
製本原稿スキャン画像の歪み補正技術

Distortion Correction Technique for Scanned Book Images

荒木 禎史*

Tadashi ARAKI

関 海克

Haike GUAN

小島 啓嗣

Keiji KOJIMA

要 旨

製本原稿をフラットベッドスキャナーで読み取った画像は、綴じ部付近の形状が歪み、加えて輝度の低下や文字ボケの発生が見られる。我々はこのような製本原稿スキャン画像の劣化をきれいに補正する技術を開発した。そこでは、特別な機器を用いることなく、製本原稿画像のページ外形や、文字、罫線等、画像中の情報のみを利用して形状、輝度、文字ボケの補正ができる。本方式にて補正した画像を主観評価した結果、ほとんど歪みが気にならない程度に補正された。さらに、補正前後の原稿画像のOCR認識結果を比較すると、日本語文献で約9.3ポイントの認識率向上が見られた。

Abstract

Book images scanned by a flat bed scanner have shape distortion, shading, and blurring around the book center. We have developed a technique to correct such distortion. It can correct the scanned book image by using the information in the image itself such as page edges, characters, and ruled lines without any dedicated book-scanner appliance. The subjective test suggests that the distortion of the corrected images is not a problem and the OCR rate is improved 9.3 point between the original images and the corrected images.

* ソフトウェア研究開発本部 画像アプライアンス研究所
Image Appliance Lab, Software R&D Group

1. はじめに

製本原稿をフラットベッド型のスキャナーで取り込む場合、原稿の綴じ部付近がスキャナーのコンタクト面から浮き上がることで、画像の綴じ部付近の形状が歪み、かつ、黒い影と文字のボケが生じる。このような劣化は単に見栄えが悪いだけでなく、綴じ部付近の文字や図表が読みにくくなり、また、OCRを施した場合の認識率も低下してしまう。そこで、劣化を補正するための技術がこれまでもいくつか提案されてきた[1][2][3][4]。

これらのうち[1]、[2]、[3]の方法は、何らかの方法で製本原稿の断面形状（スキャナーのコンタクト面と原稿のページ表面間の距離）を求めた後に補正する。例えば、Wada他の研究[1]では、スキャナーの光源と製本原稿の表面の反射特性等を基に光学モデルを構成し、反復型の非線形最小2乗法によって製本原稿の断面形状を算出して補正している。大削他の研究[2]では、製本原稿スキャン画像のページ外形の形状とスキャナーのパラメーターを利用して断面形状を算出し、歪んだ形状を復元するための伸張率を求めて画像を伸張/復元している。浮田他の研究[3]では、スキャン画像の陰影パターンと製本原稿の断面形状の組み合わせを予め学習しておき、固有空間法を用いて、新たに入力したスキャン画像の陰影情報から断面形状を推定して伸張率を求めている。

ところが、これらには、特定のスキャナーに依存し、かつ、製本原稿の綴じ部とスキャンの主走査方向が垂直である場合に補正できない[1][2][3]、処理時間が長くなる[1]、ページ外形がスキャン画像中に無い場合は補正できない[2]、学習用のデータを予め取得する必要がある[3]、等々の問題点があり、実用面では大きな制約があった。

一方、宮崎他の研究[4]では、スキャン画像中の文字の縦成分/横成分を抽出し、それらの間隔の変化を利用して画像の伸張率を求めている。この方法では上記の諸問題点は減じられているものの、対象が主にアルファベットで書かれた原稿であるため、日本語や中国語の原稿には不適であった。

また、製本原稿スキャン画像を補正する機能をもった製品としては、ミノルタの「BookPro7000D」がある[5]。これは、原稿台の上に製本原稿を上向きに載置し、その上方から原稿画像を読み取ると同時に、側方から原稿の断面形状を読み取って補正するスキャナーである。このスキャナーは特殊

かつ高価であるばかりでなく、外光の影響により断面形状の認識が不安定になるという問題点がある。

さらに、デジタルカメラで撮影した製本原稿画像を補正するソフトウェアとして、富士通の「表OCR for ExcelV4.0」他がある[6]。これは、上向きに載置した製本原稿を斜め上方から撮影した後、補正するものであり、一般のスキャナーにより入力された画像を対象とするものではない。

本論文では、これらの問題点を解決できる、より汎用的かつ実用的な製本原稿スキャン画像の歪み補正技術を提案する。その方式の基本的なコンセプト（特徴）を以下に示す。

- ・汎用のフラットベッドスキャナーでスキャンした製本原稿画像を対象とする
- ・ページ外形の存在を必要としない
- ・日本語/英語ともに対応可能
- ・製本原稿の綴じ部とスキャンの主走査方向が平行/垂直である場合の両方に対応可能
- ・処理時間が短い

本方式の概要は、スキャン画像中のページ外形以外に罫線や文字（行）を抽出し、その形状の歪みから画像の補正（伸張）を行うというものである。補正前の製本原稿スキャン画像と本方式により補正した画像とを、主観評価やOCR認識率によって比較した結果、本方式の有効性が確認でき、かつ、処理時間も実用レベルであることがわかった。

以下、第2章で製本原稿スキャン画像の特徴、第3章で提案方式の詳細、第4章で主観評価、OCR認識率評価および処理時間評価の結果と考察、第5章でまとめと今後の展開をそれぞれ述べる。

2. 製本原稿スキャン画像の特徴

2-1 画像の劣化

前述のように、劣化の原因は製本原稿の綴じ部がスキャナーのコンタクト面から浮き上がることにある。そのため、スキャン画像にはFig.1に示すような劣化が生じる。即ち、

- ① 形状の歪み
- ② 影
- ③ 文字のボケ



Fig.1 Distorted Scanned Book Image (1).

である。しかも、形状の劣化は製本原稿の綴じ部とスキャンの主走査方向の位置関係によって異なる。以下、両者が平行および垂直に位置する場合のそれぞれについて説明する。

2-2 綴じ部と主走査方向が平行な場合

Fig.2により説明する。Fig.2の左上に綴じ部と主走査方向が平行となるように製本原稿をコンタクト面上に載置した状態を上方からみた様子を示す（ここで、「光軸の軌跡」とは、スキャナーレンズの光軸の移動方向の軌跡である）。左下の図はこれを手前から見た状態であり、さらにそれを左上図の

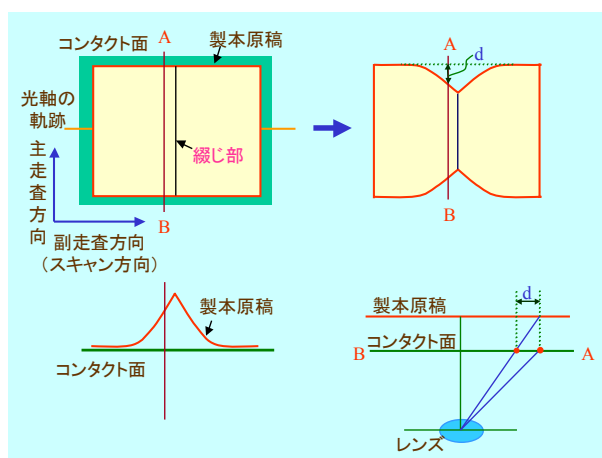


Fig.2 Parallel Case.

ラインABで切った断面図が右下の図である。製本原稿面がコンタクト面から浮き上がっているため、原稿のページ外形が本来の位置より距離dだけ主走査方向の内側に食い込んで投影されてしまう。その結果、右上図に示すような歪んだ形状の画像が出力されるのである。加えて、綴じ部の浮き上がった部分では、副走査方向にも画像が縮んで投影される。先のFig.1の例は、「綴じ部と主走査方向が平行な場合」である。

2-3 綴じ部と主走査方向が垂直な場合

Fig.3により説明する。Fig.3の左上は綴じ部と主走査方向が垂直の場合を上方から見た様子である。ここではラインABでの断面図のみを右下に示している。

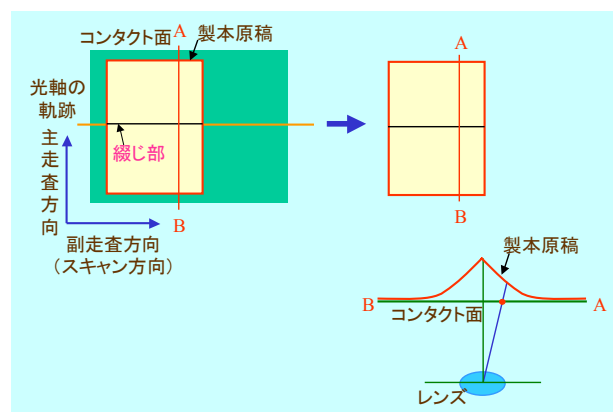


Fig.3 Perpendicular Case.

Fig.2の場合と異なり、ページ外形が「副走査方向の」内側へ食い込んで投影されることはなく、単に、綴じ部の浮き上がった部分で「主走査方向に」画像が縮むだけである。このような出力画像の例をFig.4に示す。

3. 提案する補正方式

3-1 概要

補正は大きく「形状補正」、「輝度補正」、「文字ボケ補正」からなるが、特に重要で従来方式と異なるのが形状補正であり、本稿ではこれを主体に述べる。また、以下の説明では、「綴じ部と主走査方向が平行」な場合で、かつ、多値の



Fig.4 Distorted Scanned Book Image (2).

スキャン画像を補正することを前提とする（垂直な場合や二値の場合について述べる際はその旨を注記する）。

形状補正では、スキャン画像中のページ外形や罫線、文字（行）を抽出して画像を補正する。主走査方向については、ページ外形＞罫線＞文字行、をこの優先順位で抽出し、その歪み形状を基に画像の伸張率を求めて補正（伸張）する。ここでは補正精度を高めるため、より安定して抽出できるものの優先順位を高くしている。そのため、まず、ページ外形や罫線の存在や文字行の縦横の別を判定する（スキャン画像が二値の場合は、綴じ部付近でページ外形が不安定なので罫線と文字行のみ利用する）。

副走査方向は文字外接矩形のアスペクト比（横の長さ/縦の長さ）を利用する。綴じ部付近の矩形のアスペクト比が平坦部分のそれと等しくなるように補正（伸張）する。

さらにこれらの前処理として、綴じ部の抽出処理とスキュー補正処理がある。以上の形状補正の手順をFig.5に示す。以下の節で補正処理の詳細を説明する。

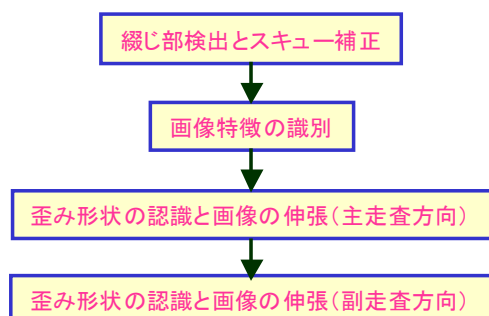


Fig.5 Shape Correction Process.

3-2 形状の補正

3-2-1 綴じ部抽出とスキュー補正

主走査方向の各ライン上での最大輝度値の中で、副走査方向において最小値を与える位置を最小輝度の代表位置として決定する（Fig.6）。

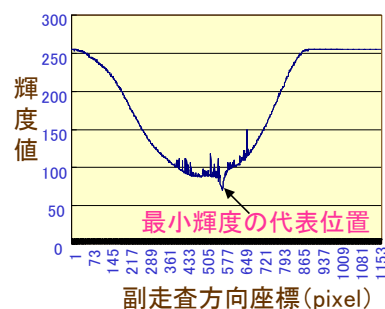


Fig.6 Representative Minimum Intensity Point.

次に画像を副走査方向に貫く中心線との交点を起点（Fig.7の点O）とし、画像の上方と下方の各方向に向かって、各主走査方向位置における最小輝度位置を追跡し、その軌跡（Fig.7の矢印）を綴じ部とする。さらに、求めた軌跡と主走査方向のなす角度（Fig.7の θ ）をスキュー角度とし、これを基にスキュー補正する。

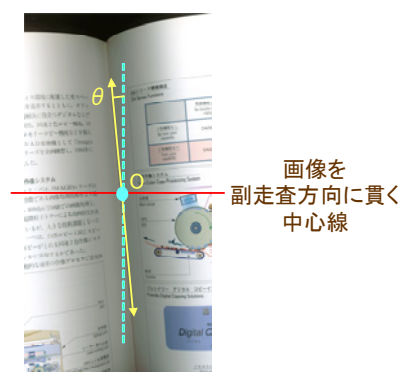


Fig.7 Search for the Minimum Intensity Location.

3-2-2 画像特徴の識別

画像を適応二値化し、主走査方向に対する黒画素数ヒストグラムを、綴じ部の左右でそれぞれ構成する（Fig.8）。

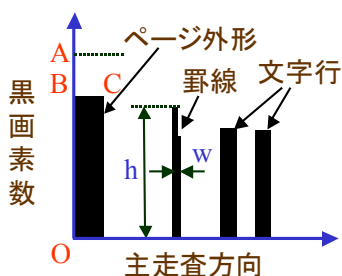


Fig.8 Histogram of the Number of Black Pixels.

Fig.1の例のようにページ外形が存在する場合、画像の上端もしくは下部に黒い帯が現れるので、画像の上端もしくは下端付近のヒストグラム値は大きくなる。Fig.8でAOは綴じ部から画像の左側もしくは右端までの距離で、BOはヒストグラム縦棒の高さである。その比率BO/AOが予め定めた閾値より大きければページ外形が存在すると判断する。

罫線が存在する場合は、ヒストグラムに細いピークが現れる。そこで、ピークの高さ h とその半分の位置での幅 w を求める (Fig.8)。 h が予め設定した閾値より高く、かつ、 w が予め設定した別の閾値より小さいピークの中で、 h が最も大きいピークの位置を罫線の候補位置とする。

文字行の縦行/横行の判別は、主走査方向と副走査方向のそれぞれに対する黒画素/白画素の反転数のヒストグラム値の分散を比較する。主走査方向に対するヒストグラムの分散のほうが大きい場合は横行、逆の場合は縦行と判別する。

3-2-3 歪み形状の認識と画像の伸張 (主走査方向)

綴じ部の左右それぞれの画像上部と下部の計4箇所から、ページ外形もしくは罫線、文字行を抽出する。その際、3-2-2節においてページ外形があると判断された場合は、それを最優先で抽出する。ページ外形が無く、かつ、罫線候補が存在すると判定された場合は罫線抽出を行うが、途中で切断点が多い場合は真の罫線ではないと判断して抽出を中止する。ページ外形も罫線も無いと判断された場合 (と罫線抽出を中止した場合) は文字行を抽出する。その際、文字行が縦行か横行かによって抽出方法を変える。以下にこれを詳述する。

[ページ外形の抽出]

ページ外形の抽出は、適応二値化した画像ではなく多値スキャン画像に対して行う (Fig.9)。入力画像上端 (もしくは下端) から、主走査方向に画像を探索し、輝度値が大きく

変化する位置をページ外形とみなす。その際、綴じ部付近に近づくにつれてページ外形の内外の輝度値の差が小さくなるので、単一の閾値によらず適応的に輝度変化を検出する。

ページ外形を抽出したら、その直線部分の近似直線をハフ変換を利用して求める。Fig.9に近似直線を求めた例を示す。



Fig.9 Detection of the Page Edge.

[罫線の抽出]

罫線の抽出は適応二値化した画像で行う (Fig.10)。3-2-2節で求めた罫線の候補位置 (主走査方向) と、副走査方向の予め定めた位置の交点を探索開始点とし、画像の左右方向に黒画素を探索していき、罫線の位置を求める。切断点 (黒画素が途切れた点) が予め定めた値を超えた場合は、前述のように罫線でないと判断して探索を中止する。

罫線抽出後、ここでもハフ変換により近似直線を求める。Fig.10のAB部分は、近似直線を画像左端に延長したものである。逆に綴じ部付近のCD部分については、罫線の多項式近似曲線を延長して補う。

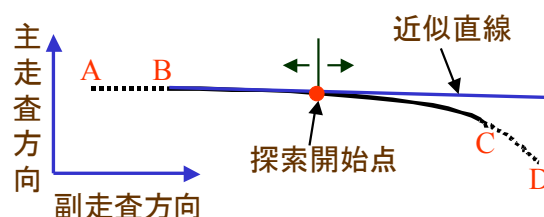


Fig.10 Detection of the Ruled Line.

[横文字行の抽出]

これも適応二値化画像に対して行う (Fig.11)。まず、黒画素の連結性を基に文字の外接矩形を求め、次に矩形同士的位置関係を基に同一行に含まれる可能性の高い矩形を副走査方向に統合したものを文字行とする。一般に文字行は複数個存在するが、その中で、綴じ部付近まで食い込み、長く、かつ、画像の上端もしくは下端に近いものを優先的に選択する。

適当な文字行を選択した後、それを構成する各文字外接矩形の中心位置を基に、ハフ変換による近似直線と、多項式近似曲線を求める。

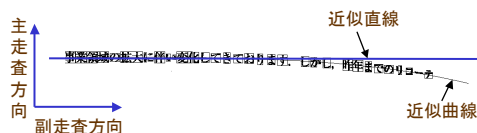


Fig.11 Detection of the Horizontal Character Line.

〔縦文字行の行頭/行末の抽出〕

縦文字行の場合は、適応二値化画像から縦文字行の外接矩形を抽出し、その行頭（もしくは行末）を副走査方向に連結する近似曲線を求める（Fig.12）．なるべく多くの行頭（行末）を利用する方が近似曲線が安定するので、Fig.12に示すように、副走査方向の短冊領域の中で、最も多くの行頭（行末）を含む領域に着目する（Fig.12の α ）．その中で最も画像左端に近い行（a）から始めて、綴じ部へ向かって行頭位置が近接した行を選択していき（Fig.12の赤行），その行頭位置を基に多項式近似曲線，および、ハフ変換による近似直線を求める．

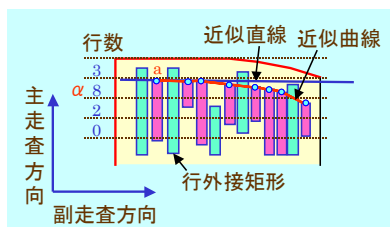


Fig.12 Detection of the Top of the Vertical Character Line.

〔主走査方向の画像の伸張〕

以上のように抽出したページ外形、罫線、文字行、もしくはその近似曲線が、ハフ変換により算出した近似直線に一致するように、多値スキャン画像を主走査方向に伸張する（Fig.13）．即ち、副走査方向の位置 x において点 P が P' に、点 Q が Q' に、それぞれ補正される。

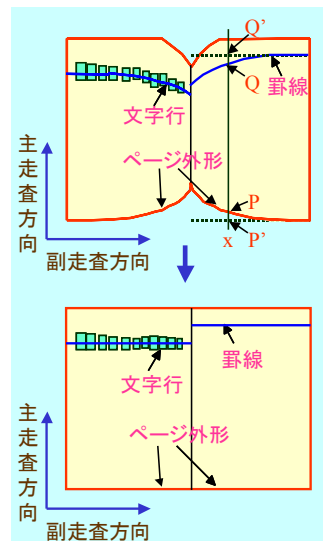


Fig.13 Expansion in the Main-Scanning Direction.

3-2-4 歪み形状の認識と画像の伸張（副走査方向）

主走査方向に伸張した適応二値化画像から、文字外接矩形を抽出し、そのアスペクト比を求める．そして、主走査方向の各短冊領域に含まれる矩形のアスペクト比の平均が等しくなるように、多値スキャン画像を副走査方向に伸張する（Fig.14）．

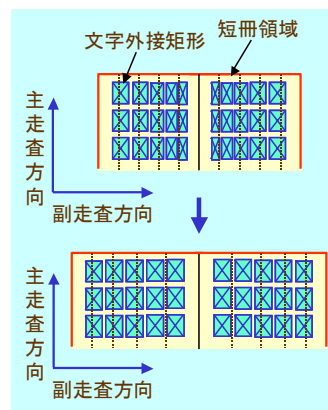


Fig.14 Expansion in the Sub-Scanning Direction.

なお、「綴じ部と主走査方向が垂直」の状態でスキャンした画像（Fig.4）の場合は、Fig.14の主走査/副走査方向を読み替え、これと全く同様の方法により形状の補正ができる（この場合は3-2-3節に相当する補正は不要）．

3-3 輝度の補正

本方式では、スキャン画像の地肌は白と仮定している。そこで、主走査方向のライン毎に、ライン上の最大輝度値が白となる比率を各輝度値に乗じて輝度を補正する (Fig.15)。

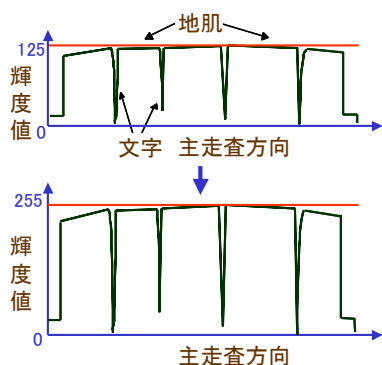


Fig.15 Intensity Correction.

3-4 文字ボケの補正

文字のボケはエッジ強調フィルタ (アンシャープフィルタ) を用いて補正する。製本原稿のコンタクト面からの浮きの高さが高いほどボケが大きくなる。そこで、3-2-3節にて求めた近似直線と近似曲線の差が大きいほど、フィルタの強度を強くする。Fig.16に文字ボケ補正の例を示す。

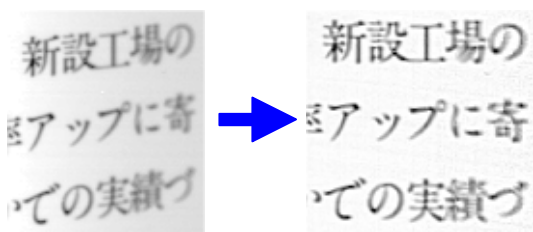


Fig.16 Blur Correction.

4. 補正結果と評価

4-1 補正結果

製本原稿スキャン画像を、3章で述べた方式を利用して補正した例を以下に示す。これらは8bitグレースケールもしくは24bitフルカラーの画像で、Fig.17~Fig.19は「綴じ部と主

走査方向が平行」の場合で、Fig.17に横書き画像、Fig.18に縦書き画像、Fig.19に罫線が存在する画像の例をそれぞれ示す。Fig.20は「綴じ部と主走査方向が垂直」の場合である。



Fig.17 Image Example with Horizontal Lines.

Fig.18 Image Example with Vertical Lines.

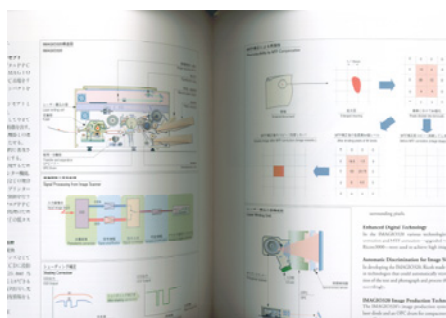


Fig.19 Image Example with Ruled Lines.

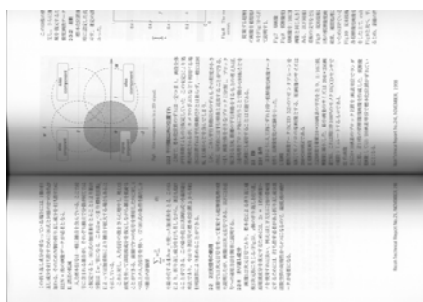


Fig.20 Image Example in the Perpendicular Case.

4-2 主観評価

製本原稿を、200dpi、8bitグレースケールでスキャンした補正前画像と、本方式にて補正した補正後画像を目視で比較し、下記の採点基準で主観評価した（小数第1位まで採点）。

- 5.0： 歪みが完全に補正されている
- 4.0： やや歪みが残っているが気にならない
- 3.0： 歪みが気になるが元画像よりは改善されている
- 2.0： 元画像と同程度の歪みである
- 1.0： 元画像よりも歪みが大きく不自然である

評価の対象画像は次のとおりである。

- ・綴じ部と主走査方向が平行な場合

日本語横書き7枚、日本語縦書き10枚、英語6枚

- ・綴じ部と主走査方向が垂直な場合

日本語横書き8枚、日本語縦書き5枚、英語6枚

なお、評価は異なる時期に分けて実施したので、全ての画像に対して同一の評価者ではない。平行で日本語（横）と英語の場合が10人、平行で日本語（縦）が13人、垂直の場合が7人である。Table.1にそれぞれの評価結果を示す。

Table.1 Subjective Test Result

| 綴じ部と主走査方向 が平行な場合 | 主観評価値 (平均) | 綴じ部と主走査方向 が垂直な場合 | 主観評価値 (平均) |
|---------------------|---------------|---------------------|---------------|
| 日本語(横) 7枚 | 4.13 | 日本語(横) 8枚 | 4.26 |
| 日本語(縦) 10枚 | 3.69 | 日本語(縦) 5枚 | 4.23 |
| 英語 6枚 | 3.99 | 英語 6枚 | 3.91 |

平行な場合は、日本語（横）と英語では平均的に歪みが気にならない程度に補正されているが、縦書きの場合はやや評価値が落ちる。縦行の適切な行頭/行末を抽出するのは横行の抽出よりも一般に困難なため、補正精度が若干劣り、評価値も下がったと考えられる。また、日本語（横）よりも英語が劣る理由は、アルファベットは文字ごとにアスペクト比が変化すると、隣接文字が接触するため明確な文字外接矩形が切り出しにくいためであると考えられる。

垂直な場合は、行抽出を行わずに文字外接矩形のみで補正するため、日本語の縦横の差はほとんど無く、いずれも高い値を示している。英語の評価値が低いのは、平行な場合と同様の理由であると考えられる。

4-3 OCR認識率評価

リコーシステム開発「読取物語EX V4」を用いて、補正前後のOCR認識率を比較した。対象画像は、主観評価にて「平行、かつ、日本語（横）と英語」に使用したものと同一である。認識率以外に、下記の定義による「改善率」や「効率」も算出した。

補正前（後）認識率 = (補正前（後）正認識文字数) / (全文字数)

改善率 = (補正前誤認識文字中での補正後正認識文字数) / (補正前誤認識文字数)

効率 = (補正前誤認識文字中での補正後正認識文字数) / (補正前後での認識の正誤逆転文字数)

Table.2に評価結果を示す。

Table.2 OCR Result

| 綴じ部と主走査方向 が平行な場合 | OCR認識率 | | 改善率 | 効率 |
|---------------------|--------|------|------|------|
| | 補正前 | 補正後 | | |
| 日本語(横) 7枚 | 84.7 | 94.0 | 68.7 | 89.7 |
| 英語 6枚 | 94.6 | 98.7 | 81.0 | 94.0 |

日英ともに、補正の効果が現れ、OCR認識率が上昇している。特に効率はいずれも90%前後の高い値を示している。Fig.21にOCRの認識改善例を示す（赤字は誤認識文字）。

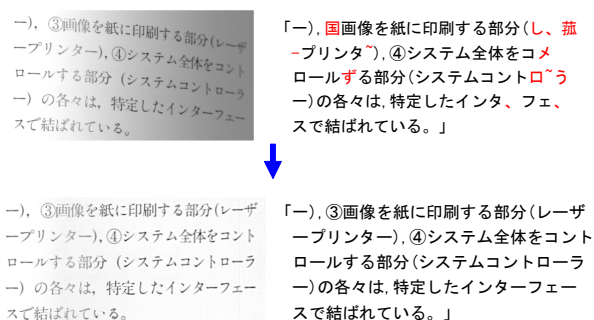


Fig.21 Example of the OCR Improvement

4-4 処理時間評価

処理時間を、CPU: Pentium 4, 2.26GHz, メモリ: 512MB, OS: Windows2000のPCで測定した。A4, 200dpi, 8bit グレースケールのスキャン画像を対象とし、スキュー補正したものとし、しないものを5画像ずつ選択した（スキュー角度が0.2°以上と認識された場合のみスキュー補正した）。さらに、これ

Table.3 Processing Time Result

| 解像度 (dpi) | 平均処理時間(秒) | |
|--------------|-----------|----------|
| | スキュー補正あり | スキュー補正無し |
| 200 | 1.80 | 0.95 |
| 400 | 6.79 | 3.59 |
| 600 | 14.32 | 8.21 |

らを画像処理ソフトウェアで400dpiと600dpiに解像度変換した画像についても測定した。各5画像の平均処理時間をTable.3に示す。200dpiではほぼ実時間で補正できており、実用レベルに達している。

5. まとめと今後の展開

特殊な機器を用いず、ページ外形や罫線、文字（行）等の画像中の情報のみを利用する製本原稿スキャン画像の歪み補正方式を開発した。補正前の製本原稿スキャン画像と本方式により補正した画像とを、主観評価やOCR認識率によって比較した結果、本方式の有効性が確認でき、かつ、処理時間も実用レベルであることが分かった。

しかしながら、スキャナーのパラメーターが既知の場合は、従来方式のように断面形状を求めた方が補正精度が上がるので、今後は両者を併用して適応的に使い分ける方式を探索する。また、画像が英文の場合は、（文字外接矩形ではなく）文字の縦成分のみを抽出する方法が有効であるとの報告がなされており[4]、その導入を試みる。他に、地肌がカラーの場合の輝度補正方法や、ページ外形の外側にハードカバーが写りこんだ場合の対策、製本原稿画像とそうでないものの自動判別等を考案して、さらなる補正精度向上を図る。

参考文献

- 1) T. Wada, H. Ukida, and T. Matsuyama: "Shape from Shading with interreflections under a proximal light source: Distortion-free copying of an unfolded book," IJCV, Vol.24, No.2, pp.125-135, Sept. (1997).
- 2) 大削弘樹, 伊藤秀隆, 隈元昭: "書籍の輪郭情報に基づくコピー機の画質補正," 画像電子学会誌, vol.27, No.4, pp.397-404, Aug. (1998).
- 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," 電子情報通信学会論文誌, Vol.J83-D-II, No.12, pp.2610-2621, Dec. (2000).
- 4) 宮崎隆, 長崎健, 川嶋稔夫, 青木由直: "紙の変形を考慮した机上文書のハンドリング," 信学技報, PRMU 2000-180, pp.143-148, Jan. (2001).
- 5) <http://www.minolta-sales.co.jp/products/office/book/index.html>
- 6) <http://pr.fujitsu.com/jp/news/2002/06/5.html>

注1) ミノルタ、BookPro7000Dはミノルタ株式会社の商標です。

注2) 富士通、表OCR for Excelは富士通株式会社の商標です。

注3) リコーシステム開発、読取物語EXはリコーシステム開発株式会社の商標です。

注4) PentiumはIntel Corporationの米国およびその他の国における商標です。