九州大学大学院
システム情報科学紀要
第7巻　第1号　平成14年3月

Research Reports on Information Science and
Electrical Engineering of Kyushu University
Vol.7, No.1, March 2002

動的計画法を用いた書棚画像中の書籍の傾き補正法

高山誠悟* ・ 平 英 二* ・ 内田誠一**・ 迫江博昭**

Slant Correction for Bookshelf Image Using Dynamic Programming

Seigo TAKAYAMA , Eiji TAIRA, Seiichi UCHIDA and Hiroaki SAKOE

(Received December 14, 2001)

Abstract: For building book database or managing book arrangement automatically, systems to extract book information, such as title, from bookshelf images have been studied. In such systems, slant of each book in bookshelf images may degrade the extraction accuracy. In this paper, a slant correction technique for bookshelf images is proposed where the slant correction problem is formulated as an optimal estimation problem of local slant angles at all horizontal positions. The optimal estimation is provided by a dynamic programming-based algorithm. In the present technique, color features are utilized for improving the slant estimation accuracy. The effectiveness of the present technique was shown by experiments.

Keywords: Slant correction, Dynamic programming, Bookshelf image

1. は じ め に

図書館など大量の書籍を扱う場所では，書架中の書籍の管理に多大の労力を必要とする．配架情報を電子化し省力化をはかる検討もなされているが，計算機内の情報と実際の配架状況の整合性を保つためには定期的に書棚を検査する必要がある．こうした現状に対して管理の自動化を図る手法として書棚に配架された各書籍名の自動認識技術が検討されている[1)]．

本研究では，書棚書籍名認識の第１段階として，書棚を撮影して得られた画像中の各書籍の傾き補正を行う手法を検討する．これは書籍の出し入れが繰り返されることにより生じる非一様な傾きを補正して，認識を容易にするという前処理技術である．従来の研究では，エッジ抽出により書籍境界を検出し，その結果から傾きを推定・補正するという方法が試みられている[2)−4)]．しかし，照明条件や書籍のデザイン，タイトル文字フォントの影響を考えると，エッジのような局所情報のみに依存するよりは，全体的な整合性も考慮して傾き補正および境界検出を行った方が安定した結果が得られると考えられる．

本論文では，手書き文字列の傾き補正法[5)]の考え方を書棚画像に応用した書籍の傾き補正法を提案する．本手法では，局所傾き角の系列として全体の傾き変化をとらえ，その最適推定問題を解くことで局所的および全体的に妥当な傾き系列を求める．具体的にはエッジに加えて背表紙部分での色彩の一様性等を局所的に評価し，これを書棚全体で統合した量を目的関数とする．さらに傾きの推移が実際の書籍列の傾き現象と整合するように制約を設けた動的計画法を用い，目的関数を最小化する傾き系列を得る．

以下では本論文の基本原理を述べ，実際に撮影した書棚画像を用いて実験を行い，本手法の有効性を確認する．

2. 書籍傾き補正法

2.1 概要

書棚画像中の書籍の非一様な傾きを補正するために，本手法では局所傾き角の最適推定を行う(**Fig.2**)．書棚画像$A = \{\boldsymbol{a}(i,j)\}$をサイズ$M \times N$のRGBカラー画像とする．$\boldsymbol{a}(i,j)$は位置$(i,j)$の画素値のR成分，G成分，B成分よりなる3次元ベクトルとする．局所傾き角を画像の第1行の列座標値iごとに推定するものとし，第i列の局所傾き角は画像の第1行の列座標値iに対応する第N行の列座標値p_iで表す．以下，第N行を上辺，第1行を下辺とし，上辺の点(p_i, N)と下辺の点$(i, 1)$とを結ぶ線分を補正線分と呼ぶ．

局所傾き角の系列$p_1, \ldots, p_i, \ldots, p_M$の最適推定問題を次の評価関数の最小化問題として定式化する．

$$\sum_{i=1}^{M} \{f_i(p_i|p_{i-1}) + \rho(p_i|p_{i-1})\} \tag{1}$$

ここで関数$f_i(p_i|p_{i-1})$は第$(i-1)$列の局所傾き角p_{i-1}に対しての第i列の局所傾き角p_iの妥当性を評価する関数であり，妥当性が高いほど小さい値を与える．関数$f_i(p_i|p_{i-1})$の詳細は**2.2**で述べる．一方，関数$\rho(p_i|p_{i-1})$は局所傾き角p_iとp_{i-1}の変化量を表す関数で，次式で与える．

平成 13 年 12 月 14 日受付

* 知能システム学専攻修士課程
** 知能システム学部門

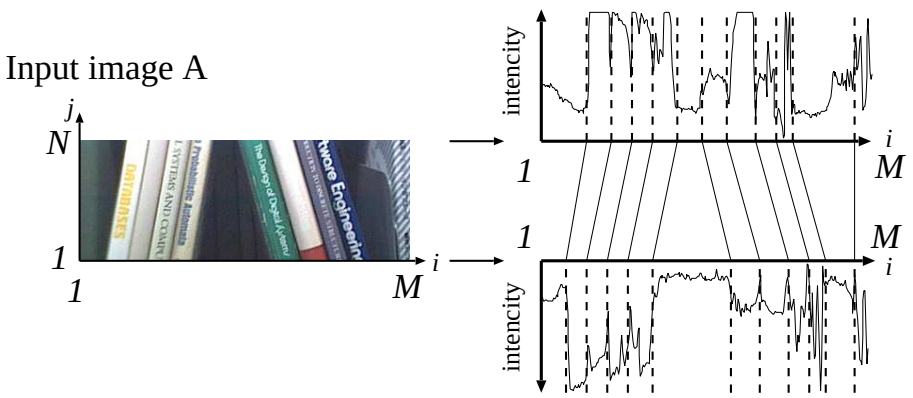

Fig.1　Matching between top and bottom lines.

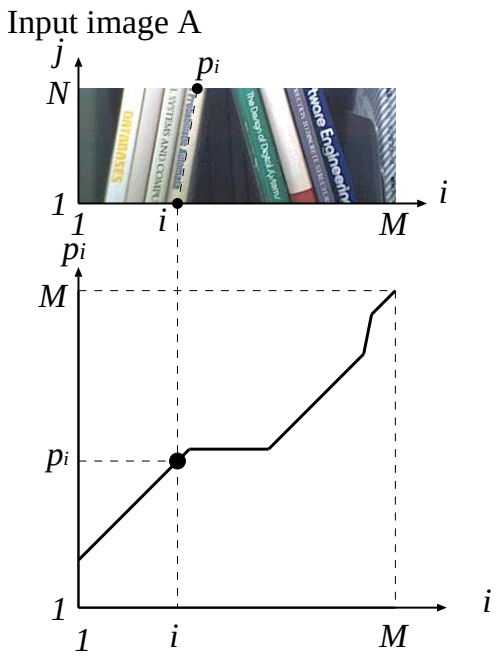

Fig.2　Nonuniform slant.

$$\rho(p_i|p_{i-1}) = \omega_\rho |p_i - p_{i-1} - 1| \tag{2}$$

ここでω_ρは重み係数である．これにより局所傾き角の系列は滑らかに変化する．

ただし，最小化は次の2つの制約条件の下で行う．

第1の条件は単調性条件

$$p_i \geq p_{i-1} \tag{3}$$

である．これは傾き補正で画素位置が反転しないための条件である．

第2の条件は境界条件であり，書棚画像の左右の端部での傾き角を

$$p_1 = 1, \quad p_M = M \tag{4}$$

とする．これは書棚画像の左右の端部に存在する一部しか見えていない書籍の影響を抑制するための条件である．

傾き補正は，推定した局所傾き角p_iが示す補正線分上の画素を垂直にマッピングすることにより行う．

2.2　傾き角の妥当性評価関数

評価関数(1)の第1項は次式で与える．

$$f_i(p_i|p_{i-1}) = \sum_{q=\min\{p_{i-1}+1, p_i\}}^{p_i} \{\omega_s s_i(q) + \omega_c c_i(q) + \omega_l l_i(q)\} \tag{5}$$

ここで$\omega_s, \omega_c, \omega_l$は重み係数である．式(5)中の総和は，$(i,1), (p_{i-1}, N), (p_i, N)$を頂点とする3角形の全領域について評価するためのものである．

関数$s_i(p_i)$は補正線分の上下両端の色の相違度を表し，その値は次式で与える．

$$s_i(p_i) = ||\boldsymbol{a}(p_i, N) - \boldsymbol{a}(i, 1)||^2 \tag{6}$$

関数$s_i(p_i)$のみについての最小化問題を考えれば，上辺と下辺の色変化パターン間の1次元マッチングになる(**Fig.1**)．書籍の表紙色はタイトル文字が目立つように単色で塗られていることが多い．そのため，マッチングにより得られる下辺を基準とした上辺の列座標値の系列は理想的な局所傾き角の系列を近似していると考えられる．

関数$c_i(p_i)$は補正線分に含まれる画素の色の一様性を示す関数である．具体的には次の手順で関数の値を与える．

1. RGB各成分をそれぞれ5bitに量子化して$32 \times 32 \times 32 = 32768$色のそれぞれの出現頻度を求める．ここで出現頻度は同じ色を持つ画素の数とする．
2. 出現頻度の高い5色について，量子化した色空間でのユークリッド距離が2以下の色があれば上位のもの

に統合して一色とみなす．

3. 上位5位までの色が全て1位の色に統合されたとき，統合された色の出現頻度を$c_i(p_i)$とする．上位5位までの色の中に他の色が混ざってるとき，1位の色の出現頻度から次に多い色の出現頻度を引いたものを$c_i(p_i)$とする．

手順1，2で表紙色の候補を絞り，3で表紙色の特定とその一様性を評価する．結果的に補正線分に含まれる画素の色がほとんど同じ場合に$c_i(p_i)$は大きくなり，表紙色が一様な書籍を通る補正線分の傾き角を強調する．評価関数を最小化するため，重み係数ω_cは負値とする．

関数$l_i(p_i)$は補正線分に含まれるエッジの長さ(エッジ点の個数)を与える関数である．書棚画像では，隣接する書籍間の境界や，書籍間の影と書籍の境界，背景と書籍の境界でエッジが現れるため，エッジを多く含む補正線分は書籍の境界を表していると考えられる．エッジ抽出には広井ら[3]と同様にRGBカラー画像からHSY変換し，各要素のエッジ画像の中で強度が最大となるものを集めてエッジ画像とする手法を用い，入力画像Aの2値エッジ画像をBとする．関数$l_i(p_i)$は非エッジ画素の数を与えることで画像Bにおけるp_iが示す補正線分上にエッジ部分を多く含むほど小さな値を与えるようにする．また，非エッジ画素の数が$N/2$以上のものは書籍境界のエッジとして不適切と考え，$l_i(p_i) = N$とする．

2.3 DPアルゴリズム

評価関数(1)の最小化問題を効率的に解くために，動的計画法(Dynamic Programming:DP)に基づく以下のアルゴリズムを用いる．本手法で用いるDPアルゴリズムを以下に示す．

1. 初期化
   $g_1(1) = f_1(1|1)$
2. DP漸化式
   以下の計算を$i = 2, \ldots, M$の順に全てのp_iについて行う．
   $g_i(p_i) = \min_{q \in [1, p_i]} \{g_{i-1}(q) + f_i(p_i|q) + \rho(p_i|q)\}$
3. 終了
   評価関数の最小値は
   $g_M(M)$
   で与えられる．

本アルゴリズムの計算量は$O(M^3N)$である．ただし，局所傾き角をある範囲$W(0 < W << M)$に抑える整合窓条件

$$|p_i - i| \leq W \tag{7}$$

を加えることで計算量を$O(MNW^2)$に削減できる．

終了後，バックトラック処理を行うことで最適な局所傾き角の系列が求められる．

3. 傾き補正実験

3.1 実験試料

本実験で用いる書棚画像として，書棚をデジタルカメラで撮影して5種類の画像を作成した．ただし，書籍の間以外には背景部は含まないものとし，背景の色はほぼ一定と仮定する．このような画像を対象とする理由は，妥当性評価関数として用いる$s_i(p_i)$が補正線分の上下両端の色を評価するので，同一書籍に対応する場所が補正線分の上下両端に無ければならないためである．また背景の色が一様でない場合も，作成したエッジ画像にノイズが混ざるなど評価関数に悪影響を及ぼす．なお現実にこのような仮定を満たす画像を自動的に作成するのは容易と思われる．

3.2 結果と考察

3.1で作成した画像に対して傾き補正実験を行った．本実験では評価関数$f_i(p_i|p_{i-1})$を構成する$s_i(p_i), c_i(p_i), l_i(p_i)$のバランスを次のように変える．

(a) ： 補正線分の両端の色差$s_i(p_i)$のみ
$(\omega_s = 1/128, \omega_c = 0, \omega_l = 0)$

(b) ： $s_i(p_i)$にエッジの長さ$l_i(p_i)$を追加
$(\omega_s = 1/128, \omega_c = 0, \omega_l = 256)$

(c) ： $s_i(p_i), c_i(p_i), l_i(p_i)$すべて使用
$(\omega_s = 1/128, \omega_c = -64, \omega_l = 256)$

また，全ての場合で$\omega_\rho = 64$とした．

実験結果を**Fig.3**に示す．図の各行が一つの画像の処理結果を表し，図中の(a)，(b)，(c)は上記と呼応する．

Fig.3(c)より，本手法では非一様に傾いた書籍を含む書棚画像に対しても適切な傾き補正を行えることがわかる．また，際だったエッジが検出されない部分についても問題なく傾き補正が行えることがわかる．これは傾き角を推定するときに用いる評価関数において画像の横方向の色変化パターンの関数$s_i(p_i)$を中心に評価しているためである．十分な書籍境界のエッジが得られるところではエッジ情報の関数$l_i(p_i)$が有効に働いていることが(a)と(b)の比較からわかる．また，エッジの長さを与える関数$l_i(p_i)$は書籍境界以外の領域で大きな値を持つため$\rho(p_i|p_{i-1})$と同様に傾き変化を抑制する働きがある．この効果は(5)の(a)の中央部にあるタイトル文字が歪んだ書籍で確認できる．(a)では書籍のタイトル文字が歪んでいるが，(b)では書籍自体の傾きの補正量を同じにしたままタイトル文字の歪みだけ取ることができた．反面副作用として書籍境界以外の全てに対して傾き変化を抑制してしまうので，(4)，(5)のように(a)で補正できていた左側の書籍が(b)ではまた傾いてしまうことがある．表紙色の一様性を示す関数$c_i(p_i)$は，$l_i(p_i)$のこのような働きに対し

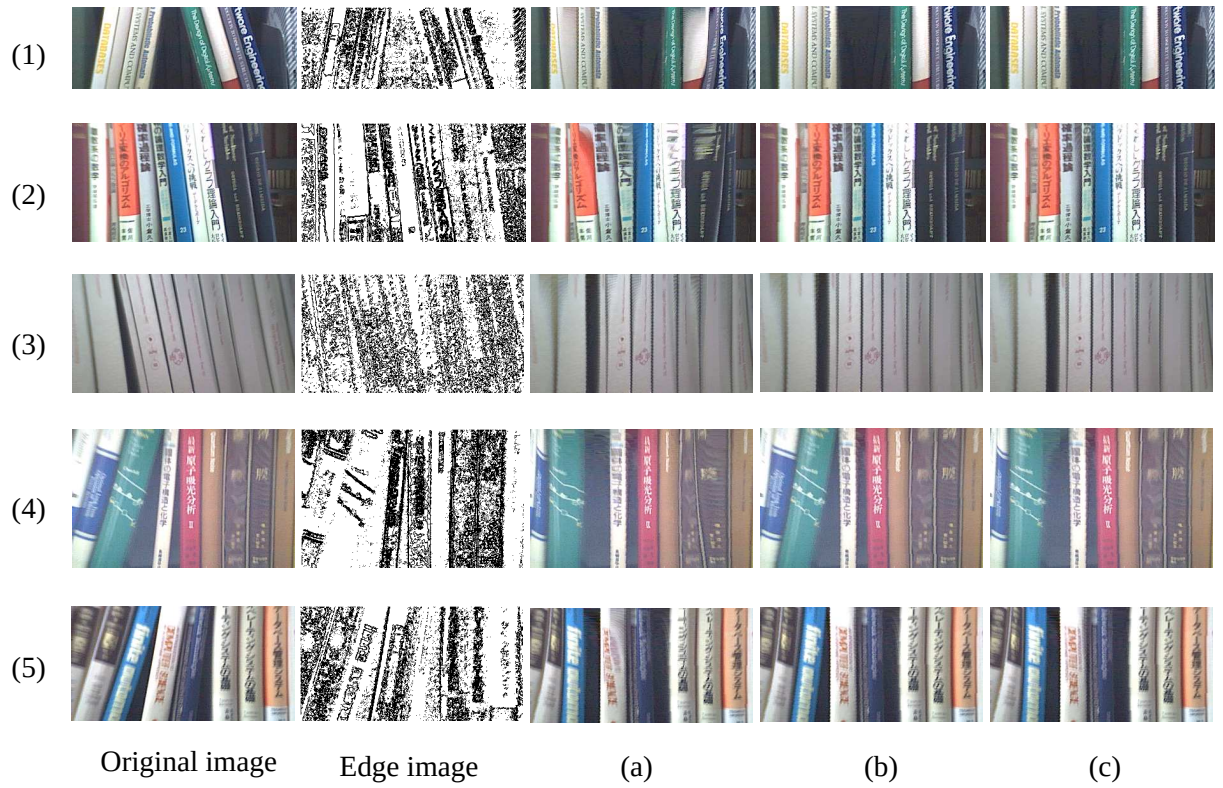

Fig.3　Examples of slant correction. (a)Slant correction result with function $s_i(p_i)$. (b)Slant correction result with function $s_i(p_i)$ and $l_i(p_i)$. (c)Slant correction result with function $s_i(p_i), c_i(p_i)$ and $l_i(p_i)$.

て色が一様な背表紙の背景部として適切な領域と色変化がある文字等を含む部分で差をつける働きがある．この働きにより(4)，(5)の(b)で傾いていた左側の書籍を(c)では傾きが補正できているのがわかる．以上より，関数$s_i(p_i), c_i(p_i), l_i(p_i)$を組み合わせて用いることの有用性がわかる．

Fig.3(4)，(5)を見ると，どの場合でも左端の書籍について傾き補正ができていない．また，**Fig.3**(3)では右側の書籍が補正できていない．この根本的な原因は書棚画像の左右の端部に一部しか見えていない書籍が存在することにある．したがって実用的には書籍とオーバーラップして撮像し，各画像の補正結果について中心部分のみを採用すればよい．

4. お　わ　り　に

本論文では書棚画像中の書籍のタイトル文字認識で問題となる書籍の傾きを補正する手法として動的計画法を用いた手法を提案した．本手法では色情報も用いているのでエッジが検出しにくい書籍についても補正が可能である．実験により本手法の有効性を確認した．

今後は評価関数を改善し，高速化と高機能化をはかる．また，書籍境界を探索する手法についても検討する．

参　考　文　献

1) 岩元隆二, 伊藤稔: 書棚画像からの書籍タイトル文字領域の抽出, 電子情報通信学会情報・システムソサイエティ大会, D-343, September 1996.
2) 長尾智晴, 安居院猛, 中嶋正之: 書棚画像からの書籍の背文字領域抽出に関する研究, テレビジョン学会技術報告, Vol. 15, No. 42, pp. 27-32, July 1991.
3) 広井敬輔, 藤山雅晶, 三浦竜哉, 伊藤稔: 書棚画像からの書籍境界の抽出, 画像の認識・理解シンポジウム講演論文集I, MIRU I-271, July 2000.
4) 澤木美奈子, 村瀬洋, 萩田紀博: 劣化推定に基づいた辞書の自動選択による本棚画像中の文字認識, 映像情報メディア学会誌, Vol. 54, No. 6, pp. 881-886, June 2000.
5) 平英二, 石田敏之, 内田誠一, 迫江博昭: 手書き文字列の非一様な傾き補正法, 電子情報通信学会技術研究報告, PRMU 2000-207, pp. 47-52, March 2001.