

# 大局的最適化に基づく文書画像の湾曲歪み補正法

江崎 弘記<sup>†</sup> 内田 誠一<sup>†</sup> 浅野 晃<sup>††</sup> 迫江 博昭<sup>†</sup>

<sup>†</sup>九州大学大学院システム情報科学府知能システム学専攻

〒 812-8581 福岡県福岡市東区箱崎 6-10-1

<sup>††</sup>広島大学総合科学部数理情報科学講座

〒 739-8521 広島県東広島市鏡山 1-7-1

E-mail: <sup>†</sup>{ezaki,uchida,sakoe}@human.is.kyushu-u.ac.jp, <sup>††</sup>asano@mis.hiroshima-u.ac.jp

あらまし スキャナ等で取得された書籍の文書画像には、湾曲歪みが生じることがある。本論文では、歪みの生じた各行に 3 点を通る 3 次スプライン曲線をフィッティングすることで画像の歪みを表現し、補正を行う。ただし従来法とは異なり、紙面全体での整合性を考慮しながら補正を行う。このため、数式などの短い行や、図形などを含む文書画像であっても安定したフィッティングが可能となる。様々な種類の画像に対して実験を行った結果、本手法の有効性が示された。また、十分に補正できなかった画像に対しては、画像を分割して補正範囲を狭めることで、相対的に歪みモデルの表現能力の向上を図り、対処した。この有効性については実験により確認できた。

キーワード 歪みモデル, 制御点, 3 次スプライン曲線, フィッティング, 評価関数, 動的計画法 (DP)

## Dewarping of document image by global optimization

Hironori EZAKI<sup>†</sup>, Seiichi UCHIDA<sup>†</sup>, Akira ASANO<sup>††</sup>, and Hiroaki SAKOE<sup>†</sup>

<sup>†</sup> Graduate School of Information Science and Electrical Engineering, Kyushu University

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

<sup>††</sup> Faculty of Integrated Arts and Sciences, Hiroshima University

1-7-1 Kagamiyama, Higashi-Hiroshima-shi, Hiroshima, 739-8521 Japan

E-mail: <sup>†</sup>{ezaki,uchida,sakoe}@human.is.kyushu-u.ac.jp, <sup>††</sup>asano@mis.hiroshima-u.ac.jp

**Abstract** This paper proposes a dewarping technique for scanned document images of bounded volumes. This technique estimates the warp of a document image by fitting cubic splines to each text line and space between text lines, and then creates a dewarped image by straightening each fitted spline. Differing from conventional techniques, the proposed technique fits cubic splines to the warped document image while considering their global optimality. Thus, it is possible to dewarp document images including short text lines, figures, and so on without misfitted splines. The experimental results on several document images including short text lines, and figures showed the effectiveness of the proposed technique.

**Key words** warping model, control node, cubic spline, fitting, criterion function, dynamic programming(DP)

### 1. はじめに

書籍の各ページをスキャナやカメラなどを用いて電子化する場合、書籍の厚み等により綴り付付近で紙面とスキャナ面との間に隙間が生じることがある。その結果、取得文書画像には歪みが生じる。その画像を直接人間が読む場合、この歪みはさほど問題とはならないが、OCR で処理する場合には大きな問題となることが多い。特に横書き文書においては、ページの各行が曲線状に変形すると、OCR は傾き補正処理や行切り出し処理に失敗し、その認識性能は著しく低下する。これは、一般的に

行が直線的であることが OCR の前提となっているためである。

この問題を解決するために、歪みの生じた文書画像に対して何らかの画像処理を施すことで歪みを補正する手法が検討されている。これら従来法は、次の 3 種に大別される。

(i) ステレオ計測や 3 次元計測に基づく手法 [1] [2] [3]

(ii) Shape from shading の原理に基づく手法 [4] [5] [6]

(iii) モデルフィッティングに基づく手法 [7] [8] [9]

上記 (i) の一手法として、川原子ら [1] の手法がある。この手法は、デジタルカメラによるステレオ計測により紙面の 3 次元形状を測定し、湾曲部の補正を行っている。上記 (ii) の代表

的な手法として、和田ら [4] [5] [6] の手法がある。この手法は、スキャナにより撮影された書籍表面画像の陰影情報から、書籍の3次元形状を推定し、画像中の陰影や歪みを補正するというものである。具体的には、光の強度減衰や光源の移動、書籍表面の反射特性などを考慮した光学モデルを構成し、非線形二乗法によって書籍表面の形状を推定している。しかしこれら (i), (ii) の手法は、特別な装置や書籍の配置に制約を必要とするため、容易に補正処理を行えないという問題がある。さらに、和田らの手法では照明条件や書籍の反射特性などの事前知識が必要である。

上記 (iii) の手法は、曲線を文字行にフィッティングすることで各文字行の歪みを表現する手法である。上記 (i), (ii) とは違い、このタイプの手法は書籍の配置の制約や事前知識を必要としないため、最も広く容易に適用できるという利点がある。Zhang ら [7] は、歪みを含む画像を鮮明な領域と陰領域とに分け、鮮明な領域の各行を直線で、陰領域の各行を2次関数でモデル化して補正を行っている。Wu ら [8] は、歪みを含む画像上で文字行をトラッキングすることで、行の歪みを表現するメッシュを作成して補正を行っている。

これらのモデルフィッティングに基づく従来法は、局所的な評価に基づいてモデル化が行われている。そのため、数式やタイトルのような短い文字行や、図形などを含む文書画像では誤ったフィッティングが行われる可能性がある。実際に Wu らは後処理として、ある基準に基づき適切にトラッキングしていない行を除外している。しかし、除外される行が多いと全体の歪みを適切に表現できないという問題が残る。

本研究では、各行の湾曲および傾きに応じて曲線をフィッティングすることで各行の歪みを推定し、その結果を用いて各行の湾曲歪みおよび傾きを補正する手法を提案する。ただし、従来法とは異なり各行独立ではなく紙面全体での整合性を考慮する。すなわち、局所的な評価に基づいてフィッティングを行うのではなく、大局的な評価に基づいてフィッティングを行う。これにより、例えば曲線が安定してフィッティングできないような短い行や図形などを含む文書画像であっても、全体としては安定した歪み補正が可能となる。また、より安定したフィッティングを行うために文字行間にある空白行にも曲線フィッティングを行う。本手法では行の湾曲歪みが3次スプライン曲線でモデル化できることが唯一の仮定であり、特殊な装置や書籍の配置の制約、事前知識を必要としない。

以下、第2章では歪みモデルを定義し、湾曲歪み補正問題を最適化問題として定義する。次に、この最適化問題を解くための動的計画法 (DP) に基づくアルゴリズムについて述べる。第3章では歪みを含む文書画像の補正実験を行い、その結果を考察する。第4章では今後の課題を含め結論を述べる。

## 2. 大局的最適化に基づく湾曲歪み補正法

### 2.1 歪みモデル

原画像のサイズを横  $W$  画素、縦  $H$  画素とする。本手法では、各行の湾曲に応じて  $H$  本の曲線からなる歪みモデルをフィッティングすることで各行の歪みを推定し、その結果を用いて各

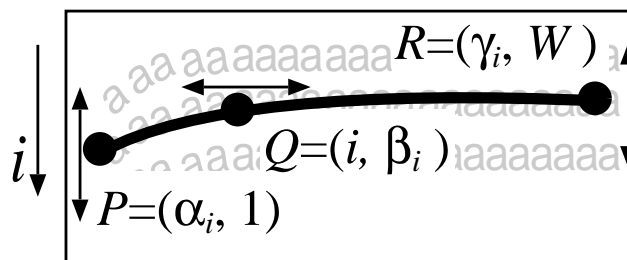


図1 3つの制御点を通る第  $i$  行目のスプライン曲線

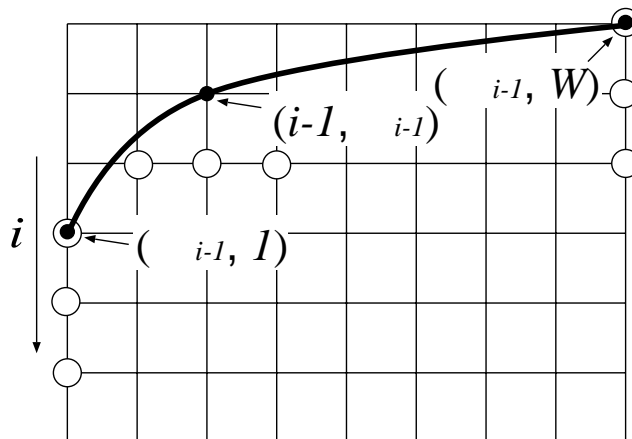


図2  $i-1$  行目の制御点に対する  $i$  行目の制御点の候補  
(  $\bullet$  :  $i-1$  行目の制御点,  $\circ$  :  $i$  行目の制御点の候補)

```

/* DP recursion */
1: for  $i := 1$  to  $H$  do begin
2:   for all possible  $\theta_i = (\alpha_i, \beta_i, \gamma_i)$  do begin
3:      $g(\theta_i|i) := \{f(\theta_i|i) + \omega\lambda(\theta_i, \theta_{i-1})\}$ 
           +  $\min_{\theta' \in \text{Prev}(\theta_i)} g(\theta'|i-1)$ 
4:      $b(\theta_i|i) := \arg \min_{\theta' \in \text{Prev}(\theta_i)} g(\theta'|i-1)$ 
5:   end
6: end
/* Back-tracking */
7:  $\theta_H^o := \arg \min_{\theta_H} g(\theta_H|H)$ 
8: for  $i := H-1$  down to 1 do
9:    $\theta_i^o := b(\theta_{i+1}^o|i+1)$ 

```

図3 DP アルゴリズム ( $\{\theta_i^o \mid i = 1, \dots, H\}$  は最適解を導出するパラメータの組)

行の湾曲歪みおよび傾きを補正する。この歪みモデルは、 $H$  本の3次スプライン曲線 (以下、スプライン曲線) の集合で定義され、原画像を2値化したものにフィッティングされる。

図1に示すように第  $i$  行目のスプライン曲線は両端および中央付近の3点 (以下、制御点と呼ぶ) を通るスプライン曲線で表現される。これら3つの制御点は左端制御点  $P = (\alpha_i, 1)$  が上下動, 中央制御点  $Q = (i, \beta_i)$  が左右動, 右端制御点  $R = (\gamma_i, W)$  が上下動することを許される。これら3つのパラメータの組  $\theta_i = (\alpha_i, \beta_i, \gamma_i)$  により第  $i$  行目のスプライン曲線は一意に決まる。

本手法では、隣接行のスプライン曲線が極端に変化せず、か

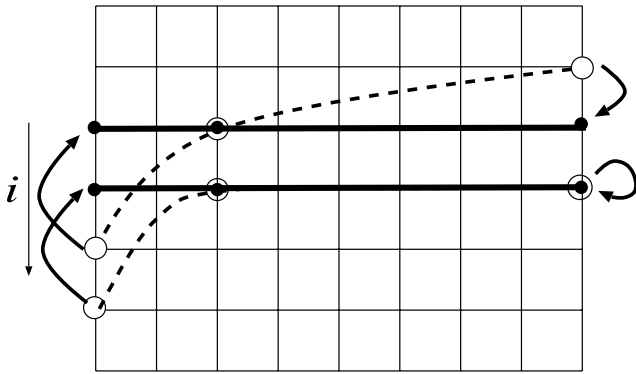


図 4 補正画像の生成 (点線：フィッティングされたスプライン曲線，実線：補正後のスプライン曲線)

つ交差しないようにするために，以下の制約条件 (図 2) を設ける．

$$\begin{cases} 0 \leq \alpha_i - \alpha_{i-1} \leq 2 \\ -1 \leq \beta_i - \beta_{i-1} \leq 1 \\ 0 \leq \gamma_i - \gamma_{i-1} \leq 2 \end{cases} \quad (1)$$

この制約条件により，各行独立にフィッティングを行うのではなく，歪みモデル全体として，すなわち， $H$  本のスプライン曲線全体として最適なフィッティングを行うことになる．このため，数式やタイトルなどの短い文字行や，図形などを含むような文書画像であっても安定したフィッティングが可能となる．なお本手法では， $H$  本のスプライン曲線をフィッティングするので，非常に密にフィッティングを行うことになる．従って，空白行にもフィッティングを行うことになるが，以下に述べるように本手法ではこのことが非常に重要な役割を担う．

## 2.2 歪みモデルの最適フィッティング問題の定式化

本手法では，各行にスプライン曲線をフィッティングする際，空白行に対するフィッティングを重要視する．これは，空白行が文字行と比べて安定した歪みを表現するためである．すなわち，文字行には単語間の空白や，数式やタイトルなどの短い行が存在することがあるが，空白行は連続的で長いため，空白行にフィッティングを行った方が安定したフィッティングを行うことができるためである．

歪みモデルのフィッティングを評価する問題は，制約条件 (1) の下で，次の評価関数  $J$  の最小化問題として定式化できる．

$$J = \sum_{i=1}^H \{f(\theta_i|i) + \omega\lambda(\theta_i, \theta_{i-1})\} \quad (2)$$

ここで， $\omega$  は非負の重み定数である．

関数  $f$  は歪みモデルと実際の文書画像の歪みの違いを表現する関数であり，本手法では，以下のような非常に簡単な式で表現される．

$$f(\theta_i|i) = \begin{cases} 0 & (\text{if total}(i) = 0) \\ 1 & (\text{otherwise}) \end{cases} \quad (3)$$

ここで  $\text{total}(i)$  は第  $i$  行目のスプライン曲線上の黒画素数を表す．すなわち，スプライン曲線が空白行にあるならば関数  $f$  は

0，逆にスプライン曲線が文字行を通れば関数  $f$  は 1 を返す．

この関数  $f$  による評価だけでは，次の二つの問題が生じる．第一の問題は，曖昧性である．例えば関数  $f$  単独では，スプライン曲線が不等間隔にフィッティングされた場合も等間隔にフィッティングされた場合も，評価関数  $J$  の値が同一になってしまう．一般的な文書画像の歪みを想定すれば，スプライン曲線が等間隔にフィッティングされる方が適したフィッティングといえるので，こうした曖昧性を解消し，なるべくスプライン曲線が等間隔にフィッティングされるような配慮が必要となる．第二の問題は，関数  $f$  単独では，文字行を避けるようにスプライン曲線がフィッティングされるということである．これにより，文字行の歪みが適切に表現できないということが問題となる．

以上の問題を解決すべく (2) では正則化関数  $\lambda$  を導入している．関数  $\lambda$  は具体的には以下の式で表現される．

$$\lambda(\theta_i, \theta_{i-1}) = |\alpha_i - \alpha_{i-1} - 1| + |\beta_i - \beta_{i-1}| + |\gamma_i - \gamma_{i-1} - 1| \quad (4)$$

関数  $\lambda$  により，スプライン曲線が不等間隔にフィッティングされる場合は，評価関数  $J$  の値が大きくなる．このため，等間隔にスプライン曲線がフィッティングされ易くなる．

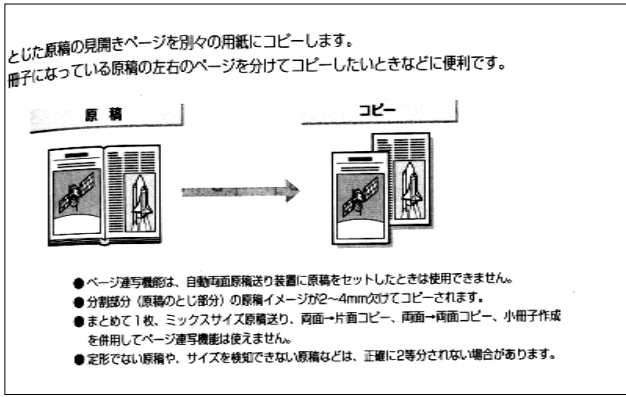
以上の評価は，文書画像に陰影を含まないことを仮定している．陰影への対処については将来的な課題となるが，例えば次のような解決法が考えられる．第一の解決法としては，前処理としてある領域ごとに 2 値化の閾値を決定する局所 2 値化法の利用が挙げられる．また他の解決法として，評価関数の改良が挙げられる．例えば，綴り付近に生じる縦方向に連続的な陰影については，評価しようとしている黒画素からある閾値よりも長く上下に黒画素が延びている場合，その黒画素は評価の対象に入れないようにするといった対処が考えられる．

## 2.3 DP アルゴリズム

前節で定式化した最小化問題において，関数  $f$  は各  $i$  で独立に定義され，また制約条件 (1) と正則化関数  $\lambda$  は隣接行間，すなわち  $i, i-1$  行間において定義されるので，この最小化問題は， $i$  を段とする多段決定過程により表現される．従って，この問題を動的計画法 (DP) で解くことができる．図 3 に，本手法における DP アルゴリズムを示す． $\text{Prev}(\theta_i)$  は  $\theta_i = (\alpha_i, \beta_i, \gamma_i)$  において制約条件 (1) を満たす  $\theta_{i-1} = (\alpha_{i-1}, \beta_{i-1}, \gamma_{i-1})$  の集合である．図 3 中の Step 3 を  $i=1$  から  $i=H$  まで繰り返すことによって，評価関数  $J$  の最小値が  $\min_{\theta_H} g(\theta_H | H)$  として得られる．また，図 3 中の Step 4 でバックポインタを保持することで最適フィッティングとなる歪みモデルのパラメータの組  $\{\theta_i^o = (\alpha_i^o, \beta_i^o, \gamma_i^o) | i=1, \dots, H\}$  を得ることができる．

## 2.4 補正画像の生成

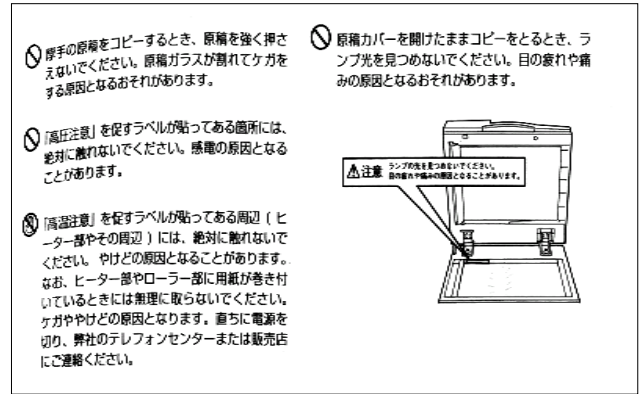
2.2 節で定式化される最小化問題を解いた結果，フィッティングされるスプライン曲線を，図 4 のように各行直線にすることで補正画像を生成できる．すなわち，第  $i$  行のスプライン曲線上の画素を， $y$  座標が  $i$  となる位置に移動することで補正画像を生成できる．



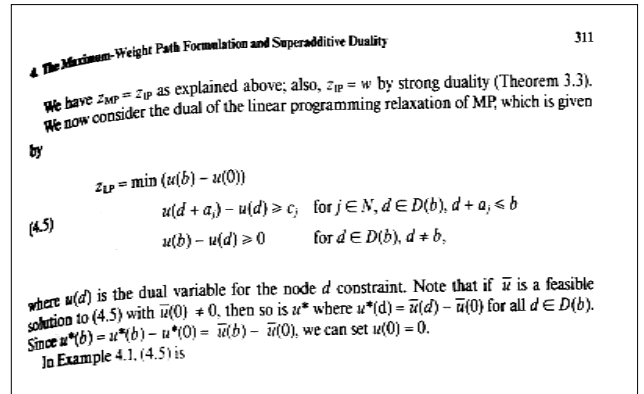
(a)

units in the network. We consider this example to be a simplified case. It is possible to imagine hidden units which respond to patterns among the input units. In the general case, we, of course, recognize that hidden units would be required to give different coalitions enough coherence. As we have pointed out elsewhere in the book (cf. Chapter 2 and Chapter 8), multilayer systems containing hidden units are sometimes required to carry out certain computations. In the present instance,

(c)



(b)



(d)

図5 原画像

### 3. 文書画像の補正実験

この章では、実際に行った文書画像の補正実験および結果について述べる。なお、定数  $w$  は幾つかの事前実験により 1 に定めた。

図5, 図6, 図7はそれぞれスキャナを用いて取得した文書画像(以下、原画像と呼ぶ)、最適フィッティングされた歪みモデル、補正画像である。図5の原画像にはそれぞれ次のような特徴がある。(a)は図形を含む文書画像、(b)は段組みおよび図形を含む文書画像、(c)はどの文字行も端から端までである文書画像、(d)はタイトルや数式など短い文字行を含む文書画像である。

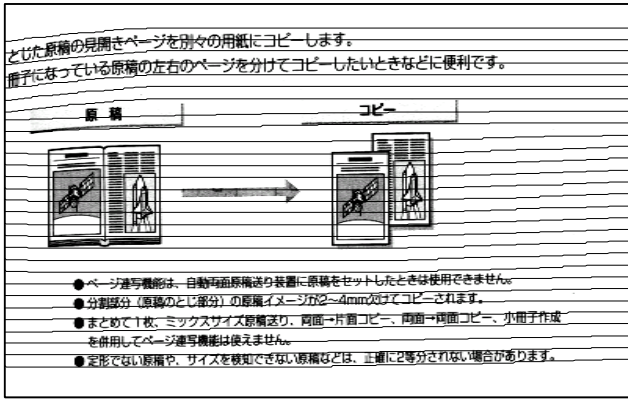
#### 3.1 実験結果

実際に図5(a)と図7(a), 図5(b)と図7(b)をそれぞれ比べると湾曲歪みが若干残っているが、全体としては適切に補正されていることが分かる。図5(a)の場合、挿入されている図形の上下にある文字行間の空白行には、関数  $f$  の効果によりその湾曲や傾きに応じてスプライン曲線が適切にフィッティングされるのに対し、図形の存在する行では、関数  $f$  の値が同一になるスプライン曲線が多数存在するため、関数  $f$  単独では適切にフィッティングされるとは限らない。しかし正則化関数  $\lambda$  の効果により、結果的には図形の存在する行についても適切なフィッティングが行われたものと考えられる。すなわち、図形の存在する行では関数  $\lambda$  の効果により、図形の存在しない行にお

けるフィッティングが伝播し、等間隔でかつ実際に生じている歪みに応じたフィッティングが行われたと考えられる。局所的な基準に基づいてフィッティングを行う従来法(例えば[8])で補正を行う場合、図形や段組みが挿入されているためトラッキングが適切に行われなことが予想されるが、本手法では大局的な評価に基づいてフィッティングを行うため、適切なフィッティングを行うことができたと考えられる。また従来法とは異なり、空白行のフィッティングを重要視していることも安定したフィッティングを行うことができた要因の一つといえる。

図5(c)と図7(c), 図5(d)と図7(d)をそれぞれ比較すると、補正はされているものの依然湾曲歪みが残っており、十分に補正ができたとはいえない。このような結果が得られた最も大きな原因としては、歪みモデルの表現能力の限界が挙げられる。実際に、図5(c)の一番上の文字行に最も適切にフィットするように左端制御点および右端制御点を固定し、中央制御点のみを手動で移動させ、求めたスプライン曲線の例を図8に示す。この結果から、制御点をどのように動かしたとしても、3点を通る3次スプライン曲線では実際の文書画像の歪みを完全には表現できないことがわかる。

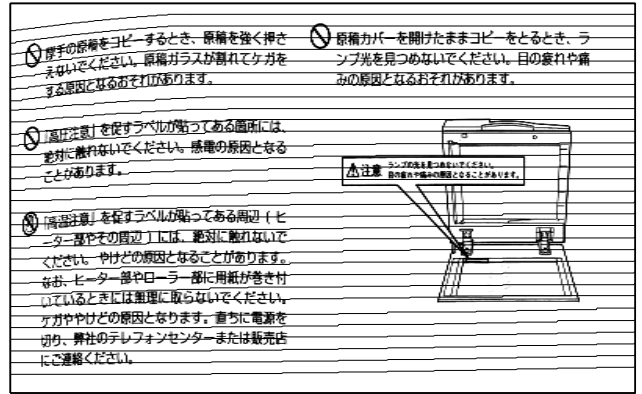
この問題を解決する方法としては次の二つが考えられる。第一の方法は、湾曲部分に制御点を増やす、あるいは画像を分割して補正する範囲を狭くすることによって、3次スプライン曲線の表現能力を向上することである。次節では、その検討結果について述べる。第二の方法は、文献[9][10]にあるような3次



(a)

units in the network. We consider this example to be a simplified case. It is possible to imagine hidden units which respond to patterns among the input units. In the general case, we, of course, recognize that hidden units would be required to give different coalitions enough coherence. As we have pointed out elsewhere in the book (cf. Chapter 2 and Chapter 8), multilayer systems containing hidden units are sometimes required to carry out certain computations. In the present instance,

(c)



(b)

4 The Maximum-Weight Path Formulation and Superadditive Duality 311

We have  $z_{MP} = z_{LP}$  as explained above; also,  $z_{LP} = w$  by strong duality (Theorem 3.3). We now consider the dual of the linear programming relaxation of MP, which is given by

$$z_{LP} = \min (u(b) - u(0))$$

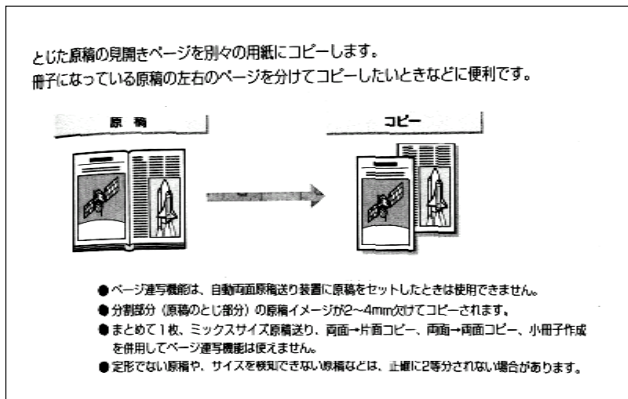
(4.5)  $u(d + a_j) - u(d) \geq c_j$  for  $j \in N, d \in D(b), d + a_j \leq b$

$$u(b) - u(d) \geq 0$$
 for  $d \in D(b), d \neq b.$ 

where  $u(d)$  is the dual variable for the node  $d$  constraint. Note that if  $\bar{u}$  is a feasible solution to (4.5) with  $\bar{u}(0) \neq 0$ , then so is  $u^*$  where  $u^*(d) = \bar{u}(d) - \bar{u}(0)$  for all  $d \in D(b)$ . Since  $u^*(b) = u^*(b) - u^*(0) = \bar{u}(b) - \bar{u}(0)$ , we can set  $u(0) = 0$ . In Example 4.1, (4.5) is

(d)

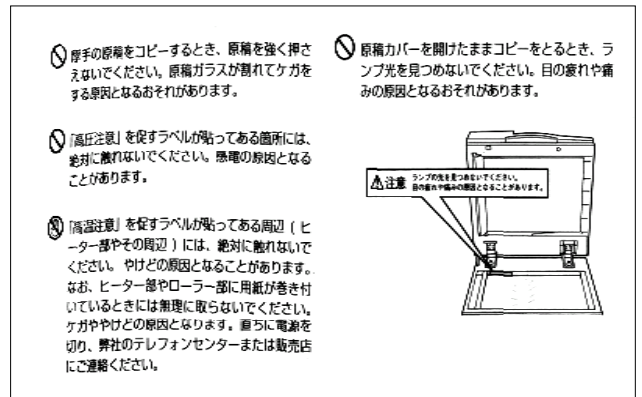
図6 最適フィッティングされた歪みモデル (見易さのために15行毎に表示している。)



(a)

units in the network. We consider this example to be a simplified case. It is possible to imagine hidden units which respond to patterns among the input units. In the general case, we, of course, recognize that hidden units would be required to give different coalitions enough coherence. As we have pointed out elsewhere in the book (cf. Chapter 2 and Chapter 8), multilayer systems containing hidden units are sometimes required to carry out certain computations. In the present instance,

(c)



(b)

4 The Maximum-Weight Path Formulation and Superadditive Duality 311

We have  $z_{MP} = z_{LP}$  as explained above; also,  $z_{LP} = w$  by strong duality (Theorem 3.3). We now consider the dual of the linear programming relaxation of MP, which is given by

$$z_{LP} = \min (u(b) - u(0))$$

(4.5)  $u(d + a_j) - u(d) \geq c_j$  for  $j \in N, d \in D(b), d + a_j \leq b$

$$u(b) - u(d) \geq 0$$
 for  $d \in D(b), d \neq b.$ 

where  $u(d)$  is the dual variable for the node  $d$  constraint. Note that if  $\bar{u}$  is a feasible solution to (4.5) with  $\bar{u}(0) \neq 0$ , then so is  $u^*$  where  $u^*(d) = \bar{u}(d) - \bar{u}(0)$  for all  $d \in D(b)$ . Since  $u^*(b) = u^*(b) - u^*(0) = \bar{u}(b) - \bar{u}(0)$ , we can set  $u(0) = 0$ . In Example 4.1, (4.5) is

(d)

図7 補正画像

units in the network. We consider this example to be a simplified case. It is possible to imagine hidden units which respond to patterns among the input units. In the general case, we, of course, recognize that hidden units would be required to give different coalitions enough coherence. As we have pointed out elsewhere in the book (cf. Chapter 2 and Chapter 8), multilayer systems containing hidden units are sometimes required to carry out certain computations. In the present instance,

units in the network. We consider this example to be a simplified case. It is possible to imagine hidden units which respond to patterns among the input units. In the general case, we, of course, recognize that hidden units would be required to give different coalitions enough coherence. As we have pointed out elsewhere in the book (cf. Chapter 2 and Chapter 8), multilayer systems containing hidden units are sometimes required to carry out certain computations. In the present instance,

units in the network. We consider this example to be a simplified case. It is possible to imagine hidden units which respond to patterns among the input units. In the general case, we, of course, recognize that hidden units would be required to give different coalitions enough coherence. As we have pointed out elsewhere in the book (cf. Chapter 2 and Chapter 8), multilayer systems containing hidden units are sometimes required to carry out certain computations. In the present instance,

units in the network. We consider this example to be a simplified case. It is possible to imagine hidden units which respond to patterns among the input units. In the general case, we, of course, recognize that hidden units would be required to give different coalitions enough coherence. As we have pointed out elsewhere in the book (cf. Chapter 2 and Chapter 8), multilayer systems containing hidden units are sometimes required to carry out certain computations. In the present instance,

図 8 中央制御点を手で移動させた例

スプライン曲線ではない、より精度の高い歪みモデルを構築することである。いずれも、実際の文字行の歪みをより適切に表現できるような歪みモデルを構築することで、この問題を解決しようとする方法である。

### 3.2 歪みモデルの表現能力を向上させて行った補正実験およびその結果

ここでは3次スプライン曲線の表現能力を向上させることを目的として、画像を分割し、補正する範囲を狭めた場合の実験結果について述べる。具体的には、先の手法で十分に補正できなかった文書画像を、図9のように2本の垂直線によって幅が1:1:2の3つの部分画像に分割する。ただし、それぞれ左右に数十画素の重複部を設ける。そして、それぞれの画像において、3点を通るスプライン曲線をフィッティングする。重複部を設けるのは、補正画像を生成する際に、画像の接合部分を滑らかにするためである。この場合、3つの部分画像でフィッティングされたそれぞれのスプライン曲線から各画素における行の歪みを求める。ただし画像が重なり合っている箇所は、各画素におけるスプライン曲線の歪みの平均をその画素における歪みとする。

先の手法で十分に補正できなかった文書画像(図5(c), (d))に対し、以上の手法を適用した結果が図10である。図10の左側の画像は、分割後の3つの画像から最適にフィッティングされた歪みモデルを、15行おきに表示したものである。図10の右側の画像は、それぞれの補正後の画像である。図7(c), (d)と比べても湾曲歪みがほとんどなくなっていることがわかる。この結果から、画像を分割して補正する範囲を狭くすることによる、3次スプライン曲線の表現能力を向上させることの有効性が示された。

## 4. むすび

本論文では、大局的最適化に基づく文書画像の湾曲歪み補正法について提案した。本手法は従来法とは異なり、各行独立ではなく隣接行間に制約を設け、紙面全体の整合性を考慮しながら

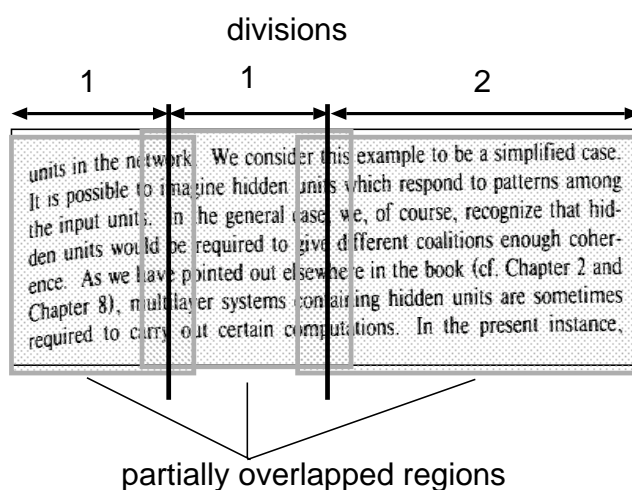


図 9 画像を3分割する様子

ら歪み補正を行う。このため、数式やタイトルなどの短い行や、図形などを含む文書画像でも適切な補正を行うことが可能である。本手法では、各行の湾曲歪みが3次スプライン曲線でモデル化できることが唯一の仮定であり、特殊な装置や書籍の配置の制約、事前知識を必要としない。

実際に歪みの生じた文書画像を用いた実験により、本手法の有効性が示された。数式やタイトルなどの短い行や、図形を含む画像を補正する場合、局所的な評価に基づいてフィッティングを行う従来法では、適切に補正できないと予測される。一方、本手法では、部分的にこうした不規則性が存在しても (i) 制約条件、(ii) 正則化、ならびに (iii) 文字行よりも安定した湾曲抽出が可能空白行に注目した評価、を用いながら、大局的最適化を行うために安定したフィッティングが可能となる。適切に補正できなかった画像もあったが、その原因としては歪みモデルによる表現能力の限界が挙げられる。この問題の解決法としては歪みモデルの表現能力の向上が挙げられる。具体的には、制御点の数を増やす、あるいは画像を分割して補正する範囲を

units in the network. We consider this example to be a simplified case. It is possible to imagine hidden units which respond to patterns among the input units. In the general case, we, of course, recognize that hidden units would be required to give different coalitions enough coherence. As we have pointed out elsewhere in the book (cf. Chapter 2 and Chapter 8), multilayer systems containing hidden units are sometimes required to carry out certain computations. In the present instance,

units in the network. We consider this example to be a simplified case. It is possible to imagine hidden units which respond to patterns among the input units. In the general case, we, of course, recognize that hidden units would be required to give different coalitions enough coherence. As we have pointed out elsewhere in the book (cf. Chapter 2 and Chapter 8), multilayer systems containing hidden units are sometimes required to carry out certain computations. In the present instance,

4 The Maximum-Weight Path Formulation and Superadditive Duality 311

We have  $z_{MP} = z_{IP}$  as explained above; also,  $z_{IP} = w$  by strong duality (Theorem 3.3). We now consider the dual of the linear programming relaxation of MP, which is given by

$$z_{LP} = \min \{u(b) - u(0)\}$$

$$(4.5) \quad \begin{aligned} u(d+a_j) - u(d) &\geq c_j && \text{for } j \in N, d \in D(b), d+a_j \leq b \\ u(b) - u(d) &\geq 0 && \text{for } d \in D(b), d \neq b, \end{aligned}$$

where  $u(d)$  is the dual variable for the node  $d$  constraint. Note that if  $\bar{u}$  is a feasible solution to (4.5) with  $\bar{u}(0) \neq 0$ , then so is  $u^*$  where  $u^*(d) = \bar{u}(d) - \bar{u}(0)$  for all  $d \in D(b)$ . Since  $u^*(b) = u^*(b) - u^*(0) = \bar{u}(b) - \bar{u}(0)$ , we can set  $u(0) = 0$ . In Example 4.1, (4.5) is

4 The Maximum-Weight Path Formulation and Superadditive Duality 311

We have  $z_{MP} = z_{IP}$  as explained above; also,  $z_{IP} = w$  by strong duality (Theorem 3.3). We now consider the dual of the linear programming relaxation of MP, which is given by

$$z_{LP} = \min \{u(b) - u(0)\}$$

$$(4.5) \quad \begin{aligned} u(d+a_j) - u(d) &\geq c_j && \text{for } j \in N, d \in D(b), d+a_j \leq b \\ u(b) - u(d) &\geq 0 && \text{for } d \in D(b), d \neq b, \end{aligned}$$

where  $u(d)$  is the dual variable for the node  $d$  constraint. Note that if  $\bar{u}$  is a feasible solution to (4.5) with  $\bar{u}(0) \neq 0$ , then so is  $u^*$  where  $u^*(d) = \bar{u}(d) - \bar{u}(0)$  for all  $d \in D(b)$ . Since  $u^*(b) = u^*(b) - u^*(0) = \bar{u}(b) - \bar{u}(0)$ , we can set  $u(0) = 0$ . In Example 4.1, (4.5) is

図 10 歪みモデルの表現能力の向上結果 (左：最適フィッティングされた歪みモデル，右：補正画像)

狭めることによる，3次スプライン曲線の表現能力の向上が挙げられる．実際に実験から，画像を分割し補正する範囲を狭めることによって，3次スプライン曲線の表現能力を向上させることの有効性が示された．

今後の課題としては，補正処理の高速化ならびにより精度の高い歪みモデルの構築が挙げられる．また，陰影を含む画像に対しても局所2値化法や評価関数の改良を行うなどの対処をし，実験を行っていくことも今後の課題としたい．

謝辞 本研究の一部は，(財)大川情報通信基金の研究助成金ならびに科研費(若手(B)，No. 17700198)によった．

### 文 献

- [1] 川原 淳，山下 淳，金子 透，三浦 憲二郎，“ステレオビジョンシステムを用いた紙面形状復元・補正”，信学技報，PRMU2002-226，pp. 19-24，2003-03.
- [2] M. S. Brown, W. B. Seales, “Image restoration of arbitrarily warped documents,” IEEE Trans. PAMI, vol. 26, no. 10, pp. 1295-1306, 2004.
- [3] A. Doncescu, A. Bouju, and V. Quillet, “Former books digital processing: image warping,” Proc. Workshop on Doc. Image Anal., pp. 5-9, 1997.
- [4] 和田 俊和，浮田 浩行，松山 隆司，“イメージスキャナを用いた書籍表面の3次元形状復元(I)-近接照明下でのShape from Shading-”，信学論(D-II)，vol. J77-D-II，no. 6，pp. 1059-1067，1994-06.
- [5] 和田 俊和，浮田 浩行，松山 隆司，“イメージスキャナを用いた書籍表面の3次元形状復元(II)-相互反射を考慮した近接光源下のShape from Shading-”，信学論(D-II)，vol. J78-D-II，no. 2，pp. 331-320，1995-02.
- [6] 浮田 浩行，小西 克信，和田 俊和，松山 隆司，“固有空間法を用いた陰影情報からの書籍表面の3次元形状復元”，信学論(D-II)，vol. J83-D-II，no. 12，pp. 2610-2621，2000-12.
- [7] Z. Zhang, C. L. Tan, “Correcting document image warping based on regression of curved text lines,” Proc. Int. Conf.

- Doc. Anal. Recog., pp. 589-593, 2003.
- [8] C. Wu, G. Agam, “Document image de-warping for text/graphics recognition,” SSPR&SPR 2002, LNCS 2396, pp. 348-357, 2002.
  - [9] H. Cao, X. Ding, and C. Liu, “Rectifying the bound document image captured by the camera: a model based approach,” Proc. ICDAR, pp. 71-75, 2003.
  - [10] T. Kanungo, R. M. Haralick, and I. Phillips, “Global and local document degradation models,” Proc. ICDAR, pp. 730-734, 1993.