

機関番号：57301

研究種目：基盤研究 (C)

研究期間：2008～2010

課題番号：20500501

研究課題名 (和文) 文字認識技術を用いた実環境内の文字情報提供システムの開発

研究課題名 (英文) Development of a character information service system for a real world by using character recognition.

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・准教授

研究者番号：00235516

研究成果の概要 (和文)：

実環境内の文字を認識するため、(1) 全方位カメラ、ステレオカメラ、通常のカメラを用いた撮影システムの開発、(2) 全方位画像からの特定マーク検出方法の開発、(3) 斜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。特に、(3)においては斜めから撮影した文字列や湾曲している文字列も認識できるように、文字列の補正法と文字の正規化法を開発した。

研究成果の概要 (英文)：

To recognize the characters in a real world, we have studied the topics described below. (1) A camera system consists of a stereo camera, omnidirectional camera and normal camera was developed. (2) A method of mark detection from panoramic images was developed, (3) Methods of recognizing character strings that receive perspective distortions were developed.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2008年度	2,700,000	810,000	3,510,000
2009年度	500,000	150,000	650,000
2010年度	500,000	150,000	650,000
年度			
年度			
総計	3,700,000	1,110,000	4,810,000

研究分野：文字認識

科研費の分科・細目：人間医工学・リハビリテーション科学・福祉工学

キーワード：文字認識、斜影ひずみ

1. 研究開始当初の背景

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。従来のスキャナで読み込んだ紙文書を対象とする文字認識技術と違い、実環境文字認識は、文字を斜めから撮影することによる射影ひずみ(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。

申請者はこれらの問題に対し、新しい文字抽出および文字列認識方法を開発した。この方法は、JSTの平成16年度革新技術開発研究事業「情景画像からの文字情報抽出技術の研究」による研究補助(3年間)、国内の有力文字認識メーカーで構成される(社)電

子情報技術産業協会・認識形入力方式標準化グループでの招待講演「一般情景における文字認識（非整備環境文字認識）」など注目されており、さらに現在発売中の携帯電話テキストリーダーやデジカメ用文字認識ソフトの一部に採用されるなど、実用的な性能も認められている。

一方、日本の視覚障害者は約30万人といわれ、そのうち点字が読めるのは3万人程度といわれている。また70歳以上が視覚障害者の半数を占め、糖尿病による失明者が多い。糖尿病では指先の感覚が失われるため、点字による情報の取得ができないという問題がある。

視覚障害者にとって、目的とする建物の近くまで行けるが、最後までたどり着けない場合や、家庭内の薬や飲食物などが分からない場合などがある。これらに対し、ボランティアにテレビ電話（携帯電話）でアクセスし、周囲の動画を見せることで居場所や飲食物の名前を教えてもらうサービスが行われている。これらは、実環境文字認識により、周囲の看板や案内板あるいは飲食物や薬の文字を認識することにより、自分自身で解決できるようになる。

また街中を自律的に行動するロボットに対する文字情報の提供にも、実環境文字認識の応用が可能である。

2. 研究目的

現在の実環境文字認識では、利用者が文字にカメラを向けて画像を取得することを想定している。例えば、日本に旅行中の外国人が、日本語で書かれている看板を撮影して、英語に翻訳することを想定している。これに対し、視覚障害者あるいは自律ロボットでは、風景のどこにカメラを向ければいいのか自体が分からないため、周囲を同時に観測する必要がある。さらに利用者に文字情報を提供する場合、近距離にある文字と遠方にある文字では、情報の重要性も変わってくる。

そこで本研究では、全方向カメラ（周囲360度のパノラマ画像を撮影可能）、ステレオカメラ（対象物までの距離を画素値とする距離画像を取得）、および通常のカメラの3種類のカメラによる画像撮影システムを構成する。そして、この画像撮影システムで得られた画像向けに、申請者がこれまで開発してきた実環境文字認識を発展させることにより、視覚障害者向け文字情報提供システムを開発する。具体的には、次の2点である。(1)パノラマ画像、距離画像、通常の画像に対する実環境文字認識を発展させ、実際の場面での文字情報を提供できるシステムを構築する。さらに本システムを使って100%の精度で利用者に情報を提供するために、文字

列仕様（文字の最小サイズ、文字間隔、文字フォント、撮影角度の範囲など）を整備する。(2)本システムを、実際のヒューマンインターフェースとした場合の問題点（どの程度の認識精度が必要か。複数の文字情報が存在する場合どのような優先順位で提示するか、など）を明らかにする。

3. 研究の方法

本研究では、以下のテーマについて主に取り組んだ。

- (1) 撮影システムの構築
- (2) 全方位画像からの特定マーク検出
- (3) 斜影ひずみを受けている文字列の認識

(1)については、全方向カメラ、ステレオカメラ、および通常のカメラの3種類のカメラによる画像撮影システムを構築した。全方位カメラは固定、ステレオカメラと通常のカメラはパソコン制御の雲台により可動とした。

本システムでは、まず図1(a)の全方位カメラにより、現地点の360度のパノラマ画像を得る。このパノラマ画像から、文字らしい領域を検出する。パノラマ画像は、文字認識を行うには解像度が低いが、文字らしい領域を検出することは可能である。次に、図1(b)のパソコン制御の雲台を使ってステレオカメラとIEEEカメラを検出された文字らしい領域に向け、その領域を撮影する。ステレオカメラを使う理由は、距離を使って、文字領域と背景領域を分離したり、射影ひずみ（3次元歪み）を補正したりするためである。さらに、IEEEカメラでは高解像度な画像を取得できるため、小さい文字や文字の微細な形状をとらえて、高い認識精度を達成することができることを想定している。



(a)全方位カメラ



(b)パソコン制御
ステレオカメラ

図1 カメラシステム

(2)については、風景の中の文字列を読み取るために、あらかじめ読み取り対象の文字列の近傍に「特定マーク」を付加しておくことを想定する。そこで、パノラマ画像から特定マークを検出することで文字列の大まかな位置を求め、雲台に乗せたステレオカメラと通常のカメラを文字列位置に向け、文字列を撮影する方法を開発した。

(3)のカメラで斜めから撮影した文字列を認識する技術は、携帯電話やデジカメによる街角の看板文字列の認識にも応用できるため、非常に重要な技術である。斜めから見た文字列には射影ひずみが生じている。そこで、この問題に対し、次の4つの研究を行った。

①一般的な風景に存在する文字列の抽出方法を開発した。この方法では、注目文字を含む文字列の一部分(部分文字列)に外接する平行四辺形をハブ変換により求め、この平行四辺形から得られる文字列高さ、方向、文字の傾斜をもとに文字列全体を抽出している。
 ②射影ひずみを受けている文字列を、正面から見たように補正する方法として、図2に示すような方法をすでに開発している(方法1と呼ぶ)。この方法では、文字の傾斜は補正できるが、文字の大きさの補正はできなかった。そこで新しい方法として、図3に示すように、文字列の水平方向と垂直方向の消失点をそれぞれ求め、これらの2点をそれぞれ通過する2対の直線による囲まれた四辺形を、長方形に変換する方法を開発した(方法2と呼ぶ)。

③曲面上に書かれた湾曲ひずみを受けている文字列の認識方法を開発した。図4に示すように、この方法では、文字列の局所直線性を仮定して、注目している文字の形状補正を行い認識している(方法3と呼ぶ)。

④前述した②、③により、歪んでいる文字の認識精度が向上するが、文字の起こしすぎによる副作用も生じる。そこで、図5に示すような、文字単体の正規化を新たに導入する。具体的には本来の文字形状ではなく、文字の垂直(あるいは水平)ストローク強度が最大になる基準で文字の形状を揃える。辞書用の学習サンプルおよび傾斜補正後の文字列から切り出された文字を、本方法で正規化した後に、認識を行う。

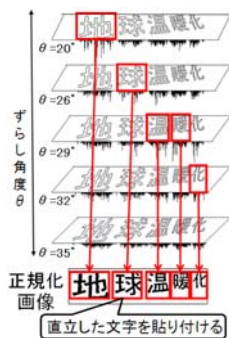


図2 射影ひずみ補正(方法1:従来法)

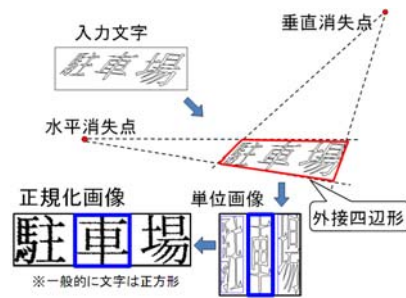


図3 射影ひずみ補正(方法2:提案法)

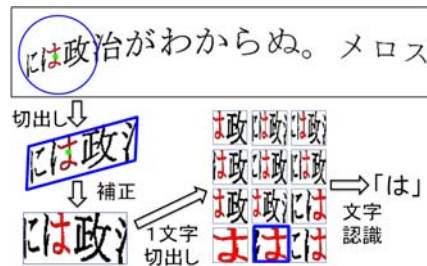


図4 湾曲文字列の認識(方法3:提案法)

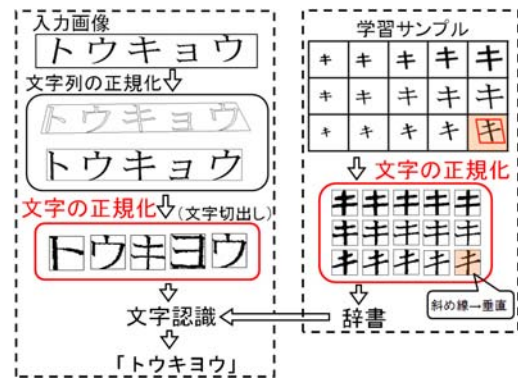


図5 文字単体の正規化(正規化:提案法)

4. 研究成果

(1) 撮影システムの構築

パソコン制御雲台に乗せたカメラおよびステレオカメラを用いて物体を追尾するシステムを開発した。図6は赤い点を追尾した例である。赤い点が、常にカメラで入力した画像の中央に位置するように、雲台を制御している。このときカメラにステレオカメラを用いることで、赤い点までの距離も取得でき、赤い点の3次元位置を計測できる。実験より、センチオーダーの精度で物体にカメラを向けることができることを確認した。これにより、情景中の対象とする文字にカメラを自動的に向ける制御が可能であることがわかった。

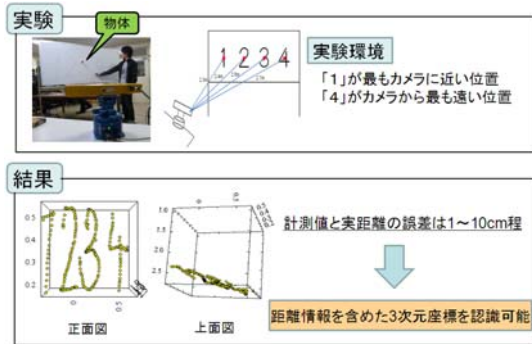
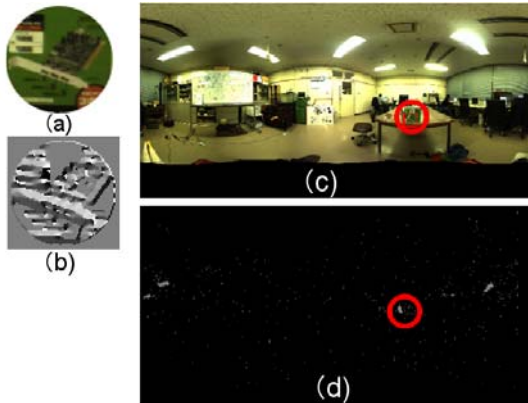


図6 パソコン制御ステレオカメラの実験結果

(2) 全方位画像からの特定マーク検出

特定マークを検出するため、アクティブ探索法を改良した方法を開発した。具体的には、特定マークを円形のテンプレートとし、色相によるマッチング、エッジの方向頻度を特徴としたざらしマッチングの採用により、マークの大きさや回転角度の変化、照明変化による明度変化に頑健で、かつ回転角度検出ができるように改良した。図7に全方位カメラで撮影したパノラマ画像から特定マークを検出した結果を示す。これにより、屋外の看板に文字情報を表す特定マークを付けた場合にも、ある程度の照明変化にロバストにマークを検出することができるようになった。



(a: テンプレート, b: エッジ, c: 入力画像, d: 結果)

図7 マーク探索の処理の一例

(3) 斜影ひずみを受けている文字列の認識

開発した方法を評価するため2つのデータセット(セットA, B)を用意した。図8, 9に示すように、セットAは平面に書かれた文字列をさまざまな方向から撮影した文字列画像1250枚(5,660文字)、セットBには厚い本を斜めから撮影した画像から手動切り取った文字列画像587枚(7,276文字)と全方位カメラで撮影したパノラマ画像から手動で切り取った文字列画像14枚(112文字)が含まれている。

図10に、セットAに対する方法2(提案法)

および正規化法(提案法)の実験結果を示す。ただし、文字列方向はあらかじめ手動で水平になるように変換している。実験により以下のことを明らかにした。

- ・補正をしない場合や方法1(従来法)に比べて、方法2を採用したほうが、文字認識率が高くなっている。

- ・提案した正規化を方法1, 2と組み合わせることで、認識率がさらに向上している。

図11にセットBに対する方法3(提案法)および正規化法(提案法)の実験結果を示す。

この実験では文字列方向の変換は行っていない。実験より以下のことを明らかにした。

- ・補正をしない場合は全く認識できなかった文字列を、方法3の適用により76%程度の精度で認識できるようになった。

- ・提案した正規化法を方法3と組み合わせることで、認識率がさらに向上している。

図10, 11に示すように、本研究で開発した文字列の補正法、文字の正規化法が斜影ひずみを受けている文字列の認識に対して有効であることを明らかにした。

本研究の成果についてまとめる。

まず、文字認識率を向上させるため、撮影条件を改善する方法と、読み取り対象の文字列にマークをつける方法について検討した。具体的には、全方位カメラ、ステレオカメラ、通常のカメラを用いた撮影システムと、情景画像からマークを高速・高精度に検出する方法を開発した。これらの技術を利用することにより文字認識率の向上が考えられるが、コスト面と手間の面から現実的にどの程度受け入れられるかはさらに検討の余地がある。

次に撮影装置や環境に依存しない問題として、さまざまな条件で撮影された文字列を認識する方法の開発を行った。具体的には斜めから撮影した場合に生じる射影ひずみの問題、曲面や全方位カメラで撮影した場合に生じる湾曲ひずみの問題に注目した。開発した文字列補正・認識法により、上記のひずみを受けた文字列が従来よりも高い精度で認識できるようになった。今後実用レベルにするためには、さらなる認識率の向上を行う必要がある。

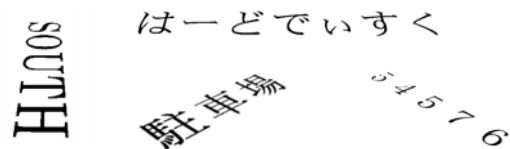
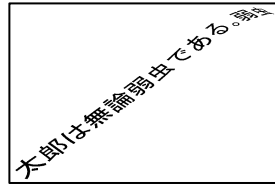
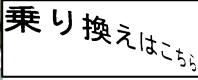


図8 セットA(平面)



(a)



(b)

図9 セットB (a:曲面、b:全方位カメラ)

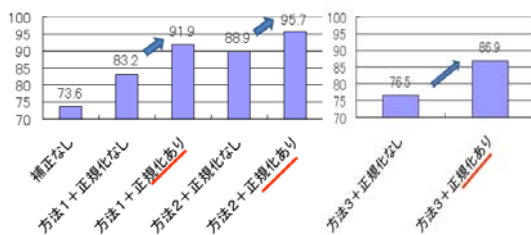


図10 実験結果1 図11 実験結果2

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 林信吾、黒田英夫、志久修、玉井敬一、福本博文、射影ひずみを受けた活字文字列認識の一方法 ～斜めから撮影した文字を認識する～、画像ラボ、査読有り、vol. 19, No. 6, pp. 26-29(2008-6)

[学会発表] (計5件)

- ① 松本祐介・志久修、傾斜文字を認識するための正規化法、第63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-07, (2010.9.26)
- ② 下川創・志久修、湾曲した文字列の認識、第63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-08, (2010.9.26)
- ③ 福本洋平・志久修、他7名、2番目、ダイヤモンド砥粒の固着工具における切れ刃の3次元計測-第2報:画像処理による切れ刃の解析-、第63回電気関係学会九州支部連合大会講演論文集、査読無し、01-2P-03, (2010.9.26)
- ④ 下川創・志久修、湾曲した文字列の認識、第62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-16, (2009.9.29)

- ⑤ 志久修、林信吾、文字列画像に対する射影歪み補正の一手法、第61回電気関係学会九州支部連合大会講演論文集、査読無し、02-1A-07, (2008.9.24)

[産業財産権]

○取得状況 (計3件)

名称: 文字列の傾斜補正方法及び装置
 発明者: 志久修
 権利者: 独立行政法人国立高等専門学校機構
 種類: 特許
 番号: 2005-295237
 取得年月日: 2011.1.26
 国内外の別: 国内

名称: 文字成分抽出装置
 発明者: 志久修
 権利者: 同上、オムロン株式会社
 種類: 特許
 番号: 2004-125907
 取得年月日: 2010.9.17
 国内外の別: 国内

名称: 文字認識方法および文字認識装置
 発明者: 志久修
 権利者: 同上、オムロン株式会社
 種類: 特許
 番号: 2002-331726
 取得年月日: 2009.3.6
 国内外の別: 国内

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・准教授
 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: