

手書き文字認識における字形変動処理に関する一検討 : ダッチロールワープ

石田, 敏之

九州大学大学院システム情報科学研究科知能システム学専攻 : 修士課程

内田, 誠一

九州大学大学院システム情報科学研究科知能システム学専攻

迫江, 博昭

九州大学大学院システム情報科学研究科知能システム学専攻

<https://doi.org/10.15017/1513737>

出版情報 : 九州大学大学院システム情報科学紀要. 5 (1), pp.99-104, 2000-03-24. 九州大学大学院システム情報科学研究所

バージョン :

権利関係 :

手書き文字認識における字形変動処理に関する一検討

－ ダッチロールワープ －

石田敏之*・内田誠一**・迫江博昭**

An Efficient Elastic Image Matching and Its Application to Handwritten Character Recognition – Dutch Roll Warping

Toshiyuki ISHIDA, Seiichi UCHIDA and Hiroaki SAKOE

(Received December 10, 1999)

Abstract: An efficient elastic image matching technique is investigated with application to off-line handwritten character recognition. In the present technique, each column of an image is mapped to another by controlling the mappings of the two endpoints of the column and linearly interpolating between them. Thus, the technique can adjust nonlinear deformation of images, with local slant and translation. The dynamic programming (DP) based algorithm searches for the optimal mapping with a reasonable amount of computation. The effectiveness of the present technique was indicated by recognition experiments on handwritten English alphabets.

Keywords: Dynamic programming, Two-dimensional warping, Handwritten character recognition

1. はじめに

文字認識の最も基本的な手法である単純重ね合わせは、雑音には強いものの、手書き歪みなどによる文字の変形に対しては極めて弱い。この弱点を克服するため入力パターン画像を標準パターン画像に最も近づけるように変形する写像(2次元ワープ)を介して両パターンの類似度を評価する方法が検討されている。

2次元ワープを実行する最適化法の1つとして、動的計画法(DP)がある。DPを用いることにより、総当たりの最適化に比べ大幅に少ない計算量で最適なワープを求めることができる。DPを用いた2次元ワープの文字認識への応用の歴史は古く、音声認識における時間軸歪み吸収問題のDPによる効率的な解法¹⁾に注目した研究者らにより、1970年代より検討がなされている。これらの従来法の多くは、1次元的なワープを組み合わせたものであり²⁾³⁾⁴⁾⁵⁾、手書き文字に含まれる歪みの吸収のためには依然として検討の余地が残されている。

DPを用いた本格的な2次元ワープ法に単調連続2次元ワープ⁶⁾がある。この手法はDPを真に2次元的に拡張して用いているため、パターンの位相を保存する範囲の歪みならば原理的に吸収できるという性質をもつ。その反面、計算量が画像サイズの指数オーダーとなる欠点を持つので実用的には何らかの対策が必要であった。

単調連続2次元ワープの計算量低減には2つの方向が

考えられる。1つは、ビームサーチ(beam search)などの解探索過程の効率化法⁶⁾の導入である。この場合、ワープの最適性は保証されなくなる。もう1つの方向は、単調連続2次元ワープの自由度を部分的に制限することで解空間そのものを縮小し、計算量を低減するものである。制限されたワープの自由度と対象固有の変形特性とを合致させることができれば、この方向は極めて有望と考えられる。

本論文では、この第2の方向に従ったワープ法としてダッチロールワープ^{†1)}と呼ぶ手法を提案し、その文字認識への応用を検討する。単調連続2次元ワープではすべての画素に自由度を持たせてワープの制御を行うのに対し、ダッチロールワープでは画像の列を単位として各列の両端にのみ自由度を持たせて制御を行い、両端以外の画素の像は線形補間により定める。ダッチロールワープの自由度は単調連続2次元ワープに比べて低くなるものの、計算量を大幅に低減することができる。

本論文ではさらに、ダッチロールワープに関して2種類の拡張型を検討する。具体的には、1)ダッチロールワープを水平、垂直順に逐次的に行う方法、2)ダッチロールワープの単位となる列の軸方向に1次元非線形ワープを用いる方法の2種類である。これは、いずれもダッチロールワープの自由度をより高くし、変形吸収能力を向上させることを目的としている。

以上のダッチロールワープおよびその拡張型について

平成11年12月10日受付

* 知能システム学専攻修士課程

** 知能システム学専攻

†1 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。

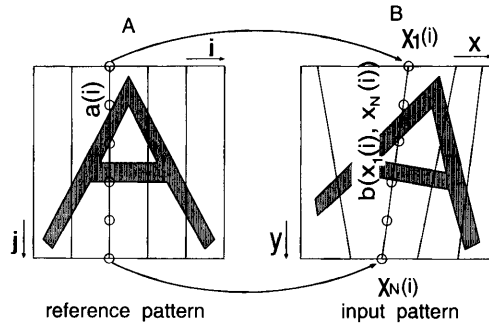


Fig.1 Dutch roll warping.

英大文字26文字種を対象とした手書き文字認識実験を行い、それらの基本的な特性を検討した。また、従来法との比較によりそれらの有効性を確認した。

2. ダッチロールワープ

2.1 原理

標準パターン画像を $A = \{a(i, j) | i, j = 1, \dots, N\}$ 、入力パターン画像を $B = \{b(x, y) | x, y = 1, \dots, N\}$ で表す。ここで、 $a(i, j), b(x, y)$ は画素の特徴量を表す。また、標準パターンの第 i 列を N 個の座標の組 $\mathbf{a}(i) = [(i, 1), \dots, (i, j), \dots, (i, N)]$ で表す。ダッチロールワープは、 $\mathbf{a}(i)$ の両端の画素 $(i, 1), (i, N)$ をそれぞれ B 上の画素 $(x_1(i), 1), (x_N(i), N)$ に対応づけ、その間を線形補間することにより残りの画素に対応づける写像である (Fig.1)。よって、ワープは $x_1(i), x_N(i), (i = 1, \dots, N)$ を制御変数として最適化される。

ダッチロールワープは次の単調連続性条件を満たす。

$$\left. \begin{aligned} 0 \leq x_1(i) - x_1(i-1) \leq 2 \\ 0 \leq x_N(i) - x_N(i-1) \leq 2 \end{aligned} \right\} \quad (1)$$

この単調連続性により、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。条件(1)に加え次の境界条件を与える。

$$\left. \begin{aligned} x_1(1) = x_N(1) = 1 \\ x_1(N) = x_N(N) = N \end{aligned} \right\} \quad (2)$$

また、極端な変形を防ぐ目的で次の整合窓条件を用いる。

$$\left. \begin{aligned} |x_1(i) - i| \leq w \\ |x_N(i) - i| \leq w \end{aligned} \right\} \quad (3)$$

$w (\geq 0)$ は窓半径と呼ばれる定数である。

$\mathbf{a}(i)$ の B 上での像を N 個の座標の組 $\mathbf{b}(x_1(i), x_N(i))$ で表す。 $\mathbf{a}(i)$ の j 番目の要素のワープによる像が $\mathbf{b}(x_1(i), x_N(i))$ の j 番目の要素に対応する。前述のように

$\mathbf{b}(x_1(i), x_N(i))$ は $(x_1(i), 1), (x_N(i), N)$ の線形補間により求められる。両端の画素の j 方向の座標はワープ前後で変わらず1および N なので、 $\mathbf{b}(x_1(i), x_N(i))$ の j 番目の画素の y 座標はワープ前と変わらず j となる。一方、 $\mathbf{b}(x_1(i), x_N(i))$ の x 座標は一般に実数値となるので、その点における特徴量を何らかの方法で与える必要がある。そこで本論文では最近傍補間を用いる。すなわち、

$$x = \left\lfloor \frac{x_N(i) - x_1(i)}{N-1} j + x_1(i) + 0.5 \right\rfloor \quad (4)$$

ここで、記号 $\lfloor \cdot \rfloor$ はガウス記号とする。

$\mathbf{a}(i)$ を $\mathbf{b}(x_1(i), x_N(i))$ にワープした時の距離 $d(i, x_1(i), x_N(i))$ を

$$d(i, x_1(i), x_N(i)) = \sum_{j=1}^N |a(i, j) - b'(j)| \quad (5)$$

と定義する。ここで、 $b'(j)$ は $\mathbf{b}(x_1(i), x_N(i))$ の j 番目の要素の座標における特徴量である。この $d(i, x_1(i), x_N(i))$ を $i = 1, \dots, N$ で積分したものを目的関数

$$\sum_{i=1}^N d(i, x_1(i), x_N(i)) \quad (6)$$

とし、その最小値をワープによる2画像 A, B 間の距離 $D(A, B)$ とする。ダッチロールワープによる文字認識において、入力パターン画像 B は、 $D(A, B)$ を最小にする標準パターン A のカテゴリに認識される。

2.2 DP アルゴリズム

ダッチロールワープでは、DPを用いてワープの最適化、すなわち目的関数(6)を条件(1)~(3)の下で最小化する。このとき、DP漸化式は次のようになる。

$$\begin{aligned} g(i, x_1(i), x_N(i)) \\ = \min_{p, q=0,1,2} [g(i-1, x_1(i)-p, x_N(i)-q) \\ + d(i, x_1(i), x_N(i))] \end{aligned} \quad (7)$$

ここで、 $g(i, x_1(i), x_N(i))$ は $\mathbf{a}(i)$ を $\mathbf{b}(x_1(i), x_N(i))$ にワープしたとき入力パターン画像 A の第1列から第 i 列目までの $d(i, x_1(i), x_N(i))$ の最小累積値である。ここで、

$$g(1, 1, 1) = d(1, 1, 1)$$

とする。画像間距離 $D(A, B)$ は $g(N, N, N)$ で与えられる。なお $x_1(i), x_N(i)$ に関して単調連続性条件(1)を満足するため $p, q = 0, 1, 2$ としている。

2.3 計算量

ダッチロールワープの計算量は以下のように見積もられる。まず、DP漸化式(7) 1回あたりの計算量は、

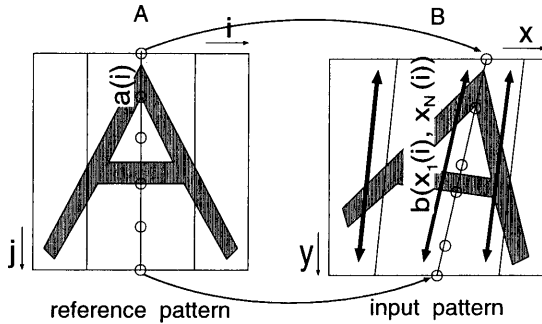


Fig.2 Dutch roll warping with intra-line warping.

$g(i, x_1(i), x_N(i))$ の最小値選択回数 9 と $d(i, x_1(i), x_N(i))$ の計算量 $O(N)$ より $O(9 + N) = O(N)$ となる。また、 $a(i)$ が対応する可能性のある $b(x_1(i), x_N(i))$ の数は、 $x_1(i)$ と $x_N(i)$ の組合せの数である $O(N^2)$ (より詳細には $O((2w + 1)^2)$)となる。DP漸化式は、各 i でこれら全ての場合について計算されるため、画像全体のDP漸化式計算回数は $O(N^2) \times N = O(N^3)$ となる。以上より、ダッチロールワープの計算量は $O(N) \times O(N^3) = O(N^4)$ となる。単調連続2次元ワープの計算量 $O(N^3 9^N)$ と比較すると、計算量が大幅に低減されているのがわかる。

3. ダッチロールワープの拡張

2.1で述べたようにダッチロールワープでは、 x 方向に対してダッチロールワープを行う場合、 x 方向の歪みは吸収できるが、 y 方向の歪みは吸収できない。よって、以下ではダッチロールワープの y 方向の歪みを同時に吸収するため2種類の拡張法を提案する。

3.1 拡張1

水平(x)方向→垂直(y)方向(もしくは垂直方向→水平方向)の順に2段階のワープによる水平、垂直両方向の歪み吸収を図る。より詳細には、まず、水平方向にダッチロールワープを行い、次に最初のワープにより変形された入力パターンを新たな入力パターンとして垂直方向にダッチロールワープを行う。これにより、水平方向、垂直方向の歪みを順次吸収する。この場合も整合窓を用いるが、ワープを行う方向ごとに窓幅は変えてよい。

この拡張法の計算量のオーダーは拡張前のダッチロールワープのオーダーと同じく $O(N^4)$ である。

3.2 拡張2

列ベクトル $a(i)$ と $b(x_1(i), x_N(i))$ とのマッチングにおいて、DPによる1次元非線形ワープを行うことで y 方向の歪みを吸収する。具体的には、 $a(i)$ と $b(x_1(i), x_N(i))$ のマッチングの評価として(5)式のかわりに次式を用いる。

$$d'(i, x_1(i), x_N(i)) = \min_{j(i)} \sum_{j=1}^N |a(i, j) - b'(j(i))| \quad (8)$$

ここで、 $j(i)$ は1次元ワープ関数であり、 $j(1) = 1$ 、 $j(N) = N$ 、 $0 \leq j(i) - j(i-1) \leq 2$ を満たすものとする。(8)式の最小化問題は、音声認識における時間歪み整合のための1次元ワープと同様DPで効率的にとけ、その計算量は $O(N^2)$ となる。これに応じて、拡張2のアルゴリズム全体の計算量は $O(N^5)$ となる。

4. 文字認識実験

本節では、ダッチロールワープとその拡張型の基本特性を文字認識を通して確認する。また、有効性を1次元ワープの組み合わせを基本とした従来法との比較により示す。

本節の実験に用いたワープ法をダッチロールワープと従来法を併せてFig.3に一覧する。いずれも上下-左右で対称な自由度を持たないワープ法であるため、転置した形態も併せて検討してある。Fig.3(a)~(f)は2節のダッチロールワープおよびその拡張1, 2である。(g)~(n)は従来試みられているワープを表す。(g),(h)は、行または列をひとまとまりとした垂直もしくは水平方向の非線形1次元ワープであり、2節で与えたダッチロールワープにおいて $x_1(i) = x_N(i)$ とした特殊型と見なすこともできる。(i),(j)は、(g),(h)を逐次的に行う方法²⁾である。(k),(l)は、各行または列において内部的に1次元非線形ワープを行う方法⁴⁾である。(m),(n)は、それぞれ(g)と(k)、および(h)と(l)を組み合わせたワープ³⁾である。

4.1 実験条件

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作成用、残りの100サンプルをテストパターンとした。全ての文字サンプルに対し前処理として、雑音除去の後、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 の2値画像とした。画素の特徴量には輝度値をそのまま用いた。標準パターンは、以上の前処理を行った文字サンプルを文字種ごとに単純平均し、ヒストグラム平坦化により作成した。実験に用いた標準パターンをFig.4に、テストパターンの例をFig.5に示す。

なお、以下の実験の結果はそれぞれ最適な窓幅を用いた場合を示してある。実験に用いた計算機のCPUはPentium II (350MHz)である。

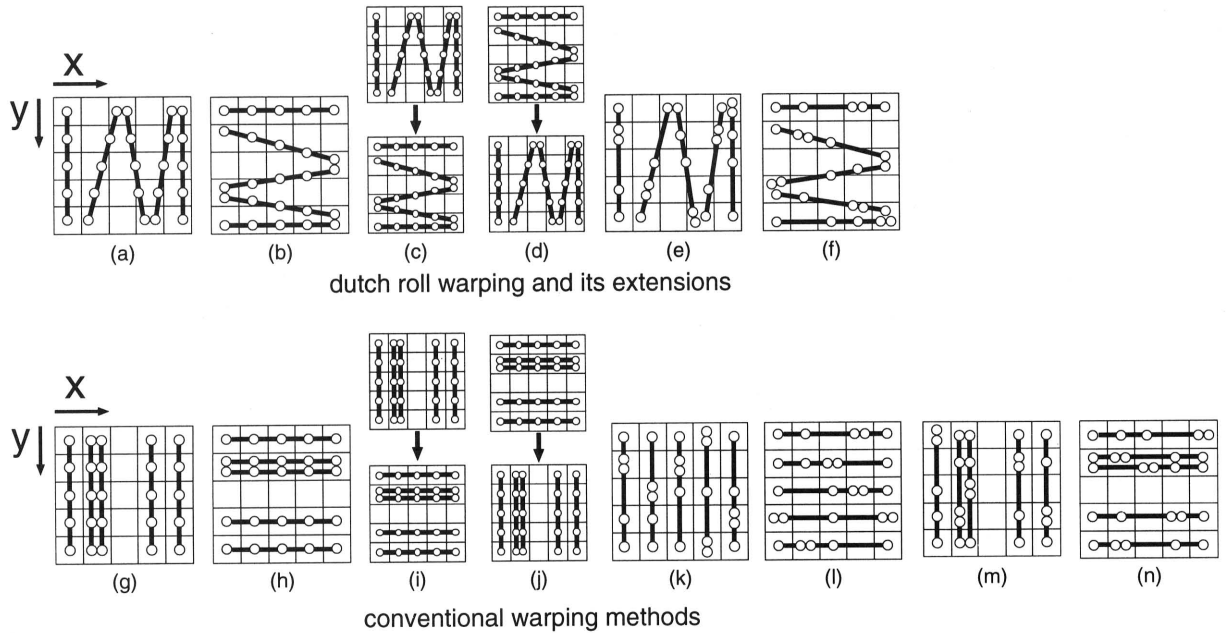


Fig.3 Warping types submitted to comparative experiment.



Fig.4 Reference characters.



Fig.5 Test character examples.

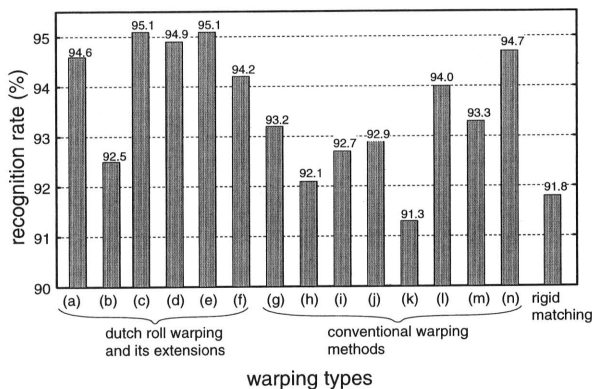


Fig.6 The recognition rate by dutch roll warping and its extensions, and conventional methods.

4.2 実験結果および考察

各手法による実験結果をFig.6に示す。グラフより、拡張1によるダッチロールワーピング(Fig.3(c)), 拡張2によるダッチロールワーピング(Fig.3(e))を用いた場合が最も認識率が高く95.1%となっている。このことから、本手法の有

Table 1 Recognition result of rigid template matching.

		recognition result																									
		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	0	0	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	0	0
C	0	0	96	0	0	0	1	0	0	0	0	1	0	0	0	2	0	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	4	0	0	0	0	0	2	
F	1	0	0	0	0	89	0	0	0	0	0	0	0	0	4	0	0	0	0	0	6	0	0	0	0	0	
G	1	0	2	0	0	0	92	0	0	0	0	0	0	0	4	0	0	0	0	1	0	0	0	0	0	0	
H	1	0	0	0	0	0	0	98	0	0	0	0	1	0	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	1	0	0	0	0	1	0	0	0	0	0	0	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	0	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	1	0	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	1	0	74	4	0	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	95	0	0	0	0	0	0	2	0	2	0	1	0	0	
O	5	2	1	0	3	0	0	0	3	0	0	0	81	0	0	0	0	1	0	2	0	0	1	1	0	0	
P	0	0	0	1	1	2	0	0	0	0	0	0	0	94	0	0	0	0	0	2	0	0	0	0	0	0	
Q	2	0	3	0	0	0	4	0	0	0	1	0	0	0	89	1	0	0	0	0	0	0	0	0	0	0	
R	1	0	0	0	1	1	0	0	0	3	0	0	0	3	0	88	0	0	0	1	0	2	0	0	0	0	
S	0	0	2	1	2	2	0	0	0	1	0	0	0	2	0	1	88	0	0	0	0	0	0	0	0	0	
T	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	94	0	1	0	1	0	3	0	0	0	
U	1	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	97	0	0	0	0	0	0	0	0	0	
V	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	2	0	89	3	0	1	0	0	0	0	0	
W	0	0	0	0	0	1	0	0	0	1	0	1	0	1	0	0	0	2	94	0	0	0	0	0	0	0	
X	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	95	3	0	0	0	
Y	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	98	0	0	0	
Z	0	0	0	0	1	0	0	0	5	0	0	0	0	0	0	0	0	0	2	0	0	0	0	92	0	0	

average 8.2% error

効性が確認される。

また、ダッチロールワーピングに基づく方法で見ると、Fig.3(a)(c)(e)の方がそれぞれ(b)(d)(f)よりも高い認識率となっている。これは、本実験で用いた文字サンプル英大文字では垂直(y)方向よりも水平(x)方向の歪みが大きく、水平(x)方向にワーピングを行った方が歪みの吸収効率がよいためと思われる。

Table 1に単純重ね合わせの場合のエラーマトリクス、Table 2にFig.3(a)のダッチロールワーピングの場合のエラーマトリクスを示す。Table 1より、単純重ね合わせでは“B”や“O”の認識率が低く、また特に“M”の認識率が著しく低く74%となっている。一方Table 2を見ると全体的に認識率が高くなっている。文字種別では

Table 2 Recognition result of basic dutch roll warping of Fig.3(a)

		recognition result																									
		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
recognition result	A	99	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	B	1	92	0	0	0	1	0	0	0	0	0	1	0	0	2	1	0	0	0	0	1	1	0	0	0	0
	C	0	0	99	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	D	0	0	2	95	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	1	1	0	0	0
	E	0	0	0	0	91	5	0	0	1	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	1
	F	1	0	0	0	0	94	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	1	0	0	2	0
	G	0	0	0	0	0	0	93	0	0	0	0	1	0	0	5	0	0	0	0	0	1	0	0	0	0	0
	H	0	0	0	0	0	0	0	97	0	0	0	1	1	0	0	0	0	0	0	0	0	0	1	0	0	0
	I	0	0	0	0	0	0	0	0	97	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	1	0
	J	0	0	0	0	0	0	0	0	0	95	0	0	0	0	0	0	0	0	0	4	1	0	0	0	0	0
	K	0	0	0	0	0	0	0	0	0	0	97	1	0	0	0	0	0	1	0	0	0	0	0	0	1	0
	L	0	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
	M	0	0	0	0	0	0	0	0	0	0	0	90	0	0	0	0	0	1	0	0	0	0	0	0	0	0
	N	0	0	0	0	0	0	0	0	0	0	0	0	0	97	0	0	0	0	0	1	0	2	0	0	0	0
	O	1	3	1	0	0	1	0	0	0	0	0	0	0	0	88	0	0	0	2	0	2	0	0	1	1	1
	P	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	98	0	0	0	0	0	0	0	0	0	0
	Q	2	0	1	0	0	0	2	0	0	0	0	1	0	0	0	0	94	0	0	0	0	0	0	0	0	0
	R	3	0	0	0	0	0	2	0	0	0	4	1	0	0	0	4	0	86	0	0	0	0	0	0	0	0
	S	0	1	1	0	1	1	0	0	1	0	0	0	0	0	4	0	2	1	88	0	0	0	0	0	0	0
	T	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	95	0	0	0	0	0	4	0
	U	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	98	1	0	0	0	0	0
	V	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	93	4	0	1	0
	W	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	2	96	0	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	97	3	0
	Y	0	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
	Z	0	0	0	0	0	0	0	0	6	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	92

average 5.4% error

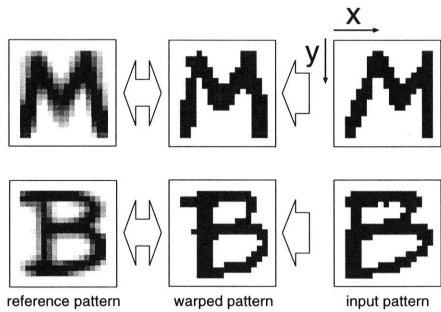


Fig.7 Examples correctly recognized by dutch roll warping of Fig.3(a).

“B”、“F”、“M”、“O”、“Q”に目立った向上が見られ、特に“M”については改善サンプル数1に対して改善サンプル数17であったため認識率も飛躍的に向上している。ダッチロールワープによる改善例をFig.7に示す。この入力パターン“M”は水平(x)方向の傾きを含んでおり、この傾きによる位置ずれが原因で単純重ね合わせでは“A”に誤認識されていた。一方、ダッチロールワープは、ストロークのある程度の傾きを吸収可能であるため傾きを補正でき、結果として正しく認識された。全体としてみると、単純重ね合わせからダッチロールワープを用いることで改善したサンプル数36、改善したサンプル数109であった。なお、このときのダッチロールワープの整合窓幅wは3であり、1画像当たりの計算時間は26.2 msecであった。

Table 3にダッチロールワープの拡張型1 (Fig.3(c))による結果を示す。また、拡張による改善例をFig.8に示す。この入力パターン文字“R”は文字の上半分に大きな歪みを持つ。Fig.3(a)のダッチロールワープは水平(x)方向の歪みしか吸収できないため、変形後のパターンを見ると垂直(y)方向の歪みが依然として大きく存在し、結果として“A”と誤認識された。一方、Fig.3(c)のワープでは、

Table 3 Recognition result of extended dutch roll warping of Fig.3(c).

		recognition result																										
		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	
recognition result	A	99	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
	B	0	96	0	0	0	1	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	1	0	
	C	0	0	98	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	
	D	0	0	2	94	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	2	0	0	0	0	
	E	0	0	0	0	92	5	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	
	F	1	0	0	0	0	95	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	1	
	G	0	0	1	0	0	0	94	0	0	0	0	0	1	0	0	2	2	0	0	0	0	0	0	0	0	0	
	H	0	0	0	0	0	0	0	100	0	0	0	0	0	0	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	0	2	0	0	0	0	1	
	J	0	0	0	0	0	0	0	0	0	1	93	0	0	0	0	0	0	0	0	0	0	5	0	1	0	0	
	K	0	0	0	0	0	0	0	0	0	1	0	97	1	0	0	0	0	0	0	0	0	0	0	0	0	1	
	L	0	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	
	M	0	0	0	0	0	0	0	0	0	11	0	0	0	87	1	0	0	0	0	0	0	1	0	0	0	0	
	N	0	0	0	0	0	0	0	0	0	0	0	0	0	0	98	0	0	0	0	0	0	0	1	0	1	0	
	O	1	1	1	0	0	1	0	0	0	0	0	0	0	0	0	93	0	0	0	0	3	0	0	0	0	1	
	P	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	99	0	0	0	0	0	0	0	0	1	
	Q	2	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	95	1	0	0	0	0	0	0	0	
	R	2	0	0	0	0	0	2	0	0	0	2	1	0	0	0	2	0	90	0	0	0	1	0	0	0	0	
	S	0	1	0	1	0	2	3	0	0	0	1	0	0	0	0	3	0	1	1	88	0	0	0	0	0	0	
	T	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	98	0	0	0	0	0	1	
	U	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	99	0	0	0	0	0	
	V	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	92	4	0	0	0	
	W	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	1	2	95	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	0	100	0	
	Y	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	100	0
	Z	0	2	0	0	1	1	0	0	10	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	85	

average 4.9% error

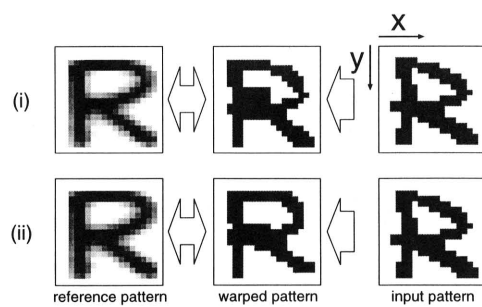


Fig.8 Example (i) incorrectly recognized by basic dutch roll warping of Fig.3(a) and (ii) correctly recognized by extended dutch roll warping of Fig.3(c).

ダッチロールワープを2回行うことにより水平(x)方向、垂直(y)方向の両方向の歪みを吸収できる。このため、変形後のパターンを見ると標準パターンに十分近づけることができ、結果的に正しく認識された。この拡張による改善は48サンプル、改善は35サンプルであった。このため、変形追従能力の向上の一方で過変形の問題が生じていることが考察される。なお、このときの拡張型1の条件は水平方向の窓幅3、垂直方向の窓幅1であり、1画像当たりの計算時間は30.4 msecであった。

Table 4にダッチロールワープの拡張型2 (Fig.3(e))による結果を示す。改善例をFig.9に示す。この入力パターンは全てのストロークが太く、また、部分的な傾きを持つ。そのため、この文字を標準パターン文字“O”に近づけるためには、傾きを吸収しつつストロークの幅を細くする必要がある。Fig.3(a)のワープ法では、傾き補正はある程度適切に行われているが、垂直(y)方向に対しては有効な変形は行われず、特に、上段の横ストロークの幅は変化がないことがわかる。一方、Fig.3(e)のワープを用いた場合は水平(x)方向、垂直(y)方向のどちらに対

Table 4 Recognition result of extended dutch roll warping of Fig.3(e).

		recognition result																										
		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	
Extended Dutch Roll Warping	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	1	93	0	0	0	0	0	1	0	0	0	0	0	0	2	1	0	0	1	0	0	0	0	0	0	1	0
	C	0	0	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
	D	0	0	2	94	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	1	1	0	0	0	0	0	0
	E	0	0	0	0	90	4	1	0	3	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1
	F	1	0	0	0	0	91	0	0	0	0	0	0	0	0	2	0	0	0	4	0	1	0	0	0	1	0	
	G	0	0	1	0	0	0	91	0	0	0	0	0	1	0	0	1	5	0	0	0	0	0	0	0	0	1	0
	H	0	0	0	0	0	0	0	99	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
	I	0	0	0	0	0	0	0	0	98	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
	J	0	0	0	0	0	0	0	0	1	94	0	0	0	0	0	0	0	0	0	4	0	0	0	0	1	0	0
	K	0	0	0	0	0	0	0	0	0	0	98	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
	L	0	0	0	0	0	0	0	0	0	0	0	99	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
	M	0	0	0	0	0	0	0	0	8	0	0	0	0	90	1	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	96	0	0	0	0	0	0	2	0	2	0	0	0	0	0
	O	0	2	1	0	0	1	0	0	1	0	0	0	0	0	92	0	0	0	0	2	0	0	0	0	1	0	0
	P	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	99	0	0	0	0	0	0	0	0	1	0	0
	Q	2	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	96	0	0	0	0	0	0	0	0	0	0
	R	2	0	0	0	0	3	0	0	0	0	3	0	0	0	0	1	0	91	0	0	0	0	0	0	0	0	0
	S	0	1	0	1	0	4	0	0	0	0	0	0	0	0	2	0	1	0	91	0	0	0	0	0	0	0	0
	T	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	97	0	0	0	0	3	0	0
	U	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	98	0	0	0	0	0	0
	V	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	1	0	0	1	0	92	4	0	0	0	0	0
	W	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	2	96	0	0	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	100	0	0	0	
	Y	0	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	
	Z	0	1	0	0	2	1	0	0	3	0	0	0	0	0	0	0	0	0	0	2	0	1	0	0	1	89	

average 4.9% error

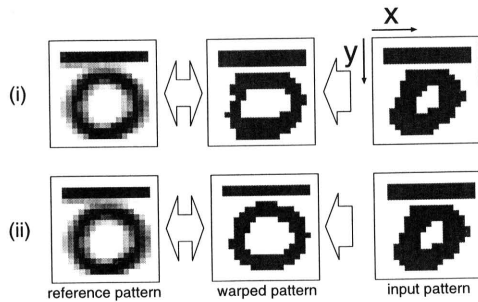


Fig.9 Example (i) incorrectly recognized by basic dutch roll warping of Fig.3(a) and (ii) correctly recognized by extended dutch roll warping of Fig.3(e).

しても有効な変形が行われているのがわかる。この拡張による改善サンプル数は45, 改悪サンプル数は32であり拡張1と同様の傾向が見受けられる。なお、このときの拡張型2の条件は $w = 4$, 各列毎の1次元ワープの整合窓幅1であり, 1画像対当たりの計算時間は42.7 msecであった。

4.3 拡大実験

ダッチロールワープおよびその拡張型(Fig.3(a),(c),(e)), および, 従来法で最も認識率が高かったFig.3(n)についてテストサンプル数を100から1000に拡大して実験を行った。追加した900×26サンプル

もETL6に含まれるものである。標準パターンなど, 他の実験条件は4.1と同様である。

その結果, Fig.3(a)のワープによる認識率は96.7%, Fig.3(c),(e)のワープによる認識率は共に97.1%, Fig.3(n)のワープによる認識率は96.8%となり, ダッチロールワープの優位性に変化はなかった。

5. おわりに

本論文では, パタンの傾きなどの2次元的な変形をある程度吸収可能としながら計算量の比較的少なくて済むワープ法としてダッチロールワープを提案した。その文字認識における有効性を実験的に検討したところ, 従来検討されている1次元ワープを組み合わせた方法に比べて高い認識率を得た。また, 自由度を高めることを目指した拡張型により計算量をほとんど増やさずに認識率を向上できた。英大文字26文字種各1000サンプルに対する認識率は97.1%であった。

謝 辞

本研究を行うにあたり, 電子技術総合研究所のETL6を利用して頂いた。ここに深く感謝する。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) 迫江博昭, 千葉成美: “動的計画法を利用した音声の時間正規化に基づく連続単語認識,” 日本音響学会誌 Vol.27, No.9, pp.483-490, Sep. 1971.
- 2) 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 Vol.57-D, No.1, pp.15-22, Jan. 1974.
- 3) 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.
- 4) 田中直樹, 塩野充, 真田英彦, 手塚慶一: “動的的方向性整合法による手書き漢字認識,” 信学論 Vol.J68-D, No.1, pp.56-63, Jan. 1985.
- 5) 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.
- 6) 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 Vol.J81-D-II, No.6, pp.1251-1258, June 1998.