

# ステレオカメラを用いた歪曲形状書籍画像補正の高速化

## Fast Correction of Distorted Document Images Using a Stereo Camera System

秋本 廉太郎, 山下 淳, 金子 透

Rentaro AKIMOTO, Atsushi YAMASHITA, Toru KANEKO

静岡大学工学部機械工学科

Department of Mechanical Engineering, Shizuoka University

E-mail: tayamas@ipc.shizuoka.ac.jp

### Abstract

書籍の情報をデジタル化するにはスキャナが広く用いられているが、書籍の綴じ目付近がスキャナ面に密着しないことによって、紙面内容の歪みや陰影が発生し、OCR（光学文字認識）の精度を低下させる原因となる。陰影や文字列の向きを利用して補正する手法などがあるが、書籍の色や内容に依存するという欠点がある。一方、カメラを用いたデジタル化の研究も行われている。撮影の手間が少ないなどの利点はあるものの、形状歪みや陰影はスキャナの場合よりも広域にわたって発生する。スキャナやカメラを含め様々な手法が提案されているが、撮影の手間や装置、処理速度などの全ての問題に対応した手法は提案されていない。

本研究ではステレオカメラシステムを用いて、湾曲に起因する形状歪みと陰影の補正を行う。また、様々な処理を工夫して組み合わせることで、高速な補正処理の手法を提案する。

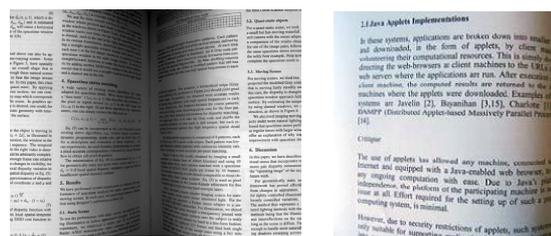
## 1 序論

近年、書籍等の既存文書のデジタル化が盛んに行われている。古書のような貴重な資料が劣化等によって損失した場合に備え、文書情報をデジタル保存しておくことが重要である [1]。例えば貴重な古書をデジタル化する専用ロボットシステム [2] が開発されているなど、文書情報のデジタル化が重要視されている。またオフィス環境等では、膨大な量の既存の資料をデジタル化することにより、省スペース化や資料検索効率の向上も期待できる。

本研究は、各種装置で撮影された書籍画像を、閲覧性や文字認識率を高めるために施す補正に関するものである。補正の対称は主に形状の歪みと陰影であり、従来から多様な研究がされてきたが、そのアプローチは基本的に2種類に分けられる。スキャナベースの研究

とカメラベースの研究である。

一般に普及しているのはスキャナベースの手法である。スキャナベースでは紙面をガラスの読み取り面に押し付けることで物理的に平面状態を実現し、形状歪みは綴じ目付近にのみ発生する（図 1(a)）。照明や撮像装置の拘束も強いので、形状歪みや陰影のような不具合が少なく、補正すべき部分が少ないという利点がある。また、一般的にカメラベースよりも解像度が高い。



(a) スキャナ

(b) カメラ

図 1 取得画像の違い

スキャナベースの代表的な手法は、Shape from shading を用いたものである [3][4][5]。スキャナで取得した文書画像の綴じ目付近には、スキャナ面から紙面までの距離に応じた陰影が生じる。そこで陰影から距離を逆算することによって紙面形状を推定し、推定した3次元形状に基づいて形状歪み・陰影の補正を行っている。また、読み取り面の一部が紙面に沿うような円筒状になっており、綴じ目付近の陰影を軽減するスキャナも存在する。しかしどの手法も、読み取り面に押し付けることに起因する綴じ目へのダメージが大きいことや、手間が多い（読み取り ひっくり返す ページをめくる ひっくり返す）といった、構造的な欠点は回避できない。

一方、カメラベースの研究も盛んに行われている [6][7]。カメラによる取得画像（図 1(b)）を見ると分かるように、形状歪み・陰影ともにスキャナベースの場合よりも広範囲に発生しており、その補正は格段に困難であることは明らかである。その代わり、書籍への

ダメージや撮影の手間などの問題はほぼ解決される。

カメラベースの研究は主に、レーザ [8][9]・Shape from shading [10][11][12]・ステレオカメラ [13][14] を用いて書籍の 3 次元形状を取得してそれを引き伸ばすものや、紙面の局所画像をモザイクングによって高解像度な画像を取得するものなどがある [15]。陰影については、照明条件を変えて複数回撮影して陰影分布を推定する手法 [14] や、紙面の濃淡値をそのまま陰影の度合いとみなす手法 [16] などがある。またスキャナベースとは違い固定された読み取り面を持たないので、対象を限定しないという大きな利点を持つ手法もある。しかしこの手法も、(1) 装置が大規模または一般的でない、(2) 処理時間が長い、(3) 紙面内容に依存する、といった欠点を 1 つ以上抱えている。特に処理速度に関しては最重要視されることは少なく、実用化に向けた大きな課題の 1 つとなっている。

## 2 研究目的

第 1 章で述べた通り、スキャナベースの手法には構造的に避けられない欠点があるのに対して、カメラベースの手法の欠点は主に、形状歪み・陰影などの不具合が多く補正が困難なことである。そこで本研究ではこの問題に加えて処理速度の課題にも対応可能な新しい手法を提案する。

本研究では従来研究 [14] と同じくステレオカメラを用い、形状歪みと陰影補正を高速に行う。基本的な処理の流れも踏襲するが、各処理で計算コストを重視し大幅な高速化を目指す。書籍の 3 次元形状は輪郭から計測・推定し、既知の照明条件を考慮することで陰影の補正を行う。また、ステレオカメラにより画像を取得するため、1 台のカメラのみを用いた手法に比べて多くの情報を得ることが可能となる。

## 3 処理の概要

### 3.1 撮影装置と前提条件

本研究で用いる装置の概要を図 2 に示す。黒いテーブルの上に見開いた書籍を置き、上方からステレオカメラで 1 回撮影する。

前提条件として、以下の事項を設ける。

- カメラと照明の位置・姿勢は既知である。
- 綴じ目に垂直な断面の、綴じ目方向の変化は直線的である。
- ステレオカメラの基線と綴じ目がなす角度は  $\pm 45^\circ$  以内である。

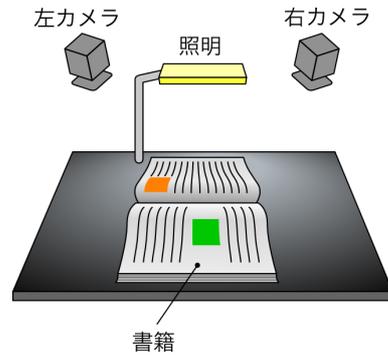


図 2 装置の概要

### 3.2 処理の流れ

本研究の処理の流れは以下の 1～5 の順に行う (図 3)。

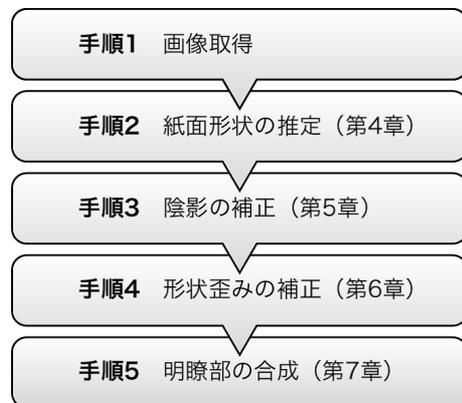


図 3 処理の流れ

#### 手順 1: 画像取得

撮影に先立って、カメラの校正を行い、カメラの像距離、基線長、姿勢などを算出する。校正後はそれらのパラメータは変えず、ステレオ画像を取得する。

#### 手順 2: 紙面形状の推定

左右の画像における紙面上端・下端に対して対応点検出を行い、三角測量の原理を用いて紙面上端・下端の 3 次元座標を算出する。そして内部は補間することによって紙面全体の形状を推定する。最終的には、点の集合から成るメッシュモデルを作成する。形状推定についての詳細は、第 4 章において述べる。

#### 手順 3: 陰影の補正

陰影の補正処理は、既知の光源情報を、推定した紙面形状情報に基づき、紙面での光の反射を計算することによって行う。まず光源情報と紙面形状を利用して、紙面の全域について、反射光強度を計算する。そして反射光強度の分布を基にして、原画像から陰影成分を取り除く。陰影補正についての詳細は、第 5 章で述べる。

#### 手順 4: 形状歪みの補正

歪み補正は、推定された紙面を平面に引き伸ばすことで行う。まず、メッシュモデルの各頂点について、引

き伸ばし後の位置を算出する．そして引き伸ばし前の位置に対応する取得画像を，引き伸ばし後の位置へと変形して当てはめ，歪み補正画像を生成する．形状歪みの補正についての詳細は，第6章で述べる．

#### 手順 5: 明瞭部の合成

ステレオカメラによる撮影では紙面とカメラの位置関係によって，紙面の同じ部分を撮影したにも関わらず一方では明瞭，他方では不明瞭となる部分が発生する．そこで，2枚の画像の明瞭な部分のみを組み合わせることによって，全域に渡って明瞭な画像を得る．明瞭部の合成についての詳細は，第7章で述べる．

## 4 紙面形状の推定

### 4.1 紙面上端・下端のステレオ計測

本研究では書籍の3次元形状を推定するにあたってステレオ計測を用いるが，計測対象は紙面上端と下端のみとする．理由の1つは，高速化である．画像全体に対してパターンマッチングを行えば紙面の大きな3次元形状は得られるが，計算時間が長くなることは避けられない．もう1つの理由は，紙面内容に依存しないことである．紙面内部に存在する対応点を探す場合，紙面に書かれている文字や絵，模様（テクスチャ）に頼ることになる．従って，テクスチャの乏しい部分の形状は得られず，曲面へのフィッティング等，別の処理が必要となってしまう．

以上の理由から本研究では，ステレオ計測によって3次元座標を取得するのは，紙面上端と下端のみとする．紙面内部の形状は，第3章で述べた前提のもと，補間により取得する．

計測に先立ってカメラの校正を行う．これによりカメラの像距離・レンズの歪曲収差といった内部パラメータや，カメラの位置，姿勢といった外部パラメータを既知として扱える．さらに，左右の画像に平行化という前処理を施す．これは輻輳ステレオにより得られた画像を，投影変換を用いて平行ステレオ画像に変換処理である．平行化した画像対において対応点は同じ水平走査線上に存在するため，後の計測処理が容易になる．

対応点の検出には単純に閾値を用いる．水平走査線上を探索し，ある画素のグレースケール値がある閾値を超えたら紙面，閾値以下なら背景であると判別する．対応点が検出できたらその座標とカメラのパラメータから，3次元座標値が算出できる．

この処理を紙面上端・下端に対してある間隔ごとに行うことで，紙面上端・下端の3次元形状が求められる．このとき，プロットする点の数は上下で等しくしておく．

### 4.2 紙面内部の補間

紙面の内部の形状は，先に計測した点を補完することで推定する．上下端の対応する点同士を結ぶ線分を  $n$  等分することで，紙面内部の3次元形状を直線的に補間する．以上の処理によって図4のような，書籍形状を表すメッシュモデルを作成する．メッシュモデルを構成する4つの頂点からなる四角形領域をパッチと呼び，以後このパッチに対して各補正処理を施す．

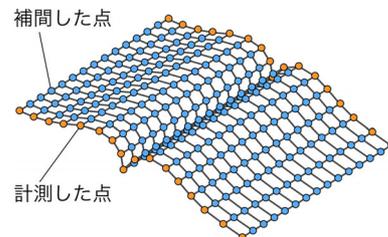


図4 メッシュモデル

## 5 陰影の補正

歪曲形状を有する紙面に照明光を照らすと，光源と紙面の位置関係および紙面の歪曲形状に起因した陰影が生じる．本研究では，照明光の紙面上での反射をモデル化し，紙面上の各位置において反射光強度値の計算を行う．そして反射光強度値の大小に基づいて画像の画素値を変化させることにより，陰影の補正を行う．

### 5.1 反射光強度分布の算出

拡散反射のみを考えた場合，紙面上の点  $P(x, y, z)$  における反射光強度  $I(x, y, z)$  は，式(1)で表される(図5)．

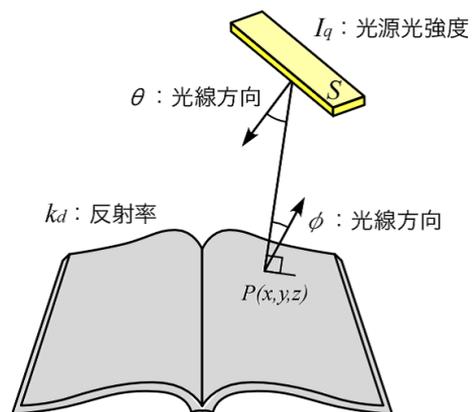


図5 光の反射のモデル化

$$I(x, y, z) = k_d I_q \iint_S \frac{\cos \theta}{r^2} \cos \phi \, ds \quad (1)$$

上式において， $k_d$  は紙面の拡散反射率， $I_q$  は光源光の強度， $S$  は光源の面積， $\theta$  は光源に垂直な方向と光線方向とのなす角， $r$  は光源と点  $P$  との距離， $\phi$  は紙面法



## 7.1 合成領域の決定

まず第4章で作成したメッシュモデルの各メッシュについて、カメラ視線と紙面法線との角度差を計算する(図8)。この2つのベクトルの角度差が小さいほど、紙面はより正面に近い方向から撮影されており、高精細なテクスチャが得られているはずである。そこで、メッシュモデルの各パッチにおいてこの角度差を求める。算出された角度差の値から、左右どちらの画像を使った方がより高精細かを各メッシュごとに判定し、合成すべき領域が決まる。

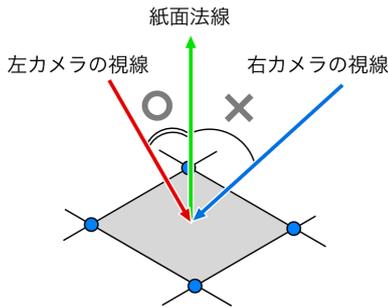


図8 パッチ法線とカメラ視線との角度差

## 7.2 合成処理

本来なら上記の方法で決定した合成領域で、単純に他方の画像へコピーすればテクスチャは完全に一致するはずである。しかしステレオ計測等の誤差によりわずかにずれが生じることがある。そこで各メッシュについて、合成基の画像と合成先の画像とでマッチングを行い、より自然な合成画像を得る。

以下に処理手順を説明する。合成領域の選別の結果、図9のように、左画像の赤い領域と右画像の青い領域が採用されたとする。つまり青い領域を左画像の下半分に描画することになるが、単にコピーしただけではテクスチャは一致しない。そこでまず合成境界付近のパッチ(図中では濃い色のパッチ)を用いてテンプレートマッチングを行い、合成前後での対応点を求める。その対応関係から、変形に必要な射影変換行列を算出し射影変換を施すことで、境界でのずれは軽減される。

さらに境界付近を描画する際には、境界からの距離に応じたグラデーションを施す。つまり境界付近では、右画像に近い画素ほど右画像の画素値の影響を多く、左画像に近い画素ほど左画像の画素値の影響を多く持たせて、合成後の境界を目立ちにくくさせる。

## 8 実験

実験は図10に示す環境で行った。テーブルに開いた書籍を置き、上方から2台のデジタル一眼レフカメラ

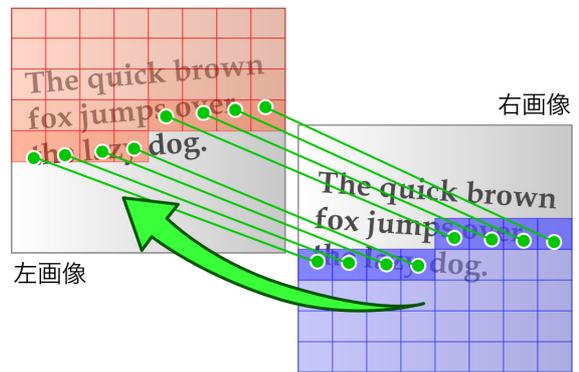


図9 合成領域の対応付け

ラ(Nikon D700)によって紙面を撮影した。光源には面光源(スタンドライト)を用いた。カメラ間の距離は約20cm、テーブルとカメラとの距離は約40cmである。この装置で撮影されるステレオ画像の例を図11に示す。

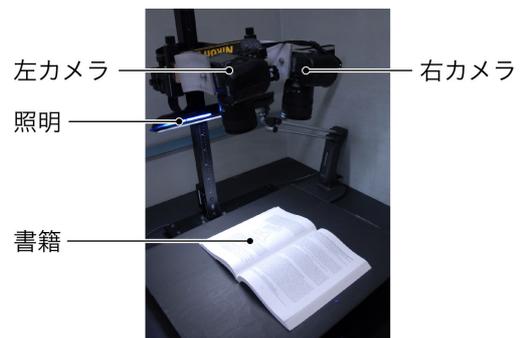


図10 実験環境

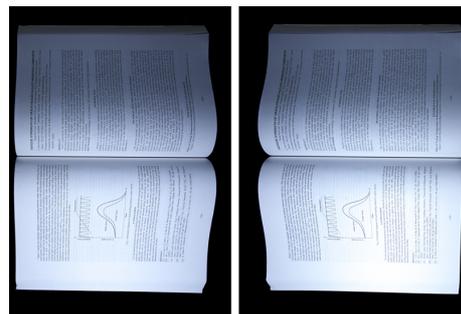


図11 ステレオ画像

画像サイズは左右それぞれ4256×2832ピクセルである。7.2節の合成領域の位置合わせをする際に使うテンプレートサイズは101×101ピクセルとし、正規相互相関係数の閾値は0.90とした。また、処理に用いたコンピュータのCPUはCore 2 Duo 3.0GHz、メモリは4GBである。

次に実際の処理結果を示す．形状歪みの補正結果を図 12 に，3 次元形状の推定結果を図 13 に，陰影分布の推定結果（推定した反射光強度をグレースケールで表現したもの）を図 14 に，左右画像の合成結果を図 15 に，最終的な補正結果を図 17 に示す．

計測したのは上下端のみであるにもかかわらず，紙面の形状と陰影分布が再現できていることがわかる．明瞭部を合成した境界も滑らかであり，実用上十分な結果となっている．

形状補正の精度は，20 mm 間隔の格子模様を印刷した書籍に補正処理を施し，その寸法を測ることで評価した（図 16）．本手法で補正した後の格子間隔の平均値は 19.7 mm，標準偏差は 0.72 mm，従来研究 [14] での平均値は 19.7 mm，標準偏差は 1.08 mm であった．

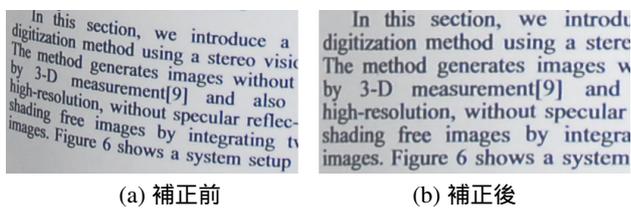


図 12 形状歪みの補正結果

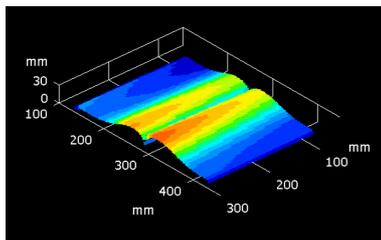


図 13 形状の推定



図 14 陰影の推定結果

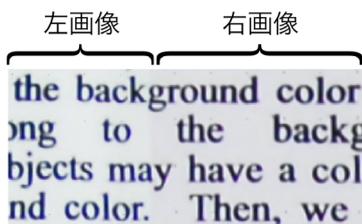


図 15 左右画像の合成結果

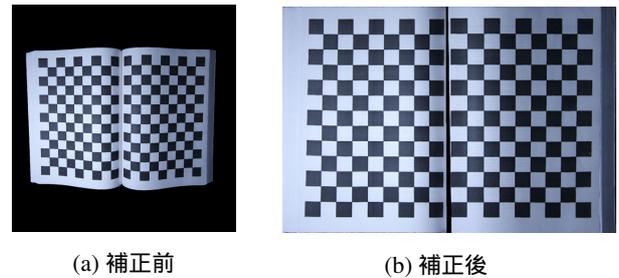


図 16 格子パターンの補正

表 1 処理速度

	本手法 [sec]	従来研究 [14][sec]
形状歪みの補正	2	25
陰影の補正	2	150
明瞭部の合成	2	5
計	6	180

また，A.I.Soft 社の OCR ソフト「読ん de!!ココ ver.13」を用いて文字認識率の評価も行った．補正前には 59.7% だった認識率が，本手法では 98.7%，従来研究 [14] では 98.9% となった．ちなみに，スキャナを使って理想的な状態（紙面を物理的に平面にした状態）で読み取りを行った画像の文字認識率は，99.9% であった．本手法が平面状態に迫る仕上がりを実現できていることがわかる．

参考のため従来研究 [14] の結果（図 18）と本研究の結果（図 19）を比較すると，本手法が従来研究と同程度の結果を残していることがわかる．ただし本研究では，従来研究のように照明条件を変えての複数回の撮影は行っていないので，鏡面反射が残っている（矢印部分）．

最後に，各処理に要した時間を表 1 に示す．形状補正の精度や文字認識率は従来研究 [14] と同程度を保ちながら，すべての処理を 10 秒以下という短い時間で実行できた．大きな要因は，対象を書籍に限定したことである．この条件を設けることによって，各処理で「ある程度疎な計測をして，その他の部分は補間する」という処理の流れが可能になり，高速化に大きく貢献した．

ただし従来研究 [14] と同様，綴じ目付近の陰影補正が正確にできておらず，明度が不自然になる場合がある．これは，陰影を推定する際に綴じ目の相互反射を考慮していないためだと思われる．

## 9 結論

ステレオカメラで撮影した紙面画像の歪みを補正する手法を提案した．

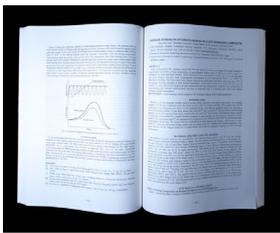
まず，ステレオ計測による紙面上端と下端のみの

形状を計測し、それを補間することで紙面形状を推定した。次に、推定した形状に対して光学的なシミュレーションを行い、反射光強度を求め、その分布に基づいて陰影の補正を行った。そして、推定した紙面形状を平面に引き伸ばすことによって形状歪みの補正を行った。最後に、ステレオカメラで取得した2視点画像の明瞭な部分同士を組み合わせることで、全域にわたって明瞭な合成画像を生成した。

今後の課題として、綴じ目の相互反射や光源の種類等を厳密に考慮して、より正確な陰影分布を推定することが挙げられる。



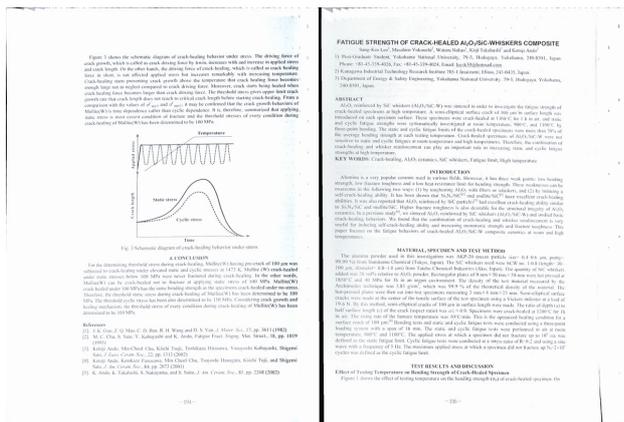
図 18 従来研究 [14] の結果



(a) 原画像 A



(b) 原画像 B



(c) 統合処理結果 A



(d) 統合処理結果 B

図 17 原画像と統合処理結果



図 19 本手法の結果

謝辞

本研究の一部は、科学技術振興機構 平成 20 年度 地域イノベーション創出総合支援事業 シーズ発掘試験の補助を受けた。

参考文献

- [1] M. Kashimura, T. Nakajima, N. Onda, H. Saito and S. Ozawa: "Practical introduction of image processing technology to digital archiving of rare books", Proceedings of International Conference on Signal Processing Application Technology, pp.1025-1029, 1998.
- [2] M. Dumiak: "Book-scanning robots digitize delicate texts", IEEE Spectrum, Vol.45, No.1, p.10, 2008.
- [3] T.Wada, H. Ukida and T. Matsuyama: "Shape from shading with interreflections under a proximal light source", International Journal of Computer Vision, Vol.24, No.2, pp.125-135, 1997.
- [4] Z. Zhang, C. L. Tan and L. Fan: "Estimation of 3d shape of warped document surface for image restoration", Proceedings of the 17th International Conference on Pattern Recognition, Vol.1, pp.486-489, 2004.

- [5] C. L. Tan, L. Zhang, Z. Zhang and T. Xia: "Restoring warped document images through 3D shape modeling", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.28, No.2, pp.195–208, 2006.
- [6] J. Liang, D. Doermann and H. Li: "Camera-based analysis of text and documents: a survey", *International Journal on Document Analysis and Recognition*, Vol.7, No.2–3, 2005.
- [7] 黄瀬浩一, 大町真一郎, 内田誠一, 岩村雅一: "カメラを用いた文字認識・文書画像解析の現状と課題", *電子情報通信学会技術研究報告*, Vol.104, No.742, pp.85–90, 2005.
- [8] M. S. Brown, M. Sun, R. Yang, L. Yun and W. B. Seales: "Restoring 2d content from distorted documents", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.29, No.11, pp.1904–1916, 2007.
- [9] 天野敏之, 安部勉, 西川修, 伊與田哲男, 佐藤幸男: "アイスキャナによる湾曲ドキュメント撮影", *電子情報通信学会誌 D-II*, Vol.J86-D-II, No.3, pp.409–417, 2003.
- [10] S. I. Cho, H. Saito and S. Ozawa: "Shape recovery of book surface using two shade images under perspective condition", *電気学会論文誌 C*, Vol.117-C, No.10, pp.1384–1390, 1997.
- [11] F. Courteille, A. Crouzil, J.-D. Durou and P. Gurdjos: "Towards shape from shading under realistic photographic conditions", *Proceedings of the 17th International Conference on Pattern Recognition*, Vol.2, pp.277–280, 2004.
- [12] L. Zhang and C. L. Tan: "Restoring warped document images using shape-from-shading and surface interpolation", *Proceedings of the 18th International Conference on Pattern Recognition*, Vol.1, pp.642–645, 2006.
- [13] A. Yamashita, A. Kawarago, T. Kaneko and K. T. Miura: "Shape reconstruction and image restoration for non-flat surfaces of documents with a stereo vision system", *Proceedings of the 17th International Conference on Pattern Recognition*, Vol.1, pp.482–485, 2004.
- [14] 鈴木優輔, 山下淳, 金子透: "ステレオビジョンシステムによる歪曲形状文書画像の歪み補正", *ビジョン技術の実利用ワークショップ ViEW2007 講演論文集*, pp.353–358, 2007.
- [15] T. Sato, A. Iketani, S. Ikeda, M. Kanbara, N. Nakajima and N. Yokoya: "Mobile video mosaicing system for flat and curved documents", *Proceedings of the 1st International Workshop on Mobile Vision*, pp.78–92, 2006.
- [16] 包躍: "ページめくり機における書物の紙面復元機械的ページめくり機に対応可能な補正手法", *画像ラボ*, Vol.14, No.8, pp.1–5, 2003.