

弾性マッチングにおける方向特徴の利用に関する一考察

石田慎二郎[†] 内田 誠一^{††} 迫江 博昭^{††}

[†] 九州大学工学部電気情報工学科

^{††} 九州大学大学院システム情報科学研究所

〒 812-8581 福岡市東区箱崎 6-10-1

E-mail: {ishida, uchida, sakoe}@human.is.kyushu-u.ac.jp

あらまし オフライン文字認識における手書き変形への対処法の1つに弾性マッチング法がある。この方法では、入力と参照の二つの文字画像間についてそれらの最大一致を与える2次元 - 2次元の非線形写像(2次元ワープ)を与える。2次元ワープが2画像間の画素対応関係を正しく与えているとすれば、ワープ最適化の目的関数として最小化された画素間距離は変形不変量になる。従って、その距離を識別関数値として用いることで高い認識精度が期待できる。本論文では、弾性マッチング精度の向上を目的とした方向特徴の利用法について考察する。具体的には、入力・参照画像間のワープを決定する過程において、入力画像側に生じている局所的な回転変形に対して、ある画素の周囲の局所的なワープの状況に応じて、その回転量を推定し、それによって方向特徴を補正する。英大文字を対象として実験を行った結果、その有効性が確認された。

キーワード 弾性マッチング, 方向特徴, 変形, 画像パターン

On the use of directional feature in the elastic matching of image patterns

Shinjiro ISHIDA[†], Seiichi UCHIDA^{††}, and Hiroaki SAKOE^{††}

[†] School of Engineering, Kyushu University

^{††} Faculty of Information Science and Electrical Engineering, Kyushu University

Hakozaki 6-10-1, Higashi-ku, Fukuoka-shi, 812-8581 Japan

E-mail: {ishida, uchida, sakoe}@human.is.kyushu-u.ac.jp

Abstract Elastic image matching, or two-dimensional warping, is a deformation-tolerant nonlinear image registration technique and has been employed in image pattern recognition tasks to obtain deformation-invariant distance between two image patterns. In order to obtain reasonable matching between two image patterns, it is required that the feature vectors of two corresponding pixels should resemble each other. In several features, especially in directional feature, however, this requirement does not be satisfied because the directional features are sensitive to the geometrical deformations appeared on the character patterns. In this paper, a technique to compensate the change of the feature vector due to the geometrical deformations is proposed, where the matching itself is utilized as the guide showing the deformations. From experimental results, it is shown that the directional features between two character images of the same category become closer using the proposed technique.

Key words elastic matching, directional feature, deformation, image pattern

1. まえがき

幾何変形を伴う画像パターンの認識における要素技術の一つに弾性マッチング法がある。この方法では、入力と参照の2つの画像パターン間についてそれらの最大一致を与える2次元-2次元の非線形写像(2次元ワープ)を与える。2次元ワープが2画像間の画素対応関係を正しく与えているとすれば、ワープ最適化の目的関数として最小化された画素間距離は変形不変量になる。従って、その距離を識別関数値として用いることで高い認識精度が期待できる。

弾性マッチングにおいては、各画素の特徴量として何を用いるかが重要となってくる。特徴量に望まれる性質としては、(i) 対応すべき2画素の特徴量が同じになる、(ii) 同じ特徴量を持つ画素は対応すべき2画素を除いて他にない、の2つが挙げられる。この(ii)は画素対応の曖昧性が少ないことを要求するものであり、ワープによる過変形の防止に役立つ。

これらの要求に対し、方向特徴の利用が考えられる。方向特徴とは各画素をどの方向のエッジ(文字パターンの場合には文字線)が通過しているかを表す特徴量である[1]。従って各画素の近傍情報も含んでおり、画素対応の曖昧性の低減に効果がある。実際、文献[2]では弾性マッチングに基づく文字認識において方向特徴を利用し、その効果を確認している。

ところで、弾性マッチングが対象としている画像パターンでは、局所的な回転変形を生じていることが多い。定義から明らかなように、基となるパターンとそれが回転したパターンでは、方向特徴が異なってくる。このことは、対応すべき画素の方向特徴量が必ずしも一致しないことを意味する。よって、2画素の方向特徴量を単純に比較する従来の方式から見れば、上記(i)の性質を完全には満足できていなかった。

そこで本研究では、弾性マッチングにおける方向特徴の利用形態の改良を試みる。具体的には、今考えている2次元ワープが各画素においてどのような回転変形を表現しているかに応じて、その画素の方向特徴も同じだけ回転させることを考える。もし2次元ワープが正しい画素対応となっているならば、対応すべき画素の方向特徴はこの回転処理によってほぼ一致することになる。よって、逆にこの仕組みをワープの最適化過程に組み込むことで、方向特徴を上記(i)(ii)の性質を満たす特徴量として機能させることができ、より対応精度の高い2次元ワープを期待できる。

本論文の実験では、弾性マッチング法としてダッチロールワープを用いる。ただし、どのような弾性マッチング

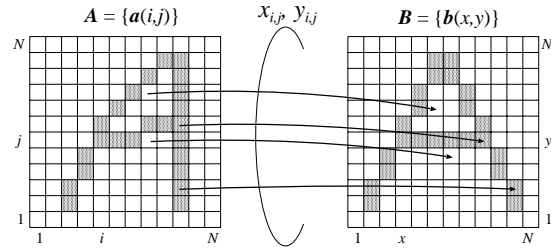


図1 弾性マッチング

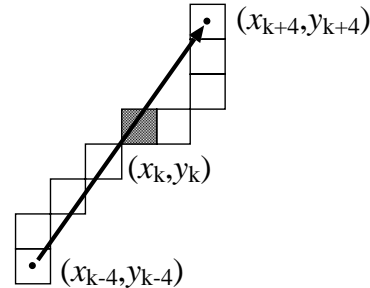


図2 各画素における文字線方向の抽出。四角で表されているのは、文字線の輪郭線を構成する画素。

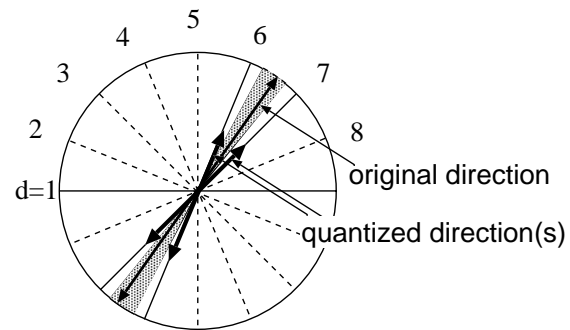


図3 方向特徴量の量子化

法についても本論文の考え方は原理的に適用可能である。また、本論文においては文字パターンを中心として議論/実験を行うが、顔パターンなど他の画像パターンにおいても適用可能である。

本論文は次のように構成される。2節では、弾性マッチング法とそれに応じた方向特徴量の回転について述べる。3節では、2節に基づいて、弾性マッチング法としてダッチロールワープを用いた場合の本手法の有効性を検証するために行った実験結果について述べる。

2. 方向特徴量の回転を考慮した弾性マッチング

2.1 弾性マッチング

弾性マッチングとは、一般に2つの $N \times N$ 画像 $A = \{a(i, j)\}$ と $B = \{b(x, y)\}$ との間に画素の最適

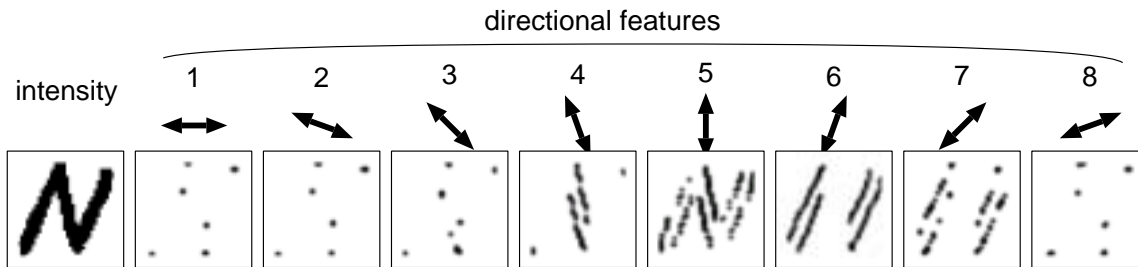


図 4 方向特徴の例

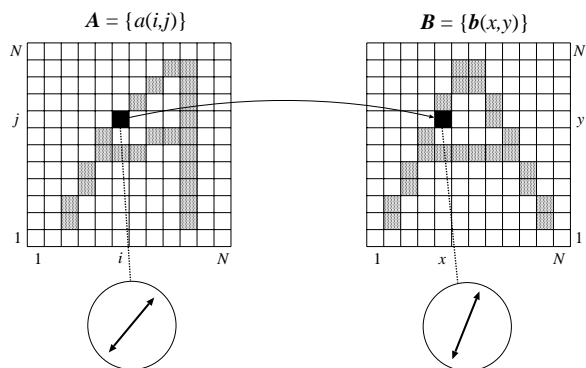


図 5 方向特徴の不一致

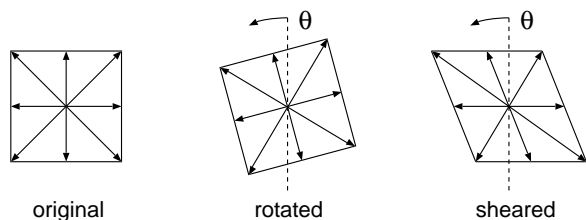


図 6 方向特徴量に対する回転 (rotation) と剪断 (shear)

対応関係 $\{(x_{i,j}, y_{i,j}) \mid i, j = 1, \dots, N\}$ を定める手法として定式化される (図 1) . ここで, $(x_{i,j}, y_{i,j})$ は \mathbf{A} 上の画素 (i, j) が対応する \mathbf{B} 上の画素を示す . 具体的には最小化問題

$$D(\mathbf{A}, \mathbf{B}) = \min_{\{(x_{i,j}, y_{i,j})\}} \sum_{i=1}^N \sum_{j=1}^N \delta(\mathbf{a}(i, j), \mathbf{b}(x_{i,j}, y_{i,j}))$$

を何らかの制約条件の下で解いたときの, その最小値を与える $\{(x_{i,j}, y_{i,j})\}$ として 2 次元ワープ, すなわち最適画素対応関係が求まる . また, 関数 $\delta(\cdot, \cdot)$ は画素特徴ベクトル間の距離を評価する関数である . 最小値 $D(\mathbf{A}, \mathbf{B})$ は 2 次元ワープにより 2 画像 \mathbf{A}, \mathbf{B} 間の最大一致を図った後の距離であり, 変形不変量となる . 従ってこの距離 $D(\mathbf{A}, \mathbf{B})$ を識別関数値として用いることで, 高い認識精度を期待できる .

2.2 方向特徴

方向特徴とは各画素におけるエッジ方向等を表現した

特徴量であり, 文字認識等で従来より利用されている [1] . 本論文で扱う方向特徴とは, 各画素を通過するエッジもしくは文字線の方向および強さを 8 次元ベクトルで表すものである . これに応じて, 画素特徴 $\mathbf{a}(i, j)$ を以下では次のように表現する .

$$\mathbf{a}(i, j) = (a_1(i, j), \dots, a_d(i, j), \dots, a_8(i, j))$$

画素特徴 $\mathbf{b}(x, y)$ についても同様に,

$$\mathbf{b}(x, y) = (b_1(x, y), \dots, b_d(x, y), \dots, b_8(x, y))$$

のように表す .

文字パターンを対象とした場合, これら方向特徴は具体的には以下の手順で求められる . まず, 文字線の輪郭座標 $\{(x_k, y_k) \mid k = 1, \dots, K\}$ をトラッキングにより抽出する . 次に, (x_{k-l}, y_{k-l}) と (x_{k+l}, y_{k+l}) を結ぶ線から位置 (x_k, y_k) でのストローク方向 (角度) を決定する . ここでパラメータ l を大きくすると, それだけ方向特徴は平滑化され, デジタル化による輪郭方向のゆらぎは抑えられる . ただしあまり大きくすると微細構造が失われる . 以下では, この妥協点として $l = 4$ を採用した (図 2) .

こうして各画素についてその文字線方向を求めた後, それを水平方向を基準に反時計回りに 0 度, 22.5 度, 45 度, 77.5 度, 90 度, 112.5 度, 135 度, 157.5 度の 8 方向に量子化する . 具体的には図 3 に示すように, 前述の角度方向のベクトル (長さ 1) が網掛けの領域内 ($d = 6$ と 7 のなす角の二等分線から ± 5.625 度まで) にある場合, そのベクトルを挟む近傍の量子化ベクトル成分に大きさを均等に分解する . また, 隣り合う 2 つの量子化ベクトル成分に挟まれるどの網掛け領域にも含まれない場合は, 最近傍の量子化ベクトル成分へ長さ 1 で割り当てる . この結果, この場合の画素の方向特徴ベクトルは $(a_1(i, j), \dots, a_d(i, j), \dots, a_8(i, j)) = (0, 0, 0, 0, 0, 0.5, 0.5, 0)$ となる . この処理をすべての画素について行うことで, すべての画素それぞれに 8 次元特徴ベクトルが与えられる . なお, 輪郭が通過していない画素の特徴ベクトルは $(0, 0, 0, 0, 0, 0, 0, 0)$ となる . 図 4

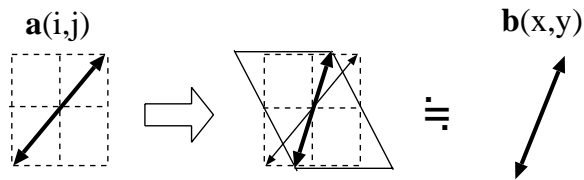


図7 ワープに応じた方向特徴量の回転 (剪断補正の場合)

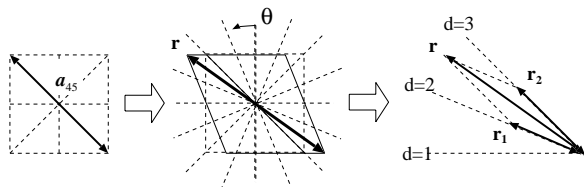


図8 方向特徴量の再量子化 (剪断補正の場合)

は以上の処理により実際に求めた方向特徴の例である。

2.3 ワープに応じた方向特徴量の回転

図5は2画像 A, B において対応すべきある2画素の方向特徴量を示している。この例から分かるように、局所的な回転変形を生じている文字においては、対応すべき画素の方向特徴量が必ずしも一致しないことが多い。よって、2画素の方向特徴量を単純に比較する従来の方式から見れば、正しく比較を行えていなかった。本研究では、この比較の際に方向特徴量に対して回転を与えることで、マッチング精度の向上を図る。その際、特徴量がどの程度回転しているかを推定する必要がある。本論文では、ワープそのものが画像の変形を表現していることに着目し、ある画素の周囲の局所的なワープの状況に応じて、回転量を推定し、それによって方向特徴を補正する。

以上では、回転変形を中心として述べたが、図5の例にもあるように、剪断変形を生じている場合もある。回転と剪断の違いは、回転はいわゆる単純回転で、どの方向面の特徴も同じ角度だけ回転するのに対して、剪断は水平方向と平行に回転する (水平方向の特徴は回転しない) ことである (図6)。

回転・剪断変形に対するマッチングの具体的な考え方は次のようになる。画素 $\mathbf{a}(i, j)$ は画素 $\mathbf{b}(x, y)$ に対して局所的に回転・剪断変形しているの、それに対してワープの表す角度 θ で特徴量 $\mathbf{a}(i, j)$ を回転・剪断変換すれば、 $\mathbf{b}(x, y)$ と一致させることができる (図7)。

実際には、方向特徴は8レベルに量子化して考えているので、参照パターン画像の各画素の方向特徴との比較のためには、その回転させた結果も再び22.5度毎の各方向面に量子化して扱う必要がある (図8)。ただし、元の方向特徴の大きさが保存されるように再量子化するものとする。この再量子化は図8の例では、図中の記号を

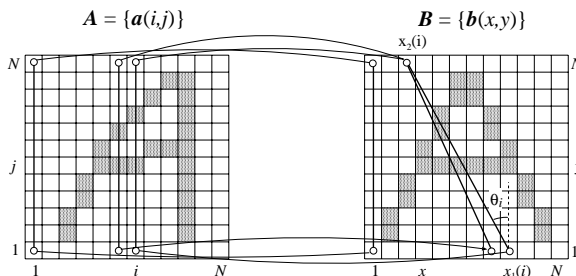


図9 ダッチロールワープ

使って次のようになる。

$$\|\mathbf{r}_{d2}\| = \|\mathbf{a}_{45}\| \times \frac{\|\mathbf{r}_1\|}{\|\mathbf{r}_1\| + \|\mathbf{r}_2\|}$$

$$\|\mathbf{r}_{d3}\| = \|\mathbf{a}_{45}\| \times \frac{\|\mathbf{r}_2\|}{\|\mathbf{r}_1\| + \|\mathbf{r}_2\|}$$

特徴補正はマッチングを最適化した後に行う考え方、マッチングの最適化過程に組み込んで行う考え方があるが、本論文ではより本質的な考え方である後者を採用する。すなわち、最適マッチングの探索過程において、距離 $\delta(\mathbf{a}(i, j), \mathbf{b}(x_{i,j}, y_{i,j}))$ を計算する際に、変数 $(x_{i,j}, y_{i,j})$ に応じて \mathbf{a} の画素特徴を補正する。

3. 実験

3.1 実験試料

本論文では、評価実験のために、電総研文字データベース ETL6 の英大文字データ (画像サイズ 64×64) を用いた。このデータベース中のサンプルについて2.2で述べた8次元方向特徴 [1] を求めた。

3.2 ダッチロールワープ

本実験では、弾性マッチング法としてダッチロールワープ [3] を用いた。ダッチロールワープとは入力パターン画像の各列の両端点をワープの制御点とし、それらの写像先を参照パターン画像上に探索する手法である (図9)。サイズ $N \times N$ の入力パターン画像 $\mathbf{A} = \{\mathbf{a}(i, j) \mid i, j = 1, \dots, N\}$ の第 i 列は、参照パターン画像 $\mathbf{B} = \{\mathbf{b}(x, y) \mid i, j = 1, \dots, N\}$ に $(x_1(i), 1)$ と $(x_2(i), N)$ を両端点とする線分に写像される。この際、同列上の画素 (i, j) ($j = 1, \dots, N$) の対応先は線形補間により定められるものとする。すなわち、

$$(i, j) \mapsto \left(\frac{x_2(i) - x_1(i)}{N - 1} j + x_1(i), j \right) \quad (1)$$

ダッチロールワープ最適化問題は次式のように定義される。

$$[\text{目的関数}] \sum_{i=1}^N \sum_{j=1}^N \delta(\mathbf{a}(i, j), \mathbf{b}(x_{i,j}, y_{i,j})) \rightarrow \min$$

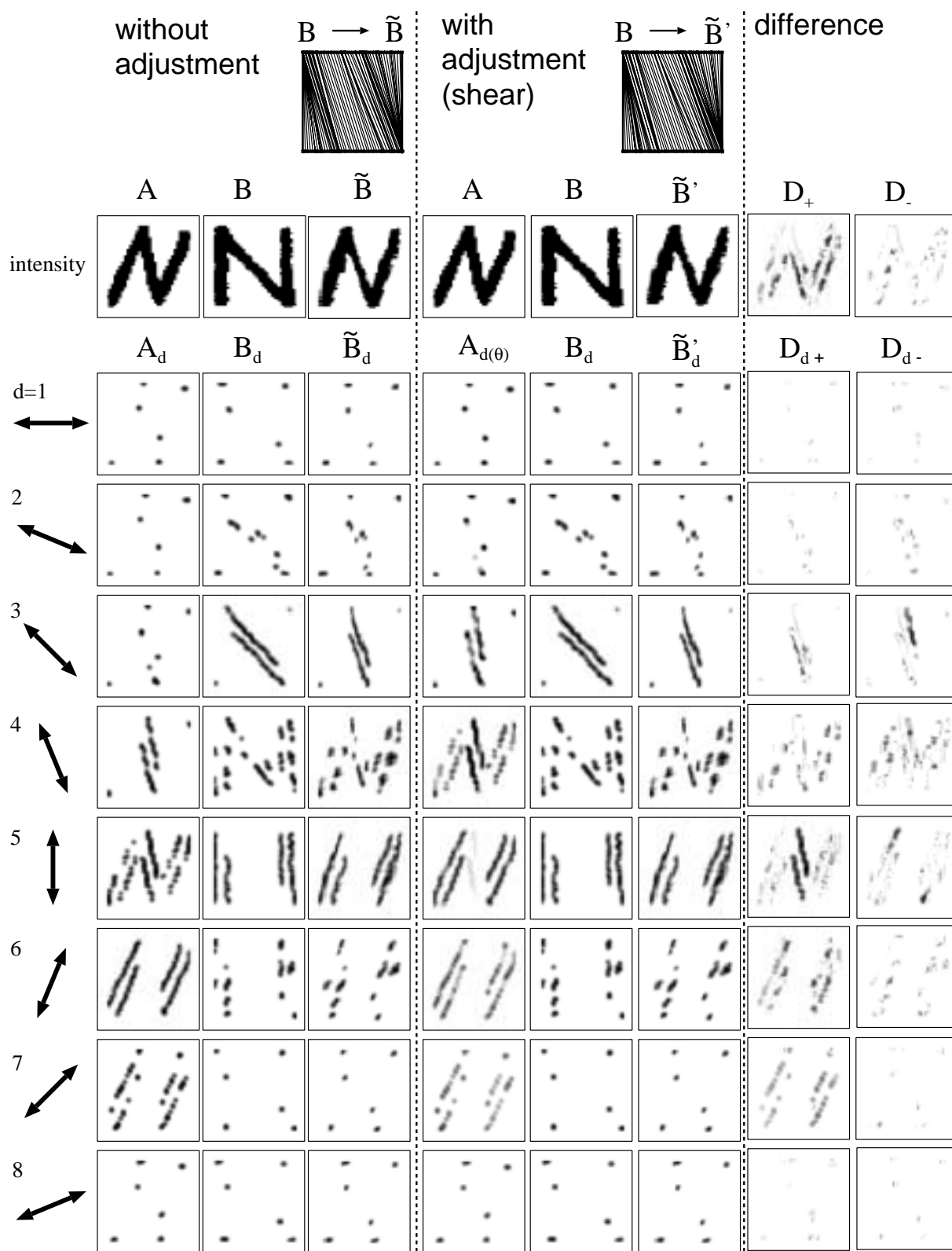


図 10 従来法および本手法によるマッチング例．各方向特徴面毎 ($d = 1, 2, \dots, 8$) に結果を示してある．右 2 列 D_{d+} および D_{d-} は、それぞれ本手法による改善箇所と改悪箇所を示している．

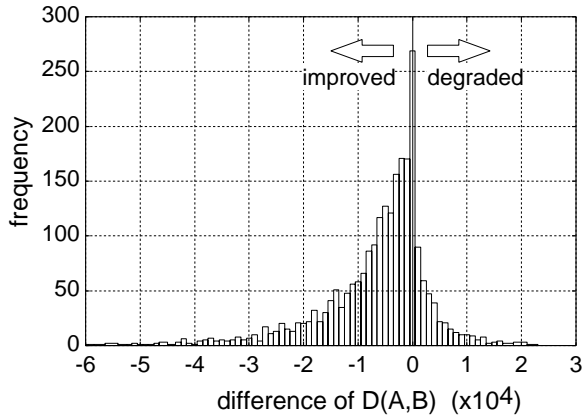


図 11 従来法と本手法との距離 $D(\mathbf{A}, \mathbf{B})$ の差のヒストグラム

$$= \sum_{i=1}^N \sum_{j=1}^N \sum_{d=1}^8 |a_d(i, j) - b_d(x_{i,j}, y_{i,j})| \quad (2)$$

$$\text{ただし} \begin{cases} x_{i,j} = \frac{x_2(i) - x_1(i)}{N-1} j + x_1(i) \\ y_{i,j} = j \end{cases} \quad (3)$$

$$[\text{制御変数}] \{x_1(i), x_2(i) \mid i = 1, 2, \dots, N\}$$

$$[\text{制約条件}] \begin{cases} \text{単調連続性} \\ 0 \leq x_p(i) - x_p(i-1) \leq 2 \\ \text{整合窓条件} \\ i - w \leq x_p(i) \leq i + w \\ \text{境界条件} \\ x_p(1) = 1, \quad x_p(N) = N \end{cases} \quad (4)$$

ここで、 p は 1 または 2 を表し、 w は正定数である。単調連続性は位相構造を保存した範囲でのマッチングを求めするための制約条件である。以上のダッチロールワーブ最適化問題の厳密解は動的計画法 (Dynamic Programming, 以下、DP) によって求めることができる。詳細は付録 1. に示す。

3.3 方向特徴の補正

図 9 からわかるように、ダッチロールワーブは基本的に剪断変形を補償対象としている。(すなわち上下方向の変動は補償できない) この理由により、以下の実験で、方向特徴の補正に剪断変形を用いることとする。

3.4 実験結果および考察

ダッチロールワーブでは、前述のようにワーブ自身が剪断変形の性質を持つ。従ってそのような変形を生じている文字に対して適していると考えられる。そこで、特に剪断変形に着目して本手法の効果を確認する実験を行った。前述した文字サンプルのマッチング結果から、特に本手法の効果が顕著に見られた例を図 10 に示す。図中の「 $\tilde{\mathbf{B}}$ 」, 「 \mathbf{A}_θ 」, 「 $\tilde{\mathbf{B}}'$ 」は、回転を考慮しない場合のワーブによって変形した画像パターン (すなわち $\{b(x_{i,j}, y_{i,j})\}$),

回転を考慮したときの剪断変換後の画像パターン \mathbf{A} , およびそれに対するワーブ後の変形した画像パターン \mathbf{B} を表す。また図の最上部に、ダッチロールワーブによって入力パターン画像の各列が対応した参照パターン画像上の線分を表した。

本手法の効果は、 \mathbf{A} と $\tilde{\mathbf{B}}$ との差異、 \mathbf{A}_θ と $\tilde{\mathbf{B}}'$ との差異の比較によって確認される。すなわち、前者の差異 ($|\mathbf{A} - \tilde{\mathbf{B}}|$) が従来の特徴補正を行わない場合のマッチングの残差であり、後者の差異 ($|\mathbf{A} - \tilde{\mathbf{B}}'|$) が特徴補正を行った場合のマッチングの残差である。この残差が小さいほどマッチングが適切に行われていることを意味する。同図右 2 列には、本手法と従来法でこれら残差に生じた違いを示している。本手法のほうが残差が小さくなった部分、すなわち改善された部分を D_{d+} 、逆に本手法により残差が増えた部分を D_{d-} で表している。

まず \mathbf{A} と \mathbf{B} の輝度特徴を見ると、 \mathbf{A} は \mathbf{B} に対して右方向に剪断変形した形になっていることがわかる。また、方向特徴を比較すると、 \mathbf{A} は $d = 5, 6, 7$ の方向特徴を多く含み、 $d = 3, 4$ の特徴量が少なく、 \mathbf{B} は $d = 3, 4, 5$ の方向特徴を多く含み、 $d = 6, 7$ の特徴量が少ないことがわかる。従って本来対応づけられる画素であっても方向特徴が異なっていることが確認できる。ワーブの過程で、特徴補正として、この剪断変形に応じた剪断変換が施されると、 \mathbf{A} の方向特徴は \mathbf{A}_θ の方向特徴に変わる。この \mathbf{A}_θ の方向特徴は、元の方向特徴に対して $d = 5, 6, 7$ の特徴量が減少し、 $d = 3, 4$ の特徴量が増加していることから、剪断変形に応じて方向特徴の補正が正しく行えていることが確認できる。また、 \mathbf{A} と \mathbf{B} の方向特徴の関係と \mathbf{A}_θ と $\tilde{\mathbf{B}}'$ の方向特徴の関係を見ると、 \mathbf{A}_θ と $\tilde{\mathbf{B}}'$ の方が互いに近い方向特徴量を持っていることがわかる。以上の改善は D_{d+} から明らかである。なお、弾性マッチングでは、全方向特徴を一括した目的関数 (2) を最小化するために、補正により若干残差が増加している方向も存在する。こうした残差が増加した箇所は D_{d-} に現れている。

図 11 は、同じカテゴリの文字間のマッチングにおいて、本手法と従来法が与えた距離 $D(\mathbf{A}, \mathbf{B})$ の差のヒストグラムである。この距離の差が負であれば、本手法のほうが距離 $D(\mathbf{A}, \mathbf{B})$ を小さくできたことに相当し、従って同じカテゴリの文字間の距離をより妥当に評価できていることになる。この実験では、英語大文字 26 カテゴリそれぞれについて 10 サンプルを準備し、各カテゴリ内でそれら 10 サンプルについて総当たりに弾性マッチングを行った。従って、各カテゴリで 90 回のマッチングを行い、全カテゴリで 2340 (= 90 × 26) 対について距離 $D(\mathbf{A}, \mathbf{B})$ を求めている。

同図より，明らかに本手法のほうがより多くの画像対について，距離 $D(\mathbf{A}, \mathbf{B})$ を小さくできていることがわかる．従って，本手法の方が同じカテゴリの文字間の距離をより妥当に評価できていると言える．

4. ま と め

本論文では，弾性マッチング精度の向上を目的とした方向特徴の利用法について考察した．具体的には，入力・参照画像間のワーブを決定する過程において，入力画像側に生じている局所的な回転変形に対して，ある画素の周囲の局所的なワーブの状況に応じて，その回転量を推定し，それによって方向特徴を補正した上で，参照画像側との最適画素対応関係を求めた．その結果，剪断変形に対しての本手法の有効性が確認された．

今後は，回転変形を生じている文字への対処法（ワーブ法，回転の考え方）についても，検討が必要と思われる．また，異なるカテゴリの画像パターンのマッチングの際に，本手法がどのような副作用を生じるかも検討するべきと思われる．

謝 辞

本研究の一部は文部科学省科学研究費補助金（若手 B No.14780293），および総務省戦略的情報通信研究開発推進制度の支援を受けた．

文 献

- [1] 斎藤，山田，山本，“手書き漢字の方向パターン・マッチング法による解析”，電子通信学会論文誌，Vol. J65-D，No. 5，pp. 550-557，1982.
- [2] 内田誠一，迫江博昭，“単調連続 2 次元ワーブ法によるオフライン手書きひらがな認識”信学技報，PRMU98-163，1998.
- [3] 石田敏之，内田誠一，迫江博昭，“手書き文字認識における字形変動処理に関する一検討 - タッチロールワーブ -”，九州大学大学院システム情報科学研究科報告，Vol. 5，No. 1，2000.

付 録

1. DP アルゴリズム

タッチロールワーブ [3] では，DP を用いてワーブの最適化，すなわち目的関数 (2) を条件 (4) の下で最小化する．このとき，DP 漸化式は次のようになる．

$$\begin{aligned} g(i, x_1(i), x_2(i)) \\ &= d(i, x_1(i), x_2(i)) \\ &\quad + \min_{\substack{p=0,1,2 \\ q=0,1,2}} g(i-1, x_1(i)-p, x_2(i)-q) \end{aligned}$$

ここで，

$$d(i, x_1(i), x_2(i)) = \sum_{j=1}^N \sum_{d=1}^8 |a_d(i, j) - b_d(x_{i,j}, y_{i,j})|$$

```

/* 初期化 */
1:  $g(1, 1, 1) := d(1, 1, 1)$ 
/* DP 漸化式 */
2: for  $i := 2$  to  $N$  do begin
3:   for  $x_1 := i - w$  to  $i + w$  do begin
4:     for  $x_2 := i - w$  to  $i + w$  do begin
5:        $g(i, x_1, x_2) := d(i, x_1, x_2)$ 
         +  $\min_{\substack{p=0,1,2 \\ q=0,1,2}} g(i-1, x_1-p, x_2-q)$ 
6:     end
7:   end
8: end
/* 終了 */
9:  $D(\mathbf{A}, \mathbf{B}) := g(N, N, N)$ 

```

図 A.1 タッチロールワーブの DP アルゴリズム

である．ただし， $(x_{i,j}, y_{i,j})$ は，式 (3) に従うものとする．DP 漸化式における $g(i, x_1, x_2)$ は入力パターン画像 \mathbf{A} の第 i 列を参照パターン画像 \mathbf{B} の $(x_1, 1)$ と (x_2, N) で結ばれる線分にワーブしたとき， \mathbf{A} の第 1 列から第 i 列目までの $d(i, x_1, x_2)$ の最小累積値である．タッチロールワーブによる変形補償後の画像間距離 $D(\mathbf{A}, \mathbf{B})$ は $g(N, N, N)$ で与えられる．なお，変数 $x_1(i)$ ， $x_2(i)$ に関して単調連続性条件を満足するため $p, q = 0, 1, 2$ としている．以上の DP アルゴリズムの擬似コードを図 A.1 に示す．