

**科学研究費助成事業 研究成果報告書**

平成 30 年 6 月 26 日現在

機関番号：57301

研究種目：基盤研究(C) (一般)

研究期間：2015～2017

課題番号：15K00258

研究課題名(和文) 曲がった文字列も対応可能な実環境文字列認識方法の開発

研究課題名(英文) Development of real environment character string recognition method that can handle curved character strings

研究代表者

志久 修 (SHIKU, OSAMU)

佐世保工業高等専門学校・電子制御工学科・教授

研究者番号：00235516

交付決定額(研究期間全体)：(直接経費) 3,600,000円

研究成果の概要(和文)：カメラで入力した歪んだり曲がったりした文字列を高精度に認識するため、(1)注目文字と隣接する他の文字の位置関係を用いて、注目文字の形状を補正し認識する方法と(2)傾いた文字を単文字の形状だけを用いて形状を補正し認識する方法の開発を行った。(1)では注目文字候補を中心とする近傍の文字(部分文字列)から文字の回転角と傾斜角を求め、注目文字の形状を補正して認識する。(2)では(1)で補正できなかった文字の傾斜を、単文字に外接する平行四辺形を正方形にする正規化法で補正して認識する。実験により開発した方法の有効性を確認した。

研究成果の概要(英文)：We developed a character normalization method for camera based character recognition. In this method, the character normalization process is executed for each character candidates extracted from images. The details are as follows. (1) a part of the character string is taken out as a substring. (2) rotation and shear normalization are applied for the substring. (3) characters are extracted from the normalized substring. (4) shear and size normalization are applied for the extracted characters. We tested the proposed method for many images. The results show that the proposed method is more accuracy than a conventional method.

研究分野：画像認識

キーワード：文字認識 文字正規化 カメラ入力

1. 研究開始当初の背景

文字認識とは、画像で表された文字を、電子テキストに変換する技術である。古くから研究されている技術で、代表例として、郵便区分機の郵便番号や宛名の読み取りがある。以前はスキャナで紙をスキャンした画像が対象であったが、スマートフォンやデジタルカメラの急速な普及に伴い、これらのカメラで撮影された文字が、新しい認識対象となってきた。カメラを使うと、風景の中の文字を簡単に撮影できるため、例えば、外国旅行中に読めない文字を撮影するだけで翻訳したり、看板の文字を撮影するだけで、インターネット検索したりできるようになる。このように便利な応用が考えられるため、カメラで撮影した文字の認識の研究が、多くの会社や大学等で活発に行われている。我々も、カメラで撮影した画像に対し、文字成分抽出、文字列抽出、文字列形状補正、文字認識の各方法の開発を行ってきた。この方法では、平面上の文字を斜めから撮影した場合にも、文字の傾斜を補正することで認識を可能にしているが、過補正（文字の起こしすぎ）の問題が残っている。また、曲面上の書かれた湾曲している文字列には全く対応できていない。そこで本研究では、傾斜した文字列山立っている文字列にも対応できるような文字認識方法の開発を行う。

2. 研究の目的

スマートフォン等のカメラで撮影した文字を高精度で認識することを目的とする。カメラで撮影した文字は、斜めから撮影することによる歪み（文字列の回転、文字の傾斜、文字の大きさの変化）、照明による影の影響、文字と背景図形の区別、ピントずれによるボケの影響、さらには文字列自体の曲がりなどの問題があり、高精度な認識が実現できていない。本研究では、歪んだり曲がっていたりする文字列を高精度に認識する技術を開発する。具体的な目的は、文字列を構成する歪んだ文字の正規化・認識技術を開発する。

3. 研究の方法

(1) 射影歪みを受けた文字列の補正

水平方向に回転補正された文字列に対し、文字列全体の正規化と個別文字の正規化の2段階の正規化で、射影歪みを補正する方法を提案する。

図1に提案する文字列認識処理の流れを示す。学習フェーズでは文字認識用の辞書を作成する。具体的には、あらかじめ用意した正立文字に対し水平方向のせん断変形を加え、これらの文字に正規化を適用する。正規化後の文字から特徴抽出及び識別辞書を作成する。認識フェーズでは、入力された文字列に対し、文字列全体の正規化を行った後、文字列から切り出された各文字に対して文字の正規化及び文字認識を行う。文字認識法として、方向線素特徴量と部分空間法を採用する。

図2に消失点抽出による文字列の正規化の説明図を示す。まず入力画像（図2(a)）に対

し、各連結成分の輪郭線を折れ線近似することで、文字列を線分データで表す（図2(b)）。次に線分データから文字列の垂直方向の消失点（垂直消失点と呼ぶ）と水平方向の消失点（水平消失点と呼ぶ）をそれぞれ推定する。そして、これらの消失点を通る直線群から、文字列に外接する四辺形を求める（図2(c)）。さらに一般的に文字は正方形であることを利用し、正規化後の文字列の縦横比を求める（図2(d)）。最後に、求めた四辺形領域と縦横比を利用して、入力画像から正規化画像を得る（図2(e)）。

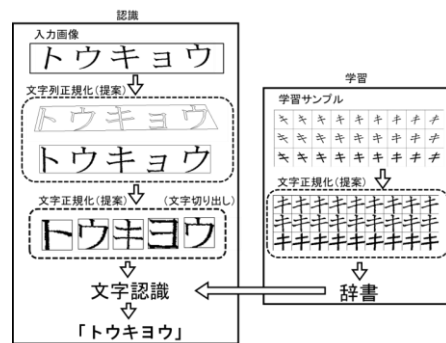


図1 処理の流れ

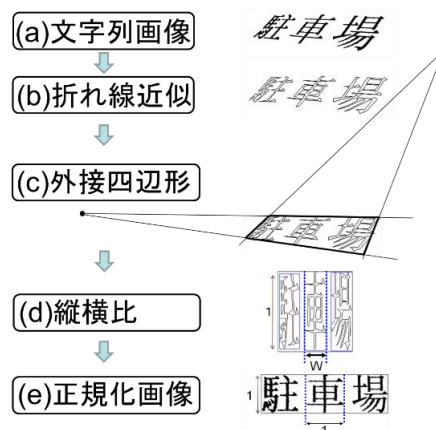


図2 文字列の正規化の流れ

(2) 曲がった文字列の補正

図3は提案する方法の概略である。まず入力濃淡画像（図3(a)）に対し、2値化処理を行い、文字らしい連結成分を文字候補成分として抽出する（図3(b)）。そして、文字候補成分の輪郭線を折れ線近似することで、文字列を線分データで表す（図3(c)）。文字候補成分の中でも、より文字らしい成分を基点成分として選び出す。さらに基点成分の近傍に存在する文字候補成分を選出する。これらを部分文字列と呼ぶ（図3(d)）の太線は基点成分、その他が文字候補成分）。この部分文字列に外接する平行四辺形を求める（図3(e)）。求めた平行四辺形が長方形になるように、文字候補成分を構成する線分データに対し回転と傾斜の変換を行う。図3(f)が回転と傾斜の変換を行った結果である。変換後の成分に外接する方形（図3(g)）を求め、上下に重な

る成分を一つの方形に統合する (図 3(h))。これらの方形において、基点成分を含む方形 (図 3(h)の太枠) と隣接する方形をマージしていく。図 3(i)に部分文字列から切り出された文字候補を示す。切り出された文字候補に単文字の正規化 (詳細は(3)) を行う。これらの処理を全ての基点成分に対して繰り返す。

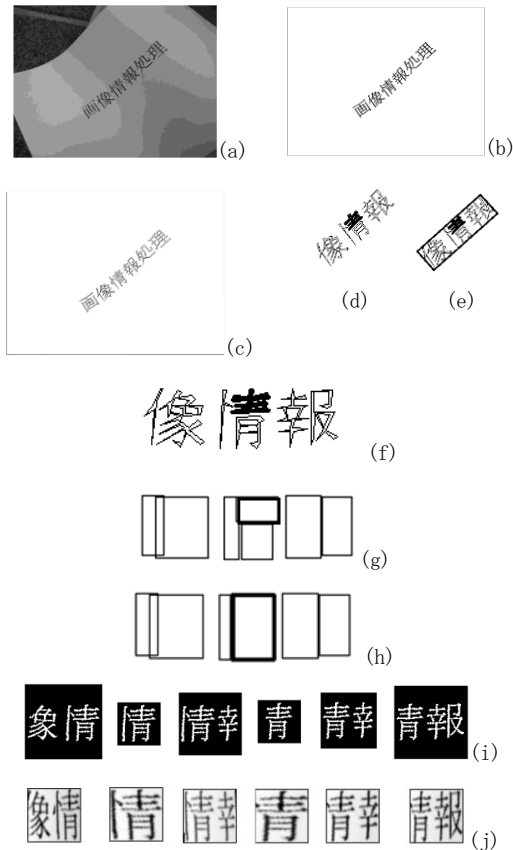


図 3 処理の概要

### (3) 傾斜した文字の認識

前記 (1) および (2) では文字列全体の形状の補正を行っているが、依然として文字に水平方向の傾斜 (せん断) が残っている場合がある。そこで、単文字の傾斜歪みを正規化する方法を提案する。提案法では、図 4 に示すように、単文字画像 (図 4(a)) から文字に外接する平行四辺形 (図 4(b)) を求め、その平行四辺形が正方形になるように座標変換を行う (図 4(c))。

平行四辺形は、単文字画像の投影ヒストグラムから求める。図 5 に、様々な方向からの x 軸に対する投影ヒストグラムを示す。ここで x 軸は水平軸 (文字列の方向) である。基本的なアイデアは、以下のとおりである。まず各方向の投影ヒストグラムにおいて、投影値の 2 乗の合計 (山の面積と呼ぶ) と投影値が 0 より大きい要素の総数 (山の幅と呼ぶ) を求める。その上で、面積と幅の比が最大となる投影方向を文字の傾斜角とする。この傾斜角と山の左右の端点から平行線を求める。文字の上下に接する平行線は、同様に y 軸へ

の投影ヒストグラムから求める。



図 4 正規化処理の例 左から (a) 文字画像、(b) 外接平行四辺形、(c) 正規化画像

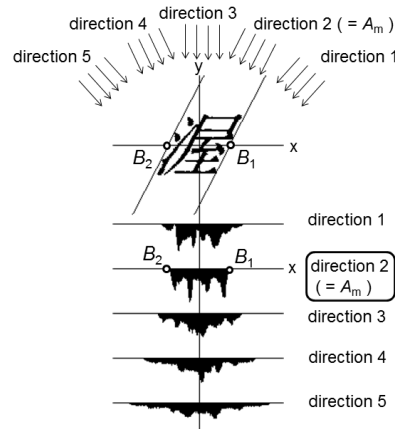


図 5 方向ごとのヒストグラム

## 4. 研究成果

### (1) 射影歪みを受けた文字列の補正

印刷した文字列 250 個 (漢字・片仮名・ひらがな・アルファベット・数字を各 50 個、明朝体・72 ポイント) をデジタルカメラで 10 方向から撮影した画像を用意する。画像での文字の大きさは約 100 画素程度である。これらの半分をパラメータ決定に用い、残りを評価用とした。評価用の文字列画像は 1250 画像 (=125 個×10 方向; 5660 文字) である。

図 6 に評価用の文字列画像 (10 方向) に対し、提案する文字列の正規化を適用した結果を示す。文字列の高さ方向の正規化は良好に行われていることがわかる。これは文字列方向には多くの線分が存在するため投票値の信頼性が高まり、水平消失点が正しく求められているためである。一方、個々の文字の傾斜については、漢字のように縦ストロークを多く含む場合は垂直消失点が正しく求められるが、カタカナのような斜めストロークを多く含む場合は垂直消失点が斜めストロークの延長線上にずれてしまい、斜めストロークを垂直にする文字の起こしすぎが生じている。

表 1 に評価用の文字列画像 1250 画像に含まれる 5660 文字に対する認識率を示す。文字の正規化に大きさの正規化を採用した場合は、提案する文字列の正規化法は文献 [5] とほぼ同等で、高い精度を得ることができていない。その理由は、図 6 に示すように文字の起こしすぎが生じてしまい、正立文字だけを学習させている通常辞書では認識に失敗したためである。文字の正規化に提案法を採用した場合、文字列の正規化の違いにかかわらず文字の認識率が向上している。これは射影歪みや文字列正規化での起こしすぎにより生じる文字の傾斜を、提案する文字の正規

化により補正し、学習サンプルに近づけているためである。

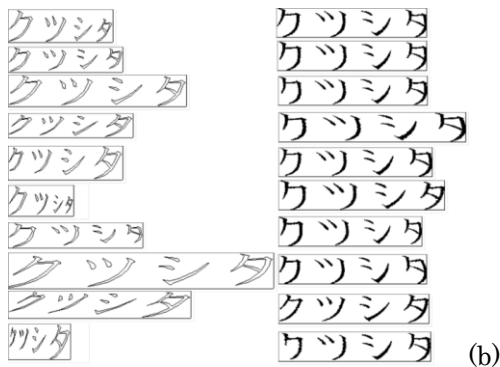


図 6 提案する文字列正規化の結果

左：入力文字列（線分）、右：正規化結果

表 1 実験結果（カッコは 3 位累積認識率）

		文字列の正規化		
		なし	従来法	提案法
文字の 正規化	大きさ のみ	71.7 (78.5)	88.4 (92.8)	88.1 (91.6)
	提案法	94.9 (97.1)	95.1 (97.3)	97.8 (99.8)

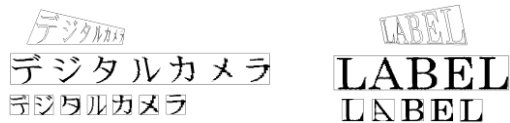
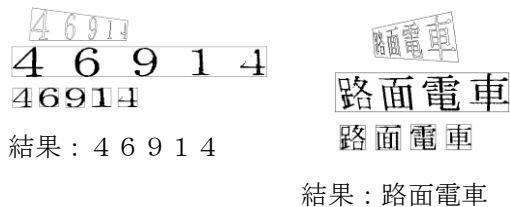


図 7 成功例 上：入力（線分）、中：文字列の正規化（提案法）、下：文字正規化（提案法）

特に提案した文字列の正規化と文字の正規化を組み合わせた場合、最も高い認識率が得られている。図 7 に認識に成功した例を示す。図 8 に提案する文字列および文字の正規化を適用しても認識できなかった例を示す。文字列の正規化による過補正や文字のかすれのために、文字の認識に失敗していた。



図 8 失敗例 上：入力（線分）、中：文字列の正規化（提案法）、下：文字正規化（提案法）

(2) 曲がった文字列の補正

直線状および円周状に並んだ文字列をデジタルカメラで様々な方向から撮影した。文字列を構成する文字は、最小サイズが 100 画素程度になるようにした。図 9 に処理結果を示す。図 9(a) は入力画像、図 9(b) は正規化結果に対して文字認識を行い、文字と判定された領域を示す。射影歪みを受けている文字が正規化され、正しく認識されていることがわかる。その反面、多くの文字以外の図形が文字と誤って判定されている。

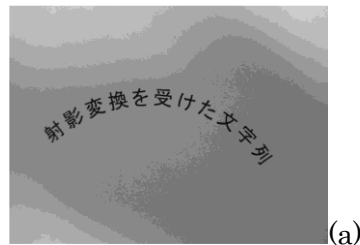


図 9 処理結果

(3) 傾斜した文字の認識

実験では、3,169 字種（漢字、かな、カナ、英数字）を対象とする。識別器の学習のため、1 字種あたり 270 個の文字画像を用意する。

これらはスキャナ（400dpi、2 値）で入力した 1 字種あたり 30 個の文字画像に、y 軸に対し -40 度から 40 度まで 10 度刻みで 9 種類のせん断変形を加えて作成する。30 個の内訳は、3 フォント（明朝体、ゴシック体、楷書体）×5 サイズ（6、8、10、12、14 ポイント）×2 種類のレーザプリンタである。

学習用の文字画像を正規化し、識別器の学習に用いる。正規化サイズは 64×64 画素とする。図 10 に学習用の文字画像（図 10(a)）と正規化画像（図 10(b)、(c)）を示す。従来システムでは大きさの正規化だけなので、正規化後の文字にせん断変形が残っていることがわかる。一方、提案システムでは正規化後の文字が本来の形状と異なる場合があるが、せん断角度によらずほぼ同じような形状あるいは少数の形状に揃っていることがわかる。

評価用の画像として、1 字種当たり 27 個の画像を用意する。これらは学習用と同じ手順で作成するが、学習用とは別に印刷した 6 ポイントの 3 フォント文字を複写機で品質を落とし、さらにせん断角度を正立（せん断角度 0 度）以外は異なる角度（-35 度から 35 度まで 10 度刻み）にしている。図 11 に評価用の文字画像を示す。

部分空間法と CNN（畳み込みニューラルネットワーク）をそれぞれ識別器としたときの評価用の文字画像の角度毎の認識率を調べた。部分空間法の次元数は、提案システムは 8、従来システムは 15 とした。また従来システムでは正立した文字だけを識別器に学習させた場合も調べた。

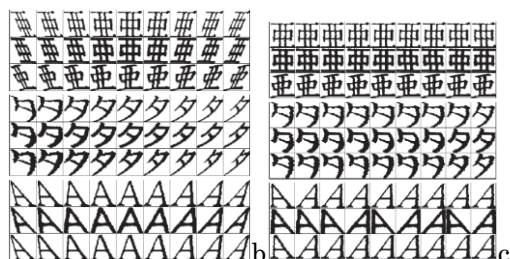
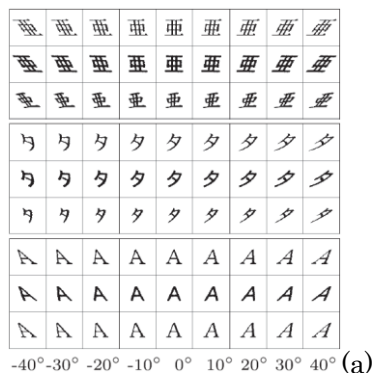


図 10 学習用の文字画像と正規化画像 a:学習画像、b:従来法、c:提案法



図 11 評価用の文字画像

表 2 角度ごとの認識率

角度	部分空間法 (次元数)				CNN	
	従来 (15)		提案 (8)	従来		提案
	正立	傾斜		正立	傾斜	
-35°	1.45 (3.60)	94.48 (98.66)	97.55 (99.53)	2.64 (5.40)	93.40 (98.22)	96.08 (98.97)
-25°	11.50 (21.59)	96.30 (99.36)	97.43 (99.51)	10.27 (18.64)	93.99 (98.39)	95.88 (99.04)
-15°	74.17 (88.99)	97.21 (99.57)	97.75 (99.61)	47.71 (64.15)	93.75 (98.40)	96.10 (99.18)
-5°	98.31 (99.89)	97.84 (99.71)	97.78 (99.59)	91.00 (96.68)	93.80 (98.47)	96.12 (99.11)
0°	99.02 (99.94)	98.17 (99.77)	98.00 (99.63)	94.27 (98.43)	94.30 (98.67)	96.68 (99.27)
5°	98.62 (99.93)	97.88 (99.68)	97.92 (99.54)	91.85 (97.19)	93.97 (98.44)	96.12 (99.17)
15°	74.54 (88.53)	97.45 (99.55)	97.91 (99.56)	45.41 (62.43)	93.80 (98.49)	96.22 (99.13)
25°	11.24 (21.14)	96.56 (99.33)	97.77 (99.60)	8.66 (16.08)	93.80 (98.30)	95.67 (98.98)
35°	1.17 (2.99)	94.23 (98.39)	97.77 (99.67)	2.60 (5.08)	92.81 (97.94)	96.14 (99.04)

表 2 に角度毎の認識率（第 3 位累積認識率）を示す。表より、両識別器において、従来システムでは正立文字だけではなく傾斜した文字も学習させることで、傾斜した文字の認識率が向上していることがわかる。同じく両識別器において、提案システムでは角度によらずほぼ一定の認識率になっているのに対し、傾斜文字を学習させた従来システムでは部分空間法においては傾斜が大きくなると認識率が低下し、CNN においてはすべての角度で認識率が低くなっていることがわかる。これより提案法を前処理とする文字認識システムが、傾斜した文字の認識に有効であることがわかる。

### 5. 主な発表論文等

（研究代表者、研究分担者及び連携研究者には下線）

〔雑誌論文〕（計 2 件）

- ① 志久 修、手島 裕詞、内田 誠一、傾斜文字認識のための正規化方法、電子情報通信学会論文誌 D 情報・システム、査読有、J100-D、2017、902~906  
DOI: 10.14923/transinfj.2017JDL8007
- ② 手島 裕詞、志久 修、金谷孝之、西尾 孝治、小堀 研一、点群データの特徴量抽出、日本設計工学会誌、査読有、52 巻、2016、107-117

〔学会発表〕（計 9 件）

- ① 坂口翔太、志久修、手島裕詞、嶋田英樹、兼田一幸、CNN を用いた傾斜文字および回転文字の認識、電気学会九州支部沖縄支所講演会、2017年12月10日、琉球大学
- ② 関航平、手島裕詞、志久修、小堀研一、点群モデルの形状変形手法、電気学会九州支部沖縄支所講演会、2017年12月10日、琉球大学
- ③ 馬場康平、手島裕詞、志久修、嶋田英樹、兼田一幸、小堀研一、畳み込みニューラルネットワークを用いたボクセルモデルの欠損補間法に関する一考察、電気学会九州支部沖縄支所講演会、2017年12月10日、琉球大学
- ④ 手島裕詞、中尾亮、志久修、小堀研一、点群データの欠損検出法とその評価、電気学会九州支部沖縄支所講演会、2017年12月10日、琉球大学
- ⑤ 手島裕詞、アタルサイハン・ガントウグス、志久修、小堀研一、点群データを対象としたモルフォロジーフィルタの開発と検証、電気学会九州支部沖縄支所講演会、2016年12月11日、琉球大学
- ⑥ 中尾亮、手島裕詞、志久修、小堀研一、点群データの欠損検出とその高速化、電気学会九州支部沖縄支所講演会、2016年12月11日、琉球大学
- ⑦ Tateno K、Shiku O、Ohtubo Y、Counting Algorithm for Sequential Optical Images of Immunoreactive Mouse Taste-Bud Cells、3rd International Conference on Computational Methods in Engineering and Health Sciences (ICCMER- 2016)、2016年12月18日、Kitakyushu、Japan
- ⑧ 志久修、手島裕詞、カメラ入力文字認識のための文字正規化方法、電気学会九州支部沖縄支所講演会、2016年12月11日、琉球大学
- ⑨ 志久修、林信吾、手島裕詞、2段階の正規化を用いた射影歪みを受けている文字列の認識、画像関連学会連合会第3回秋季大会 PC-10、2016年11月17日、京都繊維大学

## 6. 研究組織

### (1) 研究代表者

志久修 (SHIKU OSAMU)

佐世保工業高等専門学校・電子制御工学科・教授