

ロボットビジョンにおける曲面形状上からの 情報取得に関する研究

A study on information acquisition from curved surface shape in robot vision.

原 和道[†] 矢野 良和[†] 江口 一彦[†]
Kazumichi Hara Yoshikazu Yano Kazuhiko Eguchi

Abstract

When acquiring the information on the curve surface by a camera, curved surface distortion occurs on the captured image. According to this distortion, the accuracy of image matching and the OCR ratio get down. So we propose the method to reproduce planar surface image. First, obtained image is divided into local small regions which are assumed as planar ones, and is analyzed in order to obtain the 3D surface shape information. Then, local regions are transformed according to a perspective projection with estimated 3D surface shape. Each local regions are combined into one planar image.

1 はじめに

コンピュータビジョンに関する研究が進み、ロボットへの応用も検討される。ロボットはカメラから得られた情報を画像処理によって状況を判断する。実際の環境に存在する情報は、人間に最適化された場合が多い。道路表示や街頭広告、商品ラベル、書籍などが挙げられる。このように、多くの情報は湾曲面形状上に存在する場合がある。一方で多くの画像処理アルゴリズムは、平らな表面上の情報に対しての処理を前提としている。ロボットが湾曲面形状上に存在する情報を取得するのは難しい。そのため、情報の取得時に曲面上の情報を平面として処理できる形に補正する必要がある。

従来の手法は、大きく次の手法が挙げられる。ステレオ計測や 3 次元計測に基づく手法 [1][2]、Shape from shading の原理に基づく手法 [3][4]、モデルフィッティングに基づく手法 [5][6]、カメラ運動と対象形状復元に基づく手法 [7][8] の 4 種である。ステレオ計測に基づく手法 [1] では、専用撮影機器を用いて、あらかじめ取得した背景画像との差分によって、高精度な 3 次元推定を行なっている。このようにステレオ計測に基づく手法では、いずれもステレオカメラや計測機器の綿密なキャリブレーション

が問題に挙げられている。Shape from shading の原理に基づく手法の代表である浮田らの手法 [3] は、撮影された書籍表面画像の陰影情報から、書籍の 3 次元形状を推定するものである。これ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。撮影環境の照明条件が対象の 3 次元形状データ取得に影響されてしまう。モデルフィッティングに基づく手法 [6] 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曲面歪みを表現し、対象形状を推定する手法である。カメラ運動と対象形状復元に基づく手法 [7] は、複数枚の画像を用いて、カメラの運動と対象の形状を復元するものである。この手法では、文字情報などの特徴付けが難しい対象での精度に問題がある。

既存研究では、対象や撮影環境を限定することで 3 次元データ高精度な取得を行う。高精度な 3 次元形状データをもとに歪み補正を実現してきた。ロボットビジョンのように対象や撮影環境を限定できない場合、3 次元形状復元には、モデルフィッティングに基づく手法や、カメラ運動と対象形状復元に基づく手法が選択される。一方で、これらの手法による形状推定には多くの誤差が含まれる。本研究では、高精度な 3 次元形状データ取得が難しいロボットビジョンにおいても実現可能な、曲面による歪みを含んだ画像の平面化処理を提案する。

[†]愛知工業大学 工学部 電気学科 (豊田市)

2 曲面上情報の平面化处理

2.1 平面化处理

平面化の処理手法について述べる．曲面の局所領域は平面とみなすことができる．そこで，曲面を小領域ごとに分割し，各領域の面法線方向を揃えて同一平面上に再配置することで平面化画像を得る．面法線方向を揃えて同一平面上に再配置するための処理として透視投影変換を用いる．透視投影変換とは，ワールド座標系にある平面の座標を投影面のスクリーン座標系に変換して対応付けることをいう．正面から撮影できないとき，撮影した情報は透視投影される．透視投影された平面画像を正面から見たように画像を補正するために透視変換が使われる．透視変換された画像は正面から見たような画像に補正できる．

提案する処理手法は以下の手順からなる．

1. 画像取得
2. 対象曲面形状の把握
3. 取得画像に格子をフィッティング
4. 格子交点の平面変換後座標を推定
5. 局所領域ごとに透視投影変換

以上の手順により，曲面歪みを補正した平面画像を取得する．

取得した 3 次元形状データをもとに格子状に座標を取得する．図 1 に示すように，それぞれの領域で透視投影変換を行う．透視変換後の座標は，図 2 に示す方法で求めることができる．推定した断面を一定間隔で分割する．分割した間隔を積算するように，同一平面上に分割した点を配置する．これによって透視投影変換前後の対応座標を計算することができる．この座標の対応関係をもとに，透視投影変換を行うことで平面化画像を得ることができる．

2.2 透視投影変換

透視投影変換とは，ワールド座標系にある平面の座標を投影面のスクリーン座標系に変換して対応付けることをいう．対象を正面から撮影できないとき，撮影した情報は透視投影される．画像処理アルゴリズムの多くは，正面で撮影されていることを前提として処理している．透視投影された

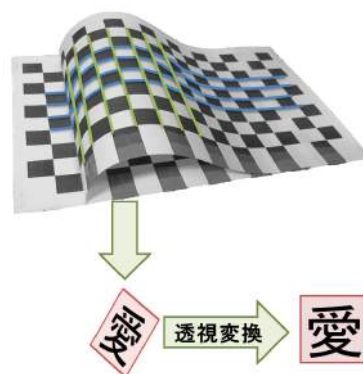


図 1: 領域分割変換

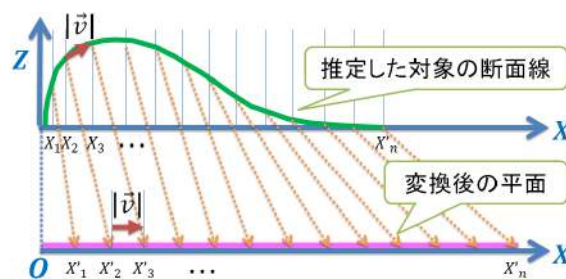


図 2: 透視変換後の座標決定手法

平面画像を正面から見たように画像を補正するために透視投影変換が使われる．透視変換された画像は正面から見たような画像に補正できる．透視投影変換を適用するには対象が平面である必要がある．透視投影画像中にある情報の頂点となる四点を選択し透視投影変換が行われる．

透視投影変換を数式モデルとして考えると，式 (1) のように表すことができる．

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = H_m \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad (1)$$

ここで， u, v は 2 次元平面座標， x, y, z は 3 次元空間座標とする． H_m は式 (1) 中の各パラメータを 1 つにまとめた 3×3 の行列で表す透視投影変換パラメータであり，次のように表現する．

$$H_m = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \quad (2)$$

H_m のみで画像の移動，回転，拡大，縮小，遠近感を全て表現することができる．

2.3 計測によらない

3次元形状情報の取得

曲面形状情報の推定方法について述べる．図3に曲面形状の取得手法を示す．本研究では，撮影する対象を基準にした三次元座標を世界座標として置く．原点を対象の左下端に位置付ける．対象の横方向に X 軸，縦方向に Y 軸，対象面の法線方向に Z 軸を置く． $X-Z$ 平面において， Y 座標が同一となる対象表面上の点 A, B の座標を取得する．このとき， Y 座標が同一となる点 A, B を得る方法として文字列の配置に対するモデルフィッティング法 [5][6] や文字情報から特徴点を得る方法 [9] が考えられる．

取得した座標 A, B から Z 軸方向に伸びる点 C, D を取得する．撮影された画像情報のみでは Z 軸方向を決めることは難しい．そこで， Z 軸方向に伸びる点は，探索によって決定する． $X-Z$ 平面における四角形 $ABCD$ は，撮影画像中では透視投影歪みによって長方形とは異なる形状となる． C, D の位置を推定により求めるが，4次元の探索空間では推定が困難となる．そこで， $AB \parallel CD$ と形状を近似する．また，底辺とのなす角 $\angle DAB, \angle CBE$ も，ともに θ で同一とした．この平行四辺形の底辺と斜辺のなす角 θ を変えながら理想的な四点をみつける．取得した四点の座標をもとに，透視投影変換を行う． Y 座標を一定とした $X-Z$ 平面上に置かれた四点座標の透視変換結果は図4のように断面形状として得られる．曲面形状推定を Y 軸方向を変えながら行う．これより曲面の3次元形状を推定する．

2.4 高画質化処理

高画質化処理について述べる．図5に示すように，カメラ光軸と対象表面法線のなす角が大きい領域では，撮影画像中で表現されるための画素数が少なくなってしまう．また，カメラと対象までの距離が離れている領域は小さく撮像される．小さな面積の領域は透視投影変換によって引き伸ばされ，間抜けした画素は補完される．そのため，変換後の画質が局所的に低下してしまう．OCR への応用を考えると 300dpi 程度の画質が要求される．

透視変換後に高い画質を維持するため，画質の落ちた座標の複数視点による補間を行う．複数の視点で撮影された画像を提案する手法によって平

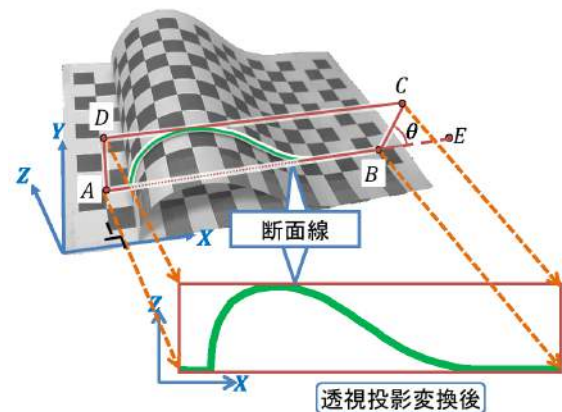


図 3: 曲面形状の取得手法

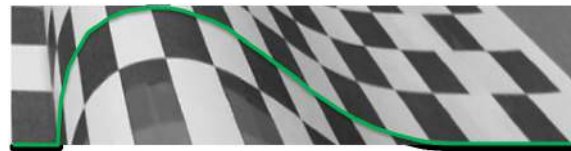


図 4: 取得できる断面形状

面化する．このとき，変換後の画像はそれぞれの撮影位置によって局所的に画質が劣化する．このときの画質変化を数値化する．複数視点のそれぞれの画像で比較し，より高画質となった画素を選択して抜き出し合成する．

多視点の複数画像から高解像画素のみを抽出するために画像内に存在する対象平面全体の画質を数値的に把握する．ここで対象平面，カメラ画像を図??(a),6(b)に示すように画素の単位に分割して考える．対象平面の画素座標 (x, y) がカメラ画像内に存在する平面の画素座標 (u, v) に対応しているとき， u, v をそれぞれ x, y の関数として次のように表す．

$$u = f(x, y)$$

$$v = g(x, y)$$

対象平面の1画素の面積を 1×1 の単位画素と考え，この画素面積が透視投影されたカメラ画像では $P_x \times P_y$ に変化しているとする．この面積の変化により画質がどのように変化しているかを数値的に把握することができる．これを局所領域解像度率 ARR (Area Resolution Ratio) として画質の指標とした．数式で表すと式 (3) のようになる．

$$ARR = P_x \times P_y \quad (3)$$

カメラ画像中に存在する平面の画素 (x, y) の横幅 P_x と画素の縦幅 P_y は，式 (4), (5) に示すように，

隣接する画素の座標値の差分から座標距離を求めることにより算出できる.

$$P_x = \sqrt{\left(\frac{d}{dx}u\right)^2 + \left(\frac{d}{dx}v\right)^2} \quad (4)$$

$$P_y = \sqrt{\left(\frac{d}{dy}u\right)^2 + \left(\frac{d}{dy}v\right)^2} \quad (5)$$

式 (4),(5) 中の各微分式は式 (2) の透視投影パラメータ H_m の各要素 h_{ij} を用いてそれぞれ,

$$\frac{d}{dx}u = \frac{(h_{11}h_{32} - h_{12}h_{31})y - h_{13}h_{31} + h_{11}h_{33}}{(h_{31}x + h_{32}y + h_{33})^2}$$

$$\frac{d}{dx}v = \frac{(h_{21}h_{32} - h_{22}h_{31})y - h_{23}h_{31} + h_{21}h_{33}}{(h_{31}x + h_{32}y + h_{33})^2}$$

$$\frac{d}{dy}u = \frac{(h_{12}h_{31} - h_{11}h_{32})x - h_{13}h_{32} + h_{12}h_{33}}{(h_{31}x + h_{32}y + h_{33})^2}$$

$$\frac{d}{dy}v = \frac{(h_{22}h_{31} - h_{21}h_{32})x - h_{23}h_{32} + h_{22}h_{33}}{(h_{31}x + h_{32}y + h_{33})^2}$$

と表される. x, y の範囲をそれぞれ,

$$-\frac{W}{2} \leq x < \frac{W}{2}$$

$$-\frac{H}{2} \leq y < \frac{H}{2}$$

W : 透視変換前の対象平面における横幅

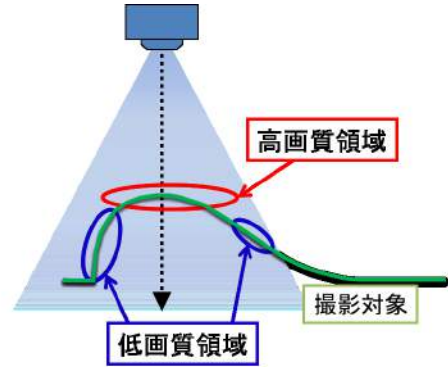
H : 透視変換前の対象平面における高さ

として ARR の計算を行うことで, カメラ画像内に存在する対象平面の座標毎の面積を求められる.

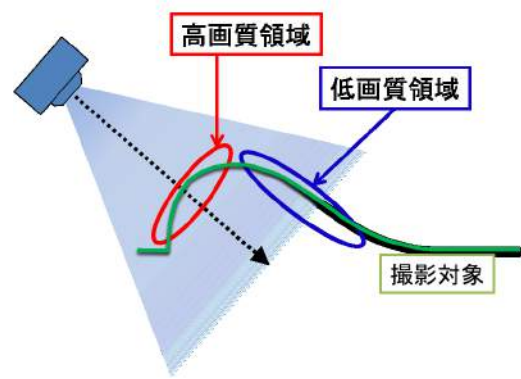
ARR の値が高い領域はより画質が高いといえる. ARR を用いて複数視点の画像中でより画質の高い部分を選択して使用することで高画質化を実現する.

3 実験結果および考察

提案する手法によって断面形状を推定し, 平面化を行う. 撮影には USB カメラ (ELECOM UCAM-DLK130T) を使用し, 画像サイズは 640×480 pixel とした. 推定した形状をもとに平面化処理を行う. 各局所領域の大きさは断面線を 16 等分するような大きさで設定した. 撮影で得られた図 7 に対し, 提案する手法による平面化処理を行う. また, 文書情報の書かれた図 8, 9 に対し平面化と高画質化処理を行う. 得られた平面画像の有用性を評価するため, OCR を適用した識字率評価を行う. 本稿では OCR として Google ドライブのドキュメント変換機能を使用した. 評価の対象として, フラットベッ

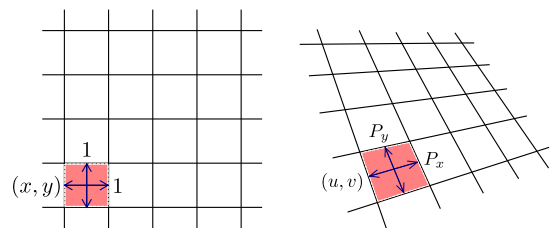


(a) 対象正面から撮影



(b) 対象左側から撮影

図 5: 画質変化の原因



(a) 対象平面

(b) カメラ画像

図 6: 画素面積の変化

ドスキャナで読み取った画像と, 平面の対象を撮影した画像を用意した. フラットベッドスキャナで取り込んだ画像は, 物理的な紙面から画像情報を形状歪みなく取り込んだ理想画像といえる. 平面の対象を撮影した画像は, 曲面展開を行っていないため, 提案手法での平面状態の再現に対する性能比較を行う対象とした.

提案手法による形状推定を用いた平面化処理を行ったものを図 10 に示す. SfM (Structure from Motion) [7][8] による形状推定を用いた平面化処理を行ったものを図 11 に示す. 図 10 を見ると, 視覚的に曲面形状による歪みが少ない画像が得られた.

ロボットビジョンにおける曲面形状上からの情報取得に関する研究

局所的な画質の低下や、鏡面反射によるコントラストの変化がみられた。SfMを用いた平面化画像である図11と比較する。図10には、僅かながら不連続な領域がある。図11には、明らかに不連続な領域がある。これは、どちらも3次元形状の推定に誤差が含まれた結果だと考えられる。ただし、この結果から提案手法の3次元形状推定はSfMを用いた推定と比べても適当な結果が得られたといえる。

図8,9に対し平面化を行った結果を図14,15を示す。この2枚から提案手法によって高画質化した画像を図16に示す。比較対象として、スキャナから読み取った画像を図12、平面の情報を撮影した画像を図13に示す。これらの画像からOCRによって識字した結果を図17に示す。OCRによる識字結果を表1に示す。図中の下線は正しく識別できた文字であることを示す。理想的な状態であるはずの図17(b),(c)でも一部誤認識があることがわかる。図8,9の画像をOCRにかけると、文字として認識されることがなかった。曲面画像を平面化した画像の認識結果である図17(d)をみると、53.9%の文字は正しく認識されていることがわかる。曲面上に存在する情報に比べ、OCRで認識できる程度まで可読性が向上したことがいえる。ただし、半分以上で認識しているものの、画質の低い左側領域で認識精度が悪くなっていることがわかる。図中の二重下線は、高画質化によって新たに認識できた文字であることを示す。図17(e)をみると図17(b)に比べて平面画像の左側領域で正しく認識された文字が増していることがわかる。誤って識字された文字は、曲面歪みを含んでいない対象から識字した結果である図17(c)でも誤って識字されていた。また、図17(c)でも誤った識字をしていた文章冒頭の“其”の文字を正しく識字した。53.9%だった識字率は提案する高画質化処理によって61.7%まで向上した。補完を行った左側領域では、高画質化前には誤識別された文字が28字あった。高画質化後の誤識別された文字は14字となった。誤識別された文字のうち50%の文字を正しく識別できるようになった。このように、高画質化した左側領域で識字結果の向上が見られた。図16右側領域での誤認識は、使用した2枚の画像でどちらも画質が低かったことが原因である。図16右側領域も補完できるように、対象の右側方向から撮影された画像を用意することで認識率を更に向上させることができると考える。

表 1: OCR の識字率

認識対象画像	識字率
スキャナで読み込んだ画像 (図 12)	94.5%
平面対象を撮影した画像 (図 13)	83.6%
提案手法：平面化处理画像 (図 14)	53.9%
提案手法：高画質化处理画像 (図 16)	61.7%



図 7: 入力画像 1

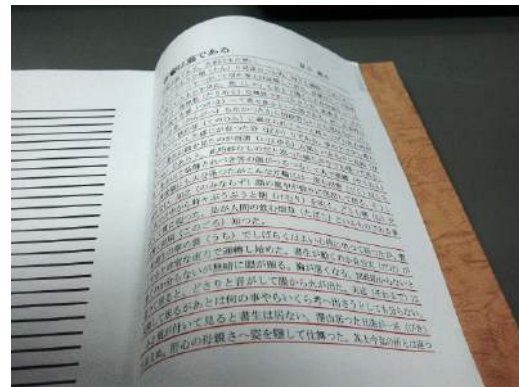


図 8: 入力画像 2

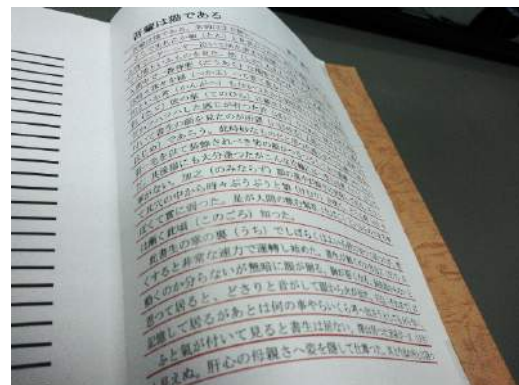


図 9: 入力画像 3



図 10: 提案手法による図 7 の平面化画像



図 11: 従来手法による図 7 の平面化画像

だ。其後猫にも大分逢つたがこんな片輪には一度も出會（でく）はした事がない。加之（のみならず）顔の眞中が餘りに突起して居る。そうして其穴の中から時々ぶうぶうと烟（けむり）を吹く。どうも咽（む）せぼくて實に弱つた。是が人間の飲む烟草（たばこ）といふものである事

図 12: スキャナによって読み取った画像

だ。其後猫にも大分逢つたがこんな片輪には一度も出會（でく）はした事がない。加之（のみならず）顔の眞中が餘りに突起して居る。そうして其穴の中から時々ぶうぶうと烟（けむり）を吹く。どうも咽（む）せぼくて實に弱つた。是が人間の飲む烟草（たばこ）といふものである事

図 13: 平面の対象を撮影した画像

だ。其後猫にも大分逢つたがこんな片輪には一度も出會（でく）はした事がない。加之（のみならず）顔の眞中が餘りに突起して居る。そうして其穴の中から時々ぶうぶうと烟（けむり）を吹く。どうも咽（む）せぼくて實に弱つた。是が人間の飲む烟草（たばこ）といふものである事

図 14: 図 8 の平面化画像

だ。其後猫にも大分逢つたがこんな片輪には一度も出會（でく）はした事がない。加之（のみならず）顔の眞中が餘りに突起して居る。そうして其穴の中から時々ぶうぶうと烟（けむり）を吹く。どうも咽（む）せぼくて實に弱つた。是が人間の飲む烟草（たばこ）といふものである事

図 15: 図 9 の平面化画像

だ。其後猫にも大分逢つたがこんな片輪には一度も出會（でく）はした事がない。加之（のみならず）顔の眞中が餘りに突起して居る。そうして其穴の中から時々ぶうぶうと烟（けむり）を吹く。どうも咽（む）せぼくて實に弱つた。是が人間の飲む烟草（たばこ）といふものである事

図 16: 提案手法による高画質化画像

ロボットビジョンにおける曲面形状上からの情報取得に関する研究

た。其後猫にも大分逢つたがこんな片輪には一度も出會(でく)はした事がない。加之(のみならず)顔の眞中が餘りに突起して居る。そうして其穴の中から時々ふうふうと烟(けむり)を吹く。どうも咽(む)せぼくて實に弱つた。是が人間の飲む烟草(たばこ)といふものである事

(a) 正解のテキスト

だ。其後猫にも大分逢つたがこんな片輪には離度も出會(でく)はした事がない。加之(のみならず)顔の眞中が餘りに突起して居る。そうして其穴の中から時々ふうふうと烟(けむり)を吹く。どうも咽(む)せぼくて實に弱つた=是が人間の飲む爛華(たばこ)といふものである事

(b) スキャナで読み込んだ画像

柱 ”其後猫にも大分違つたがこんな片籠には=度も川縁(でく)はした麦がない。加之(のみならず)翼の眞中が離りに突起して居る。そうして共穴の中から聴々ふうふうと胸(けむり)を吹く。どうも咽(む)せぼくて翼に弱うた。ご延が人蘭の飲む焔雄(たばこ)といふものである。嚼

(c) 平面上の情報を撮影した画像

’鷺賛一にも大分違つたがこんな片臍には優も幽鬱(でく)備した笠ず青ざい。加之(のみならず)幽の翼中が驚りに従忠して触る。そうして審注含一ゆから時々ふうふうと烟(けむり)を吹く。ピラ薰離ぎと蝉襲二願つた ”運が人肌の飲む鱗たばこ2一旦掘含饗g疫あ牽牽

(d) 平面化画像

、其後猫にも大分違つたがこんな片臍には 優も幽鬱 (でく) 備した
・麦がない。加之(のみならず)幽の翼中が驚りに従忠して触る。そうして共穴の中から時々ふうふうと烟(けむり)を吹く。ピラ薰離ぎと
ぼくて翼に弱うた。是が人間の飲む鱗たばこ2一旦掘含饗g疫あ牽牽

(e) 高画質化画像

図 17: OCR でのテキスト化の結果

4 まとめ

平面化における歪み補正精度の向上のため, 座標調整による補正手法を提案した. 平面化を行った画像に対して提案する手法によって歪み補正精度の向上を行った. 対象に描かれた線の直線性によって, 歪み補正手法の有効性を示した.

参考文献

- [1] 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”, 信学論誌 D-II, Vol.86, No.3, pp.409-417, 2003.
- [2] 田中 友, 鈴木 優輔, 山下 淳, 金子 透: “ステレオ計測を用いた見開き書籍画像の歪み補正”, 電子情報通信学会技術報告書, Vol.106, No.144, pp.31-36, 2006.
- [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, Vol.83, No.12, pp.2610-2621, 2000.
- [4] L.Zhang, 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.
- [5] Z.Zhang, C.L.Tan: “Correcting document image warping based on regression of curved text lines”, Document Analysis and Recognition, Proceedings.Seventh International Conference on.IEEE, 2003.
- [6] C.Wu, G.Agam: “Document image de-warping for text/graphics recognition”, SSPR2002, LNCS 2396, pp.348-357, 2002.
- [7] 満上育久: “Structure from Motion for Un-ordered Image Collections(私の研究開発ツール)”, 映像情報メディア学会誌, Vol.65, No.4, pp.479-482, 2011.
- [8] C.Tomasi, T.Kanade, “Shape and motion from image streams under orthography A factorization method”, Int.J.Comput.Vision, pp.137-154, 1992.
- [9] 中居 友弘, 黄瀬 浩一, 岩村 雅一: “特徴点の局所的配置に基づく文書画像モザイクの実験的検討”, 信学技報 (PRMU), pp.157-162, 2009.

(受理 平成 27 年 3 月 19 日)