

画像認識への応用を目的とした非一様な傾き補正法

高山誠悟* 平英二* 内田誠一** 迫江博昭**

(*九州大学大学院システム情報科学府, **九州大学大学院システム情報科学研究院)

1 はじめに

物体が複数並んでいるような画像について、各物体の傾きはそれらの境界を検出する場合に悪影響を及ぼす。その対応策として、推定された平均傾きを基に一括して傾き補正を行う方法や、傾きに依存しない特徴を抽出する方法が検討されている。本稿では前者を改良し、複数の物体がそれぞれ独立に傾いているような画像も補正できる非一様な傾き補正法を提案する。さらに具体的な応用として、書棚画像と手書き文字列に対する傾き補正について述べる。

2 非一様な傾き補正法

2.1 非一様な傾き補正のモデル

画像 A を水平方向に非一様に傾いた部分パターンからなるサイズ $M \times N$ の画像とする。本手法では、非一様な傾きを図 1(a) のように各列における局所傾きとして表し、その補正のために局所傾きの最適推定を行う。ここで第 i 列の局所傾きは 2 点 $(i, 1), (p_i, N)$ を結ぶ線分で表す。局所傾きの最適推定問題は系列 $p_1, \dots, p_i, \dots, p_M$ の最適推定問題と等価である (図 1(b))。但し、 p_i は整数値とし、

$$0 \leq p_i - p_{i-1} \leq 2, \quad |p_i - i| \leq W \quad (1)$$

を満たすものとする。ここで、 W は正の定数とする。左式は画像の単調連続性を保持するための条件であり、右式は局所傾きの最大値を制限することで計算量を低減させるための条件である。

以下に述べる手法を用いて p_i の系列を最適推定した後、 $(i, 1), (p_i, N)$ を結ぶ線分上の画素を第 i 列に垂線としてマッピングすることで傾き補正処理後の画像を得る。

2.2 系列の最適推定問題の定式化

系列 $p_1, \dots, p_i, \dots, p_M$ の最適推定問題を次のような評価関数の制約 (1) 付き最大化問題として定式化する。

$$\sum_{i=1}^M [f_i(p_i) + \rho_i(p_i|p_{i-1})] \Rightarrow \max \quad (2)$$

関数 $f_i(p_i)$ は第 i 列での p_i の妥当性を評価する関数であり、関数 $\rho_i(p_i|p_{i-1})$ は傾き変化する p_i と p_{i-1} の変化の滑らかさを評価する関数である。関数 $f_i(p_i)$ は補正する対象の特性を反映した関数であり、その具体的な役割については 3 節で紹介する。

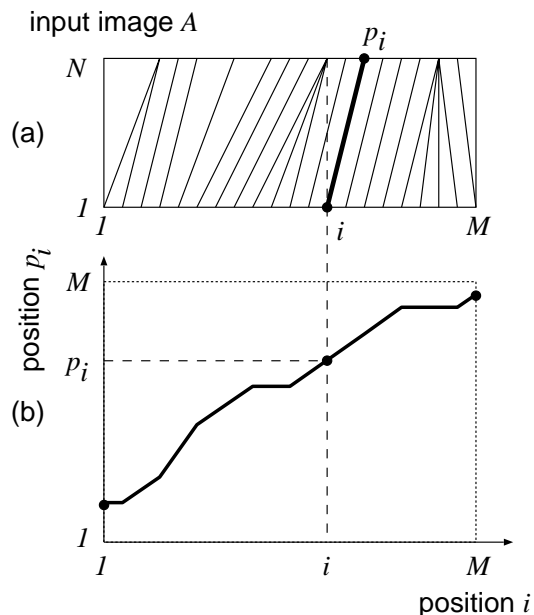


図 1: 非一様な傾きのモデル (a) とその関数表現 (b)。

2.3 動的計画法による解法

p_1 から p_M への逐次決定過程を考えると、制約 (1) および評価関数 (2) の定義により p_i の決定は p_{i-1} のみに依存するので、この過程はマルコフ過程となる。よって系列 $p_1, \dots, p_i, \dots, p_M$ の最適推定は動的計画法 (DP) を用いることで効率的に解くことができる。具体的には次の DP 漸化式を $i = 1$ から M まで順に全ての可能な p_i について計算する。

$$g_i(p_i) = f_i(p_i) + \max_{p_{i-1}=p_i-\{0,1,2\}} [g_{i-1}(p_{i-1}) + \rho_i(p_i|p_{i-1})] \quad (3)$$

ただし、 $i = 1$ においては右辺第 2 項を 0 とする。ここで、 $g_i(p_i)$ は $[f_i(p_i) + \rho_i(p_i|p_{i-1})]$ の値を第 1 列から第 i 列まで累積したものの最大値となっている。図 2 は DP による最適系列の決定過程を表したものである。漸化式 (3) を図 2 の矢印に従って繰り返し計算し、 M 段目 ($i = M$) において最大値となる $g_M(p_M)$ を求めれば、それが (2) の最大値となる。その後、バックトラック処理により最適系列を得る。

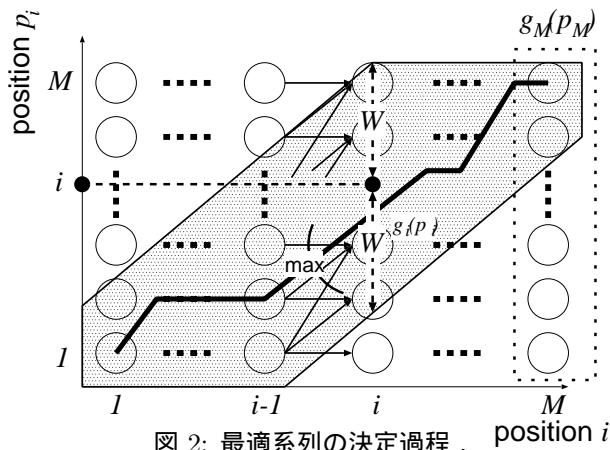


図 2: 最適系列の決定過程 . position i

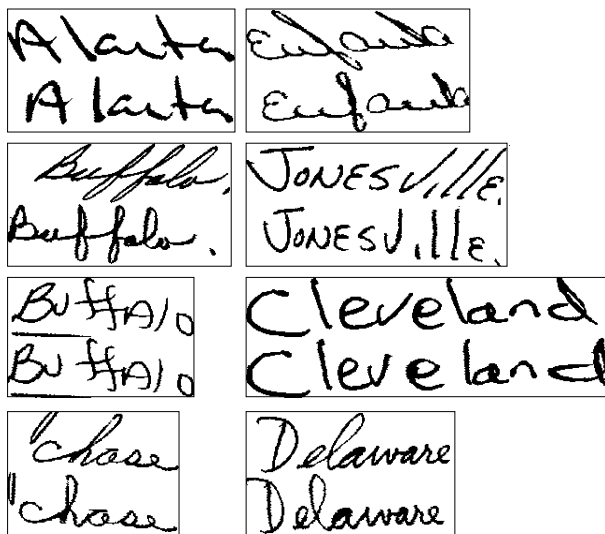


図 3: 手書き文字列の傾き補正結果 . 上: 原画像 . 下: 補正画像 .

3 応用例

3.1 手書き文字列の傾き補正 [1, 2]

入力画像として CEDAR CDROM(city and state words) に含まれる手書き英単語画像を用いて本手法による傾き補正実験を行った . 関数 $f_i(p_i)$ は, p_i が示す線分上に存在する縦ストロークの高さに相当し, 長い縦ストロークの傾きを強調する役割を担う . これは長い縦ストロークの傾きはそのストロークを含む文字の傾きを強く反映しているという観察結果に基づいている .

図 3 に傾き補正結果を示す . 本手法により各文字ごとに適切な傾き補正が行われ, 非一様な傾きを持つ文字列についても各文字がほぼ直立した形に補正されていることがわかる .

3.2 書棚中の書籍の傾き補正 [3]

入力画像としてデジタルカメラで撮影した書棚画像を用いて本手法による傾き補正実験を行った . 補正する対象

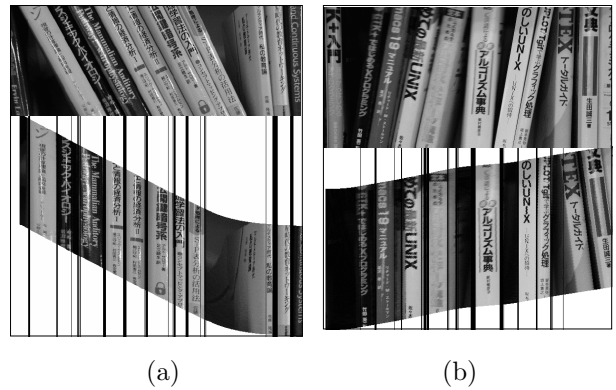


図 4: 書棚中の書籍の傾き補正結果 . (a) と (b) 共に上が原画像 . 下が補正画像 . 書籍境界の検出結果を画像上に黒線で示す .

は書籍の背表紙部分とする . 関数 $f_i(p_i)$ は, p_i が示す線分が背表紙部分の特徴に合致する場合に高い値を与える . こうすることで, 得られる系列は非一様に傾いている書籍の背表紙部分に沿うことになり, 書籍の傾きとして適切な系列が推定される .

図 4 に傾き補正結果を示す . 本手法により書籍が非一様に傾いている書棚画像でも補正できた . また, 補正と同時に書籍境界を検出した結果, ほとんどの書籍境界を検出することができた . なお書棚画像中には書籍のひら部分が見ることがあり, 図 4(b) の結果では書籍背表紙の境界を検出する際に誤検出の原因となっている .

4 まとめ

傾き補正問題を局所傾きの最適推定問題として定式化し, DP に基づいたアルゴリズムにより最適解を求める非一様な傾き補正法を提案した . 手書き文字列と書棚書籍についてその傾きの補正実験を行い, 本手法の有効性を確認した .

今後は得られた補正画像からの領域抽出, 認識について検討する . また, 他の応用対象に関しても本手法の適用可能性を検討する .

謝辞: 本研究の一部は文部科学省科学研究費補助金 (No.14780293) によった .

参考文献

- [1] 平, 石田, 内田, 迫江, “手書き文字列の非一様な傾き補正法,” PRMU2000-207, Mar 2001.
- [2] S. Uchida, E. Taira, H. Sakoe, Nonuniform slant correction using dynamic programming, Proc. 6th ICDAR, pp. 434-438, Sep 2001.
- [3] 高山, 平, 内田, 迫江, “動的計画法を用いた書棚画像中の書籍の傾き補正法,” 九州大学システム情報科学紀要, 7 巻, 1 号, pp. 37-40, Mar 2002.