

氏名（本籍）	太田 貴大（和歌山県）
学位の種類	博士（工学）
学位授与番号	甲 第53号
学位授与日付	平成24年3月23日
専攻	システム工学専攻
学位論文題目	模倣に基づく識別器の高速化と工業用OCRへの応用
学位論文審査委員	（主査）教授 和田 俊和 （副査）教授 呉 海元 准教授 中村 恭之

論文内容の要旨

1. はじめに

FA 分野では、製造ラインの高速化、品質管理の厳格化が進んでいる。このため、産業用インクジェットプリンター（IIJP: Industrial Ink Jet Printer）による印字の品位管理が徹底されるようになり、近年の OCR（Optical Character Reader, 以下 IIJP-OCR）では、高速かつ高精度な能力が要求されている。IIJP-OCR では、高次元な特徴を対象とするために計算コストを要し、低速ではあるが高精度な識別ができるもの、単純な計算によって高速に識別できるが低精度なものが存在するが、高速性が優先されるために後者の選択を余儀なくされる傾向がある。そこで本論文では、1) 高精度であるが低速な識別器の入出力関係を線形回帰木によって模倣し、これを識別器として用いることで元の識別器を高速化する、2) 識別器の高速性を利用することでトップダウン型の探索手法を構築し、ボトムアップ型の従来手法における文字切り出し性能を向上させる、3) 局所特徴を用いたボトムアップ型の文字切り出しとトップダウン型の文字探索を融合することで、従来の IIJP-OCR では対応できなかった、文字が二次元的変動をもつ、かつ背景に照明変化やノイズがある対象に対して、現在の IIJP-OCR に求められる速度、精度を実現する手法を提案する。

本論文の構成は図 1 に示す通りであり、第 1 章では、従来の IIJP-OCR における問題点をあげ、解決法を考察することで研究指針について解説する。第 2 章では、まず種々の識別器について精度を比較し、最も精度が高い識別器を選択する。つぎに、PaLM-tree（Partially Linear Mapping tree）を用いた非線形画像学習によって識別器の模倣を行うことで、識別器を高速化する手法について述べる。また、学習データの自動生成方法について述べ、学習データの収集コストの低減だけでなく、パターンの摂動によって識別精度が向上することを確認する。第 3 章では、高速な識別器を用いたトップダウン型文字認識手法、CCS（Classification based Character Segmentation）の構築方法について説明する。さらに、CCS の前処理として、局所特徴を用いたボトムアップ型のセグメンテーションによって予め探索候補領域を抽出することで、文字の二次元的変動に頑健な文字認識を実現する手法 MCCS（Modified CCS）を提案する。最後に、第 4 章で本研究について総括した後、今後の研究について述べる。

以下では、本論文における提案手法の概要を説明する。

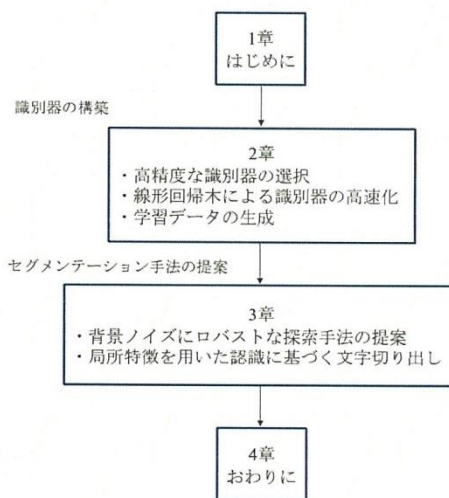


図 1. 論文の構成

2. 高精度な識別器の選択

ドット文字は手書き文字とは異なり、定型のマトリクスにドットを配置するという性質上、フォントサイズが小さくなるほど表現力が低下し、類似クラスが増加する。このような類似クラスの誤認識を軽減するため、識別器の選定および最適化を行った。実験では、最近傍識別法による従来手法、単純類似度法、複合類似度法、混合類似度法の比較から、混合類似度法が最も高精度であった。さらに、混合類似度法の拡張成分で計算する次元数を固定すること、平均パターンを IIJP の基準パターンとすることで、クラスごとに異なるクラス内変動の大小によるペナルティの不安定さをキャンセルした。次章では、この混合類似度法を高速化の対象とする。

3. 線形回帰木による識別器の高速化

識別器の振る舞いを線形回帰木によって模倣し、元の識別器と置き換える。本論文では、この手法を Classifier Molding と呼ぶ（図 2）。本研究では回帰木を構築するために、取り扱う学習データの次元数、クラス数に制約が無い点で自由度が高い PaLM-tree を採用した。計算モデルとして、線形回帰木は一般的な識別器よりも高速である。これは、識別計算が入力ベクトルの二分探索と、回帰係数を用いた線形計算による出力ベクトルの算出のみであるためである。ただし、構築した線形回帰木を識別

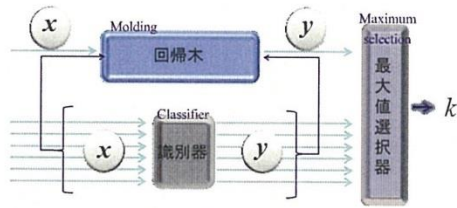


図 2. Classifier Molding

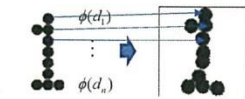


図 3. 変動パターンの生成

器として用いるためには、推定される出力において、誤差の総和が小さいだけでなく、要素間の大小関係が保存されていることが重要であるため、誤差評価方法を改良した。実験では、元の識別器に対して、1,500 倍と大幅な高速化を実現し、本手法の有効性を示した。

4. 学習パターンの生成

本研究では、IIJP の基準フォントパターンに対し、正規乱数に従った変動をドット単位またはパターン全体に与えることで摂動をシミュレートし、パターンを生成する方法をとった (図 3)。実験では、比較に用いた識別器全てにおいて実画像を用いるよりも精度の向上がみられ、データ収集コストの低減だけでなく精度の向上にも寄与していることが判り、生成学習の有効性が確認できた。

5. 背景ノイズにロバストな探索手法の提案

従来の手法では画素の連結成分に依存した抽出法をとるため、背景ノイズや文字の連結などによって図 4 (a)のようにセグメンテーションエラーが発生する。この問題を解決するために、Classifier Molding の高速性を利用し、認識に基づいて文字を探索する手法、CCS の開発を行った。CCS は、探索位置から得られる類似度の合計を最大化することを目的とし、ヒューリスティック関数の導入によって、一定の値を下回らないことを保証しつつ文字を探索することで、類似度の低い位置での探索を枝刈りすることができる。CCS は 1) 最優良探索による位置の抽出、2) バックトラッキングによる位置の補正、3) タイムアウトによる時間制限、によって構成され、タイムアウト時に得られた情報を文字位置として出力する。実験では、従来手法の適用が不可能な画像から、図 4 (b)のように文字が抽出できるだけでなく、NTSC ビデオレート以下での高速処理が実現できることを確認した。

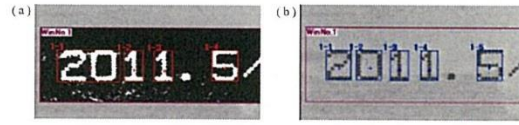


図 4. 文字切り出し結果: (a) 二値化&ラベリング, (b) CCS

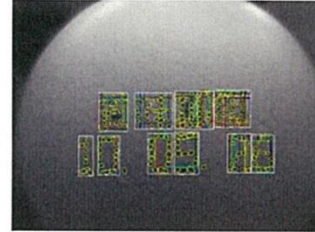


図 5. MCCA による文字切り出し (卵殻表面の印字)

6. 局所特徴を用いた認識に基づく文字切り出し

変形やノイズに対して頑健であることが知られている局所特徴は、検出した特徴点が記述する局所記述子を用いたマッチングによる特徴点の対応付けから、参照画像を入力画像から探索するために用いられることが多い。本研究では、特徴点の位置関係を利用し、距離に基づいた特徴点のグルーピングによって得られる領域の最小矩形を文字領域候補として抽出することとした。つまりボトムアップ的に文字の位置を推定するため、文字のサイズや位置などの情報が既知である必要はない。なお、特徴検出器には本研究で開発した Fast-Hessian-Affine Regions を用いた。つぎに、抽出した文字領域候補に対して、CCS を実行することでトップダウン的に文字を探索し、真の文字領域を得る (図 5)。本論文では、以上による文字探索を MCCA と呼ぶ。実験では、文字の二次元的変動や照明変動をもつ様々なテストデータを用いて MCCA の有効性を明らかにした。

7. おわりに

トップダウン型文字認識手法 MCCA によって、文字の幾何学的変形や移動、背景ノイズに頑健かつ高速な文字認識を実現することができた。このままでも十分に実用価値があるが、製品として適用範囲の拡大を考えると、IIJP の基準フォントパターンからではなく、学習パターンからの摂動パターン生成方法の検討や、未学習データに対して上昇してしまう類似度の抑制方法の検討、さらに、ヘッセ行列式の値に対する閾値をコントラストに応じて動的に求め、濃淡レベルの変化に対して頑健な局所特徴検出器の構築法について検討する予定である。

論文審査の結果の要旨

標記論文内容に関して1月30日に公聴会を開催し、約1時間にわたり、論文内容について太田氏からの発表が行われた。論文内容は、高精度かつ高速な光学文字読取り（OCR）アルゴリズムを工業用インクジェットプリンタの印字確認の為に開発するという研究である。このような用途で使われるOCRアルゴリズムとしては、高速な製造ラインでの使用に耐えうる高速性と、高い精度を持つものは現存せず、高精度なものは低速であり、高速なものは低精度であるという問題点がある。本研究では、高精度で低速な識別器を一旦作成し、その挙動を線形回帰木で模倣するというアイデアで、この問題を解決する方法を示している。また、いかに高速な識別器が実現できても、画像から文字領域が適切に切り出されなければ正しく文字読み取りが行えない。この問題を解決するため、ボトムアップとトップダウン処理を融合し、A*アルゴリズムに基づき、文字らしさを最大化する画像領域を実時間で求める方法を示している。これら2つの内容は、審査委員会において、パターン認識、画像認識の方法として斬新であると同時に高い実用性を持つものとして評価された。また論文の完成度も高く、学会発表等の実績も合格のレベルに達していると判定された。

最終試験の結果の要旨

上記公聴会の後、画像処理、微分幾何および組み合わせ最適化に関する口頭試問を行い、全ての質問に完全に回答することができた。公聴会における発表内容、質疑に対する応答、口頭試問のいずれにおいても、明らかに合格基準を上回る結果を出しており、博士の学位に値することを審査委員全員で合意した。