研究速報

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

内田　誠一[†] (正員)　　迫江　博昭[†] (正員)

A Handwritten Character Recognition Experiment Using Monotonic and Continuous Two-Dimensional Warping

Seiichi UCHIDA[†] and Hiroaki SAKOE[†], *Members*

[†] 九州大学大学院システム情報科学研究科，福岡市
Graduate School of Information Science and Electrical Engineering, Kyushu Univ., 6–10–1 Hakozaki, Higashi-ku, Fukuoka-shi, 812-8581 Japan

あらまし　単調連続 2 次元ワープの手書き文字認識への応用について実験的検討を行なった．単調連続 2 次元ワープは画像の弾性マッチングの一種であり，位相同型の範囲で 2 次元的な歪みを吸収できる．このため文字の手書き歪みの大部分を吸収できると期待される．ひらがな認識実験を通してその文字認識における有効性を確認し，また問題点を明らかにした．

キーワード　単調連続 2 次元ワープ，パターンマッチング，動的計画法，オフライン文字認識

1. まえがき

オフライン手書き文字認識においては，手書き歪みが誤認識の原因となる．その対処法の一つに，2 画像の一方をもう一方に最も近付くように変形することで歪みを吸収する，いわゆる 2 次元弾性マッチングがある．そこでは 2 画像間の最大一致を与える 2 次元-2 次元写像 (2 次元ワープ) を求める最適化問題を解くことになる．

動的計画法 (DP) は最適化法の一種であり，解の最適性や安定性，計算効率のよさの他に，扱える制約条件や目的関数の多様性という特徴を持つ．音声認識の分野では時間軸非線形整合 (1 次元弾性マッチング) の手法として確立しており，その流れから，DP に基づく 2 次元弾性マッチングによる文字認識法も幾つか検討されている [1] ~ [5] ．それらはいずれも 1 次元的な変形もしくはその組合せで 2 次元的な手書き歪みを吸収しようとするものであり，特定方向の変形には対処できるが，任意方向の傾きやストロークの部分的な非線形変形については改良の余地があった．

本論文では，単調連続 2 次元ワープ [6], [7] の手書き文字認識への応用について，実験的検討を行なった結果を述べる．単調連続 2 次元ワープは DP に基づく 2 次元弾性マッチングの一種であり，位相同型の範囲ならば 2 次元的な歪みでも吸収できるという能力を持つ．極端でない手書き歪みの多くは位相同型写像によるものと見なせるので，本手法により手書き歪みの大部分を吸収できると期待される．

以下，**2.**では単調連続 2 次元ワープの概略を述べる．**3.**では ETL8B [8] のひらがな 46 文字種を対象とした認識実験の結果について述べる．また，単純重ね合わせ法および摂動法 (ずらしマッチング) [9] との比較検討を行なう．

2. 単調連続 2 次元ワープ [6], [7]

$N \times N$ 画像 $\boldsymbol{A} = \{\boldsymbol{a}(i,j)\}$ および $\boldsymbol{B} = \{\boldsymbol{b}(x,y)\}$ を考える．ここで $\boldsymbol{a}(i,j)$，$\boldsymbol{b}(x,y)$ はそれぞれ画素 (i,j)，(x,y) の特徴量とする．このとき単調連続 2 次元ワープは

$$\sum_{i=1}^{N}\sum_{j=1}^{N} \delta(\boldsymbol{a}(i,j),\ \boldsymbol{b}(x(i,j),y(i,j))) \qquad (1)$$

の最小値を与えるワープ関数 $x = x(i,j)$，$y = y(i,j)$ として定義される (図 1)．ここで $\delta(\cdot,\cdot)$ は 2 つの画素特徴間の距離を与える関数である．目的関数 (1) の最小値は，2 次元ワープによる $\boldsymbol{B}$ の変形画像 $\widetilde{\boldsymbol{B}} = \{\boldsymbol{b}(x(i,j),y(i,j)\}$ のうちで $\boldsymbol{A}$ に最も近いものと画像 $\boldsymbol{A}$ との重ね合わせ距離に相当する．

目的関数 (1) の最小化に際し，ワープ関数は次の 3 つの制約条件に従う．第一の制約はワープに位相保存性を持たせるため単調連続性条件である．具体的には，画素の上下 / 左右関係および近傍関係をワープ後も保存させる制約である (図 2)．第二は整合窓条件であり，画素とその像の座標値の差をある範囲 $w(>0)$ 内に抑える制約である．第三は境界条件であり，$\boldsymbol{A}$ の周縁上の画素を $\boldsymbol{B}$ の周縁上に写像させる制約である．

以上の制約条件の下での目的関数 (1) の最小化問題は，文献 [6] もしくは [7] の動的計画 (DP) アルゴリズムで解くことができる．本論文では，計算量の点で有利な文献 [7] の DP アルゴリズムを基本とし，それにビームサーチを組み込んで高速化を図った近似アルゴリズムを用いる．

近似アルゴリズムにより目的関数 (1) を直接最小化すると，ビームサーチによる近似誤差や単調連続性条件の局所性に起因して，極端な変形や折り返しのある不自然なワープが生じることがある．これを避けるためにペナルティ [6] を用いる．具体的には，ワープによる変形量を評価する関数 P_1 と，折り返しの発生に対し正の値を返すペナルティ関数 P_2 を目的関数 (1) に加え，それらをまとめて最小化する．

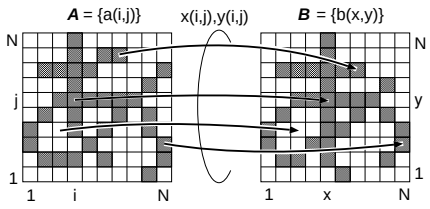

図 1　2 次元ワープ
Fig. 1　Two-dimensional warping (2DW).

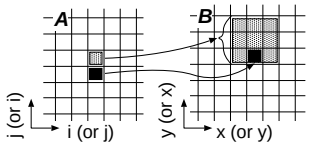

図 2　単調連続性条件
Fig. 2　Monotonicity and continuity condition.

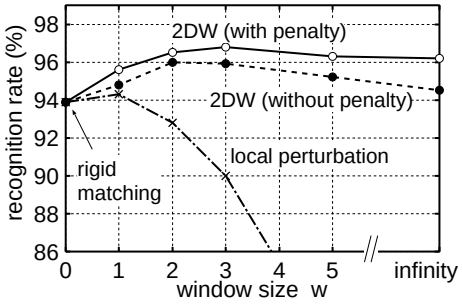

図 3　単調連続 2 次元ワープ，単純重ね合わせ，および摂動による認識率
Fig. 3　Recognition rates of monotonic and continuous 2DW, rigid matching, and local perturbation.

3. 手書きひらがなの認識実験

3.1 実験条件

実験試料として手書き文字データベース ETL8B [8] に含まれるひらがな 46 カテゴリ× 160 セットを用い，以下の手順で前処理を行なった．まず文字の大きさを線形に正規化した後，「┗」「＼」「｜」「／」の 4 方向について線素方向特徴 [10] を抽出した．この結果，画素特徴 $\boldsymbol{a}(i,j)$ および $\boldsymbol{b}(x,y)$ は輝度値を含めて 5 次元ベクトルとなる．このように画素特徴を高次元化することで，ワープ精度の向上が期待できる．線素方向特徴の抽出後，縮小処理により画像サイズを $N = 16$ とした．こうして得られた文字パターンのうち，偶数番 80 セットを入力パターン ($\boldsymbol{A}$，全 3680 サンプル) とした．残る奇数番の 80 セットについてはそれらを単純平均して各カテゴリにつき 1 つの標準パターン ($\boldsymbol{B}$) を作成した．

実験では次式の画素特徴間距離 δ を用いた．

$$\begin{aligned}&\delta(\boldsymbol{a}(i,j),\ \boldsymbol{b}(x,y))\\ &= |a_I(i,j) - b_I(x,y)|\\ &\quad + \eta \sum_{k=1}^{4} |a_D^k(i,j) - b_D^k(x,y)| \qquad (2)\end{aligned}$$

ここで $a_I(i,j)$，$b_I(x,y)$ はそれぞれ $\boldsymbol{a}(i,j)$ と $\boldsymbol{b}(x,y)$ の輝度値成分であり，$a_D^k(i,j)$，$b_D^k(x,y)$ $(k = 1,2,3,4)$ はそれらの線素方向特徴量成分である．$\eta(\geq 0)$ は線素方向特徴の重みを表す．

重み η を変えながら単純重ね合わせによる認識実験を行なった．距離 $\sum_i \sum_j \delta(\boldsymbol{a}(i,j),\ \boldsymbol{b}(i,j))$ に基づく最短距離法により識別を行なった結果，$\eta = 0.5$ の時に最高認識率 93.9%を得た．以下の実験ではこの η の値を用いた．

3.2 単調連続 2 次元ワープによる認識実験

次に単調連続 2 次元ワープによる認識実験を行なった．ペナルティ関数 (P_1, P_2) については使用 / 不使用の両方を検討した．識別は，得られた目的関数 (1) の最小値 (ペナルティ使用の場合はそれを含む) をそのままパターン間距離とした時の最短距離法によった．なお，一対の文字パターン間にワープを求めるための計算時間はおよそ 3 秒であった (Sun Ultra 2 ; SPECint_95: 12.3, SPECfp_95: 20.2)．

認識率を図 3に示す．ペナルティ関数使用，整合窓幅 $w = 3$ の時に単純重ね合せより 2.9%高い認識率 96.8%が得られた．このとき，単純重ね合わせで誤認識であった 223 サンプルのうち 131 サンプルが単調連続 2 次元ワープで正しく認識された．また，単純重ね合わせで正しく認識されていたサンプルのうち，27 サンプルがワープにより誤認識となった．これら改善例を図 4(a) に，改悪例を同図 (b) に示す．ここで $\boldsymbol{B}_1$，$\boldsymbol{B}_2$ はそれぞれ単純重ね合わせおよび単調連続 2 次元ワープで認識されたカテゴリの標準パターンであり，よって図 4(a) では $\boldsymbol{B}_2$，図 4(b) では $\boldsymbol{B}_1$ が正解カテゴリの標準パターンである．同図 (a) より，ストロークの位置ずれなどを，標準パターン $(\boldsymbol{B}_2)$ をワープにより変形することでうまく吸収できていることがわかる $(\widetilde{\boldsymbol{B}}_2)$．一方，同図 (b) より，異カテゴリの標準パターン $(\boldsymbol{B}_2)$ が過度に変形され $(\widetilde{\boldsymbol{B}}_2)$，入力パターン $(\boldsymbol{A})$ に近付き過ぎたことが誤認識の一因として考えられる．

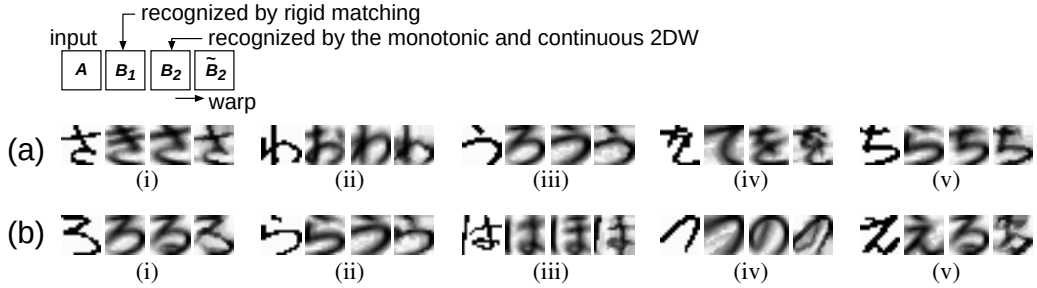

図 4　単調連続 2 次元ワープによる改善例 (a) と改悪例 (b)
Fig. 4　Samples correctly recognized (a) and misrecognized (b) by the 2DW.

実際，単調連続 2 次元ワープで誤認識された全 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 22 サンプル，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 41 サンプル，この過変形が 56 サンプルであった．次節で述べるように本手法では単調連続性制約によりある程度過変形が回避されているが，以上の考察から一層の過変形回避が今後の課題であると言える．

3.3　摂動法との比較

本手法の有効性を確認するための他の実験として，摂動法 (ずらし)[9] との比較を行なった．摂動法は各 (i, j) 独立に整合窓幅$\pm w$ の範囲内で最大一致する画素 (x, y) を探索する方法である．本手法と単調連続性制約すなわち位相保存性が無い点で異なる．

単調連続 2 次元ワープ (ペナルティなし) と摂動法の認識率を同じ w で比較すると，すべての w について本手法の方が認識率が高く，単調連続性制約の文字の歪み吸収に対する有効性が示されている．また w の増加に対する認識率の変化を見ると，摂動法では変形能力の向上よりも過変形の影響の方が顕著になったため認識率は急激に低下している．一方，本手法では 比較的安定した認識率が得られており，単調連続性制約に過変形抑止効果があることが確認される．

4. む す び

DP に基づく 2 次元弾性マッチング法の一種である単調連続 2 次元ワープの手書きひらがな認識への応用について，実験的検討を行なった結果を述べた．単純重ね合わせおよび摂動法との比較から，本手法の有効性を確認した．また，異カテゴリのパターンの過変形について対策を要することがわかった．

謝　　辞

本研究を行なうにあたり，電子技術総合研究所の ETL8B を利用させて頂いた．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．

文　　献

[1] 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.
[2] 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.
[3] 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.
[4] E. Levin and R. Pieraccini, “Dynamic planar warping for optical character recognition,” Proc.of ICASSP, Vol. 3 of 5, pp.149–152, 1992.
[5] 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.
[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 (D-II)，Vol. J81-D-II, No. 6, pp. 1251–1258, June 1998.
[7] 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.
[8] 森　俊二，山本和彦，山田博三，斎藤泰一，“手書教育漢字のデータベースについて，” 電子技術総合研究所彙報，Vol. 43, No. 11,12, pp. 60–81, 1979.
[9] 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.
[10] 水上嘉樹，古賀和利，“線素方向特徴を用いた変位抽出を行う手書き漢字認識システム，” 信学技報，PRU96-188, 1997.

(平成 11 年 9 月 7 日受付，XX 年 XX 月 XX 日再受付)