

科学研究費助成事業 研究成果報告書

平成 26 年 6 月 13 日現在

機関番号：54601

研究種目：基盤研究(C)

研究期間：2009～2013

課題番号：21500244

研究課題名(和文)文字スポッティングを導入した情景中の文字情報探索システムの開発

研究課題名(英文)Development of the character information search system in the scene image.

研究代表者

松尾 賢一 (Matsuo, Ken'ichi)

奈良工業高等専門学校・その他部局等・教授

研究者番号：10259913

交付決定額(研究期間全体)：(直接経費) 2,500,000円、(間接経費) 750,000円

研究成果の概要(和文)：我々は、情景中で文字情報が存在する高い可能性の領域を見つける文字スポッティング手法を提案した。提案手法を基にして着目領域へのズームアップや接近する行動を実現させることで、情景中の文字情報をOCRで認識可能にさせるシステムを開発を実施した。成果として、移動経路変更により文字情報の認識結果をフィードバックさせる文字駆動型ロボットを開発した。

研究成果の概要(英文)：We aimed at the development of the character information search system in the scene in these expenses. One of the systems to develop is a character information driven robot. This character information driven robot recognizes a character pattern in the scene with OCR, and it follows the meaning of the recognition result, and decides the movement course of the robot. Specially, a character information driven robot must be moved to the character candidate place to recognize a distant character in the scene. The search algorithm of the distant character candidate place in this scene was proposed in this research. Furthermore, the prototype of the character information drive type robot was developed by using the search algorithm.

研究分野：総合領域

科研費の分科・細目：図書館情報学・人文社会情報学

キーワード：文字スポッティング 文字情報探索 情景画像 システム開発

1. 研究開始当初の背景

オフライン文字認識の基本技術は、パターン理解・認識・文書画像処理の分野に属し、様々な研究結果が国内外で報告されている。実際に、F A (Factory Automation) や OCR (Optical Character Reader) が実用レベルに達している状況から見ても、オンライン文字認識同様に文字認識の基本技術を応用された汎用製品例は数多い。図1に一般的なOCRの処理手順を示す。

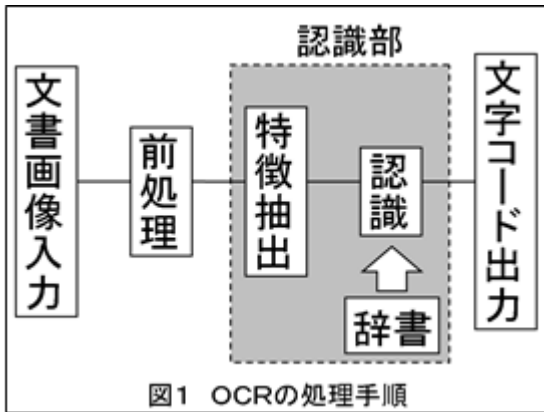


図1 OCRの処理手順

現状のOCRでは、入力となる文書画像に必ず文字パターンが存在していることを前提としている。そして、文字パターンだけを抜き出して切り出す処理、正規化、2値化が前処理で行なわれ、文字パターンは、認識部に送られる。この認識部においては、自由手書き文字のレベルにおいても、認識率が得られるようになったが、これはあくまで辞書に登録された学習パターンに対してであり、辞書にない未知パターンが入力されると、大幅に認識率が低下する。また、OCRでは、文字でないパターンが認識部に入力されても、何らかの文字コードを出力してしまう問題がある。

つまり、学習パターンで、かつ、必ず文字パターンのみ入力されることが、OCRを実用的に利用するための制約条件である。これに関連して、ある程度の制約条件を与えることで、情景中の文字をOCRで認識する研究も行なわれている。

この情景中の文字をOCRで認識するには、前述した前処理における文字の切り出し、正規化の性能が認識結果に大きく影響する。しかも、文書と異なり、文字切り出し以前において文字の存在位置の確定、文字以外の雑音パターンの除去、文字パターンに類似した雑音パターンの排除等の処理も必要となる。これに対して、研究代表者(松尾)は、OCRを利用した情景中の文字情報の認識の実現に向けて、前処理における切り出し以前の処理である「情景画像からの文字パターン抽出手法」を提案し、良好な抽出結果を得ている。その処理手順の概略を図2に示す。

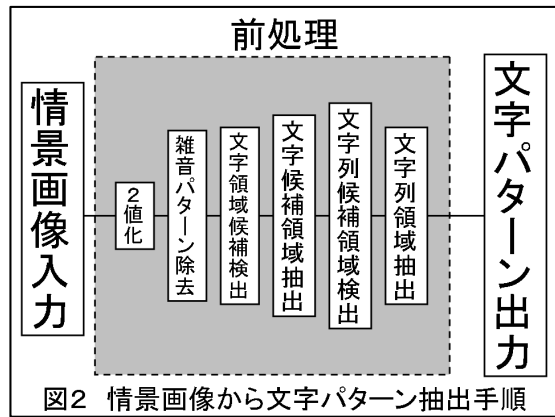


図2 情景画像から文字パターン抽出手順

この手法の提案によって、図1の前処理部分に改良を加えることで、情景画像中の文字をOCRで認識させることを可能にした。しかしながら、この手法においても、OCRの認識を前提とするならば、情景画像中に必ず文字が存在しているという文書画像と同様の制約条件を課す必要がある。それ以外に、抽出精度の面からでは、情景画像の質が出力結果である文字パターンに大きく影響を与えるため、以後の認識部において高い認識率が得られない問題があった。この問題は、他の研究においても同様の(入力された情景画像に対するシステムのロバスト性)問題として結論付けられていることが多い。

ここで、誤認識を増加させる情景画像の質の低下の原因として、入力時のぶれ、焦点ボケ、シェーディング、解像度不足等があげられる。また、高解像度に撮影しても画像の見かけ上、遠方に存在するため文字パターン自体の大きさが微小であるときも誤認識を誘発しやすくなる。

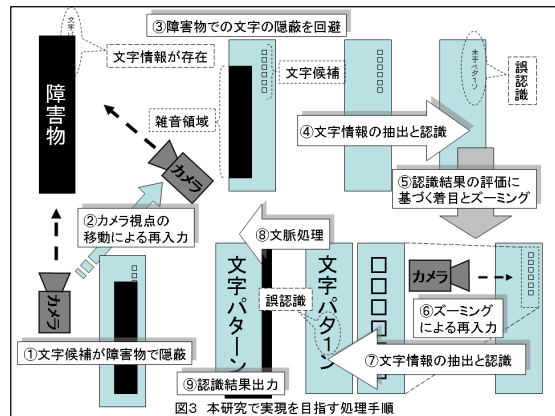


図3 本研究で実現を目指す処理手順

これに対して、情景中に文字情報が存在するか否かの判定も含め、文字情報が存在するであろうと考えられる位置や領域に対して着目する技術の開発の着想に至った。図3に示すような文字情報への着目技術(図中ではからの基本手順であるが、とを独立に処理しても良い)が実現すれば、着目領域に対する認識結果を受けて入力装置(図中ではカメラ)の制御にフィードバックできるため、OCRで誤認識となったパターンに対し

て、ズームングやカメラ視点の変更を行い、画像や映像を再入力することで詳細なパターンを再取得できる。これによって、情景画像の質の影響によるOCRの認識精度の低下を回避させることができる。

## 2. 研究の目的

我々の日常の行動において、何気ない情景中から文字情報を見つけ出し、それを理解して利用することが多い。また、文字情報が障害物や遠方で読むことが困難である場合、視点を変え、見えるところまで移動して文字情報を取得することもある。

本研究では、我々が文字情報を取得する何気ない行動プロセスを、OCRを用いた情景中の文字パターンを認識するシステムに導入することを目的とする。この研究目的に対して、以下の5つの課題について研究を推進する。

(1) 情景に文字が存在しうるか否かを判定する手法を提案する。

(2) 情景中で見かけ上遠方に存在するため認識処理を導入できない文字パターンを精度良く着目する(文字スポッティング)エンジンを開発する。

(3) (2)で着目された文字パターンの認識精度を高めるための入力装置に対する校正(calibration)プロセスを実現する。

(4) OCRで得られた認識結果である文字コードから構成される文字列に対し、後処理となる文脈解析の結果に基づいた認識処理部へのフィードバックを施し、認識結果に対する高精度化を図る。

(5) (1)から(4)の技術を集大成した情景中の文字情報探索システムの開発とその有効性の調査を実施する。

(1)から(5)の研究課題についての基礎的研究は、(1)、(2)において情景中に必ず文字情報が存在している画像においては、すでに良好な結果が得られているが、文字情報を含まない、あるいは、文字パターン自体が画像中の見かけ上遠方に存在する画像に対応できるように改良を加えることが課題となる。また、(3)、(4)については、装置や既存のシステムを流用しながら、(1)、(2)の手法を改良させて得られた結果に基づいて具現化させる課題である。については、(1)から(4)の成果を統合してシステム化し、評価する課題である。研究の目的である課題(1)から(5)までを年次進形で研究を遂行する。

## 3. 研究の方法

まず、研究本経費で導入したパンチルトズームカメラを用いて、研究目的で述べた課題(1)(2)について取り組む。

課題(1)(2)については、情景中で遠方に存在することで、文字パターンか否か判断できないため、即時にOCRが適用できない文字情報に対して、文字情報が存在する可能性の高い候補領域を検出し、候補領域に対して光学ズームによって得た高品質な文字パターンに対してOCRを適用することで、最終的に文字情報がもつ意味情報を得る手法を提案する。

課題(3)は、本経費で導入したカメラ付き(ズーム機能なし)移動用ロボットを用いて、文字情報が存在する可能性の高い候補領域に向かって移動させ、文字情報に接近することで、高品質な文字パターンを得るとともに、遮蔽された文字情報や文字情報を正面から得られない状況下を回避するために、OCRで文字情報の認識を容易にする位置まで移動させるプロセスの実現を目指す。

最後に、課題(4)(5)で、得られた文字情報に対してOCRによって文字情報から意味情報を得ることで、ロボットの移動経路に対してその意味情報を活用させられるかを明らかにすると共に、最終的には文字情報探索システムとして、文字駆動型ロボットの開発を目指す。

これらの課題とは別に、情景中の文字情報が存在する領域が確定できたとき、その文字情報の切出しが困難である状況が数多く見られる。その代表的な事例として、文字同士が接触あるいは重なっている文字の切出しと認識、多色な複雑背景および文字領域の切出しがあげられる。一般的には、これらは、OCRで認識をする以前の前処理が抱える問題となるが、課題(1)(2)の実現を妨げる問題でもある。

この問題についても、課題(1)(2)を実現するための課題として、別テーマとして取り組む。ここで、重ね書き文字の一例として、採点記号を取り上げ、採点記号の分離、抽出、認識を実現する答案採点支援システムの開発、そして、情景中での多色な複雑背景および文字領域を携帯端末カメラを用いて入力を容易にするアプリの開発についても研究を実施する。

## 4. 研究成果

課題(1)では、情景画像中で文字が存在する領域を判定させる手法を提案した。(関連発表: 18, 20, 22)

提案手法は、カメラで得た情景画像に対して、濃度差が顕著である領域を二分し、二分した領域を取り囲む外接矩形を抽出し、その外接矩形の配置状況から文字らしい配置を

評価することで、文字領域候補を抽出する。配置の評価は、文字の普遍的特徴として、外接矩形に対して、面積一致度、高さ一致度、幅一致度、等間隔度、直線度の5つを定義し、これらの5つの尺度を総合して文字列らしさとして外接矩形の配置状態を評価した。

提案手法は、文字情報を含んだ情景画像としてICDAR2003の公開データセットを用いて評価した。その結果、文字列を完全に取得したことを示す文字列領域候補検出率は74.2%であった。今後は、この提案手法をベースとしたアルゴリズムを「文字列候補領域抽出エンジン」として使用することにした。

ここで、情景画像中で文字パターン自体が高解像度で得られる場合、文字列候補領域抽出エンジンで得られた外接矩形に対して、順次OCRを適用して認識結果を得ることで文字パターンがもつ意味情報の取得は可能である。しかしながら、前述したように情景中で見かけ上遠方に存在するため、情景画像中では認識処理を導入できない文字パターンに対しては、その文字パターンが存在する領域に着目する文字スポッティングを実施し、そのスポッティング領域に対して、OCRで認識可能な状態の文字パターンが得られるまでズームアップを施す必要がある。

そこで、課題(2)では、本経費で導入したパンチルトズームカメラを用いて、情景画像に対する文字スポッティングエンジンの開発を目指した。まず、ズームアップが必要な領域を検出するために、画像を局所領域に分割し、メッシュ領域を作成する。この各メッシュは、カメラのズーム倍率を変化することに作成し、このメッシュ内で、濃度差、色差、周波数成分、複雑度の4つの尺度を算出する。

この4つの尺度を文字が存在する判定の評価値に用いて、文字が存在しうるメッシュ領域を決定する手法を提案(関連発表:10,11,14,16,17,19)した。この提案手法では、非文字領域の除去には、有効性が見られるが、文字スポッティング精度においては、メッシュ領域サイズの設定、入力された情景画像の性質によって着目範囲が広範囲になることがあった。

このことから、課題(3)以降において、すべての文字情報をスポッティングするのではなく、文字情報を認識することで得られた意味情報を移動経路にフィードバックさせる(関連発表:9)には、看板のように意味を知らせるための文字情報に現段階では対象を限定することにした。

次に、本経費で導入したOCRを用いて、パンチルトズームと認識精度の関係について調査(関連発表:3,5)した。一般的な室内にアルファベットの文字列が存在する用紙を壁に貼り付け、文字列領域抽出エンジンによって文字情報が存在する位置にカメラを向けさせ、ズームを前後させてOCRを用いて文字パターンの認識実験を実施した。そ

の結果、ズーム倍率によって、認識率が大幅に変化することと、用いるOCRごとに文字パターンのサイズの違いにより認識精度が異なることが明らかになった。

課題(4)(5)では、本経費で導入したカメラ付き移動ロボットを用いて、文字情報が存在しうる場所に自立移動し、カメラで取り込んだ情景内に認識可能な文字パターンが存在するとき、その文字パターンをOCRで認識し、その文字情報の意味を理解し、以後のロボットの移動経路に対してフィードバックを与え行動する文字駆動型ロボットの開発を目指した。

これまでの課題(1)から(3)の知見に基づき、文字パターンに対して接近する距離の決定には、ロボットのカメラで情景を取り込んだ際に、その情景画像から抽出した文字パターンサイズがOCRで最高の認識率を出力する距離となるよう事前に設定し、認識対象の文字パターンは、用紙に印刷されたアルファベット文字列に限定して開発を進めた。

当初は、文字列候補領域抽出エンジンを用いて、文字が存在しうる場所を検出し、ロボットにその領域を着目させて、アプローチ処理させたが、移動時の振動によりロボットが揺れることで着目領域を見失う、あるいは、違う経路に歩行するなどの二足歩行におけるハード側の問題が生じたため、着目領域に向けて直接歩行する方式から、縦歩きと横歩きを併用した直線的な歩行経路方式に改良(関連発表:1,6)した。あと、文字駆動型ロボットに市販のOCRを認識エンジンに用いた場合、切り出された文字パターンに歪が生じた場合に認識率の低下や、誤った認識結果が出力されることで、アプローチ処理が中断することが見られた。

これに対して、OCRでなく認識対象である文字パターンに歪を加えてマルチテンプレート化した辞書を用いた認識処理に変更したところ、認識率の低下が避けられ、良好な文字パターンへのアプローチ処理の実現できた。このことから、文字パターン付近にアプローチした際に、ロボットに微小な移動を加えた時に得られるOCRの認識結果の変化から最終的な認識結果を得るように課題(3)に対する精度向上が今後の課題と言える。

最後に、OCRでの認識を困難にする文字情報として、文字同士が重なったパターンや多色な背景や文字線をもつパターンがあげられる。これらを文字駆動型ロボットの開発での課題として課題(4)に入れ込むには、難易度が高いため、ロボットとは別にOCRを用いた情景中の文字パターンを認識するために解決すべき課題として、接触や重ね書きが多く発生する答案用紙上の採点記号および採点文字の抽出と認識を実現するアルゴリズムの提案(関連発表:12,13,15,21)、ならびに、携帯端末のカメラから取り込んだ

多色な背景や文字線をもつ文字パターンをOCRで認識可能とするアプリの開発(関連発表:2,4,7,8)を実施した。

これまで主にスキャナ等で取り込まれた文書中の文字を読み取り対象としていたOCRを、情景中の文字情報にまで適用範囲を拡張させることを目指した。

現在のOCRの認識精度は、高精度な結果が得られるが、基本的には文字当て技術に過ぎず、入力されるパターンも文字であることが前提である。

本研究では、OCRが文字当て技術であるならば、文字のみを入力させる前処理および文字当てを繰り返し、その認識結果を入力環境にフィードバックさせる一手法を提案することができた。提案手法をベースとして開発中の文字駆動型ロボットについては、発表等で数多くの賛否やご意見・ご助言をいただいた。特に、移動精度や認識精度の向上の可能性の点を除けば、開発コンセプトと今後の開発に対して、期待の声をいただくことが多かった。

今後は、残された課題(4)を早急に解決し、文字駆動型ロボットの完成を目指す。

#### 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 0 件)

[学会発表](計 22 件)

(1)川口誠一, 森下紀明, 松尾賢一: "文字駆動型ロボットの開発~文字列パターンへのアプローチ処理の実現", 2014年電子情報通信学会 総合大会 学生ポスターセッション, 2014年3月20日, 新潟大学 五十嵐キャンパス

(2)福岡秀治, 松尾賢一: "情景画像中の活字文字認識精度を向上させる前処理アプリの開発", 電子情報通信学会関西支部学生会 第19回学生会研究発表講演会 講演論文集, D3-4, 2014年2月28日, 同志社大学 京田辺キャンパス

(3)向井田一平, 松尾賢一: "ズーム倍率変化に伴う文字着目精度の調査", 電子情報通信学会関西支部学生会 第19回学生会研究発表講演会 講演論文集, D2-1, 2014年2月28日, 同志社大学 京田辺キャンパス

(4)福岡秀治, 坂口峻平, 松尾賢一: "携帯端末を用いた文字認識システムの開発 情景中の文字パターンに対する前処理の実現", 2013年 電子情報通信学会 総合大会 学生ポスターセッション, 2013年3月21日, 岐阜大学

(5)植松裕介, 松尾賢一: "パンチルトズームカメラでの文字スポッティングに向けた着目領域判定手法の改良", 電子情報通信学会関西支部学生会 第18回学生会研究発表講演会 講演論文集, D7-5, 2013年3月6日, 大阪電気通信大学 寝屋川キャンパス

(6)森下紀明, 寺脇温晃, 松尾賢一: "文字認識精度を向上させるロボット移動方法の提案", 電子情報通信学会関西支部学生会 第18回学生会研究発表講演会 講演論文集, D6-5, 2013年3月6日, 大阪電気通信大学 寝屋川キャンパス

(7)福岡秀治, 坂口峻平, 松尾賢一: "携帯端末を用いた文字認識システムの開発 情景中の文字パターンに対する前処理の実現", 電子情報通信学会関西支部学生会 第18回学生会研究発表講演会 講演論文集, D6-4, 2013年3月6日, 大阪電気通信大学 寝屋川キャンパス

(8)福岡秀治, 松尾賢一: "携帯端末を用いた手書き文字認識システムの開発" 電子情報通信学会関西支部学生会 第16回学生会研究発表講演会, 2012年3月9日, 立命館大学びわこ・くさつキャンパス

(9)結崎嵐稀, 松尾賢一: "SIFTを用いた情景動画中のガソリン価格看板領域の検出" 2012年電子情報通信学会総合大会学生ポスターセッション, 2012年3月22日, 岡山大学津山キャンパス清水記念体育館

(10)岡田亜沙美, 森下紀明, 松尾賢一: "パンチルトズームカメラを用いた文字スポッティングの実現" 2012年電子情報通信学会総合大会学生ポスターセッション, 2012年3月22日, 岡山大学津山キャンパス清水記念体育館

(11)森下紀明, 松尾賢一: "パンチルトズームカメラを用いた文字スポッティングシステムの開発" 2012年電子情報通信学会総合大会学生ポスターセッション, 2012年3月22日, 岡山大学津山キャンパス清水記念体育館

(12)宮本弘志, 松尾賢一: "ストローク交差箇所方向性に着目した採点記号の分離抽出" 平成22年電気関係学会関西連合大会ポスターセッション 2010年11月14日, 立命館大学びわこ・くさつキャンパス

(13)宮本弘志, 松尾賢一: "重畳採点記号に対する分離抽出の試み" 電子情報通信学会関西支部学生会 第16回学生会研究発表講演会, 2011年3月1日, 大阪府立大学中百舌鳥キャンパス

(14) 森下紀明, 松尾賢一: "パンチルトズームカメラを用いた文字スポッティングシステムの開発" 電子情報通信学会関西支部学生会第16回学生会研究発表講演会, 2011年3月1日, 大阪府立大学中百舌鳥キャンパス

(15) 宮本弘志, 松尾賢一: "重畳採点記号の分離抽出と認識" 2011年電子情報通信学会総合大会, 2011年3月16日, 東京都市大学世田谷キャンパス

(16) 寺脇温晃, 松尾賢一: "文字の普遍的特徴を用いた情景画像中の文字スポッティング" 2011年電子情報通信学会総合大会, 2011年3月16日, 東京都市大学世田谷キャンパス

(17) 岡田亜沙美, 寺脇温晃, 森下紀明, 松尾賢一: "パンチルトズームカメラを用いた文字スポッティング手法" 2011年電子情報通信学会総合大会, 2011年3月14日, 東京都市大学世田谷キャンパス

(18) 寺脇温晃, 松尾賢一: "情景画像中の文字スポッティングに向けた"文字らしさ"の検討" 電子情報通信学会2010年総合大会学生ポスターセッション, 2010年3月16日, 東北大学 川内キャンパス

(19) 岡田亜沙美, 越田直輝, 寺脇温晃, 松尾賢一: "文字スポッティングに向けたパンチルトカメラ制御プログラムの開発" 電子情報通信学会関西支部学生会 第15回学生会研究発表講演会, 2010年3月10日, 大阪大学 吹田キャンパス

(20) 寺脇温晃, 松尾賢一: "情景画像中の文字スポッティングに向けた"文字らしさ"の検討" 電子情報通信学会関西支部学生会 第15回学生会研究発表講演会, 2010年3月10日, 大阪大学 吹田キャンパス

(21) 宮本弘志, 松尾賢一: "ストロークの方向性を用いた重ね書き採点記号の分離抽出" 電子情報通信学会関西支部学生会 第15回学生会研究発表講演会, 2010年3月10日, 大阪大学 吹田キャンパス

(22) 寺脇温晃, 越田直輝, 松尾賢一: "文字領域検出に向けた文字の普遍的特徴分析" 平成21年電気関係学会関西支部連合大会, 2009年11月7日, 大阪大学吹田キャンパス

〔図書〕(計0件)

〔産業財産権〕  
出願状況(計0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0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  
<http://www.info.nara-k.ac.jp/~matsuo/MYSELF/myself.html>

#### 6. 研究組織

(1) 研究代表者  
松尾 賢一 (MATSUO Ken-ichi)  
奈良工業高等専門学校・情報工学科・教授  
研究者番号: 10259913

(2) 研究分担者  
( )

研究者番号:

(3) 連携研究者  
( )

研究者番号: