

ステレオ計測を用いた見開き書籍画像の歪み補正

田中 友[†] 鈴木 優輔[‡] 山下 淳[†] 金子 透[†]

[†] 静岡大学工学部 〒432-8561 静岡県浜松市城北3丁目5番1号

E-mail: [†] r5545008@ipc.shizuoka.ac.jp

あらまし 本研究ではステレオカメラを用いて、見開いた分厚い書籍などの歪曲した形状を有するドキュメントをデジタル化する場合に発生する歪みを補正する方法を提案する。提案手法は、ステレオ計測により得られるドキュメント形状を基に歪みを補正する。また、ステレオ画像中の精細な部分同士を組み合わせて補正を行うことで、一方の画像のみを用いた場合よりも鮮明な補正結果が得られる。実験では、歪み補正前の画像と歪み補正後の画像について文字認識ソフトによる文字認識精度を比較し本手法の有効性を示す。

キーワード ステレオ計測, 書籍電子化, 歪み補正

Restoration of a Distorted Image of Document Using Stereo Measurement

Yuu TANAKA[†] Yusuke Suzuki[†] Atsushi YAMASHITA[†] and Toru KANEKO[†]

[†] Faculty of Engineering, Shizuoka University 3-5-1 Johoku, Hamamatsu-shi, Shizuoka, 432-8561 Japan

E-mail: [†] r5545008@ipc.shizuoka.ac.jp

Abstract In this paper, we propose a method for correcting image distortions occurring when digitizing thick documents. The method corrects geometrical distortion using the document shape acquired by stereo measurement. Moreover, it obtains a fine restoration result by combining the better detailed portions in a stereo image pair. In an experiment, we show the validity of the method by investigating character recognition accuracy.

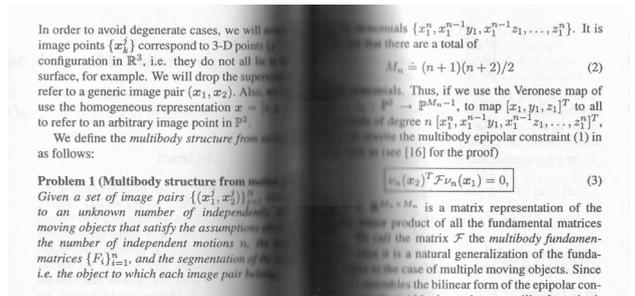
Keyword Stereo measurement, Document digitization, Restoration of distortion

1. 序論

近年、オフィス環境などにおいて資料や文書、書籍といったドキュメントを複写・デジタル化するための複写機やフラットベッドスキャナが広く普及し、必要不可欠な物となっている。さらに、フラットベッドスキャナにより画像情報としてデジタル化されたドキュメントを、文字認識ソフトを用いて電子文書化することも行われている。複写機やフラットベッドスキャナは共に、計測対象を計測面に密着させて計測を行い、ドキュメントを複写・デジタル化する装置である。そのため、見開いた状態の書籍などの立体形状物体では、ガラス面に接触しない部分において歪みや明るさの低下が生じ文字認識に支障をきたすなどの問題が起きる(図1)。これらを軽減するために平らに引き伸ばそうとして計測対象を計測面に押し付けても、分厚い書籍などでは完全には問題が解決しない場合が多い。さらに、希少な書籍などでは、無理に引き伸ばそうとすると対象を痛める恐れがあるため、引き伸ばしたり押し付けたりすることはできない。また、書籍を連続して複写・デジタル化する場合でも、ページをめくるたびに所定の位置に平らに引き伸ばして置き直す必要があ



(a) 見開いた書籍の入力



(b) 取得画像

図1 フラットベッドスキャナによる入力例

り、容易な作業ではない。

書籍画像の歪みを補正する方法としては Shape from shading を用いる方法[1][2], アクティブステレオを用いる方法[3][4][5], 画像中の文字列や書籍の輪郭を用いる方法[6][7]がある。

[1][2]の方法では、形状推定のために新たな装置を必要としないが、照明条件や対象表面の反射特性などの多くの基準データを必要とする。また、フラットベッドスキャナにより画像を取得する為、毎回所定の位置に書籍を置き直す手間が必要となる。

[3][4]や[5]の方法では、多くの基準データや、ページをめくるたびに書籍を置き直す手間を必要としない。しかし、[3]や[4]の方法では小型レンジファインダやLCD プロジェクタといった装置が必要であり、[5]の方法では形状計測を行う画像と補正を行う画像を同時に取得することができず撮影の手間が増える。

[6]や[7]の方法では、多くの基準データや、ページをめくるたびに書籍を置き直す手間を必要としない。しかし、書籍の上端部と下端部に折れや変形がないものに限られるといった制限や、自由な角度で配置された書籍に対して用いることができないといった制限があり、汎用性に乏しい。

以上より、文字認識ソフトによる電子文書化が可能な画像が取得でき、ドキュメントの形状や配置に制限がなく、多くの基準データや特殊な装置を必要としないシステムが望まれている。

多くの基準データや特殊な装置を必要としない画像補正法として、ステレオ計測を用いて歪みを補正する方法が提案されている[8]。この方法は、ステレオ計測を利用しているため、多くの基準データや特殊な装置を必要とせずに形状計測を行うことが可能である。さらに、対象物を上向きに配置できるためページをめくるたびに書籍を置きなおす手間を必要としない。しかし、書籍の配置については、書籍の綴じ目の方向が既知であるという制限があった。そのため、[8]の方法を拡張し、書籍の綴じ目の方向が未知である状況に対応可能な方法も提案された[9]。しかし、[9]の方法は、取得したステレオ画像の一方の画像のみを用いて歪みの補正を行っていた。そのため、綴じ目付近などで補正結果が不鮮明になる場合があった。

そこで、本研究では[9]の方法を拡張し、取得したステレオ画像の精細な部分同士を組み合わせ高精密な歪みの補正を行う方法を提案する。

2. 処理概要

本研究では、見開いた書籍などのドキュメントを図2に示すようなステレオカメラを用いて撮影し、取得

画像に生じる歪みを補正することで鮮明なデジタル画像を得ることを目的とする。

本研究ではドキュメントを配置する面上に世界座標系の原点をおく。対象を配置する面に対して垂直な方向にZ軸をとり、カメラの基線と平行にX軸をとる(図2)。

提案手法は以下の手順からなる。

- ①ステレオ画像を取得し、その画像からドキュメント部分を抽出する。ドキュメント部分は、あらかじめ撮影しておいた背景画像との差分をとることで抽出する。
- ②ステレオ計測を行い対象表面の3次元座標値を算出する。
- ③対象表面の3次元座標値に曲面の式をフィッティングし曲面を生成する。
- ④生成された曲面を元に、取得画像の精細な部分を用いて歪み補正を行い、対象表面が平面となるような画像を生成する。

以上の手順により歪みを補正した画像を得る。

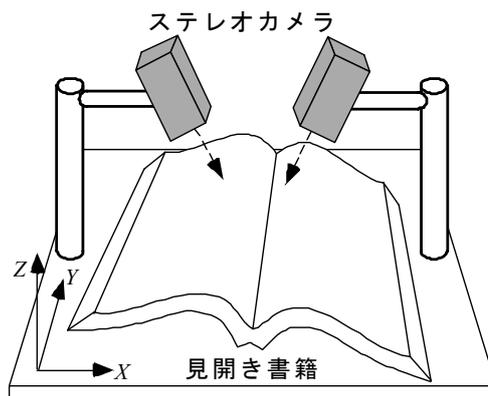


図2 装置概略

3. ステレオ計測

ステレオ計測では左右の画像間の対応点を検出することにより、その点の3次元座標値を算出することができる。

本研究では歪曲形状を有する書籍を対象としており、取得画像には明るさの変化や歪みが含まれている。そこで、明るさの変化に強いとされる正規化相互相関によるテンプレートマッチングを用いて対応点を検出する。また、対応点は正規化相互相関の相関値の2次近似によりサブピクセル精度で求める。

正規化相互相関によるテンプレートマッチングでは、テンプレート内の明るさの変化が乏しい部分で誤検出を起こしやすい。そこで、テンプレート内の明る

さの分散が小さい部分については、対応点検出を行わない。

なお、ステレオ計測においては一般的にオクルージョンが問題となるが、本研究で扱う状況では対象表面の高低差が小さいため、オクルージョンは発生しないという利点を有する。

4. 曲面生成

ステレオ計測により求めた対象表面上の点の3次元座標値から曲面を生成する。対象表面の形状を復元するためには、形状を何らかの数理モデルとして表現する必要がある。

形状を復元する方法として、対象表面をポリゴンメッシュで表現する方法[10]、NURBS曲面で表現する方法[8][9]がある。これらの方法ではメッシュ密度を高くすることで、精度良く曲面を表現できる。しかし、メッシュ密度が高いほど計算コストが高くなる。そのため、表面特性に応じてどのようにメッシュ密度を調整するかといった問題がある。また、メッシュ密度が高いほどステレオ計測結果のノイズの影響を受けやすいという問題もある。

そのため、計算コストが低く、ノイズの影響を受けにくい形状復元方法が求められている。そこで本研究では、最小二乗法を用いて多項式をステレオ計測結果へフィッティングすることにより、対象表面の復元を行う。

4.1. 変曲部の検出

対象表面には綴じ目や折れ目といった変曲部が存在しており、そのまま曲面による近似を行うと変曲部も滑らかな曲面として復元されてしまう。このため、あらかじめ変曲部を検出しておく必要がある。

変曲部では図3に示すように傾きの変化を表す角度 θ の絶対値が大きくなる。そこで、対象表面の傾きの変化角度 θ を算出し、傾きの変化角度 θ の絶対値が大きい部分を変曲部とする。

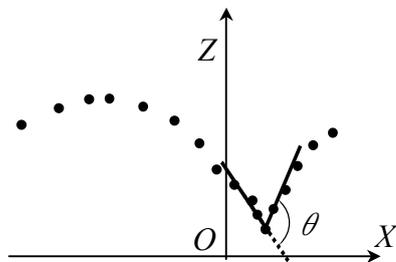


図3 変曲部の検出

ステレオ計測の結果、画像上の位置 (u,v) での3次元座標値が (x,y,z) であったとする。この時、傾きの変化

角度 θ は (u,v) の近傍の2点 (u',v) 、 (u'',v) の3次元座標値 (x',y',z') 、 (x'',y'',z'') から式(1)を用いて算出できる。

$$\theta = \tan^{-1}\left(\frac{z'-z}{x'-x}\right) - \tan^{-1}\left(\frac{z-z''}{x-x''}\right) \quad (1)$$

ここで $x'' < x < x'$ である。

しかし、近傍の2点から傾きの変化角度 θ を正確に算出することは、対応点の3次元座標値の誤差による影響のため困難である。このため、本研究では (u,v) から n 画素離れた点 $(u+n,v)$ と m 画素離れた点 $(u-m,v)$ を用いて対象表面の傾きの変化 θ を求める。さらに、計算に用いる (u,v) 、 $(u+n,v)$ 、 $(u-m,v)$ の点の3次元座標値を、各々の点の近傍の3次元座標値の平均値とすることで誤差の影響の低減を図る。

変曲部の検出後、書籍の綴じ目を検出する。 θ がある値以上となる変曲部は綴じ目である可能性がある。そこで、綴じ目を検出するための閾値を定め、 θ が閾値以上となる変曲部の点群を X - Y 平面上に投影したのに対してHough変換による直線検出を行い、綴じ目を検出する。Hough変換を用いることで綴じ目以外の点を検出結果に与える影響を抑えることができる。

綴じ目検出後、綴じ目の Z 座標値が一定となるような回転を行い、綴じ目の法線と Z 軸が平行になるように揃える。この処理は、後の画像補正処理を容易にするために行う。詳しくは、5章にて述べる。

まず、 Z - Y 平面上に綴じ目を投影し、最小二乗法を用いて綴じ目の直線を得る。次に、得られた直線の Z 座標が一定となるような X 軸回りの回転角 ϕ を求め、計測結果に回転を加える。

4.2. 表面形状の多項式表現

対象形状を多項式により表現することで3次元形状を復元する。最小二乗法により、多項式をステレオ計測結果へフィッティングすることで対象形状の復元を行う。

フィッティングする多項式は式(2)とする。

$$z = \sum_{i=0}^n \sum_{j=0}^m a_{i,j} x^i y^j \quad (2)$$

最小二乗法を用いて $a_{i,j}$ を求めれば、対象形状を多項式により表現できる。

紙面形状は、綴じ目を境に2つの曲面から成っている。そこで、綴じ目での曲面の連結について考える。

現在得られている計測結果は、4.1節の処理により綴じ目部分で Z 座標が一定である。ここで、綴じ目部分で $x=0$ となるように Z 軸回りの回転 ψ と X 軸方向の並進 T を計測結果に加える。すると、綴じ目部分に

において式(2)は、 $i=0$ 以外の項が0となる。また、綴じ目部分のZ座標は一定であるため、式(3)と書ける。

$$z = \sum_{j=0}^m a_{0,j} x^0 y^j = \text{const.} \quad (3)$$

ここで、 $a_{0,j}=0$ ($1 \leq j \leq m$) とおけば、 $a_{0,0}$ が定まり、綴じ目でZ座標が一定となる曲面は式(4)と書ける。

$$z = a_{0,0} + \sum_{i=1}^n \sum_{j=0}^m a_{i,j} x^i y^j \quad (4)$$

式(4)を用い、2つの曲面について最小二乗法を行えば、綴じ目を境とする2つの曲面が得られる。

5. 補正画像生成

多項式によりフィッティングした書籍形状をもとに、取得画像から歪みを補正した画像（補正画像）を生成する。本研究では、曲面を平面となるように引き伸ばすことで、取得画像の歪みの補正を行う。曲面上の距離を算出し、この距離をもとに取得画像を引き伸ばすことで、歪みが補正された補正画像を生成する。

また、補正画像の生成時には、ステレオ画像中の精細な部分同士を組み合わせて高精密な歪み補正画像を生成する。

書籍を平面へと引き伸ばすには、綴じ目と直行する方向に曲面を引き伸ばせばよい。4.2節で得られた曲面は、綴じ目部分で $x=0$ 、Z座標一定となる曲面である。そのため、曲面はX軸に沿って引き伸ばせばよい。

5.1. 精細な画像の選定

ステレオ画像中の精細な部分を判別するために、両カメラの傾きと紙面の各位置での傾きを比較する。

カメラの傾きと紙面の傾きの差が小さいほど、カメラは紙面を精細に撮影できる。そのため、カメラの傾きと紙面の各位置での傾きの差が小さくなる方のカメラの画像をより精細であると判別する（図4）。

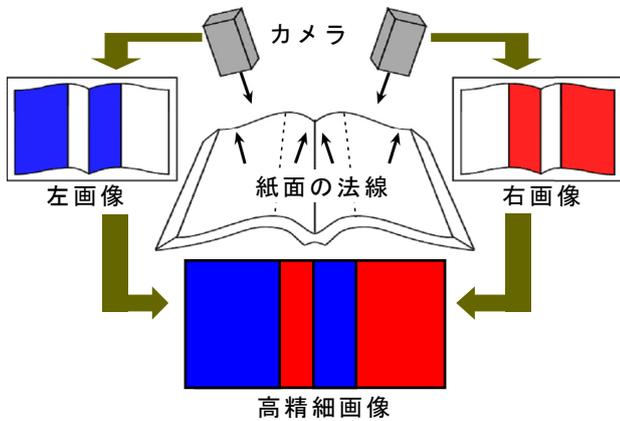


図4 精細な画像の選定

カメラの視線方向はカメラキャリブレーションにより得られる。紙面の傾き α は曲面上である距離だけ離れた2点の3次元座標 (x,y,z) 、 (x',y',z') から、式(5)により求めることができる。

$$\alpha = \tan^{-1} \left(\frac{z'-z}{x'-x} \right) \quad (5)$$

ここで左右のカメラの傾きを ζ_L, ζ_R とすれば、左右のカメラに対する紙面の傾き φ_L, φ_R は、式(6)(7)により求まる。

$$\varphi_L = |\alpha - \zeta_L| \quad (6)$$

$$\varphi_R = |\alpha - \zeta_R| \quad (7)$$

式(6)(7)により算出した、カメラに対する紙面の傾きが小さい方の画像を精細な画像として選定し、これらを用いて補正画像を生成する。

5.2. 補正画像の生成

曲面上の長さを算出する。書籍表面の3次元形状は、多項式により表現されている。実際には、書籍表面は多くの点の集合として表現されており、曲面の長さを算出するためには、点群間の距離を求める必要がある。Y座標が一定での曲線を点群間の距離を保ったまま、 $Z=Z_0$ にすることで直線にする（図5）。X-Z平面における点群間の距離 L_i は、座標値を (x_i, z_i) 、 (x_{i+1}, z_{i+1}) とすると式(8)により求まる。

$$L_i = \sqrt{(z_{i+1} - z_i)^2 + (x_{i+1} - x_i)^2} \quad (8)$$

次に、曲面を引き伸ばした場合の位置 x'_n を求める。引き伸ばす前の位置が x_n である点の引き伸ばし後の位置 x'_n は、式(9)により求めることができる。

$$\begin{aligned} x'_n &= \sum_{i=0}^{n-1} L_i \\ &= \sum_{i=0}^{n-1} \sqrt{(z_{i+1} - z_i)^2 + (x_{i+1} - x_i)^2} \end{aligned} \quad (9)$$

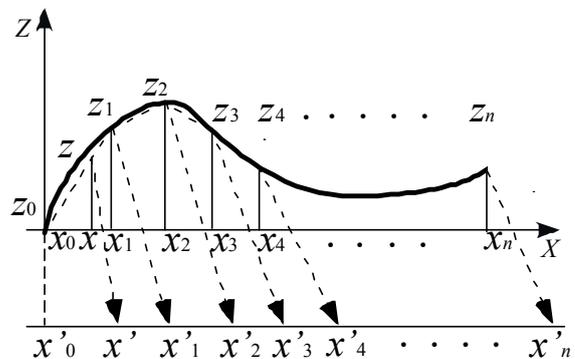


図5 歪み補正

この位置 x'_n に、引き伸ばす前の取得画像の画素値を当てはめることで、引き伸ばした画像を得ることができる。

これを Y の値を変えて順次行うことで歪みが補正された補正画像を生成することができる。

引き伸ばした結果、 x'_{i-1} と x'_i の間が大きく開いてしまい補正画像に孔ができることがある。そのため、 $x'_{i-1} < x' < x'_i$ である x' の画像 (u, v) との対応を求める必要がある。 x' の画像 (u, v) との対応は式(10)~(15)により求めることができる。

$$x = \frac{(x_i - x_{i-1})(x' - x'_{i-1})}{x'_i - x'_{i-1}} + x_{i-1} \quad (10)$$

$$z = \frac{(z_i - z_{i-1})(x' - x'_{i-1})}{x'_i - x'_{i-1}} + z_{i-1} \quad (11)$$

$$x_\psi = (x - T) \cos(-\psi) - Y \sin(-\psi) \quad (12)$$

$$Y_\psi = (x - T) \sin(-\psi) + Y \cos(-\psi) \quad (13)$$

$$u = \frac{f x_\psi}{Y_\psi \sin(-\phi) + z \cos(-\phi)} \quad (14)$$

$$v = \frac{f(Y_\psi \cos(-\phi) - z \sin(-\phi))}{Y_\psi \sin(-\phi) + z \cos(-\phi)} \quad (15)$$

ここで T は曲面生成時に行った X 軸方向の並進、 ψ は曲面生成時に行った Z 軸回りの回転、 f はカメラの像距離 (カメラのレンズ中心と撮像面の距離)、 ϕ は綴じ目を検出した後に行った回転の X 軸回りの回転角である。

6. 実験

本手法の有効性を検証するため、見開いた状態の書籍に対して、形状計測・形状復元・画像補正を行った。また、歪み補正前の画像、左画像のみを引き伸ばして歪みを補正した画像、左右画像の精細な部分同士を組み合わせ歪みの補正を行った画像、平面形状紙面をスキャナで取り込んだ画像それぞれについて文字認識ソフトによる文字認識精度を比較した。

図 6 に実験に用いたステレオ画像を示す。画像のサイズは 2048×1536 画素であり、カメラの基線長は 137.8mm である。図 7 にステレオ計測結果を示し、図 8 に形状復元結果を示す。今回の実験では 6 次の多項式でフィッティングした。

図 9 に歪み補正結果を示す。(a)は歪み補正結果である。(b)は歪み補正前の図、(c)は左画像のみを引き伸ばして歪みの補正を行った図、(d)は左右画像を組み合わせ歪みの補正を行った図である。(d)の図の中央付近

に左右画像の境界があるがほとんど目立たない。

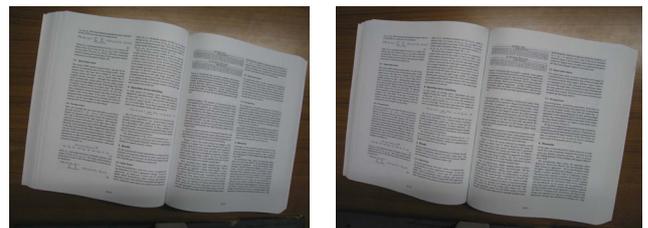
表 1 に平面形状紙面をスキャナで取り込んだ画像、歪み補正前の画像、左画像のみを引き伸ばして歪みを補正した画像、左右画像の精細な部分同士を組み合わせ歪みの補正を行った画像それぞれについて文字認識を行った結果を示す。平面形状紙面の画像では、2854 文字中の 2819 文字を正しく認識した。補正前の画像では、2854 文字中の 2566 文字を正しく認識し、左画像のみで補正した画像では、2854 文字中の 2700 文字を正しく認識した。また、左右画像を用いて補正した画像では、2854 文字中の 2797 文字を正しく認識した。本手法を用いることで、補正前の場合より正答率が 8.1%向上している。また、平面形状紙面の場合と比較しても 0.7%の差しかない。この結果から、本手法が歪みの補正に有効であることが判る。

7. 結論

本論文において、書籍などの立体形状を有するドキュメントをデジタル化する方法を提案した。綴じ目部分を検出することによりドキュメント配置を検出し、立形状物体に対応するため、ステレオ計測と多項式曲面表現を利用して形状復元を行い、左右画像の精細な部分同士を組み合わせ歪みの補正を行った。

実験では、ドキュメント配置の検出に成功し、歪曲形状により生じる画像上の歪みの補正にも成功した。また、本手法により文字認識ソフトの文字認識の結果が向上した。以上のことより、本手法が書籍などの歪曲形状を有するドキュメントの電子化に対して有効であることが示された。

今後は、多項式の最適な次数の自動的な設定方法などを検討する必要がある。



(a) 左画像 (b) 右画像

図 6 ステレオ画像

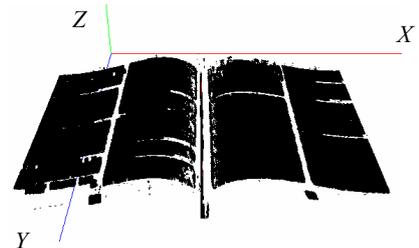


図 7 計測結果

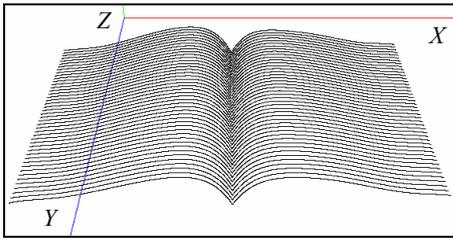
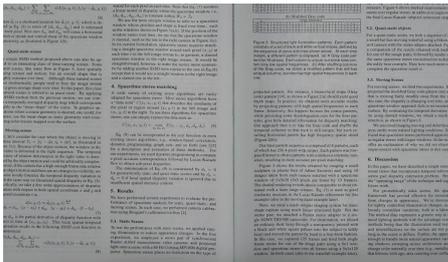


図 8 多項式曲面表現



(a) 補正結果

Eq. (9) can be incorporated as the cost function in most existing stereo algorithms, e.g., window-based correlation, dynamic programming, graph cuts, and so forth (see [15] for a description and evaluation of these methods). For our experiments, we used dynamic programming to compute a pixel-accurate correspondence followed by Lucas-Kanade flow to obtain sub-pixel disparities.

(b) 歪み補正前

Eq. (9) can be incorporated as the cost function in most existing stereo algorithms, e.g., window-based correlation, dynamic programming, graph cuts, and so forth (see [15] for a description and evaluation of these methods). For our experiments, we used dynamic programming to compute a pixel-accurate correspondence followed by Lucas-Kanade flow to obtain sub-pixel disparities.

(c) 左画像のみ使用した補正結果

Eq. (9) can be incorporated as the cost function in most existing stereo algorithms, e.g., window-based correlation, dynamic programming, graph cuts, and so forth (see [15] for a description and evaluation of these methods). For our experiments, we used dynamic programming to compute a pixel-accurate correspondence followed by Lucas-Kanade flow to obtain sub-pixel disparities.

(d) 左右画像を用いた補正結果

図 9 補正結果

表 1 文字認識結果

使用画像	文字数	正答文字数	正答率
平面形状紙面	2854	2819	98.7%
補正前の左画像	2854	2566	89.9%
左画像のみで補正した画像	2854	2700	94.6%
左右画像を用いて補正した画像	2854	2797	98.0%

参考文献

- [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(II)ー相互反射を考慮した近接光源下の Shape from Shadingー”, 電子情報通信学会論文誌 D-II, Vol.J78-D-II, No.2, pp.311-320, 1995.
- [2] Hiroyuki Ukida, Katsunobu Konisho, Toshikazu Wada and Takashi Matsuyama: “Recovering Shape of Unfolded Book Surface from a Scanner Image using Eigenspace Method”, Proceedings of IAPR Workshop on Machine Vision Applications, pp.463-466, 2000.
- [3] 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”, 電子情報通信学会論文誌 D-II, Vol.J86-D-II, No.3, pp.409-417, 2003.
- [4] Michel S.Brown and W.Brent Seals: “Image Restoration of Arbitrarily Warped Documents”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.26, No.10, pp.1295-1306, 2004.
- [5] 佐藤康弘, 長谷川雄史, 北澤智文, 青木伸, 北口貴史: “デジタルカメラを用いた 2D/3D デスクトップ画像入力システムの開発”, 第 8 回画像センシングシンポジウム講演論文集, pp.185-190, 2002.
- [6] 包躍, 吉開敬治: “画像処理を用いた書籍の歪み補正”, 映像情報メディア学会技術報告, Vol.26, No.54, pp.13-16, 2002.
- [7] Huaigu Cao, Xiaoqing Ding and Changsong Liu: “A Cylindrical Surface Model to Rectify the Bound Document Image”, Proceedings of 9th IEEE International Conference on Computer Vision, pp.228-233, 2003.
- [8] Atsushi Yamashita, Atsushi Kwarago, Toru Kaneko and Kenjiro T.Miura: “Shape Reconstruction and Image Restoration for Non-Flat Surfaces of Documents with a Stereo Vision System”, Proceedings of 17th International Conference on Pattern Recognition, Vol.1, pp.482-485, 2004.
- [9] 田中友, 山下淳, 金子透: “ステレオ計測と NURBS 曲面表現を利用した歪曲形状書籍画像の歪み補正”, ビジョン技術の実利用ワークショップ講演論文集, pp.251-256, 2005.
- [10] Maurizio Pilu: “Undoing Paper Curl Distortion Using Applicable Surfaces”, Proceedings of 8th IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol.1, pp.482-485, 2004.