

手書き英単語の局所傾き推定と補正

非会員 Yimei Ding (三重大学)
非会員 若林哲史 (三重大学)
非会員 木村文隆 (三重大学)
正員 三宅康二 (三重大学)

Local Slant Estimation and Correction for Handwritten English Words

Yimei Ding, Non-member, Tetsushi Wakabayashi, Non-member, Fumitaka Kimura, Non-member, Yashiji Miyake, Member (Mie University)

Handwritten words often have slant which need to be estimated and corrected before character segmentation and recognition. Several methods have been proposed for average slant estimation and correction. However, average slant estimation has the problem such that local slant will be overestimated or underestimated when the slant in a word varies from character to character. To solve the problem, this paper proposes three methods for local slant estimation, which are simple iterative method, high speed iterative method and 8-directional chain code method. The experimental results show that the proposed methods can estimate and correct local slant more accurately than the average slant correction.

キーワード：手書き単語認識，傾き推定，チェーンコード法，8方向量子化

1. ま え が き

手書き英単語は筆記者の書体と筆記運動の特性によって、イタリック体のように傾いていることが多い⁽¹⁾。文字の切り出しと単語認識の前に、傾きを補正するかしないかについて意見が二つに分かれている。文献^{(2)~(4)}には傾きを補正しないまま、手書き文字を認識する手法が報告されている。しかしこれらの方法では、大量の傾いた文字サンプルを学習させる必要がある。

一方、傾いた文字の切り出しを容易かつ高精度化し、文字の認識率の向上を図るために、傾きを推定・補正する手法も幾つか提案されている^{(5)~(10)}。以下に傾きを推定する代表的な手法について紹介する。

投影法⁽⁷⁾では、異なる複数方向に投影した文字画像のヒストグラムを分析し、ヒストグラムの導関数（差分）が正の最大値をとる角度を平均傾きの推定量として求める。

ランレンゲス解析法⁽⁵⁾⁽⁶⁾では、長い水平ランを含む水平ストロークを除去して、残った文字ストロークに外接する小ウィンドウにおいて、上半分の重心と下半分の重心から傾きを推定し、すべての小ウィンドウの傾きの平均を単語の傾きとする。しかしこの方法は文字の太さに影響を受けやすく、細い文字に対する傾き推定精度が悪い。

極値点対応法⁽⁸⁾では、英単語の上輪郭線の極大点と下輪

郭線の極小点の合理的な対を生成して、これらの対の角度から単語の傾き角度を推定する。

筆者らは境界画素のチェーンコードを利用して、平均傾き推定の手法（チェーンコード法）を提案した⁽⁹⁾。この方法は単語の平均傾きをより容易に推定できる一方、傾きの絶対値が大きくなると過小評価される欠点があった。そのような欠点を解決するために、さらにチェーンコードの単純反復法と高速反復法を提案し、その有効性を報告した⁽¹⁰⁾。

しかし、単語を構成する個々の文字の傾きが様でない場合、平均傾き補正では十分補正できないことがある。図1に示す例において、上段は原画像、下段は平均傾き補正の結果である。局所傾きが平均傾きより小さい文字が過剰補正され、平均傾きより大きい文字が過小補正されていることがわかる。本論文では、局所的に変動する傾きを補正

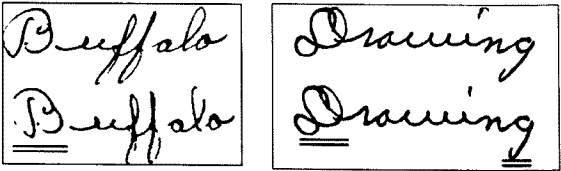


図1 平均傾き補正の問題点
Fig. 1. The Problem of average slant correction

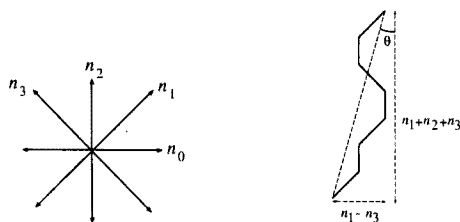


図2 チェインコードの方向指数と平均傾き
Fig. 2. Average slant of a chain code sequence

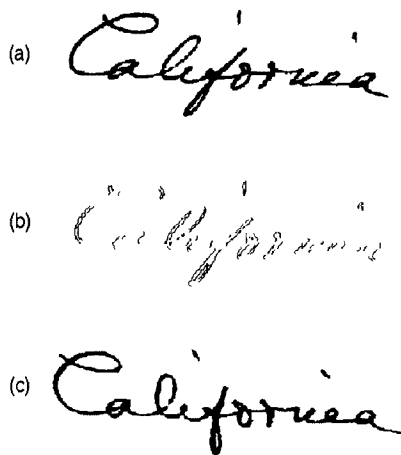


図3 入力画像(a), 非水平境界画素(b), 傾き補正された画像(c)
Fig. 3. Input image(a), non-horizontal border pixels(b), and slant corrected image(c)

するために、局所的なチェインコードの度数分布を利用して、局所傾きを推定、補正する手法について検討する。

2. 平均傾きの推定と評価

〈2・1〉 チェインコード法による平均傾きの推定 チェインコード法⁽⁹⁾による傾き推定では、垂直あるいは対角方向のチェインコードを用いて、平均傾きの推定量を次式により定義する。

$$\theta = \tan^{-1} \left(\frac{n_1 - n_3}{n_1 + n_2 + n_3} \right) \dots \dots \dots (1)$$

ここで、 n_1, n_2, n_3 はチェインコードの成分数を方向別に計数したものである。図2にチェインコードのセグメント(単純連鎖)の傾きが式(1)で計算できることを示す。この例では、 $n_1 = 3, n_2 = 3, n_3 = 1$ である。同様に単語全体の平均傾きも式(1)で推定できることを図3の例で示す。単語の境界線のチェインコードから水平成分を除くと図3(b)に示すような複数のチェインコードのセグメントが得られる。式(1)はこれらのセグメントを平行移動して連結した一つの長いセグメントの傾きを表しており、単語

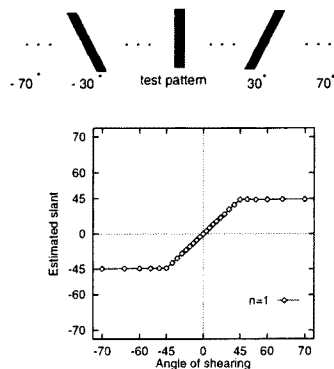
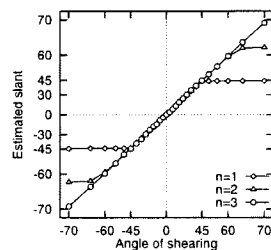


図4 テストパターンにおける推定精度の評価
Fig. 4. Evaluation of estimation accuracy for a test pattern



$$\begin{aligned} y &= 0.56x - 0.00, \quad r = 0.932 \quad \text{for } n = 1 \\ y &= 0.87x + 0.00, \quad r = 0.991 \quad \text{for } n = 2 \\ y &= 0.98x + 0.00, \quad r = 1.000 \quad \text{for } n = 3 \end{aligned}$$

図5 反復法による推定精度の向上
Fig. 5. Accuracy improvement by iterative method

の平均傾きのよい推定値となる。なお、式(1)はチェインコードの水平成分 n_0 には依存しないので、チェインコードから実際に水平成分を除去したり、個々のセグメントの傾きを個別に求めてそれらを平均する必要はない。

〈2・2〉 平均傾き推定の精度の評価 傾き推定の精度を評価するために、 -70° から 70° まで 5° ずつ傾けた入力パターンに対して、推定された角度を計測する。

テストパターンを対象にして、傾けた角度と推定された角度を図4に示す。

この結果から、 $[-45^\circ, 45^\circ]$ の範囲では入力パターンの傾きと推定角度の線形性が高く、入力パターンの傾きが正確に推定される。一方、入力パターンの傾きの絶対値が 45° 以上になると正しい傾き推定ができなくなることがわかる。その理由は式(1)により $|\tan \theta| \leq 1$ からも明らかである。このような問題は傾きの推定と補正を反復して推定精度を向上する反復法によって解決することができる⁽¹⁰⁾。

〈2・3〉 反復法による推定精度の向上 テストパターン(図4)を対象として、傾きの推定と補正を反復した場合の入力パターンの角度と推定角度の関係を図5に示す。

この結果から反復回数が増えると、線形性が改善され推

定できる傾きの範囲が広がることわかる。図5にはまた、最小2乗法による回帰直線と相関係数 r を示す。反復回数が増えるとともに相関係数が1に近づき、回帰直線の傾きも1に近づくことがわかる。

3. 局所傾きの推定と補正

単語を構成する個々の文字の傾きが一律でない場合に、平均傾き補正を利用すると、局所的に過剰補正或は過小補正になる欠点がある(図1)。このような欠点を改善するため、チェーンコードの累積度数分布から局所傾きを求める手法について以下に述べる。

〈3・1〉 局所傾きの推定 水平座標 x の列における局所傾きの推定値を次式により定義する。

$$\theta(x) = \tan^{-1} \left[\frac{n_1(x) - n_3(x)}{n_1(x) + n_2(x) + n_3(x)} \right] \dots\dots\dots (2)$$

ただし、 $n_i(x)$ ($i = 1, 2, 3$) は区間 $[x - \delta x, x + \delta x]$ の $i \times 45^\circ$ 方向のチェーンコードの数であり、 δx は画像の高さの定数倍(実験では0.5~1.0倍)とする。

$n_i(x)$ ($i = 1, 2, 3$) は次の式で計算することができる。

$$n_i(x) = s_i(x + \delta x) - s_i(x - \delta x - 1) \dots\dots\dots (3)$$

ここで、 $s_i(x)$ は $[0, x]$ における $i \times 45^\circ$ 方向のチェーンコードの累積度数である。

さらに、求められた $\tan \theta(x)$ に対して、近隣3点の平均値を取る平滑化処理を反復する。

ある入力パターンにおける推定された局所傾き $\tan \theta(x)$ 、とそれを10回平滑化した曲線の例を図6に示す。

〈3・2〉 局所傾きの補正 推定された局所傾きを用いて次式により入力画像を座標変換(shear変換)して、局所傾きを補正する。

$$\begin{cases} x' = x + y \tan \theta(x) \\ y' = y \end{cases} \dots\dots\dots (4)$$

ただし、 x, y は変換前の座標、 x', y' は変換後の座標である。

式(2)において $|\tan \theta(x)| \leq 1$ となるために、平均傾き推定と同様に局所傾き推定にも過小評価の問題が生じる。つまり局所傾きの絶対値が 45° 以上になると正しい推定ができないことがわかる。このような問題を解決するために、以下に局所傾き推定の精度を向上するための改良方法について述べる。

4. 局所傾き推定の精度向上

〈4・1〉 単純反復法

- (1) (入力)画像を輪郭追跡して輪郭線のチェーンコードを求める。
- (2) チェインコードの累積度数分布を利用して画像の局所傾きを推定し、平滑化する。

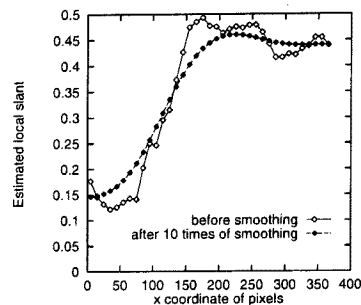


図6 平滑化前後の局所傾き $\tan \theta(x)$
Fig. 6. Local slant $\tan \theta(x)$ before and after smoothing

- (3) 全画素に対して Shear 変換によって、局所傾き補正を行う。
- (4) 輪郭線に生じたノイズを除去するために 3×3 の平均フィルタで画像を平滑化する。
- (5) (1) から (4) を指定された回数反復する。

実験では、局所傾きの推定と補正を2, 3回反復することによって推定精度が向上した。

しかし、単純反復法では、全画素に対して、局所傾き推定、補正と平滑化のための画像処理を反復するため、反復回数に比例して処理時間が増加する。

この問題を解決するために、局所傾きを推定する過程において、傾き補正と平滑化を画像のかわりにチェーンコードに対して反復する高速反復法について以下に述べる。

〈4・2〉 高速反復法

- (1) (入力)画像を輪郭追跡して輪郭線のチェーンコードを求める。
- (2) チェインコードの累積度数分布を利用して画像の局所傾きを推定し、平滑化する。
- (3) (1) から (5) を指定された反復回数反復したら6へ。
- (4) チェインコードを局所傾き補正する。
- (5) チェインコードを平滑化する。2へ戻る。
- (6) 反復過程で求められた $\tan \theta(x)$ を合成し、全画素を対象にして、局所傾き補正を行う。
- (7) 輪郭線に生じたノイズを除去するために画像を平滑化する。

ステップ(5)で行うチェーンコードの平滑化の必要性と方法について以下に述べる。

図7にチェーンコードの傾き補正における問題点を示す。同図(a)のチェーンコードを傾き補正して得られる同図(b)のチェーンコードは、不要な4連結成分が増加して、

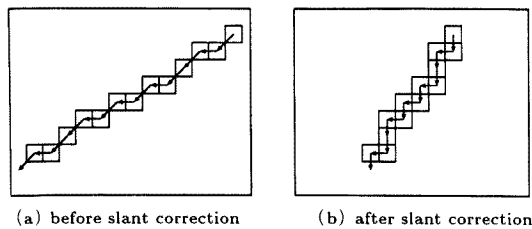


図 7 傾き補正前後のチェーンコード
Fig. 7. Slant correction for a chain code

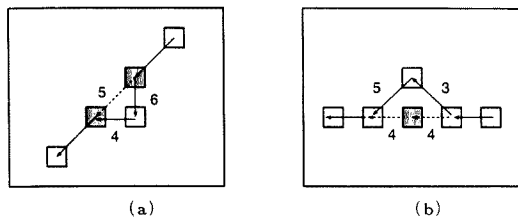


図 8 偶数直角成分と奇数直角成分
Fig. 8. Even orthogonal pair (a) と Odd orthogonal pair (b)

対角方向のチェーンコードの数 n_1 , n_3 が減少し、逆に水平、垂直方向のチェーンコードの数 n_0 , n_2 が増加することを示している。このような理由によって傾きが過小評価されることを防ぐために、以下の方法でチェーンコードの平滑化を行う。

チェーンコード上で連続する要素の対 $a_i \cdot a_{i+1}$ に対して、

$$\text{mod}_4(|a_i - a_{i+1}|) = 2 \dots\dots\dots (5)$$

が成立つとき、 $a_i \cdot a_{i+1}$ は直角成分を成す。ただし $\text{mod}_4(n)$ は、 n を 4 で割った余りである。このような直角成分に対して次式

$$p = \begin{cases} \text{mod}_8(a_{i+1} + 1) & \text{if } \text{mod}_8(a_{i+1} - a_i) = 6 \\ \text{mod}_8(a_{i+1} - 1) & \text{if } \text{mod}_8(a_{i+1} - a_i) = 2 \end{cases} \quad (6)$$

を計算し、 $a_i \cdot a_{i+1}$ が奇数直角成分の時には $p \cdot p$ 、偶数直角成分の時には p に置換える (図 8)。

〈4・3〉 8 方向チェーンコード法による局所傾き推定 手書き英字はその構造に曲線成分を多く含むため、境界画素のチェーンコードを 4 方向に量子化することが、傾き推定精度の低下の一つの原因と考えられる。この問題の対策として、筆者らは 8 方向チェーンコード法による平均傾き推定法を提案する。

境界画素を一点置きに追跡することにより、チェーンコードを 8 方向に量子化することができる (図 9)。 $n_i (i = 1 \sim 7)$ は方向 i のチェーンコードの数である。

8 方向のチェーンコードを用いる平均傾きの推定量を次式により定義する。

$$\tan \theta = \frac{(2n_1 + 2n_2 + n_3) - (n_5 + 2n_6 + 2n_7)}{(n_1 + 2n_2 + 2n_3) + 2n_4 + (2n_5 + 2n_6 + n_7)} \quad (7)$$

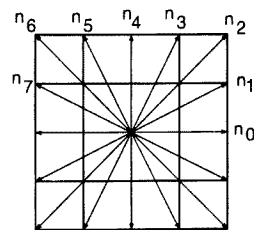


図 9 チェインコードの 8 方向量子化
Fig. 9. 8-directional quantization of chain code

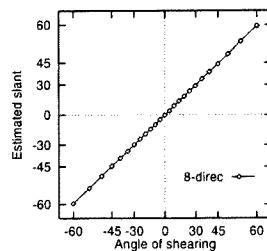


図 10 テストパターンにおける 8 方向チェーンコード法の推定精度の評価
Fig. 10. Evaluation of estimation accuracy of 8-directional method for the test pattern

ここで、 $(2n_1 + 2n_2 + n_3)$ は方向 1, 2, 3 のチェーンコードの水平成分の総数、 $(n_5 + 2n_6 + 2n_7)$ は方向 5, 6, 7 のチェーンコードの水平成分の総数、分母は方向 1~7 のチェーンコードの垂直成分の総数である。

図 10 にテストパターンにおける 8 方向チェーンコード法の平均傾き推定特性を示す。この結果から、チェーンコードの方向量子化数を 8 方向に増加すると、推定特性の線形性が高く、推定できる範囲が $[-\tan^{-1}(2), \tan^{-1}(2)]$ になり、4 方向推定法より広くなることが分かる。

これらの結果に基づいて、境界画素の 8 方向チェーンコードを利用する局所傾きの推定量を次式から求める。

$$\tan \theta = [(2n_1(x) + 2n_2(x) + n_3(x)) - (n_5(x) + 2n_6(x) + 2n_7(x))] / [(n_1(x) + 2n_2(x) + 2n_3(x)) + 2n_4(x) + (2n_5(x) + 2n_6(x) + n_7(x))] \dots\dots\dots (8)$$

ただし、 $n_i(x) (i = 1 \sim 7)$ は区間 $[x - \delta x, x + \delta x]$ の i 方向のチェーンコードの数で、式 (3) と同様に計算することができる。

5. 比較実験

〈5・1〉 局所傾き推定の比較 図 11 に単純反復法、高速反復法、8 方向チェーンコード法による局所傾き推定結果の比較を示す。

(b) には、単純反復法による 1 回目の推定値 (平滑化前

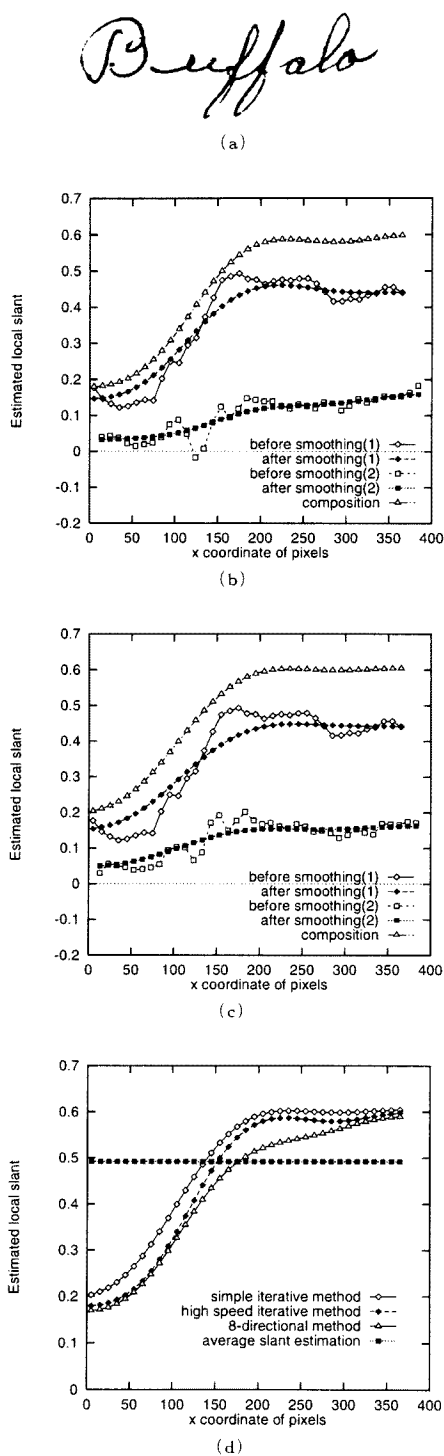


図 11 局所傾き推定値 $\tan \theta(x)$ (a) 原画像 (b) 単純反復法 (c) 高速反復法 (d) 3 手法と平均傾き推定結果の比較

Fig. 11. Estimated local slant $\tan \theta(x)$ (a) input image (b) simple iterative method (c) high speed iterative method (d) Comparison of three methods and average slant estimation

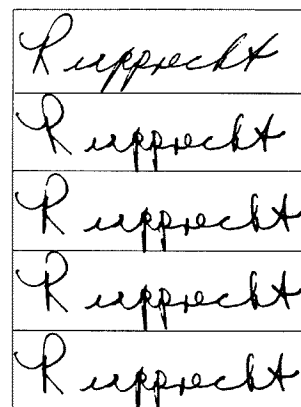
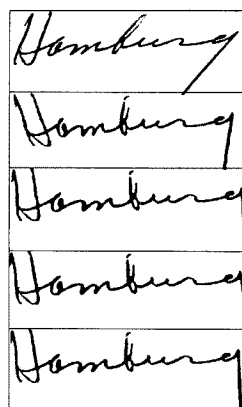
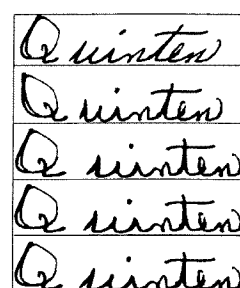
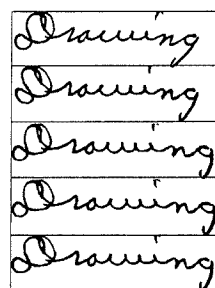
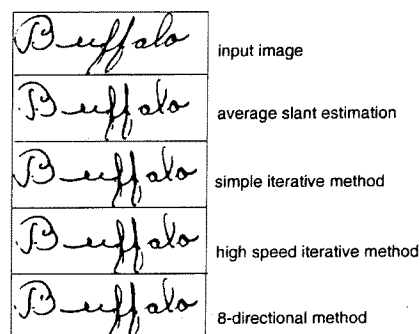


図 12 局所傾き補正の例

Fig. 12. Examples of local slant correction

後), 2 回目の推定値 (平滑化前後) と 2 回の合成値の曲線を示す。(c) には, 高速反復法による 1 回目の推定値 (平滑化前後), 2 回目の推定値 (平滑化前後) と 2 回の合成値の曲線を示す。(d) には, 単純反復法 (simple iterative method) の 2 回推定合成値, 高速反復法 (high speed iterative method) の 2 回推定合成値, 8 方向チェーンコード法 (8-directional method) の推定値と平均傾き (average slant estimation) 2 回推定値の比較を示す。

(d) の結果から平均傾き推定に比べて, 局所傾き推定の 3 手法により, 傾きが局所的に推定されることがわかった。また, 3 手法の推定特性から, 8 方向局所推定法の反復数 1 における推定精度は単純反復法や高速反復法の反復数 2 における推定精度に近いことがわかった。

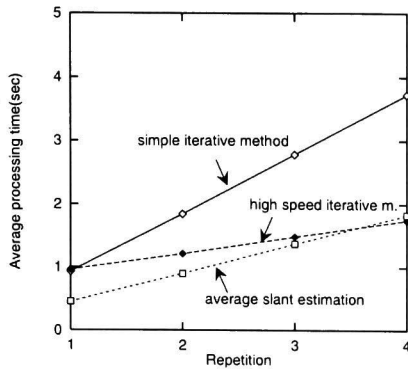


図 13 平均処理時間の比較

Fig. 13. Processing time v.s. number of iteration

(5・2) 補正結果の比較 図 12 に傾き補正結果の比較を示す。それぞれの例において、一行目は原画像、二行目は平均傾き補正（反復数 2）の結果、三行目は単純反復法による局所傾き補正の結果（反復数 2）、四行目は高速反復法による局所傾き補正の結果（反復数 2）、5 行目は 8 方向局所傾き推定法の補正結果である。

これらの結果から、平均傾き補正より、局所傾き補正のほうがより正確に局所的に変動する傾きを補正できることがわかる。

さらに、高速反復法と 8 方向局所傾き推定法のほうが全画素に対する傾き補正を 1 回しか行わないため、単純反復法より輪郭線の画質がよくなることがわかった。

(5・3) 処理速度の評価と比較 単純反復法、高速反復法、8 方向局所傾き推定法の処理速度を比較するために、入力画像を -50° から 50° まで 5° おきに傾けて、傾き推定と補正にかかる時間を計測した。

20 個の手書き英単語に対する平均処理時間と反復回数の関係を図 13 に示す。使用した CPU は hyper SPARC 125MHz である。

図 13 から、単純反復法による局所傾き推定では、全画素に対する局所傾き推定、補正と平滑化のための画像処理を反復するため、反復回数に比例して処理時間が増加する。また、高速反復法では、全画素に対する局所傾き補正と平滑化を一回しか行わないため、単純反復法より処理時間が短縮された。さらに、8 方向局所傾き推定法の処理時間は、反復処理が不要なため、単純反復法や高速反復法の一回の処理時間とほぼ同じである。

6. ま と め

局所的なチェーンコードの度数分布を利用して局所傾きを推定・補正することにより、局所的に変動する傾きを補正することができるようになった。また、8 方向局所傾き推定法の反復数 1 における推定精度は単純反復法や高速反復法の反復数 2 における推定精度に近いことがわかった。単純反復法は反復回数に比例して処理時間が増加する。高速

反復法のほうが単純反復法より処理時間が短縮された。さらに、8 方向局所傾き推定法の処理時間は、反復処理不要なため、単純反復法や高速反復法の一回の処理時間とほぼ同じであることがわかった。

今後の課題として、

- (1) g , j , y など手書き文字が、余分に書かれた右上がりのストロークのため、傾きが過剰推定されることがある。このような問題を解決するための傾き推定手法の改良、
 - (2) 平均傾き補正や局所傾き補正の手書き英単語の認識システムへの応用と評価、
 - (3) 日本語文書において、右上がりに書かれた漢字に対する傾きの推定と評価、
- について検討する予定である。

(平成 12 年 03 月 27 日受付, 同 12 年 07 月 03 日再受付)

文 献

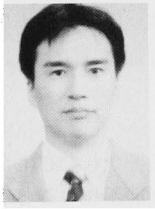
- (1) L.Simoncini, Zs.M.Kovacs-V, "A System for Reading USA Census'90 Hand-Written Fields", Proc. of the 3th ICDAR, Vol. II, pp. 86-91, Aug. 1995.
- (2) M.K. Brown and S.Ganapathy, "Preprocessing Techniques for Cursive Word Recognition", IEEE Trans.Pattern Recognition, vol.16, no.5, pp. 447-458, 1983.
- (3) M.J.J. Holt, M.M. Beglou, and S. Datta, "Slant Independent Letter Segmentation for Off-line Cursive Script Recognition", From PixelstoFeatures2, S.Impedovo and J.C.Simon, eds., pp.41-47, Elsevier Science, 1992.
- (4) T.Caesar, J.m. Gloger and E.Mandler, "Estimating the Baseline for Written Material", Proc. 3th ICDAR, pp.382-385, Montreal, 1995.
- (5) R.M. Bozinovic and S.N. Srihari, "Off-line Cursive Script Word Recognition", IEEE Transactions on PAMI, vol. II, No. 1, pp. 68-83, Jan. 1989.
- (6) A. W. Senior, F.Fallside, "An Off-line Cursive Script Recognition System Using Recurrent Error Propagation Networks", Proc. of 3th IWFHR, PP. 132-141, May 1993
- (7) Didier Guillevic, Ching Y. Suen, "Cursive Script Recognition: A Sentence Level Recognition Scheme", Proc. of The 4th IWFHR, pp. 216-223, Dec.1994.
- (8) Atul Negi, K.S.Swaroop, Arun Agarwal, "A Correspondence Based Approach to Segmentation of Cursive Words", Proc. of the 3th ICDAR, vol. II, pp. 1034-1037, Aug. 1995.
- (9) F.Kimura, M.Shridhar and Z.Chen, "Improvements of a Lexicon Directed Algorithm for Recognition of Unconstrained Handwritten Words", Proc. of the 2th ICDAR, pp. 18-22, Oct. 1993.
- (10) Yimei Ding, Fumitaka Kimura, Yasuji Miyake, Malayappan Shridhar, "Evaluation and Improvement of Slant Estimation for Handwritten Words", Proc. of the Fifth ICDAR, pp. 753-756, Sep. 1999.

Yimei Ding (非会員) 1991 中国・東南大学・工・計算機科学卒。



平 11 三重大・工・大学院博士前期課程了。同年同大学院博士後期課程入学，現在に至る。手書き文字認識，文書理解，画像処理の研究に従事。

若 林 哲 史 (非会員) 昭 60 三重大・工・電子卒。昭 62 同
大大学院修士課程了。平 3 三重大・工・助手。平
10 同助教授，現在に至る。その間平 10～11 レ
ンセレア工科大学 (RPI) 客員研究員を務める。手
書き文字認識，文書理解，画像処理，コンピュー
タグラフィックスの研究に従事。情報処理学会会
員。工博。



木 村 文 隆 (非会員) 昭 48 名大・工・電気卒。昭 53 同大
大学院博士課程了。同年同大・工・助手。昭 58
三重大・工・助教授。平 10 同教授，現在に至る。
その間平 1～3 ミシガン大客員助教授を務める。
文字・パターン認識，画像処理，コンピュータグ
ラフィックスの研究に従事。情報処理学会，日本
ME 学会，人工知能学会各会員。工博。



三 宅 康 二 (正員) 昭 35 名大・工・電気卒。昭 40 同大大学
院博士課程了。同年同大・工・助手，昭 43 同講師。
昭 43 同助教授。昭 53 三重大・工・教授，現在に
至る。コンピュータグラフィックス及び医用画像
処理の研究に従事。情報処理学会，日本 ME 学
会，日本人工知能学会，日本ロボット学会，電気
学会，日本シミュレーション学会等各会員。工博。

