

## ステレオビジョンシステムによる表面湾曲書籍の文字認識

Character Recognition for Books with Curved Surface  
using a Stereo Vision System川原子 淳<sup>†</sup>, 山下 淳<sup>††</sup>, 正会員 金子 透<sup>††</sup>, 三浦 憲二郎<sup>††</sup>Atsushi Kawarago<sup>†</sup>, Atsushi Yamashita<sup>††</sup>, Toru Kaneko<sup>††</sup> and Kenjiro T. Miura<sup>††</sup>

あらまし 既存書籍のデジタル化は重要な課題である。具体策としてスキャナ入力法があるが、分厚い書籍の場合には十分な画質が得られない。そこで、デジタルカメラにより紙面を撮影し、画像処理により紙面の湾曲を補正した後に文字認識を行う方法を提案する。

キーワード：画像処理，文書解析，書籍のデジタル化，表面湾曲書籍，ステレオビジョン，文字認識

## 1. 序 論

コンピュータを用いた文書画像処理の目的の1つに、文書情報のデータベース化がある。既存の書籍をデジタルファイルに変換する方法としては、次の2つの方法がある。紙面をそのまま画像ファイルに変換する方法と、スキャナ等を用いて入力し、文字認識することによりテキストファイルに変換する方法である。ファイル容量の小ささや編集のしやすさ、出力の美しさから後者の方がデータベース化に適している。ほとんどの場合、入力にはスキャナが利用されるが、分厚い書籍を読み込むときには、綴り部分が湾曲してスキャナ面から離れてしまい、影が生じるために文字認識ができないことが多い。

そこで、画像中の陰影情報から紙面形状を推定し補正した上で、文字認識を行う手法も提案されている<sup>1)~3)</sup>。しかし、この手法では、明るさや紙質などに大きく影響ため多くの基礎データが必要となる。また、ページをめくる毎に書籍を置き直すといった手間が必要となる。

そこで本手法では、書籍の種類や紙面の明るさによらず文字認識を可能にすることを目指す。また、書籍を上向きに置くことで、動作はページをめくるのみとする。具体的には、高精細デジタルカメラによるステレオ計測により、紙面の3次元形状を測定する。そして、画像変換を行い湾

曲を補正し、明るさに左右されないようにグレースケール画像を用いて文字認識を行う方法について検討する。

## 2. 処 理 手 順

本手法は、次に示す5つの手順で処理を行う。

- (1) ステレオ計測
- (2) 画像変換
- (3) 文字テンプレート作成
- (4) 1文字領域抽出
- (5) 文字認識

今回の文字認識では、入力画像中のアルファベットの小文字のみを処理対象とする。また、書籍はカメラの基線に対して平行に置くものとする。

## 2.1 ステレオ計測

本手法では、平行ステレオにより文書画像を取得する。3次元空間中の任意の点の座標値は、左右の画像上での座標値から、ステレオ計測により3次元座標を算出することが可能である。

紙面形状は、ステレオ計測により紙面上のすべての点について、3次元座標値を算出することで復元することができる。そのためには、取得した左右の画像上で対応点を検出する必要がある。しかし、すべての点について対応点を検出するには処理時間がかかる。

そこで、画像上の代表的な特徴点についてのみ対応点検出を行い、それらの点のステレオ計測を行う。ここでは、画像中の文字のみを処理対象としており、紙面は局所的に急激な変化がないことから、文字領域の中心線上の各点を特徴点とし対応点検出を行う。

始めに、取得画像から文字領域のみを抽出する。まず、取得画像を2値化・膨張処理により小領域に分割する。

その後、小領域毎に輪郭線抽出を行い、領域の大きさお

2001年10月19日，マルチメディア情報処理研究会で発表  
2001年12月14日受付，2002年1月21日再受付，2002年2月14日採録  
† 静岡大学大学院 理工学研究所

(〒432-8561 浜松市城北3-5-1, TEL 053-478-1070)

†† 静岡大学 工学部

(〒432-8561 浜松市城北3-5-1, TEL 053-478-1070)

† Graduate School of Science and Engineering, Shizuoka University  
(3-5-1, Johoku, Hamamatsu-shi, Shizuoka 432-8561, Japan)

†† Department of Machinery Engineering, Shizuoka University  
(3-5-1, Johoku, Hamamatsu-shi, Shizuoka 432-8561, Japan)



図 1 文字領域  
Character region.

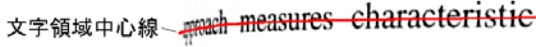


図 2 文字領域中心線  
Centerline of character region.

よび形状から文字領域を抽出する(図1)。

以上より抽出された文字領域は、次の性質を有する。

- (1) 文字のみが含まれる。
- (2) 2行以上にはならないが、1行すべてとは限らない。
- (3) スペースがあっても区切りになるとは限らない。

次に、文字領域から中心線を検出する。中心線検出は、文字領域輪郭線の垂直方向の中心点集合を候補とする。中心点集合から最小2乗法を用いて3次曲線を求め、文字領域の中心線とする(図2)。

対応点検出は、グレースケールの取得画像での正規化相互相関(NCC)を用いたテンプレートマッチング法により行う。一方の画像の文字領域中心線上の各点を中心としたテンプレートを作成し、他方の画像とマッチングを行う。平行ステレオであるため、特徴点と同じ走査線上で検出を行う。特徴点と対応点の座標値からステレオ計測を行い、文字領域中心線の3次元座標値を求める。

以上の処理を各文字領域毎に行うことで、文字領域の紙面形状を復元することが可能である。

## 2.2 画像変換

文字領域中心線の3次元座標値から紙面が平面となる画像に変換する。このとき、より鮮明な画像に変換するため、左右の画像のうち文字幅の広い方の画像を使用する。また、変換後の各画素の再配列処理には、逆変換に基づいた内挿法を用いる。

## 2.3 文字テンプレート作成

文字認識に用いる文字テンプレートは、変換画像から手動で切り出して作成する。なお、今回の主目的は、湾曲補正とNCCの有効性を検討することであるため、テンプレート用フォントの作成方法やその種類(大文字や数字、記号等)について詳細な設計を行わず、小文字26文字の文字テンプレートを作成した。一例を図3に示す。ここで、 $M$ は文字幅、 $N$ は文字高さである。

## 2.4 1文字領域抽出

変換画像から文字毎の文字の大きさを検出するため、1文字が存在する領域(1文字領域)を抽出する。認識精度の向上および処理時間の短縮を図るために、文字幅および文字高さにより、表1のように文字を5種類に分類する。

文字幅の検出は、文字の最小幅から最大幅までの検索範囲内で行う。小文字のみを対象としているので、最小幅



図 3 文字テンプレート  
Character template.

表 1 文字の分類  
Classification of characters.

		高さ	
		小	大
幅	狭	f r s	i j l t
	広	a c e n o u v x z	b d g h k p q y
		m w	

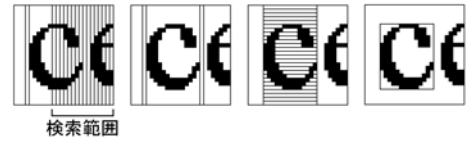


図 4 文字幅および文字高さの検出  
Detection of character width and character height.

は文字 $i$ の文字幅となり、最大幅は文字 $w$ の文字幅となる。まず検索範囲の最初の1画素目の1列について調べる。1画素毎の濃度 $g(i, j)$ と背景濃度 $\tilde{g}$ との差の2乗和 $W_i$ を式(1)より算出する。ここで背景濃度とは、文字領域以外の領域の平均濃度である。これを検索範囲の最後の列まで順次算出する。算出後、検索範囲内で背景濃度との差 $W_i$ が最小となる列 $i_{min}$ を文字の切れ目とし、文字幅として検出する(図4)。

$$W_i = \sum_{j=1}^v \{g(i, j) - \tilde{g}\}^2 \quad (1)$$

ここで、 $i$ : 列番号、 $v$ : 文字領域の高さである。

同様にして、文字高さについても1文字毎の文字幅を検出範囲として検索を行い、文字高さを検出する。

## 2.5 文字認識

一般的な文字認識では、2値化された取得画像に対して、テンプレートマッチングや構造解析を行う。しかし、画像内の位置によって明るさの変化の影響がある場合には、正しく認識することができない。

そこで、明るさの変化に対応するため、グレースケールの取得画像を用いて、NCCによるテンプレートマッチング法により文字認識を行う。

相互相関係数 $C$ は、式(2)で与えられる。

$$C = \frac{\sum_{j=1}^N \sum_{i=1}^M \{I(i, j) - \mu_I\} \{T_c(i, j) - \mu_T\}}{\sqrt{\sum_{j=1}^N \sum_{i=1}^M \{I(i, j) - \mu_I\}^2 \{T_c(i, j) - \mu_T\}^2}} \quad (2)$$

$$\text{ただし, } \mu_I = \frac{1}{MN} \sum_{j=1}^N \sum_{i=1}^M I(i, j),$$

$$\mu_T = \frac{1}{MN} \sum_{j=1}^N \sum_{i=1}^M T_c(i, j)$$

ここで、 $I(i, j)$  は取得画像、 $T_c(i, j)$  は文字テンプレート画像 (図3) である。

文字認識は、マッチング領域内の各位置で各文字テンプレートについてマッチングを行う。マッチング領域は、前節で求めた文字幅および高さを持つ領域にオフセットを加えて、誤差を補償した領域である。テンプレートマッチングにより、最も相関が大きい文字を認識結果とする。

### 3. 提案手法の評価

実験では、カメラの個体差による影響を除外するため、1台の高精細デジタルカメラ (432万画素) を書籍に対して平行に移動させて撮影した。

#### 3.1 湾曲した書籍

図5に示す湾曲した書籍について実験を行った。図5に示した書籍をスキャナ等で読みむと、綴目部分に影が生じ文字認識は困難である。実験では、取得画像 (2400 × 1800画素) の部分画像 (580 × 380画素) (図6) を使用して処理を行った。なお、取得画像サイズで A4見開きの画像を1度に取得することが可能である。

ステレオ計測による文字領域中心線の3次元座標値の算出結果を図7に示す。紙面形状の定量的な比較は行っていないが、およその形状は復元されていることが確認された。

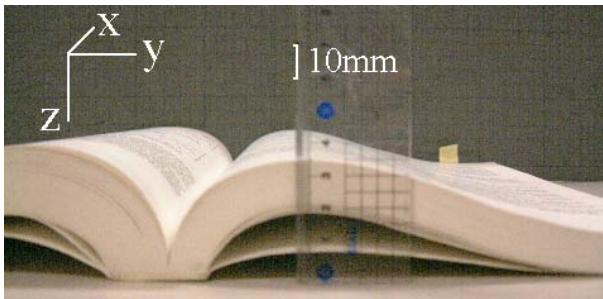


図5 表面湾曲書籍  
A book with curved surface.

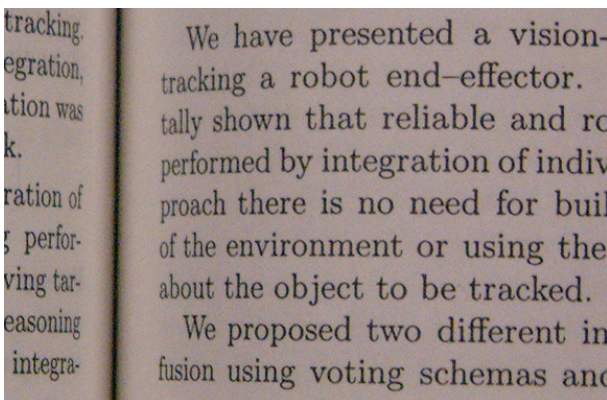


図6 取得画像 I  
Original image I.

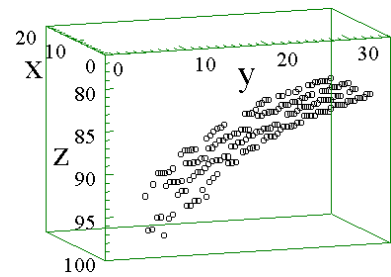
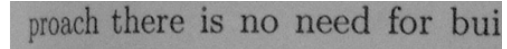
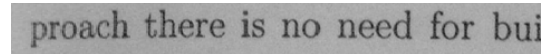


図7 紙面形状  
Shape of the book surface.



(a) 画像変換前  
Before transformation.



(b) 画像変換後  
After transformation.

図8 画像変換結果 I  
Result of transformation I.

表2 文字認識結果 I  
Result of character recognition I.

	総数	認識	認識率
(a) 市販 OCR (湾曲補正なし)	215	180	83.7 %
(b) 市販 OCR (湾曲補正あり)	215	188	87.4 %
(c) 本手法 (湾曲補正なし+NCC)	215	181	84.2 %
(d) 本手法 (湾曲補正あり+NCC)	215	200	93.0 %

画像変換の結果の一部を図8に示す。画像の再配列処理には、最近隣内挿法を用いた。画像変換の結果は、目視においては良好である。

本手法の有効性を検証するため、市販 OCR との比較を行った (表2)。表2の結果は、本手法との統合性を取るためすべて大文字を含まない結果とした。用いた市販 OCR は、取得画像を自動で2値化し文字認識を行うものである。表2において、(a) は、取得画像を用いて市販 OCR で認識した結果である。(b) は、本論文で提案した湾曲補正を行った画像を用いて、市販 OCR で文字認識した結果である。(c) は、NCC による湾曲補正前の文字認識結果である。(d) は、NCC による湾曲補正後の文字認識結果である。

表2において (a) と (b)、および (c) と (d) から湾曲補正を行うことで認識率が向上することが確認された。(a) と (b) では、湾曲補正後もそれほど認識率は向上していないが、湾曲部分の認識率に限れば 63.4% から 75.6% へ 12% ほど向上していることから、湾曲補正が適切に行われていることがわかる。

(a) と (c) では、ほとんど同じ認識率であるために NCC の効果がわかりにくくなっている。これは、市販 OCR の場合は多数のフォントに対応しているため、湾曲部分の幅の狭くなっている文字に対しては、文字幅の狭いフォントを用いることである程度認識ができることに対して、NCC で

は、1種類のフォントしか作成していないため、幅の狭くなっている文字が認識できないためである。実際に、NCCでは、文字幅の誤認識以外の原因で文字を誤認識することはほとんどなかった。

(b)と(d)より、湾曲補正にNCCを組合せることで、認識率が大幅に向上することがわかる。NCCは明るさの変化に強いのため、湾曲補正を行い正規の文字幅に変換を行うことで、2値化を行う市販OCRの方法より認識率が大幅に向上する。

これらの結果、(a)と(d)より、湾曲補正を行いNCCで文字認識を行う本手法は、市販OCRと比較して9.3%も認識率が向上しており、有効性が確認された。

本手法で認識することができなかった文字と、その原因を以下に示す。

(1) 大文字に続く文字(2文字)

小文字の文字テンプレートしか作成していないため、誤認識を引き起こす。

(2) *f* と *f* に続く文字(9文字)

*f* は、次に続く文字と接しているために、文字幅の検出が正しく行えないことがある。

(3) *m* と *w* (3文字)

文字幅が広いので2文字に分割されることがある。

(4) その他(1文字)

(1)については、大文字テンプレートを作成することで誤認識はなくなることが予想される。ただし、大文字テンプレートを使用する場合には、文字幅検出の際、大文字1文字を小文字2文字などと誤検出する可能性がある。これについては、文字幅の候補を数種類使用して文字認識を行うことで対応可能であると考えられる。

また、(2)と(3)は、市販OCRでも誤認識を引き起こす

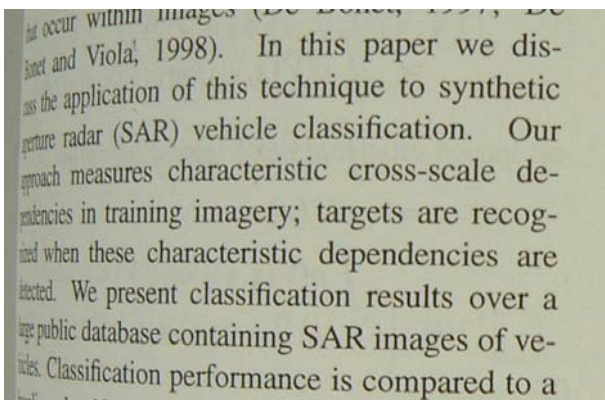


図9 取得画像 II  
Original image II.

表3 文字認識結果 II  
Result of character recognition II.

	総数	認識	認識率
(a) 本手法(湾曲補正なし+NCC)	20	2	10%
(b) 本手法(湾曲補正あり+NCC)	20	14	70%

dependencies in training imagery;

(a) 画像変換前  
Before transformation.

dependencies in training imagery;

(b) 画像変換後  
After transformation.

図10 画像変換結果 II  
Result of transformation II.

紛らわしい文字であった。

3.2 大きく湾曲した書籍

辞書や論文集などの中には、非常に分厚い書籍も多い。そこで、図6に示した画像よりもさらに湾曲が大きい書籍(紙面の最大高低差50mm以上)について実験を行った。湾曲の非常に大きい書籍の取得画像を図9に示す。

画像変換結果の一部を図10に示す。

湾曲が最も大きい部分での、特に文字形状が類似している文字*c*、*e*の文字認識結果を表3に示す。

表3より、湾曲補正を行うことにより、飛躍的に認識率が向上することが確認された。また、本論文では平行ステレオカメラにより画像を取得したが、湾曲部分がより鮮明になるよう、カメラの角度を少し傾けたシステムを用いることにより、認識率は実用レベルに近づけることが可能であると考えられる。

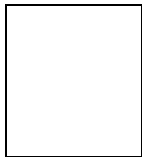
4. 結論

本論文では、高精細デジタルカメラにより撮影した紙面画像から、画像処理により本の湾曲を修正し、文字認識を行う方法について検討した。具体的には、ステレオ撮影により取得した画像からの対応点検出、3次元座標値からの画像変換、グレースケール画像を用いたNCCによる文字認識を行った。実際の書籍による実験を行い、本手法によって湾曲の修正および文字認識が可能であることを示した。

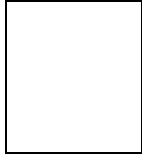
今後の課題としては、現在手動で作成している文字テンプレートの自動作成方法の検討、およびカメラの角度を傾けたシステムを用いることによる認識率の向上が挙げられる。

〔文 献〕

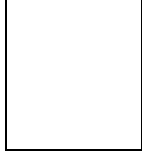
- 1) 和田俊和, 浮田浩行, 松山隆司: “イメージスキャナを用いた書籍表面の3次元形状復元(Ⅰ)—近接照明下でのShape from Shading—”, 信学論誌 D-II, J77-D-II, 6, pp.1059-1067, (1994).
- 2) 和田俊和, 浮田浩行, 松山隆司: “イメージスキャナを用いた書籍表面の3次元形状復元(Ⅱ)—相互反射を考慮した近接光源下のShape from Shading—”, 信学論誌 D-II, J78-D-II, 2, pp.311-320, (1995).
- 3) 浮田浩行, 小西克信, 和田俊和, 松山隆司: “固有空間法を用いた陰影情報からの書籍表面の3次元形状復元”, 信学論誌 D-II, J83-D-II, 12, pp.2610-2621, (2000).



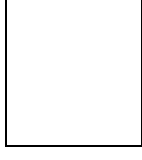
かわらご あつし  
**川原 淳** 2001年，静岡大学工学部機械工学科卒業．同年，静岡大学大学院理工学研究科機械工学専攻修士課程入学．主として画像処理の研究に従事．



やました あつし  
**山下 淳** 2001年，東京大学大学院工学系研究科精密機械工学専攻博士課程修了．同年，静岡大学工学部機械工学科助手．画像処理，ロボット群の動作計画，ロボットの移動機構の研究に従事．博士（工学）．



かねこ とおる  
**金子 透** 1974年，東京大学大学院工学系研究科物理学専攻修士課程修了．1974年，日本電信電話公社（現NTT）入社．1997年，静岡大学工学部教授．画像処理，ロボットビジョンの研究に従事．工学博士．正会員．



みうらけんじ りゅう  
**三浦憲二郎** 1984年，東京大学大学院工学系研究科精密機械工学専攻修士課程修了．同年，キャノン（株）入社．1991年，コーネル大学機械工学科博士課程修了．1993年，会津大学コンピュータ理工学部コンピュータソフトウェア学科助教授．1997年，静岡大学工学部助教授．Ph.D．曲線・曲面の設計，CAD/CAM，要素自動分割の研究に従事．