

科学研究費助成事業 研究成果報告書

平成 26 年 5 月 31 日現在

機関番号：17102

研究種目：基盤研究(B)

研究期間：2011～2013

課題番号：23300072

研究課題名(和文)部分は全体を語る 断片化による手書き文字認識ブレイクスルー

研究課題名(英文) Part-based character recognition

研究代表者

内田 誠一 (Uchida, Seiichi)

九州大学・システム情報科学研究科(研究院・教授)

研究者番号：70315125

交付決定額(研究期間全体)：(直接経費) 15,800,000円、(間接経費) 4,740,000円

研究成果の概要(和文)：文字を断片化して認識するという、従来の文字認識技術からみれば逆転の発想とも言える手法により、手書き単文字(約95%程度の認識率、以下同)だけでなく、手書き文字列(65%)、そして飾り文字フォント(75%)が機械認識できることを実証した。断片化して認識するという事は、文字の全体的な構造特徴を利用しないことを意味している。従って、文字の全体的構造が変化してしまったような大きな変形を伴った文字でも認識できる。例えば、手書き文字列の場合、隣接文字が接触してしまったようなケースでも認識できる。この成果に基づき、さらに一般的な全体的構造を併用することで、極めて頑健な認識手法が実現できると期待される。

研究成果の概要(英文)：In this research project, part-based methods for various character recognition have been proposed. Different from previous trials of character recognition, it can be applied to different character recognition tasks, such as handwritten characters, scene characters, degraded document images, etc. The experimental results proved that in handwritten character recognition, the various fonts in scene character and document image recognition. The part-based methods are also robust against the degradation of the character images, which is also proved by the experiments on degraded test sets. Moreover, the part-based method can be applied to character string and document image recognition without any specific segmentation process. In summary, the part-based methods are the good solutions for the problems in character recognition.

研究分野：総合領域

科研費の分科・細目：情報学・知覚情報処理・知能ロボティクス

キーワード：手書き文字認識 局所特徴 手書き文字列認識 セグメンテーション 情景内文字認識 パターン認識

1. 研究開始当初の背景

(1) 手書き文字認識に対するニーズは劇的に拡大している。例えば、米国 Evernote 社や日本のデータエントリ企業のように、ノート等の手書き文字を含む画像をアップロードしておけば、その中の文字をタグとして自動付与してくれるサービスが登場し、大ヒットしている。画像の洪水の中で、最も信頼のおける検索キーとして文字情報に着目した訳である。他にも、カメラ・タブレット PC・各種電子ペンなどハンディな入力デバイスの爆発的普及があり、ニーズの拡大要因は多い。手書き文字は人類にとって最も身近で生成容易な情報伝達メディアであり、ICT の進歩の中でその電子的な管理・検索を望むのは当然とも言える。

(2) ところが手書き文字認識は未完成であった。例えばノートの手書きを認識できるようなレベルの日本語 OCR は未だに存在しない。データエントリ企業では、今でも大量のオペレータ(言わば「人間 OCR」)に手書き文字を読ませて計算機に入力させている。

(3) 従って、従来とは全く違った方法論によるブレイクスルーが必要である。研究し尽された従来の方法論の延長線上には、より高い自由度を求める現在のニーズに応えられる解は期待できない。従来の常識に拘泥されない、新たな発想こそがブレイクスルーを実現する。

2. 研究の目的

(1) そこで本課題では、従来とは全く異なる方法論として、文字の大局的構造を破棄し、局所部分に断片化した上で認識することを目的とした。

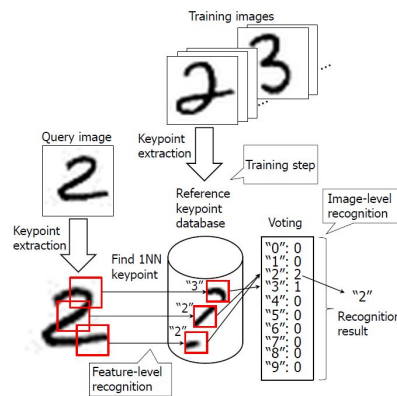
(2) 従来の常識を真っ向から否定する方法論には、従来法のボトルネックを一気に解消できるポテンシャルを持つ。

具体的な第一のブレイクスルーは、文字変形に対し比類なき頑強性を有する点である。「比較的きれいに書かれた」ことを前提としていた従来方式では、当然ながら多様な手書き変形に対応できないという限界に直面する。これに対し part-based 認識では、文字の大局的構造すら崩れたような文字(手書き文字や飾り文字フォント)も認識できる可能性を持っている。

第二のブレイクスルーは、明示的な文字切り出しが不要(セグメンテーションフリー)になる点である。文字切り出しは、従来から文字認識研究者を悩ませてきた問題である。単文字単位では認識できても、文字列では文字切り出しが必要となり一気に難しくなる。罫線上真っすぐ書かれた文字列ならば recognition-based segmentation のような方法もあるが、現在必要とされている「自由手書き」については無力である。これに対し part-based 認識ならば、適当に断片を与えれば断片単位で認識できるので、セグメンテーションをせずとも認識結果が与えられる。

3. 研究の方法

(1) 手書き単文字画像の part-based 認識:



上図が最も基本的な part-based 文字認識の原理である。入力画像を断片化し、断片(part)毎に認識を行い、認識結果を得る。そして複数の認識結果を最終的に多数決統合することで、入力画像全体としての認識結果を得る。この統合処理については、単純多数決以外にも様々なバリエーションが考えられる。これらのうち有望な方式を採用し、性能比較を行う。また文字画像に対する各種の変形・外乱について、どの程度頑強性を示せるかについて、定量的に評価する。

(2) 手書き文字列画像の part-based 認識: 単文字と異なり、複数の文字が並んで存在するため、単純な多数決では済まない。如何に断片毎の認識結果を組み合わせたかが課題となる。そこで局所認識結果に対する recognition-based segmentation を実施する。既に part に分解されているため、厳密なセグメンテーションを求めずとも認識結果が得られる点が従来法に対する優位性となる。

(3) part-based オンライン文字列認識: 断片化さえできれば、part-based 認識は、オンライン文字認識にも問題なく適用できるはずである。オンライン文字認識の場合、運筆情報が得られるために、特徴記述および認識結果統合の際に利用できる。まず単文字(農工大 KuchibueDB 利用)で実施し、その後文字列に適用する。

(4) 飾り文字フォントの part-based 認識: 様々にデザインされた飾り文字においても、時として文字の大局的構造は大きく変動し、時に破壊されている。従ってその認識は困難である。実際、情景内にはこうした飾り文字が多数存在するため、情景内文字認識タスクは非常に困難なものと考えられている。そこで、part-based 認識の考え方を飾り文字フォントにも適用し、その性能を定量的・定性的に評価する。

4. 研究成果

(1) 手書き単文字画像の part-based 認識:

part-based 文字認識の理論構築および part 特徴空間の解析については、大量の part

の分布の解析ならびに part 間の類似度の吟味、および各 part レベルの認識結果の統合法の吟味を行った。そしてその分布の特性を活かして、不要な part の除去による高速化および認識率向上、ならびにクラス固有 part の発見を行った。part が為す分布はクラス間で大きくかつ多様にオーバーラップしているという予想を、実験を通して実証した。

統合法を変えながら単文字認識実験も行った。単純な多数決型統合では 93.6% の認識率であったが、class-distance 法と呼ぶ別の統合評価手法により 97.9% という認識率が得られることを示した。この認識率は文字全体の構造を利用した認識法と同等であり、従って文字を断片化しても十分認識できること証明ともなっている。

また変形耐性を評価するべく、上下の一部が欠けた文字の認識実験も行った。その結果、通常の文字認識手法が 50% 以下に認識率が低下したのに対し、part-based 手法では依然として 80% 程度を維持できることがわかった。

(2) 手書き文字列画像の part-based 認識: 厳密なセグメンテーション不要な文字列認識に着手した。文字列から局所特徴点群を抽出し、各特徴点について単文字の場合と同様に最近傍認識を行い、その結果を、文字列画像平面内で部分統合する方法を試みた。その結果、接触文字や分離文字、激しい傾きなど、従来の文字列認識法では困難であった対象についても、65% 程度の場合において正しく構成文字を認識できることを示した。

(3) part-based オンライン文字列認識: オンライン文字認識および文字傾き補正法に適用、実験を通して、その性質および性能を定量的・定性的に評価した。オンラインの場合、文字の局所領域は単純な線分になるために、文字の特徴が失われ、認識率の低下が起こる一方で、筆順変動などの大局的な変動について著しく頑健と考えられる。漢字を対象とした実験により、この頑健性を定量的に評価した。その結果初等漢字 80 文字種に対する認識結果はおよそ 60% となった。前述のように単純な線分集合として表現されているにも関わらず、これだけの認識率が得られたということは、part-based アプローチを有望視させるものであった。また詳細解析により、筆順変動についてほとんど影響を受けないことも確認された。

(4) 飾り文字フォントの part-based 認識: 前述の通り、情景内に存在する飾り文字や多様なフォントも、パターンとしては基本の文字形状を大きく変形したものであると言える。これに対し、part-based 文字認識技術の枠組を用いれば、こうした大規模な変形が生じたとしても、部分画像に分解してしまえば、変形の悪影響は該当部分を含んだ少数の部分画像に限定される。すなわち残り大部分の部分画像については、変形のなかった場合と同様の認識が可能であり、従って頑健性が担保

される。さらに飾り文字の part をあらかじめ準備しておくことで、それを組合せてできた別の飾り文字も認識できる可能性が広がる。実際、7000 種程度のフォントを用いた英文字認識実験の結果、通常の OCR(56%) よりも高い認識率(73%)が得られることが確認された。すなわち誤認識を 2/3 にまで低減できた。こうした情景内文字の認識は、携帯カメラの普及により広く重要視されるに至っており、本研究もその進展に寄与する貢献を果たしたと言える。

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 11 件)

Muhammad Imran Malik, Marcus Liwicki, Andreas Dengel, Seiichi Uchida and Volkmar Frinken, Automatic Signatures Stability Analysis And Verification Using Local Features, Proceedings of the 14th International Conference on Frontiers in Handwriting Recognition, 採録決定, 査読有

Song Wang, Seiichi Uchida, Marcus Liwicki, Part-Based Recognition of Arbitrary Fonts, Proceedings of The 12th International Conference on Document Analysis and Recognition, pp.170-174, 2013, 査読有

Takafumi Matsuo, Song Wang, Yaokai Feng, Seiichi Uchida, Exploring the Ability of Parts on Recognizing Handwriting Characters, Proceedings of the 16th International Graphonomics Society Conference, pp.66-69, 2013, 査読有

Soma Shiraishi, Yaokai Feng, Seiichi Uchida, Skew Estimation by Parts, IEICE Transactions on Information & Systems, vol.E96-D, no.7, pp.1503-1512, July 2013.

Song Wang, Seiichi Uchida, Marcus Liwicki, Yaokai Feng, Part-Based Methods for Handwritten Digit Recognition, Frontiers of Computer Science, vol.7, no.4, pp.514-525, 2013, doi:10.1007/s11704-013-2297-x, 査読有

Song Wang, Seiichi Uchida, and Marcus Liwicki, Part-Based Method on Handwritten Texts, Proceedings of the 21st International Conference on Pattern Recognition, pp.339-342, 2012, 査読有

Wang Song, Marcus Liwicki, and Seiichi Uchida, Toward Part-based Document Image Decoding, Proceedings of The 10th IAPR International Workshop on Document Analysis Systems, pp.266-270, 2012, 査読有

Soma Shiraishi, Yaokai Feng, Seiichi Uchida, A Part-Based Skew Estimation

Method, Proceedings of The 10th IAPR International Workshop on Document Analysis Systems, pp.185-189, 2012, 査読有

Soma Shiraishi, Yaokai Feng and Seiichi Uchida, Part-Based Skew Estimation for Mathematical Expressions, Proceedings of The International Workshop on Digitization and E-Inclusion in Mathematics and Science, pp.33-41, 2012, 査読有

Song Wang, Seiichi Uchida and Marcus Liwicki, Look Inside the World of Parts of Handwritten Characters, Proceedings of The 11th International Conference on Document Analysis and Recognition, pp.784-788, 2011, 査読有

Song Wang, Seiichi Uchida and Marcus Liwicki, Comparative Study of Part-Based Handwritten Character Recognition Methods, Proceedings of The 11th International Conference on Document Analysis and Recognition, pp.814-818, 2011, 査読有

〔学会発表〕(計12件)

Muhammad Imran Malik, Marcus Liwicki, Andreas Dengel, Seiichi Uchida and Volkmar Frinken, Automatic Signatures Stability Analysis And Verification Using Local Features, The 14th International Conference on Frontiers in Handwriting Recognition, Crete Island, Greece, 2014年9月, 発表予定

Song Wang, Seiichi Uchida, Marcus Liwicki, Part-Based Recognition of Arbitrary Fonts, The 12th International Conference on Document Analysis and Recognition, Washington DC, USA, 2013年8月26日

Takafumi Matsuo, Song Wang, Yaokai Feng, Seiichi Uchida, Exploring the Ability of Parts on Recognizing Handwriting Characters, The 16th International Graphonomics Society Conference, Nara, Japan, 2013年6月12日

松尾崇史, フォン ヤオカイ, 内田誠一, part-based オンライン文字認識における特徴表現に関する検討, 電子情報通信学会パターン認識・メディア理解研究会, 電通大, 東京, 2013年3月15日

Song Wang, Seiichi Uchida, and Marcus Liwicki, Part-Based Method on Handwritten Texts, The 21st International Conference on Pattern Recognition, Tsukuba, Japan, 2012年11月12日

Wang Song, Marcus Liwicki, and Seiichi Uchida, Toward Part-based Document Image Decoding, The 10th IAPR International Workshop on Document Analysis Systems, Gold Coast, Australia, 2012年3月27日

Soma Shiraishi, Yaokai Feng, Seiichi Uchida, A Part-Based Skew Estimation Method, The 10th IAPR International Workshop on Document Analysis Systems, Gold Coast, Australia, 2012年3月27日

Soma Shiraishi, Yaokai Feng and Seiichi Uchida, Part-Based Skew Estimation for Mathematical Expressions, The International Workshop on Digitization and E-Inclusion in Mathematics and Science, Tokyo, Japan, 2012年2月8日

松尾崇史, フォン ヤオカイ, 内田誠一, part-based オンライン文字認識の試み, 電子情報通信学会パターン認識・メディア理解研究会, 長崎大, 長崎市, 2011年11月24日

白石壮馬, フォン ヤオカイ, 内田誠一, 局所特徴に基づくスキュー補正, 電子情報通信学会パターン認識・メディア理解研究会, 長崎大, 長崎市, 2011年11月24日

Song Wang, Seiichi Uchida and Marcus Liwicki, Look Inside the World of Parts of Handwritten Characters, The 11th International Conference on Document Analysis and Recognition, Beijing, China, 2011年9月20日

Song Wang, Seiichi Uchida and Marcus Liwicki, Comparative Study of Part-Based Handwritten Character Recognition Methods, The 11th International Conference on Document Analysis and Recognition, Beijing, China, 2011年9月20日

〔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://human.ait.kyushu-u.ac.jp/~uchida/mybib.html>

6. 研究組織

(1) 研究代表者

内田 誠一 (UCHIDA, Seiichi)
九州大学大学院・システム情報科学研究
院・教授
研究者番号：70315125

(2) 研究分担者

馮 堯楷 (FENG, Yaokai)
九州大学大学院・システム情報科学研究
院・助教
研究者番号：60363389

(3) 連携研究者

Marcus Liwicki (LIWICKI, Marcus)
ドイツ人工知能研究所，上級研究員