

科学研究費助成事業 研究成果報告書

平成 26 年 6 月 26 日現在

機関番号：57301

研究種目：基盤研究(C)

研究期間：2011～2013

課題番号：23500236

研究課題名(和文) 斜めから撮影した文字の高精度認識法の開発

研究課題名(英文) Development of shape normalization method of character that received perspective distortion

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) 3,900,000円、(間接経費) 1,170,000円

研究成果の概要(和文)：本研究では、斜めから撮影することにより射影ひずみを受けている文字列の認識方法の開発を行った。具体的には、(1)射影ひずみを受けている文字の形状正規化方法の開発、(2)個々の文字の認識結果(文字コード+画像中の文字位置)から、文字列辞書(文字コード列)を利用することにより、個々の文字を文字列にグループ化(文字列を認識)する方法の開発を行った。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)：In this study, we developed a method of recognizing characters that received perspective distortion by shooting from an angle.(1)we developed a shape normalization method for character undergoing perspective distortions

(2)we developed a character string recognition using linguistic knowledge for strings.The methods were applied to many images. As results, almost characters were recognized correctly.

研究分野：情報学

科研費の分科・細目：知覚情報処理・知能ロボティクスA

キーワード：文字認識 射影ひずみ 湾曲ひずみ

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、カメラ付携帯電話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。

2. 研究の目的

近年のカメラ付携帯電話やデジタルカメラの普及に伴い、これらのカメラで撮影した文字を認識する技術（実環境文字認識）の研究が活発に行われている。実環境文字認識は、

- ・ 斜めから撮影することによるひずみ（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
- などが大きな問題となる。

本研究では、斜めから撮影することによる歪みに注目し、平面及び曲面に書かれた文字列が斜めから撮影された場合にも、ロバストに認識できるアルゴリズムの開発を行う。

3. 研究の方法

本研究では、斜めから撮影した文字を認識について検討する。まず、斜めから撮影した文字のデータベースを作るための実験環境を構築する。次に、実際にカメラで撮影した風景中の文字の認識として、文字の大きさと形状に注目した文字領域抽出方法を検討する。さらに、車載カメラで撮影した道路画像に映っている行先案内板の文字に限定した認識システムと工事看板に書かれている文

字に限定したシステムを、それぞれ構築する。以下、詳細に述べる。

(1) 実験システムの構築

実験システム構築とアルゴリズム開発を行った。実験システムでは、膨大な文字画像を得るため、コンピュータ内で生成した画像を大型液晶ディスプレイに表示し、それをカメラで撮影し、認識処理の入力とした。コンピュータで様々な画像を生成することができるため、膨大な画像数を使った実験が可能となった。具体的には、①計算機内に、実際に平面上の文字を撮影した画像データ、文字フォントの組み合わせによる文字列データ（テキストデータ）を用意する。②計算機内に、文字に関する情報（文字サイズ、文字間隔、文字フォント等）を用意する。③計算機内で、文字列データ①に対し、変形パラメータ②による見え方画像を生成する。コンピュータグラフィクス技術を利用する。④計算機内で生成した画像を、大型液晶ディスプレイに表示する。⑤大型液晶ディスプレイの画面を、カメラで再撮影する。カメラは、様々な品質の画像を得られるように、高性能IEEE1394カメラ、民生用デジタルカメラ、民生用ビデオカメラ、ウェブカメラなど複数の性能のカメラを用意する。⑥撮影した画像に対し、開発する文字認識処理を行う。この文字認識処理に、文字単体の正規化、曲面上の文字列認識を組み込む。

(2) 文字の色と形状に注目した文字領域抽出

我々が開発した従来法では、濃淡画像から細い図形を文字領域として抽出したあと、直線状に並ぶ文字領域を文字列としてグループ化している。この方法では、画像中の細長い図形をすべて文字領域とするため、多くのノイズが文字領域として抽出され、それに伴い文字列が正しく抽出されないという問題がある。

本研究では、一般的に一つの文字列はほぼ同色で、ほぼ同じようなフォント形状の文字から構成されていることに注目している。そこで、本研究では文字の色と形状に注目した新たな文字領域抽出方法を提案する。また提案法により抽出した文字領域を使って文字列をグループ化するシステムについて提案する。

提案法では文字の定義を次の三つとする。

1. 同じ文字列は同色な領域で表される。
2. 同じ文字列を構成する文字領域同士は近接する。
3. 文字の形状には特徴がある。

そこで、これらの定義に従い、同色領域抽出、同色領域のグループ化、文字グループの決定の3ステップにより、文字列を構成すると考えられる領域のグループを作る。そして、これらグループに対して、それぞれ従来法の文字列抽出処理を適用していく。

図1に処理例を示す。この例をもとに説明する。図1において、図(a)はカラー入力画像、図(b)は文字グループ画像、図(c)は文字列抽出結果(青枠が文字列領域)である。

まず、カラー入力画像から同色領域を抽出し、各領域において代表色のRGB値をそれぞれ求める。次に、代表色が似ていて、かつ近接する領域を同じグループにまとめていく。最後に、各グループが、文字グループか背景グループのどちらに属するかを決定していく。文字の形状には特徴があるとして、文字形状を学習させた判別分析法を使用する。提案法では、フィッシャーの線形判別を使用し、その領域が文字か背景かを判定する。そして、文字と判定された領域が多いグループを文字グループと決定する。

図1の例では、図(a)のカラー入力画像から、図(b)に示すように4枚の文字グループが抽出されている。これらの4つの文字グループに対して、それぞれ文字列抽出処理を行った結果、2つの文字グループから、合計3つの文字列が抽出されている。

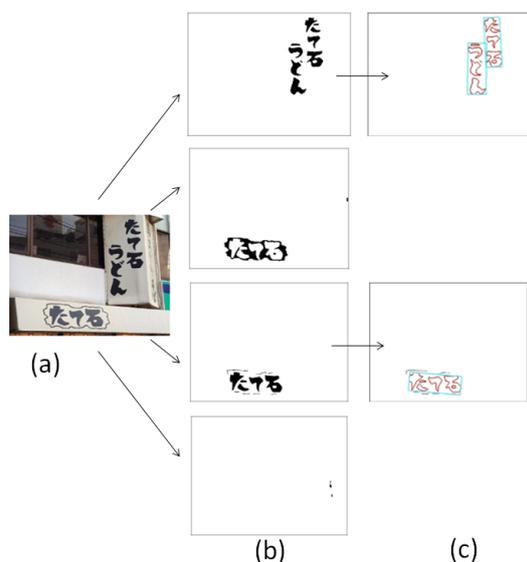


図1 文字の色と形状に注目した文字領域抽出

(3) 道路画像からの行先案内板の文字認識

本研究では、地名や施設名などを記載している経路案内標識を車載カメラで撮影し、文字認識技術により、カーナビゲーションシステム用データベースを自動的に生成することを目的とする。全国各地に存在する経路案内標識を、手動でデータベース化するのは非常に困難である。これを自動化するシステムを開発することで、非常に少ない労力でのデータベースの作成や、その更新が可能となる。

図2に処理例を示す。看板を撮影したカラー画像(図(a))を入力する。看板領域抽出では、入力画像から、HLS色空間における色相・彩度の値によって青色の領域を特定し、看板領域画像を作成する(図(b))。看板領域

画像を利用し、マスク処理によって看板画像を抽出したのち、看板領域画像に対して線領域強調を行い、文字線画像を作成する(図(c))。文字領域抽出では、文字線画像から文字線が集中している領域を文字候補領域として抽出する(図(d))。抽出したすべての文字候補領域に対して、文字認識を行う。文字認識では、抽出した領域を一定の大きさに正規化したのち、輪郭情報を抽出し、辞書と比較した結果、類似度の高いものを文字として出力する(図(e))。最後の地名特定処理では、あらかじめ用意していた地名辞書(文字コード列)と文字認識結果を照合し、全ての文字が文字認識結果の中に含まれる地名を最も適切な地名として出力する(図(f))。

(4) 斜めから撮影した工事看板の文字認識

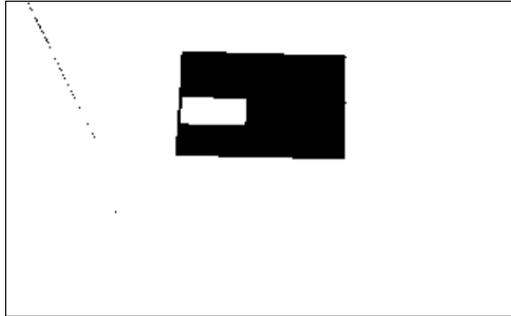
本研究では、工事看板を対象として、斜めから撮影することにより射影ひずみを受けている文字列の認識を行う。具体的には、

(1) 文字のひずみを補正するために、近接する文字の形状及び文字そのものの形状をもとに文字の正規化を行う方法の開発、(2) 個々の文字の認識結果(文字コード+画像中の文字位置)から、文字列辞書(文字コード列)を利用することにより、個々の文字を文字列にグループ化(文字列を認識)する方法を提案する。

図3に処理の一例を示す。工事看板を撮影したカラー画像(図(a))を入力する。カラー画像をグレイ画像に変換後、入力画像サイズで背景に対して明るく細い図形(図(b))と暗く細い図形(図(c))をそれぞれ文字候補として抽出する。さらに入力画像を縮小した画像から、同様に文字候補を抽出する(図(d))。抽出した文字候補ごとに、近接する文字候補の形状と注目している文字候補の形状をもとに、注目している文字候補を正立させる(図(e))。正立させた文字候補に対し、文字認識を行い、類似度がしきい値より高いものを文字として選び出す。選び出された文字を、工事看板に出現する文字列(工事中、徐行、車線変更など)と照合し、最も合致した文字の割合が多い文字列を認識結果として出力する(図(f))。



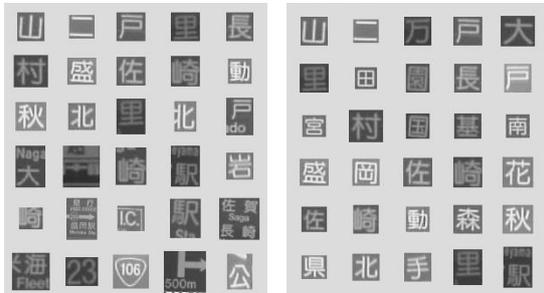
(a) 入力カラー画像



(b) 看板領域画像



(c) 文字線強調画像



(d) 文字候補領域



(e) 文字認識結果



(f) 地名認識結果

図2 道路画像からの行先案内板の文字認識



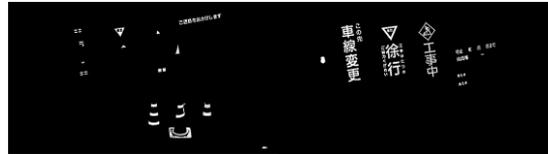
(a) 入力画像



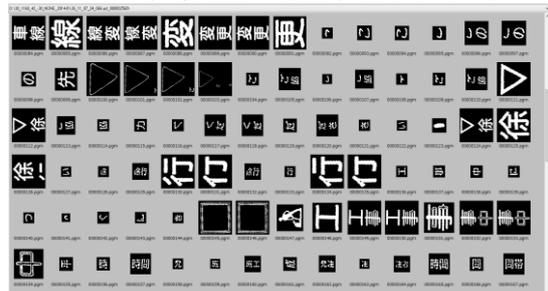
(b) 文字候補 (明るく細い図形; 原画像サイズ)



(c) 文字候補 (暗くて細い図形; 原画像サイズ)



(d) 文字候補 (明るく細い図形、暗く細い図形、縮小サイズ)



(e) 正立させた文字候補



(f) 文字列認識結果 (黒い線でつないでいる)

図3 斜めから撮影した工事看板の文字認識

4. 研究成果

(1) 実験システムの構築

実験システムを構築した。この実験システムの利点は以下のとおりである。

- ・ 膨大な画像に対する認識実験、評価を自動的に行うことができる。
- ・ 計算機内で生成された変形だけではなく、実際の撮影による現実的な変形（カメラの解像度、カメラ・プロジェクタ・スクリーンの角度、撮影時の照明等）も反映することができる。

これらのシミュレーション実験での評価をフィードバックさせながら、アルゴリズムの開発を行っていく環境を構築できた。

(2) 文字の色と形状に注目した文字領域抽出

実際の画像 100 枚（文字列総数 206 個）を用いて実験を行った。

図 4 に実験結果を示す。良好に抽出されていることがわかる。図 5 に従来法との比較を示す。従来法は色とフォント形状の情報を使わないため、近接する文字を文字列にまとめているが、提案法では正しく文字列が抽出されていることがわかる。

実験の結果、近接する文字成分の条件に色とフォント形状を使わない従来法では、再現率 85%、適合率 60%であったのに対し、色とフォント形状を条件に入れた提案法では再現率 96%、適合率 63%と、再現率、適合率ともに向上した。

今後は、さらに多くの画像を用いて、本方法の性能を調べ、さらに改良していく必要がある。

(3) 道路画像からの行先案内板の文字認識

地図製作会社から提供された道路画像に対して実験を行った。図 6(a), (b) に実験結果の例を示す。良好に抽出できていることがわかる。失敗例として、単純な形状の文字の切出し認識の精度が悪いため、例えば「○○I.C.」のように「I」を含む地名の認識に失敗した。これらの失敗に対して、「I.C.」を一つの文字とみなして認識するなどを検討していく必要がある。

(4) 斜めから撮影した工事看板の文字認識

デジタルカメラで撮影した実際の風景画像 500 枚に適用した。画像には漢字を含めた日本語、アルファベット、数字が含まれている。実験条件として、文字は 30 画素以上、認識対象の文字列はあらかじめ文字列コードとして文字列辞書に登録されていることを条件とした。また、文字認識法には方向特徴量と部分空間法を採用し、文字フォントから生成した文字認識辞書を用いた。

図 7(a), (b) に実験結果を示す。実験の結果、様々な方向からの撮影により射影ひずみを受けている文字および文字列も良好に認識することができた。

本方法では、文字列が同じ色の文字で構成

されているという特徴は利用していない。また、定量的な性能評価ができていない状況である。そこで、今後の課題として、文字列の単色性の利用及び性能を定量的に評価することである。さらに多様な画像に適用して本方法の性能を調べることである。

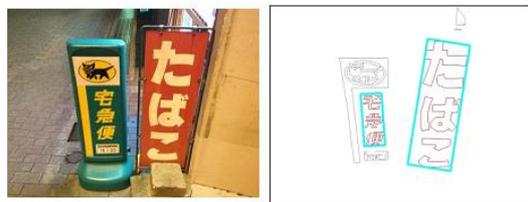


図 4 実験結果



図 5 従来方法との比較（左：入力、中央：従来法の結果、右：本方法の結果）



(a)



(b)

図 6 実験結果



(a)



(b)

図7 実験結果

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 1 件)

- ① 画像処理による砥石作業面トポグラフィの3次元計測に関する研究 第3報: 砥粒切れ刃解析の深化、川下智幸、坂口彰浩、福本洋平、松尾修二、志久修、峠直樹、樋代康広、砥粒加工学会誌, 査読有り、55号、pp. 360-365、2011. 6

[学会発表] (計 6 件)

- ① 経路案内標識からの文字切り出し、松尾俊基、志久修、手島裕詞、第21回電子情報通信学会九州支部 学生会講演会、査読無し、D-35、(2013. 9. 23)
- ② 画像処理を用いたメラノソーム輸送追跡法、濱野あゆみ、藤崎顕彰、志久修、内田誠一、画像の認識・理解シンポジウム(MIRU2012)、査読なし、IS2-76、2012. 8.
- ③ 画像処理を用いたメラノソーム輸送追跡法、濱野あゆみ、藤崎顕彰、志久修、内田誠一、画像の認識・理解シンポジウム(MIRU2012)、査読なし、IS2-76、2012. 8.
- ④ 文字の色と形状に注目した情景画像からの文字領域抽出、佐々木俊樹、志久修、手島裕詞、第65回電気関係学会九州支部連合大会講演論文集、査読無し、09-2A-11、(2012. 9. 25)
- ⑤ 電子コミックのためのシーン評価手法の一検討、濱崎亮、手島裕詞、志久修、第65回電気関係学会九州支部連合大会講演論文集、査読無し、09-2P-15、(2012. 9. 25)
- ⑥ SURF特徴量を用いた水中画像からの水中生物抽出、日和亮、志久修、手島裕詞、

第65回電気関係学会九州支部連合大会講演論文集、査読無し、09-2P-17、(2012. 9. 25)

[図書] (計 1 件)

- ① 風景の中の文字を認識する技術を目指して、志久修、中小商工業研究、査読無し、第112号、pp. 75-80、2012. 7.

[産業財産権]

○出願状況 (計 件)

名称：
発明者：
権利者：
種類：
番号：
出願年月日：
国内外の別：

○取得状況 (計 件)

名称：
発明者：
権利者：
種類：
番号：
取得年月日：
国内外の別：

[その他]
ホームページ等

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

研究者番号：00235516

(2) 研究分担者

()

研究者番号：

(3) 連携研究者

()

研究者番号：