づく従来法とほぼ同程度とした特殊形 – ダッチロールワープ – について述べる．**4**では，英大文字 26 文字種を対象とした認識実験の結果について述べる．具体的には，1 次元ワープの組合せに基づく手法や単調連続 2 次元ワープとの比較を通じて，区分線形 2 次元ワープおよびダッチロールワープの有効性を定量的に評価し，また今後の検討課題を明らかにするために誤認識の原因についても考察する．

2　区分線形 2 次元ワープ [23]

2.1　ワープ関数とピボット

2 画像 $\boldsymbol{A} = \{a(i,j) \mid i,j = 1,2,\ldots,N\}$ と $\boldsymbol{B} = \{b(x,y) \mid x,y = 1,2,\ldots,N\}$ との間のワープ $x = x(i,j)$，$y = y(i,j)$ の最適化問題を考える[1]．$\boldsymbol{A}$上のすべての画素のワープによる像を制御変数として，2 次元的な単調性や連続性の制約の下で最適化すると

[1] 本節および次節では図を含めて i,x を画像の垂直方向のインデクス，j,y を水平方向のインデクスとして説明するが，これは逆でもよい．実際，4節では両方の場合についてワープを求めた実験の結果を述べる．

いう方法 [19]–[21] (図 1(a)) は，画像サイズ N に関して指数オーダーの計算量を要し，実行困難である．

区分線形 2 次元ワープ [23] は，図 1(b) に示すように，画像$\boldsymbol{A}$の各行に $K(< N)$ 個の節点 (ピボット) を設定し，その像を制御変数として最適化を行う方法である．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．制御変数の数は計算量に対し指数的に影響するため，すべての画素の像を制御変数とする方法 [19]–[21] に比べて大幅に少ない計算量で済む．

画像$\boldsymbol{A}$の第 i 行の第 k ピボットを $(i, j_{i,k})$，そのワープによる像を $(x_{i,k}, y_{i,k})$ と表す．ピボット間の画素 $(i, j),\ (j_{i,k-1} \le j \le j_{i,k})$ の像は，ピボットの像 $(x_{i,k-1}, y_{i,k-1}), (x_{i,k}, y_{i,k})$ の線形内挿

$$\left.\begin{aligned} x(i,j) &= j' \cdot x_{i,k} + (1-j') \cdot x_{i,k-1} \\ y(i,j) &= j' \cdot y_{i,k} + (1-j') \cdot y_{i,k-1} \end{aligned}\right\} \quad (1)$$

によって定める (図 2) ．ここで $j' = (j - j_{i,k-1})/(j_{i,k} - j_{i,k-1})$ である．なお，画像$\boldsymbol{B}$は非整数座標点において定義されていないため，$b(x(i,j), y(i,j))$ の値は $(x(i,j), y(i,j))$ に最も近い (整数座標) 画素の値とする．

ピボットは謂わば各行の折り曲げ点であるから，その位置は対象とする画像パターンの変形特性を考慮して定める必要がある．逆に言えば，区分線形 2 次元ワープでは，対象 (例えば文字) の変形特性に関する事前知識に応じて，ピボットを設定することで十分な変形吸収能力を確保できる．本報告では，

$$1 = j_{i,1} < j_{i,k-1} < j_{i,k} < j_{i,K} = N \quad (2)$$

$$|j_{i,k} - j_{i-1,k}| \le 1 \quad (3)$$

を満たす範囲で，ピボットの位置を人為的に定めるものとする．これら 2 条件は，ピボットを画像$\boldsymbol{A}$上で K 本の交差しない連続的な軌跡 (ピボット軌跡) として与えることに相当する．

2.2 制約条件

区分線形 2 次元ワープでは，ワープに以下の単調連続性条件を課する．

$$-1 \le \quad j_{i,k} - j_{i-1,k} \quad \le 1 \quad (4)$$

$$0 \le \quad x_{i,k} - x_{i-1,k} \quad \le 2 \quad (5)$$

$$-1 \le \quad y_{i,k} - y_{i-1,k} \quad \le 1 \quad (6)$$

$$y_{i,k-1} < y_{i,k} \quad (7)$$

この単調連続性により，ワープの位相保存性が実現し，画素の位置順の逆転や不自然なとびを回避することができる．また特に断りのない限り，以下の境界条件を課する．

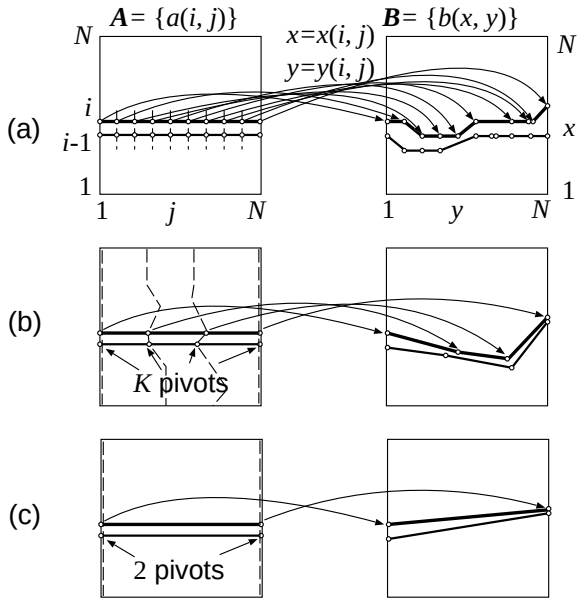

図 1: 単調連続 2 次元ワープ (a) と区分線形 2 次元ワープ (b)，および区分線形 2 次元ワープの特別な場合 (ダッチロールワープ)(c)．

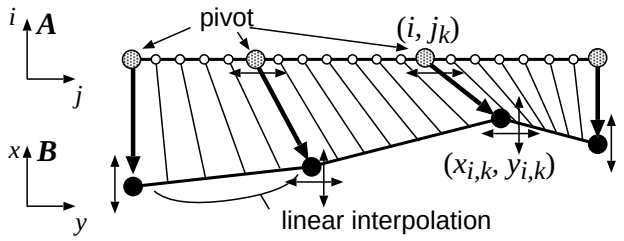

図 2: 第 i 行のワープ

$$x_{1,k} = 1,\ x_{N,k} = N \quad (8)$$

$$y_{i,1} = 1,\ y_{i,K} = N \quad (9)$$

条件 (9) は条件 (2) に対応し，これらにより画像$\boldsymbol{A}$の各行の両端ピボットの像は再び$\boldsymbol{B}$の行の両端上になるよう制限される．

パターンの変形量が少ないと仮定できれば，計算量低減および過変形回避の点で，次の整合窓条件が有効である．

$$|x_{i,k} - i| \le w, \quad |y_{i,k} - j_{i,k}| \le w \quad (10)$$

ここで w はワープの範囲を規定する非負の定数である．

表記の簡単のために，以下では$\boldsymbol{A}$の各行 i のワープを表現する $2K$ 次元ベクトル $[(x_{i,1}, y_{i,1}), \ldots, (x_{i,k}, y_{i,k}), \ldots, (x_{i,K}, y_{i,K})]$ を$\boldsymbol{xy}_K(i)$ もしくは単に

$\boldsymbol{xy}$ と表し，その集合を$\boldsymbol{XY}_K(i)$ と表記する．ここで各$\boldsymbol{xy}$は条件 (7)–(10) を満たす．さらにある$\boldsymbol{xy}$に関して条件 (4)–(6) を満たす$\boldsymbol{xy}_K(i-1) \in \boldsymbol{XY}_K(i-1)$ の集合を$\overline{\boldsymbol{XY}}_K(\boldsymbol{xy})$ と表記する．

2.3 目的関数

第 i 行のワープを

$$d(\boldsymbol{xy}_K(i) \mid i) = \sum_{j=j_{i,1}}^{j_{i,K}} |a(i,j) - b(x(i,j), y(i,j))|$$

により評価する．ここで，右辺の $(x(i,j), y(i,j))$ は (1) 式によるものとする．このとき，区分線形 2 次元ワープの解は，目的関数

$$\sum_{i=1}^{N} d(\boldsymbol{xy}_K(i) \mid i)$$

を制約条件 (4)–(10) の下で最小化する $[\boldsymbol{xy}_K(1), \ldots, \boldsymbol{xy}_K(i), \ldots, \boldsymbol{xy}_K(N)]$ として定義される．

目的関数の最小値 $D(\boldsymbol{A}, \boldsymbol{B})$ は区分線形 2 次元ワープにより 2 画像$\boldsymbol{A}$，$\boldsymbol{B}$をマッチングした時の画像間距離に相当する． 次節の文字認識実験では，この $D(\boldsymbol{A}, \boldsymbol{B})$ を基準とした識別を行う．

2.4 DP アルゴリズム

以上で定義される区分線形 2 次元ワープの最適化問題，すなわち目的関数を最小化する $[\boldsymbol{xy}_K(1), \ldots, \boldsymbol{xy}_K(i), \ldots, \boldsymbol{xy}_K(N)]$ を求める問題を解く DP アルゴリズムの一つを以下に示す．

[1] 初期状態
for all $\boldsymbol{xy} \in \boldsymbol{XY}_K(1)$ **do** $g(\boldsymbol{xy}|1) := d(\boldsymbol{xy}|1)$

[2] DP 漸化式
for $i := 2$ **to** N **do**
for all $\boldsymbol{xy} \in \boldsymbol{XY}_K(i)$ **do**
$g(\boldsymbol{xy} \mid i) := d(\boldsymbol{xy} \mid i) + \min_{\overline{\boldsymbol{xy}} \in \overline{\boldsymbol{XY}}_K(\boldsymbol{xy})} g(\overline{\boldsymbol{xy}} \mid i-1)$

[3] 終了
$D(\boldsymbol{A}, \boldsymbol{B}) := \min_{\boldsymbol{xy} \in \boldsymbol{XY}_K(N)} g(\boldsymbol{xy} \mid N)$

この DP アルゴリズムの時間計算量は，$W = 2w + 1$ とすると，$O(NW^{2K}(9^K + N))$ となる．

3 区分線形 2 次元ワープの特殊形 − ダッチロールワープ − とその拡張

区分線形 2 次元ワープにおいてピボット数を最小とした場合，すなわち $K = 2$ の場合を特に区別して，以下ではダッチロールワープ[2]と呼ぶ (図 1(c))．ピボットは画像$\boldsymbol{A}$の各行の両端に配置され，その$\boldsymbol{B}$上での像も制約 (9) により再び画像の両端となる．ダッチロールワープの計算量は$O(N^4)$ であり，1 次元ワープの組合せに基づくワープ法と同程度である．

ダッチロールワープは垂直 (i) 方向については局所的な傾きや移動を吸収できるが，(1) 式より $y(i,j) = j$ となるので，水平方向については変形吸収能力がない．そこで，これを補うため次の二つの拡張型を併せて検討する．拡張 1 は，水平→垂直の順 (もしくはその逆順) にダッチロールワープを逐次的に 2 回行うものである．拡張 2 は各行のマッチング，すなわち (2.3) 式の計算時において，内部的に j 方向の 1 次元ワープを行う方法である．いずれの拡張でも計算量の大幅な増加は無い．

4 文字認識実験

本節では，区分線形 2 次元ワープ $(K = 3)$，およびダッチロールワープとその拡張型の文字認識における有効性および問題点を実験を通じて確認する．比較のため，単純重ね合わせ，1 次元ワープベースの手法 [14]–[18]，および単調連続 2 次元ワープ [20]–[22] についても同じ試料により実験を行った．

ワープの自由度の認識率への影響に特に注目するため，自由度以外の環境を統一した．このため，従来のワープ法に関しては，用いた特徴量や前処理 (例えばぼかし処理) の点で原論文とは異なることがある．また，ワープの傾斜制限に関しても，これを $0 \leq x(i,j) - x(i-1,j) \leq 2$，$0 \leq y(i,j) - y(i,j-1) \leq 2$ に統一し，さらに (10) 式と同様の整合窓を用いた．識別法も，標準パターン$\boldsymbol{A}$とワープ後の入力パターン$\boldsymbol{B}$ との重ね合わせ距離 ($D(\boldsymbol{A}, \boldsymbol{B})$ に相当) を基準とした最短距離識別に統一した．なお，特例として，単調連続 2 次元ワープに限り，変形抑制ペナルティ (一様性ペナルティ[20]) を使用した場合を検討した．

本節の実験に用いたワープを図 3に一覧する．以下では必要に応じてこの図の記号 P_1–M_2を各手法の略称とする．$\mathsf{P}_1, \mathsf{P}_2$は区分線形 2 次元ワープ $(K = 3)$ であり，P_2は P_1の転置形である．最適化の過程も含めて考えると単調連続 2 次元ワープ以外は水平方向-垂直方向の対称性がないため，実験ではこのように転置した形態も併せて検討した．D_1–D_6はダッチロールワープおよびその拡張 1,2 である．C_1–C_8は 1 次元

[2] ワープの様子がダッチロール状態にある飛行機の翼の動きを連想させることから，このワープ法をダッチロールワープと呼んでいる．

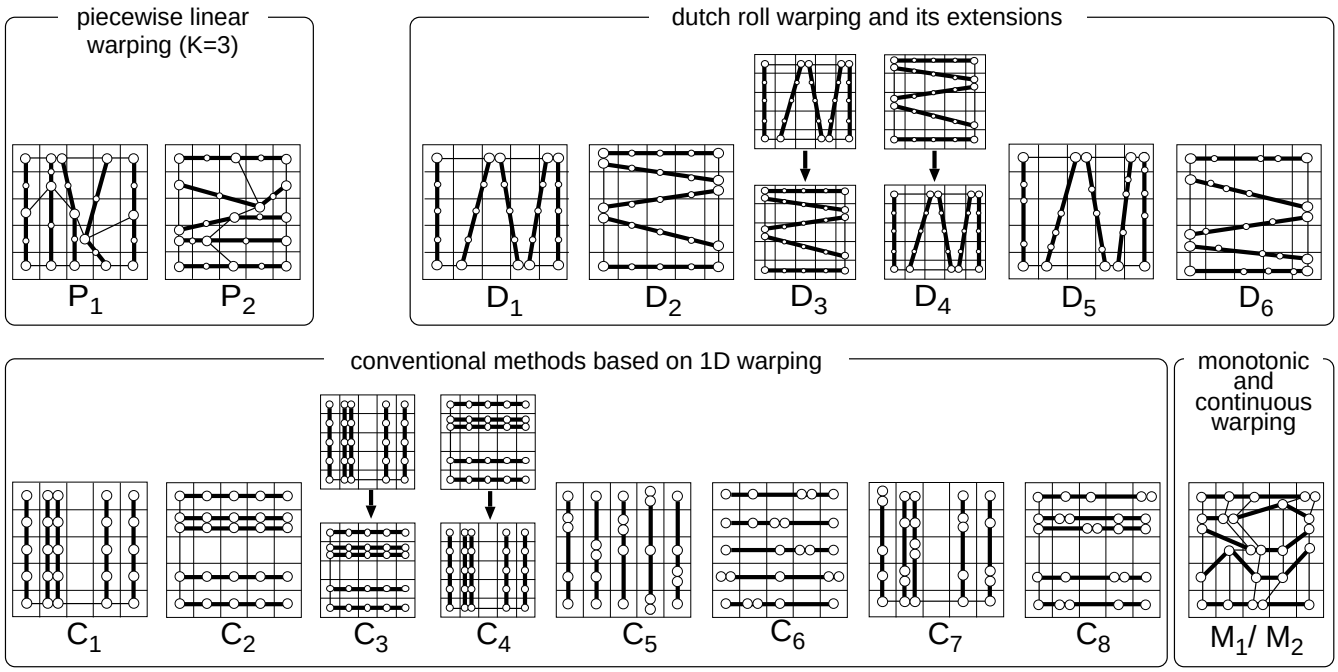

図 3: 実験で用いたワープの種類

ワープの組み合わせに基づく従来法に相当する．ここで C_2, C_4, C_6, C_8はそれぞれ C_1, C_3, C_5, C_7の転置形である．C_1は各行をひとまとまりとして扱う行方向の 1 次元ワープ [14] である．C_3は C_1, C_2を順に行う方法 [15] である．C_5は，各行において内部的に 1 次元ワープを行う方法 [16], [17] である．C_7は C_1と C_5を組み合わせたワープ [18] である．M_1, M_2は単調連続 2 次元ワープであり，M_1はペナルティ使用，M_2は不使用の場合を表す．M_1, M_2については共にビームサーチ (ビーム径$= 1000$) を導入して高速化を図った近似アルゴリズム [21] を用いてワープを求めた．

4.1 実験試料

実験に用いた文字データは電総研提供のデータベース ETL6 中の英大文字 26 文字種各 200 サンプルであり，各文字種につき最初の 100 サンプルを標準パターン ($\boldsymbol{A}$) 作成用，残りの 100 サンプルをテストパターン ($\boldsymbol{B}$) とした．全ての文字サンプルに対し前処理として，雑音除去および 2 値化を行った．さらに，外接長方形が 16×16 となるように大きさを線形正規化した後，上下左右に 2 画素づつの空白を入れた 20×20(すなわち $N = 20$) の 2 値画像とした．画素の特徴量には濃淡値をそのまま用いた．標準パターンについては，さらに文字種ごとに単純平均し，ヒストグラムを平坦化して作成した．標準パターンを図 4 に，テストパターンの例を図 5に示す．

4.2 ピボットの設定

区分線形 2 次元ワープのピボット数 K は 3 とした．すなわち標準パターン$\boldsymbol{A}$の各行は中間のピボット 1 点で屈曲され，入力パターン$\boldsymbol{B}$に重ね合わせられる．中間のピボット位置は各文字の変形特性を考慮して人為的に設定した．図 6に今回の実験で用いたピボットを示す．図 6(a),(b) はそれぞれ P_1, P_2に対応する．ダッチロールワープ D_1–D_6では，各行 2 つのピボットはそれぞれ行の両端に位置するので，このようなピボットの設定は不要である．

4.3 実験結果および考察

各手法による認識率を図 7に示す．なお，各手法において窓幅は最適化してある．手法 P_1–M_2のうち最高認識率を与えたのは区分線形 2 次元ワープ P_1であり，その値は 95.4%であった．次いで，単調連続 2 次元ワープ M_2の 95.2%，ダッチロールワープの拡張型 D_3, D_5の 95.1%となった．なお，P_2, D_2による認識率は P_1, D_1に比べて目立って低い．これから今回のタスク (ETL6 の英大文字) では，水平ストロークよりも垂直ストロークの傾きやずれが支配的であったと考えられる．

上述のように区分線形 2 次元ワープ P_1と単調連

図 4: 標準パターン

図 5: テストパターンの例

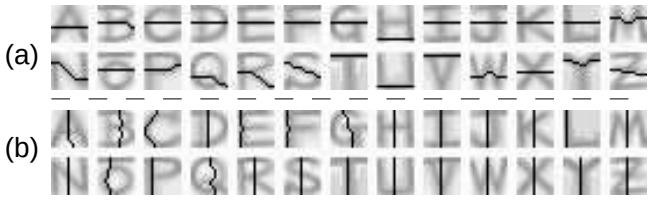

図 6: 標準パターン上に設定されたピボット: (a) 水平方向のピボット軌跡; (b) 垂直方向のピボット軌跡

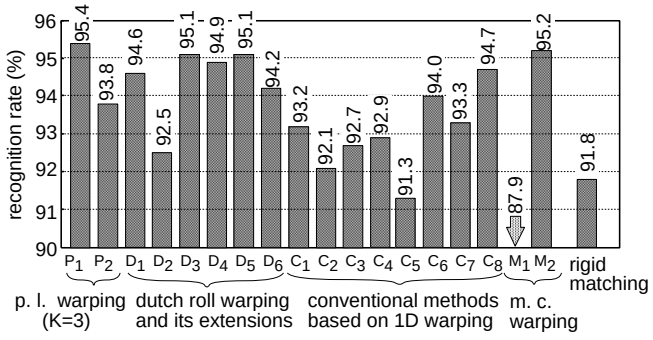

図 7: 認識実験 (各文字種 100 サンプル, 計 2600 サンプル) の結果．P_1–M_2は図 3の表記に対応．

表 1: 区分線形 2 次元ワープ P_1のエラーマトリクス

入力パターン \ 認識結果	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B	0	92	0	1	2	1	1	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	1	0
C	0	0	98	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
D	0	0	2	92	0	0	1	0	0	0	0	0	0	0	0	3	2	0	0	0	0	0	0	0	0	0
E	0	0	0	0	91	5	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	2
F	1	0	0	0	0	94	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	1	0	0	1	0
G	1	0	0	0	0	0	98	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
H	0	0	0	0	0	0	0	99	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
I	0	0	0	0	0	0	0	0	97	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0
J	0	0	0	0	0	0	0	0	0	95	0	0	0	0	0	0	0	0	0	4	0	0	0	0	1	0
K	0	0	0	0	0	0	0	0	0	0	98	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0
L	0	0	0	0	0	0	0	0	1	0	0	99	0	0	0	0	0	0	0	0	0	0	0	0	0	0
M	0	0	0	0	0	0	0	8	0	0	1	0	90	0	0	0	0	0	0	1	0	0	0	0	0	0
N	0	0	0	0	0	0	0	0	0	0	0	0	0	98	0	0	0	0	0	0	0	2	0	0	0	0
O	0	0	0	0	0	3	0	0	0	1	0	0	0	0	94	0	0	0	0	0	0	0	0	1	1	0
P	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	98	0	0	0	0	0	1	0	0	1	0
Q	3	0	2	0	0	0	1	0	0	0	0	1	0	0	0	0	93	0	0	0	0	0	0	0	0	0
R	3	0	0	0	0	1	0	0	0	0	1	1	0	0	0	2	0	90	0	0	0	0	0	1	1	0
S	1	2	0	0	0	1	0	0	0	1	0	0	0	0	1	1	0	1	92	0	0	0	0	0	0	0
T	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	97	0	0	0	0	1	0
U	1	0	1	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	96	0	0	0	0	0
V	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	91	3	0	1	0
W	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	96	0	0	0
X	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0
Y	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	99	0
Z	0	0	0	0	0	0	0	0	4	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	94

表 2: 単純重ね合わせのエラーマトリクス

入力パターン \ 認識結果	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	99	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
B	0	81	0	2	5	1	1	0	0	0	0	1	0	0	6	1	0	0	0	0	1	0	0	0	0	1
C	0	0	96	0	0	0	1	0	0	0	0	1	0	0	0	0	2	0	0	0	0	0	0	0	0	0
D	0	0	2	90	0	0	1	0	0	0	0	0	0	0	3	0	0	0	0	4	0	0	0	0	0	0
E	0	2	0	0	89	2	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	2
F	1	0	0	0	0	89	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	6	0	0	0	0
G	1	0	2	0	0	0	92	0	0	0	0	0	0	0	0	0	4	0	0	0	0	1	0	0	0	0
H	1	0	0	0	0	0	0	98	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
I	0	0	0	1	0	0	0	0	96	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1	0
J	0	0	0	0	0	0	0	0	0	96	0	1	0	0	0	2	0	0	0	0	0	0	0	1	0	0
K	0	0	1	0	1	0	0	1	0	0	97	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L	0	0	1	0	1	0	0	0	1	0	0	97	0	0	0	0	0	0	0	0	0	0	0	0	0	0
M	3	0	0	0	0	0	0	18	0	0	1	0	74	4	0	0	0	0	0	0	0	0	0	0	0	0
N	0	0	0	0	0	0	0	0	0	0	0	0	0	95	0	0	0	0	0	0	2	0	2	0	1	0
O	0	5	2	1	0	3	0	0	0	3	0	0	0	0	81	0	0	0	0	1	0	2	0	0	1	1
P	0	0	0	1	1	2	0	0	0	0	0	0	0	0	0	94	0	0	0	0	0	2	0	0	0	0
Q	2	0	3	0	0	0	4	0	0	0	0	1	0	0	0	0	89	1	0	0	0	0	0	0	0	0
R	1	0	0	0	1	1	0	0	0	0	3	0	0	0	0	3	0	88	0	0	0	1	0	2	0	0
S	0	0	2	1	2	2	0	0	0	1	0	0	0	0	2	0	1	1	88	0	0	0	0	0	0	0
T	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	94	0	1	0	0	3	0
U	1	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	97	0	0	0	0	0
V	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	89	3	0	1	0
W	0	0	0	0	0	0	1	0	0	0	1	0	0	1	0	0	1	0	0	0	0	2	94	0	0	0
X	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	95	3	0
Y	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	98	0
Z	0	0	0	0	1	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	92

続 2 次元ワープ M_2ではほぼ同程度の認識率を得た．しかし両者の計算量には大きな差異があり，後者ではビームサーチなどを導入した近似アルゴリズムを用いなければ実行できない点を考慮すると，前者の方がより実用的であると言える．実際， Pentium II (350MHz) の CPU を搭載した計算機を用いて両者が 1 画像対間にワープを求めるための計算時間を実測したところ，区分線形 2 次元ワープでは 68.3msec ， 単調連続 2 次元ワープでは近似アルゴリズムを利用していても 2360msec であった．また，単調連続 2 次元ワープ M_2の方が P_1に比べ自由度が高いのにもかかわらず認識率の点で優位性を示さなかったことから，本タスクに関しては区分線形 2 次元ワープ程度の自由度があれば，大部分の変形には対処できたと考えられる．

区分線形 2 次元ワープ P_1と単純重ね合わせによる結果のエラーマトリクスを，それぞれ表 1 , 2に示す．単純重ね合わせで特に低い認識率であった文字種 『B」『M」『O」については，いずれも区分線形 2 次元ワープによる改善が見られる．その例を図 8に示す．『M」の 2 本の垂直ストロークに生じた傾きや 『B」の垂直ストロークの位置ずれはいずれもワープにより補正され，変形後パターンは標準パター ンAと形状的に近いものとなっている．

ダッチロールワープについてはエラーマトリクスを省くが，区分線形 2 次元ワープと同様の傾向が見られた．図 9に図 8 と同じ入力パターンに対するダッチロールワープ D_1 の結果を示す．その整合能力は細部において区分線形 2 次元ワープよりわずかに劣るものの，垂直ストロークの局所的傾き，位置ずれは適切に補正されていることがわかる．英大文字のような比較的構造の簡単な文字パターンについては，ダッチロールワープ D_1の自由度でもある程度対応できることを例証している．ダッチロールワープ D_1が 1 画像対間にワープを求めるための計算時間は 1.0msec であり，

この値は 1 次元ワープベースの方法 C_7 の 1.2msec と同程度である．ダッチロールワープの拡張型 1,2 についてはいずれも D_1 よりも高い認識率を示しており，自由度向上を図った効果が認められる．

本実験では画素特徴として濃淡値のみを用いており，さらに入力パターン $\boldsymbol{B}$ は 2 値画像なので，$\boldsymbol{A}$–$\boldsymbol{B}$ 間の画素対応の曖昧性が高い．このため，過変形が起きやすい実験条件であると言える． 実際，変形抑制ペナルティを用いない単調連続 2 次元ワープ M_1 が非常に低い認識率となったのは，ビームサーチの副作用に起因するものも含めて過変形が主たる原因である．区分線形 2 次元ワープ P_1, P_2 も 1 次元ワープベースの方法 C_1–C_8 に比べ自由度が高いので過変形が危惧されたが，結果的に認識率の点では大きな問題にはなっていないと言える．しかし，区分線形 2 次元ワープ P_1 で誤認識された全 119 サンプルを誤認識の原因により大別したところ，前処理失敗が 38 サンプル (31.9%)，正解カテゴリの標準パターンの不十分もしくは不自然なワープが 27 サンプル (22.7%)，過変形が 54 サンプル (45.4%) となり，依然として過変形による誤認識は多い．過変形の例を図 10に示す．よって認識率向上のためには，不要な変形を回避するための検討は必要である．前述のように，本実験において M_2 と P_1 では同程度の認識率を得たが，区分線形 2 次元ワープ P_1 においても過変形の問題の緩和ために変形抑制ペナルティを用いて M_2 と同条件とすれば，P_1 の M_2 に対する優位性は顕在化するものと考えられる．

4.4 拡大実験

単調連続 2 次元ワープ M_1, M_2 以外について，テストサンプル数を各文字種 100 から 1000 に拡大した認識実験を行った．追加した 900×26 サンプルも ETL6 に含まれるものである．標準パターン，ピボットの設定，窓幅などの実験条件は **4.1** と同一である．

その結果，最高の認識率は区分線形 2 次元ワープ P_1 の 97.5%，それに次ぐのがダッチロールワープ D_1 およびその拡張型 D_3, D_5 の 96.7%, 97.1%,97.1%であった．このようにサンプル数を増やしても，これら手法の優位性に変化はなかった．

5 むすび

本報告では，2 次元弾性マッチング法の一種である区分線形 2 次元ワープの，手書き文字認識における有効性を実験的に検証した．区分線形 2 次元ワープ

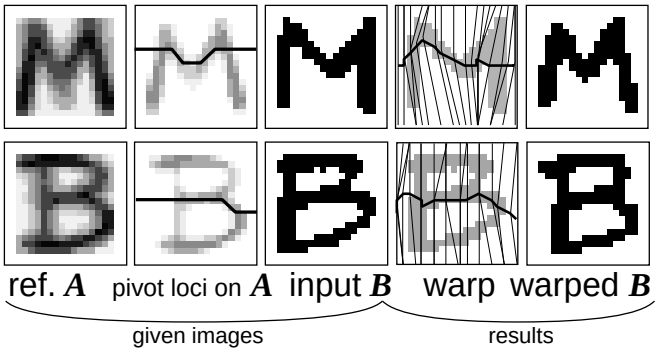

図 8: 区分線形 2 次元ワープ P_1 で正しく認識された例．単純重ね合わせではこれら 『M」 『B」はそれぞれ 『H」 『E」に誤認識された．

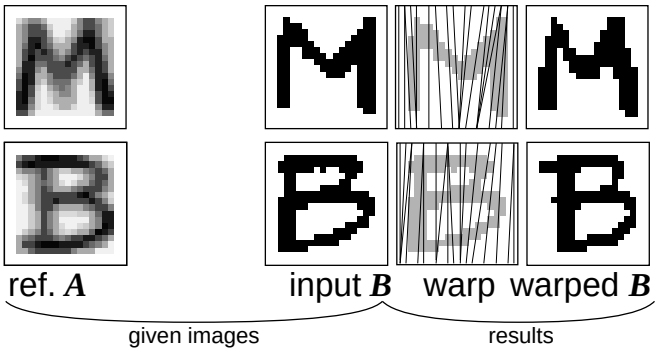

図 9: ダッチロールワープ D_1 により正しく認識された例．画像対 $\boldsymbol{A}, \boldsymbol{B}$ は図 8に同じ．

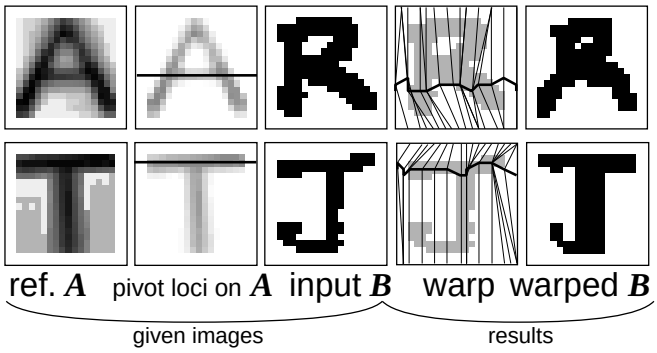

図 10: 区分線形 2 次元ワープ P_1 による過変形の例

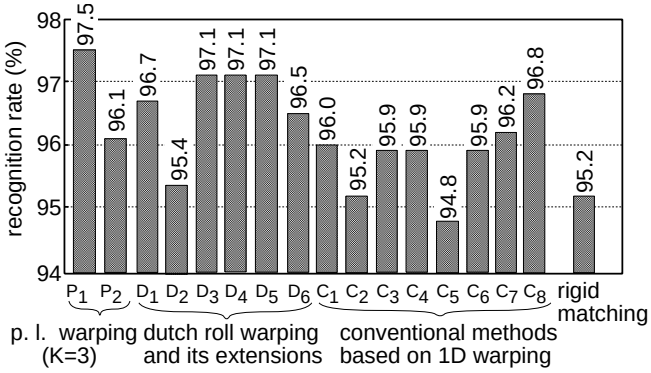

図 11: 拡大実験 (各文字種 1000 サンプル, 計 26000 サンプル) の結果．P_1–C_8 は図 3の表記に対応．

は，計算量の点で 1 次元ワープベースの手法と単調連続 2 次元ワープの中間に位置する．またその自由度すなわち変形吸収能力は，単調連続 2 次元ワープより低いものの，1 次元ワープベースの手法では吸収できない回転などの変形にも対処できる．ETL6 に含まれる英大文字 26 文字種各 1000 サンプルを対象とした認識実験では，97.5%の認識率を得た．これは 1 次元ワープベースの手法よりも高い値であった．また各文字種 100 サンプルを用いた実験では，計算量の非常に多い単調連続 2 次元ワープと同程度の認識率を得ている．これらから，手書き文字認識において，区分線形 2 次元ワープは妥当な自由度を持つ実用的なワープ法と言える．

区分線形 2 次元ワープの特殊形であるダッチロールワープも併せて検討した．これは区分線形 2 次元ワープにおいて，ワープの制御点 (ピボット) 数を最も少なくしたものであり，自由度は低くなる反面，計算量は 1 次元ワープに基づく方法と同程度となる．その認識率は区分線形 2 次元ワープに次ぐものであった．

区分線形 2 次元ワープでの誤認識の原因としては，単調連続 2 次元ワープほど深刻ではなかったが，過変形の占める割合が高かった．このため今後の検討課題としては，(1) 変形抑制ペナルティの導入，(2) 対応の曖昧性の排除のための特徴量の改善，(3) ワープの制約条件や目的関数の部分的な調整，(4) ピボット位置の最適化，が挙げられる．一方で変形不十分な場合も見られたので，最終的には文字の各部位について適切な自由度を設定する必要があると考えられる．

謝辞

実験では，電子技術総合研究所の ETL6 を利用させて頂いた．なお，本研究の一部は文部省科学研究費 (基盤研究 C(2),No.10680385) によった．

参考文献

[1] A. K. Jain, Y. Zhong, and M.-P. Dubuisson-Jolly. Deformable template models: A review. *Signal Processing*, vol. 71, no. 2, pp. 109–129, Dec. 1998.

[2] C. A. Glasbey and K. V. Mardia. A review of image-warping methods. *Journal of Applied Statistics*, vol. 25, no. 2, pp. 155–171, April 1998.

[3] H. Lester and S. R. Arridge. A survey of hierarchical non-linear medical image registration. *Pattern Recognition*, vol. 32, no. 1, pp. 129–149, Jan. 1999.

[4] A. Redert, E. Hendriks and J. Biemond, "Correspondence estimation in image pairs," IEEE SP Mag., vol. 16, no. 3, pp. 29–46, May 1999.

[5] 目黒眞一, 梅田三千雄. 階層的パターンマッチングによる手書き文字変形の抽出，信学技報，PRL77-70, 1977.

[6] 山田博三, 斉藤泰一, 森　俊二. 類似度法の一改良 – ずらし類似度 –，信学論 (D), vol. J64-D, no. 10, pp. 970–976, Oct. 1981.

[7] 泉井良夫, 原島　博, 宮川　洋. 階層的な辞書の変形を用いた手書き文字認識，信学論 (D), vol. J68-D, no. 3, pp. 361–368, March 1985.

[8] 水上嘉樹, 古賀和利, 鳥岡豊士. 変位抽出を行う手書き文字認識システム，信学論 (D-II), vol. J80-D-II, no. 1, pp. 63–72, Jan. 1997.

[9] Y. Mizukami. A handwritten Chinese character recognition system using hierarchical displacement extraction based on directional features, Pattern Recognition Letters, vol. 19, no. 7, pp. 595-604, May 1998.

[10] M. Nakagawa, T. Yanagida and T. Nagasaki. An off-line character recognition method employing model-dependent pattern normalization by an elastic membrane model, Proc. Fifth Int. Conf. Document Anal.and Recog. (ICDAR99), pp. 495–498, Sep. 1999.

[11] T. Wakahara. Shape matching using LAT and its application to handwritten numeral recognition, IEEE Trans. Patt. Anal. Mach. Intell., vol. 16, no. 6, pp. 618–629, June 1994.

[12] T. Wakahara and K. Odaka. Adaptive normalization of handwritten characters using global/local affine transformation, IEEE Trans. Patt. Anal. Mach. Intell., vol. 20, no. 12, pp. 1332–1341, Dec. 1998.

[13] A.K. Jain and D. Zongker. Representation and recognition of handwritten digits using deformable templates, IEEE Trans. Patt. Anal. Mach. Intell., vol. 19, no. 12, pp. 1386–1391, Dec. 1997.

[14] 中島　晃, 中野康明, 中田和男. 非線形伸縮マッチングによる異種フォント漢字認識, 昭 47 情報処理全大, pp.177–178, 1972.

[15] 中野康明, 中田和男, 中島　晃. 周辺分布とそのスペクトルによる漢字認識の改良，信学論 (D)，vol. 57-D, no. 1, pp. 15–22, Jan. 1974.

[16] 田中直樹, 塩野　充, 真田英彦, 手塚慶一. 動的方向性整合法による手書き漢字認識，信学論 (D)，vol. J68-D, no. 1, pp. 56–63, Jan. 1985.

[17] 津雲　淳. 方向パタンマッチング法の改良と手書き漢字認識への応用，信学技報，PRL90-20, 1990.

[18] S. Kuo and O.E. Agazzi. Keyword spotting in poorly printed documents using pseudo 2-D hidden Markov models, IEEE Trans. Patt. Anal. Mach. Intell., vol. 16, no. 8, pp. 842–848, Aug. 1994.

[19] E. Levin and R. Pieraccini. Dynamic planar warping for optical character recognition, Proc. ICASSP, vol. 3 of 5, pp.149–152, 1992.

[20] 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.

[21] 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.

[22] 内田誠一, 迫江博昭. 単調連続 2 次元ワープ法によるオフライン手書き文字認識実験. 信学論 (D-II), 掲載予定.

[23] 内田誠一, 迫江博昭. 区分線形型 2 次元ワープ法の検討. 信学技報, PRMU99-10, May 1999.