

# 文書画像における行ごとの文字列傾き補正手法

## A Method of Character String Inclination Correction Using Corner Detection

土井 猛† 鈴木 貴士‡ 辻 裕之†‡ 木村 誠聡†‡  
Takeru DOI Takasi SUZUKI Hiroyuki TSUJI Tomoaki KIMURA

† 神奈川工科大学情報学部  
Information Technology, Kanagawa Institute of Technology

‡ 神奈川工科大学情報工学専攻  
Department of Information and Computer Sciences, Kanagawa  
Institute of Technology

### 1. はじめに

光学文字認識（以下 OCR）を行う場合、紙面の傾き補正は不可欠である[1]。認識する文字は紙面に対して水平である必要があり、故に正しく傾き補正したかどうかは認識率に大きく影響する。筆者らは実際に文字認識エンジン Tesseract-OCR[2]を使用した実験において文字列の傾き角が 5 度以上である場合認識率が激しく低下する問題[3]を確認している。傾きの補正は文書画像の輪郭から角度を検出し補正を行うことが一般的であり[4]、行ごとに傾き角が変化する文書画像では補正ができない。手書き文書等で起こりうる行ごとに角度が異なる文書画像を認識するには行ごとに補正する前処理が必要である。本稿では行ごとに角度が異なる文書画像の文字列を補正する手法を提案する。

### 2. 提案する傾き推定法

提案する補正手法は以下の 8 ステップで行う。

1. 文書画像を膨張させ単語ごとに 1 つの領域を生成する。
2. 生成した領域にラベルをつけて分類する。
3. 文書画像と膨張後文章画像をラベルごとに切り出す。
4. 3 で切り出した膨張後画像のコーナーを検出する。
5. コーナー座標から回帰直線を求め角度を推定する。
6. 推定した角度から切り出した画像の傾きを補正する。
7. 3, 4, 5, 6 を切り出した全ての画像で行う。
8. 補正後の画像全てを合成し画像を作成する。

具体的には 1. においては文書画像を単語ごとに領域を生成する処理にモルフォロジー演算の黒部分を膨張させる Erosion（収縮）を使用する。この Erosion 処理を複数回適用することで最適な領域を生成する。3. ~5. におけるコーナー検出は Shi-Tomasi のコーナー検出を用いる[5]。コーナー座標から最小 2 乗法を用いて、回帰直線を算出し、文字列下部の座標を抽出する。文字列下部の座標からロバスト推定法を使用して回帰直線を算出し、傾きを推定する[6]。回帰直線の傾きから文字列の傾き角を推定し、アフィン変換によって膨張前の画像を回転することで、文字列の傾きを補正する。以上によって傾いた文字列は水平となり、その後 OCR にて認識が可能となる。

### 3. 実験方法

実験で使用する文書画像は行ごとに傾き角  $\theta$  が異なる種々の文字列がある文書画像 5 枚を使用する。文書画像の作成に Microsoft Visio2013 を用いる。作成した文書画像に Erosion を複数回適用した膨張画像を作成する(図 2)。なお、Erosion に使用した構造要素は  $3 \times 3$  の矩形を使用する。領域ごとに切り出すことで、1 行ごとに傾きを補正することが可能となる(図 3)。また、切り出した 1 行画像(図 4)からコーナーを検出し、傾き角  $\theta'$  をロバスト推定によって推定し(図 5)、傾きを補正する(図 6)。

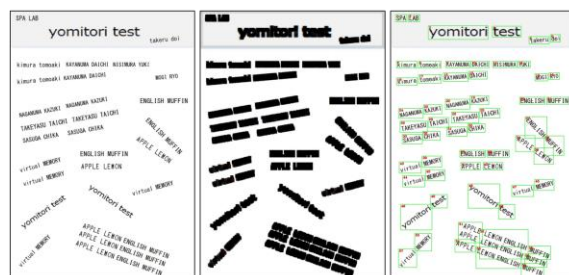


図 1: 文書画像

図 2: 膨張文書画像

図 3: 行ごとの切り出し

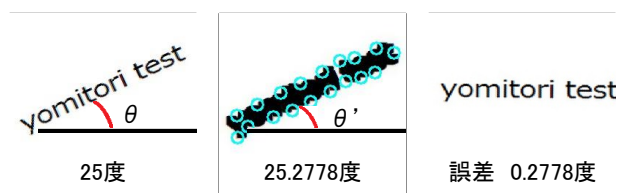


図 4: 傾き 25 度の文字列画像

図 5: コーナーからの角度推定

図 6: 補正後画像

### 4. 結果・まとめ

表 1 にロバスト推定を用いた補正手法の理想的な 1 行画像の補正精度を示す。総計 1500 行のうち平均補正誤差は 0.97 度で最大補正誤差が 7.66 度である。

提案した補正手法によって行ごとに傾き角が異なる文章画像の補正に成功した。今後は様々な文書画像でデータを取る必要がある。

表 1: 理想 1 行画像の補正精度

理想1行画像の補正精度	
総計	1500
平均補正誤差	0.97度
最大補正誤差	7.66度

### 参考文献

- [1] 林俊成, 成田誠之助, “最小 2 乗法を用いた文字列傾き手法”, 2000 年電子情報通信学会総合大会論文集, 情報・システム, D-12-28, p198, 2000/3/7.
- [2] An Overview of the Tesseract OCR Engine, <http://static.googleusercontent.com/media/research.google.com/ja//pubs/archive/33418.pdf> (2019/5/16).
- [3] 土井猛, 鈴木貴士, 辻裕之, 木村誠聡, “コーナー検出を用いた文字列傾き補正手法”, 2019 年電子情報通信学会ソサイエティ大会.
- [4] 後藤成成, “手書き答案用紙画像からの行切り出し手法の高速化”, 三重大学大学院地域イノベーション学研究所修士論文 (2015).
- [5] J. Shi and C. Tomasi, Good Features to Track. IEEE Conference on Computer Vision and Pattern Recognition, 1994.
- [6] 松原望(他), 『統計学入門』第 38 刷, 東京大学出版会, 2017.