

黄瀬浩一

公立大学法人 大阪府立大学 大学院工学研究科・教授

文字・文書メディアの新しい利用基盤技術の開発と
それに基づく人間調和型情報環境の構築

§1. 研究実施体制

(1) 黄瀬グループ

①研究分担グループ長:黄瀬 浩一 (大阪府立大学大学院工学研究科、教授)

②研究項目

- (a) ベース認識装置開発
- (b) 大規模データベース構築
- (c) 実時間文字認識
- (d) 実時間文書画像検索
- (e) 全方位認識
- (f) 検索に基づく Reading-life log
- (g) 文字・文書メディアに対する拡張現実

(2) 大町グループ

①研究分担グループ長:大町 真一郎 (東北大学大学院工学研究科、教授)

②研究項目

- (a) ベース認識装置開発
- (b) 大規模データベース構築
- (c) 付加情報に基づく文字切り出し・認識
- (d) 全方位認識

(3) 内田グループ

①研究分担グループ長:内田 誠一 (九州大学大学院システム情報科学研究院、教授)

②研究項目

- (a) ベース認識装置開発
- (b) 大規模データベース構築
- (c) 認識に基づく Reading-life log

§ 2. 研究実施内容

(文中に番号がある場合は(3-1)に対応する)

(1) 黄瀬グループ

(a) ベース認識装置開発

局所特徴量と最近傍探索を用いる文字認識手法の開発を進めた。特に今年度は高速化に注力した。具体的には、認識対象文字の認識困難さに応じて適応的に認識する手法を提案した。この手法は Anytime Algorithm の性質を満たす。実験により、同一の認識結果を得るために必要な処理時間が従来よりも小さくなることが確認できた。

(b) 大規模データベース構築

大量のデータに少ない労力で正解ラベルを付与するために、少数のラベル付きデータを用いて識別器の性能向上と同時に自動ラベル付けを実現した。従来法と異なり、文字画像を手動で切り出していない撮影したままの画像に適用可能である。また、スペイン、バルセロナの Computer Vision Center と共同で動画中のアルファベットのデータベースを作成している。このうち一部は既に ICDAR 2013 Robust Reading Competition の一部として公開済みである。

(c) 実時間文字認識

(a)のベース認識装置の項目で述べた Anytime Algorithm を日本語(漢字)認識の手法に適用し、高速化を図った。

(d) 実時間文書画像検索[A-1]

1億ページのデータベースに対して、26.8 ms/query、検索精度 98.7%を達成した。これは、目標(200ms/query, 90%以上)を大幅に上回るものである。

(e) 全方位認識

全方位認識で重要となる低解像度画像への対処を行った。具体的には、低解像度画像に対しても有用な局所特徴量の開発を試みた。その結果、同一フォントで幾何歪みがない場合であれば、比較的高速に認識可能なことが分かった。

(f) 検索に基づく Reading-life log[A-2],[A-3]

文書画像検索に基づく Reading-life log の応用として、万語計(ユーザが読んだ単語数を推定するシステム)、視線などのセンサー情報を用いて文書を読んでいるかどうか、読んでいる場合にはどのジャンルの文書を読んでいるかを推定するシステム、文書の理解度を推定するシステムを構築した。また、Reading-life log とペアになるアイデアとして Writing-life log を提案した。

(g) 文字・文書メディアに対する拡張現実

アイトラッカにヘッドマウントディスプレイを接続し、ユーザが読んでいる文書に対する補足情報を実時間で提示するシステムを構築した。

(2) 大町グループ

(a) ベース認識装置開発

動画像を用い、統計的手法により低解像度文字を高精度に認識する手法を開発した。相互部分空間法を基本とし、認識結果の候補が誤っている可能性が高い場合にのみ類似したカテゴリとの差異に着目した認識を行なうことで精度を高めることを可能とした。

(b) 大規模データベース構築

フォントの自動合成による大規模データベース構築に引き続き取り組んでいる。あるフォントと同じ特徴を持つフォントを自動的に合成するための方針について検討した。

(c) 付加情報に基づく文字切り出し・認識

付加情報を利用することで文字を高精度に認識するための方法として、色情報を活用する手法を検討した。人間が視認しづらく、かつカメラで読み取れる色の変化を与えるために、YIQ 表色系の色差信号 Q のグラデーションを変えることにより文字の言語情報を埋め込み、ある程度の精度で情報の取り出しが可能であることを確認した。また、文字フォントの自動構築法についても前年度に引き続き検討している。文字の骨格をグラフで表し、グラフのエッジ部分にパーツを貼り付けることでフォントを構築する手法を提案した。

(d) 全方位認識

全方位カメラとパーソナルコンピュータを用いたプロトタイプの実験機を用い、全方位認識の高速化・高精度化に取り組んだ。全方位カメラを用いる場合取得される画像が大きいことから、高速処理が可能なエッジを利用して文字領域を抽出し、その後文字領域か否かを検証することで高精度化を達成する手法を検討した。文字領域か否かの検討に文字領域マップとエッジを活用する手法を提案し、アルゴリズムの見直しにより高速化を図った。

(3) 内田グループ

(a) ベース認識装置開発[C-4, 12]

環境内の多様な文字を高精度に検出し、認識するためのベース認識手法について、「部品に基づく文字検出・認識」、「多重仮説に基づく文字検出・認識」、「誘目性を用いた文字検出」、「位相構造に基づく文字検出」、の4点について検討し、いずれの利点も実証した。いずれも情景内文字の全く異なる性質に着目した検出・認識法であり、組合せ効果にも期待している。

(b) 大規模データベース構築[C-8]

Ground-truth 付きシーン文字画像データベースを作成した。世界中で標準的に用いられている ICDAR2011 Robust Reading Database を含む一般情景画像 3000 枚について、画素レベルの Ground-truth(各画素が文字領域に属するか否か)を付与した。一部は、ICDAR2013 の情景内文字検出・認識コンペの課題画像セットとして利用される。このデータセットは現時点で世界最大規模である。また大規模文字データセットの統計的性質をネットワーク解析技術により明らかにする研究も同時並行して多数進めている。

(c) Reading-life log

Reading-life log の実現を目的として、認識ベース Reading-life log のプロトタイプ実装を行った。当初アイトラッカで視線付近の文字画像を得るシステムを構築していた。しかし、

視線位置の精度が不十分で、かつ OCR 処理に耐えうる画像が得られなかったため、今年度後半からはハイスピードヘッドマウントカメラにより顔方向の画像を得ることとした。結果は良好で、90%以上の認識率を得るに至った。

§3. 成果発表等

(3-1) 原著論文発表

● 論文詳細情報

- [A-1] 竹田 一貴, 黄瀬 浩一, 岩村 雅一, “大規模文書画像検索のためのメモリ効率と識別性の向上”, 電子情報通信学会論文誌 D, J95-D, 7, pp.1491-1494 (2012-7).
- [A-2] Megumi Chikano, Koichi Kise, Masakazu Iwamura, Seiichi Uchida, Shinichiro Omachi, “Recovery and localization of handwritings by a camera-pen based on tracking and document image retrieval”, Pattern Recognition Letters (Accepted)(DOI: 10.1016/j.patrec.2012.10.003).
- [A-3] Koichi Kise, Riki Kudo, Masakazu Iwamura, Seiichi Uchida, Shinichiro Omachi, “A Proposal of Writing-Life Log and Its Implementation Using a Retrieval-Based Camera-Pen”, 16th International Graphonomics Society Conference (IGS 2013) (Accepted).
- [C-1] Minoru Mori, Seiichi Uchida, Hitoshi Sakano, “Global Feature for Online Character Recognition”, Pattern Recognition Letters (Accepted).
- [C-2] Marcus Liwicki, Seiichi Uchida, Akira Yoshida, Masakazu Iwamura, Shinichiro Omachi, Koichi Kise, “More than Ink - Realization of a Data-Embedding Pen”, Pattern Recognition Letters (Accepted).
- [C-3] 森 稔, 内田誠一, 坂野 鋭, “大局的特徴に対する DP マッチング”, 電子情報通信学会論文誌(D) (Accepted).
- [C-4] Song Wang, Seiichi Uchida, Marcus Liwicki, Yaokai Feng, “Part-Based Methods for Handwritten Digit Recognition”, Frontiers of Computer Science (Accepted).
- [C-5] Takafumi Matsuo, Song Wang, Yaokai Feng and Seiichi Uchida, “Exploring the Ability of Parts on Recognizing Handwriting Characters”, 16th International Graphonomics Society Conference (IGS 2013) (Accepted).
- [C-6] Wenjie Cai, Seiichi Uchida and Hiroaki Sakoe, “An Efficient Radical-Based Algorithm for Stroke-Order Free and Stroke-Number Free Online Kanji Character Recognition”, 16th International Graphonomics Society Conference (IGS 2013) (Accepted).

- [C-7] Yutaro Iwakiri, Soma Shiraishi, Yaokai Feng and Seiichi Uchida, “On the Possibility of Instance-Based Stroke Recovery”, 13th International Conference on Frontiers in Handwriting Recognition (ICFHR2012), (DOI: 10.1109/ICFHR.2012.248)
- [C-8] Seiichi Uchida, Ryosuke Ishida, Akira Yoshida, Wenjie Cai and Yaokai Feng, “Character Image Patterns as Big Data”, 13th International Conference on Frontiers in Handwriting Recognition (ICFHR2012), (DOI: 10.1109/ICFHR.2012.190)
- [C-9] Minoru Mori, Seiichi Uchida and Hitoshi Sakano, “Dynamic Programming Matching with Global Features for Online Character Recognition”, 13th International Conference on Frontiers in Handwriting Recognition (ICFHR2012), (DOI: 10.1109/ICFHR.2012.199)
- [C-10] Seiichi Uchida, Satoshi Hokahori, and Yaokai Feng, “Analytical Dynamic Programming Matching”, Fifth Workshop on Non-Rigid Shape Analysis and Deformable Image Alignment (NORDIA'12) , (DOI: 10.1007/978-3-642-33863-2_10)
- [C-11] Song Wang, Seiichi Uchida, and Marcus Liwicki, “Part-Based Method on Handwritten Texts”, 21st International Conference on Pattern Recognition (ICPR2012)
- [C-12] Rong Huang, Shinpei Oba, Shivakumara Palaiahnakote, and Seiichi Uchida, “Scene Character Detection and Recognition Based on Multiple Hypotheses Framework”, 21st International Conference on Pattern Recognition (ICPR2012)
- [C-13] Seiichi Uchida, Masahiro Fukutomi, Koichi Ogawara, and Yaokai Feng, “Non-Markovian Dynamic Time Warping”, 21st International Conference on Pattern Recognition (ICPR2012)

(3-2) 知財出願

- ① 平成 24 年度特許出願件数(国内 0 件)
- ② CREST 研究期間累積件数(国内 3 件)