

黄瀬 浩一

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 研究代表者: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
 - ・ベース認識装置開発
 - ・大規模データベース構築
 - ・実時間文字認識
 - ・実時間文書画像検索
 - ・検索に基づく Reading-Life Log
 - ・文字・文書メディアに対する拡張現実

(2) 大町グループ

- ① 主たる共同研究者:大町 真一郎 (東北大学大学院工学研究科、教授)
- ② 研究項目
 - ・ベース認識装置開発
 - ・大規模データベース構築
 - ・付加情報に基づく文字切り出し・認識
 - ・全方位認識

(3) 内田グループ

- ① 主たる共同研究者: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
 - ・ベース認識装置開発
 - ・大規模データベース構築
 - ・認識に基づく Reading-Life Log

§ 2. 研究実施の概要

人は日々、読むことによって情報を取り入れ、知的活動を営んでいる。情報や知識という観点から見ると、“You are what you read”，すなわち人は読んだものによって形作られていると言える。このような立場から、我々は、人の読む活動の記録と、それに基づく人間調和型サービスの提供を目的として研究を行っている。

平成25年度の研究活動は、大きく以下の3点からなる。第一は、本研究の基礎となる要素技術の洗練化である。第二は、要素技術の要となる文字認識の精度向上に不可欠な、文字画像データベースの充実である。第三は、要素技術を用いた応用サービスの構築である。以下、各々について、実施概要をまとめる。

(1) 要素技術の洗練化

人が読んだものを把握する要素技術として、従来から文字認識と文書画像検索という2つの技術を開発してきた。前者は、人が読む文字を同じようにシステムが読むために必須である。一方、後者は、人が読んでいる文書を検索によって把握する試みである。今年度の成果を以下にまとめる。

(1-1) 文字認識技術の洗練化

様々な側面から洗練化に取り組んだ。具体的には、画像照合のための近似最近傍探索手法の開発、実時間文字認識の精度向上法、低解像度の文字でも高精度で認識できる手法の開発、文字大分類のための辞書構造とアルゴリズムの開発、多重仮説や視覚的顕著性に基づく情景内文字検出、部品に基づく文字検出・認識、大規模文字認識のためのネットワーク構造の考案など、多岐にわたる。これらの成果により、我々は文字認識について研究のフロントランナーとなっており、アプリケーション開発への素地が整いつつあると言える。

(1-2) 文書画像検索技術の洗練化

昨年度までは大規模化に向けて技術を開発し、一億ページのデータベースを実時間で検索できるようになっている。今年度は方向を変えて、テキストのみならず、写真や図など様々な対象を含む文書画像に対しても有効な技術を開発した。

(2) 文字画像データベースの充実

文字認識の精度を向上させるためには、大量の学習サンプルが必須となる。具体的には、多種多様なフォントの文字を広範な撮影条件で撮影した、ラベル付き文字画像が必要である。開発したデータベースは大きく以下の2つに分類できる。一つは、実画像を用いたデータベースである。これには、シーン中の文字を対象としたもの、カメラで撮影した文書画像を対象としたものの2つがある。前者については顔の隠蔽などプライバシー処理も施した。両方とも世界最大規模のデータベースとなっている。一方、もう一つのデータベースは合成により得た文字画像データベースである。この手法は、いくつかのフォントの例を与えるだけで全文字種を合成するという世界に類を見ないものであり、合成によってデータベースを充実させることに成功している。

(3) 応用サービスの構築

人の読んだものを記録し、それを応用する枠組みとして Reading-Life Log というライフログの研究開発を行っている。この研究は大きく、文書を対象とした Reading-Life Log, 情景を対象とした Reading-Life Log の2つに分類される。

まず、前者の成果を述べる。まず人の読む行動を他の行動から区別して取り出す手法を開発した。さらに、読書の対象物(新聞、論文など)を識別する手法を開発した。加えて、Reading-Life Log の基本形態として読書量を計量する「万語計」を開発した。さらに、読んだ内容を把握するために人間が見ている画像に対して文字認識を適用する手法を構築した。特に今年度は、文字認識を可能とするためのカメラシステムを完成させた。加えて、次年度に繋がる基礎的検討として、読書対象の理解度を推定する手法を検討した。また、これらの応用技術として、文書にコメントを付与したり、コメントを拡張現実によって表示したりする機構を開発した。

後者については、情景画像中のテキスト情報が我々の行動や状況理解にどの程度資するのか、すなわちテキスト情報がどの程度人間の行動理解のコンテキストとなり得るのか、について検討を開始し、情景内の文字の認識結果が人間の状態理解に有効であるという知見を得た。

§ 3. 成果発表等

(3-1) 原著論文発表

論文詳細情報(国内)

- [C-1] 武部浩明, 内田誠一, “最適 2 次元セグメンテーションによる情景内文字抽出”, 電子情報通信学会論文誌(D), Vol.J97-D, No.3, pp.667-675, 2014
- [C-2] 後藤雅典, 石田良介, 蔡文杰, 内田誠一, “最小全域木による大規模パターンの分布解析”, 電子情報通信学会論文誌(D), Vol.J97-D, No.3, pp.656-666, 2014

論文詳細情報(国際)

- [A-1] G. Bahle, P. Lukowicz, K. Kunze, K. Kise. I see you: How to improve wearable activity recognition by leveraging information from environmental cameras. Work in Progress at IEEE Pervasive Computing and Communication (PerCom) Conference Best Work in Progress.
- [A-2] K. Kunze, H. Kawaichi, K. Yoshimura, K. Kise. Towards inferring language expertise using eye tracking. Work in Progress at ACM SIGCHI Conference on Human Factors in Computing Systems.
- [A-3] Kai Kunze. Real-life Activity Recognition - Focus on Recognizing Reading Activities. Proc. 5th International Workshop on Camera-Based Document Analysis and Recognition (CBDAR 2013).
- [A-4] Dimosthenis Karatzas, Faisal Shafait, Seiichi Uchida, Masakazu Iwamura, Lluís Gomez i Bigorda, Sergi Robles Mestre, Joan Mas, David Fernandez Mota, Jon

- Almazan and Lluís Pere de las Heras, “ICDAR 2013 Robust Reading Competition”, Proc. 12th International Conference on Document Analysis and Recognition (ICDAR 2013).
- [A-5] K. Kunze, Y. Shiga, S. Ishimaru, Y. Utsumi, K. Kise. Reading activity recognition using an off-the-shelf EEG — detecting reading activities and distinguishing genres of documents. Proc. 12th International Conference on Document Analysis and Recognition (ICDAR 2013).
- [A-6] Takumi Toyama, Wakana Suzuki, Andreas Dengel, Koichi Kise. Wearable Reading Assist System: Augmented Reality Document Combining Document Retrieval and Eye Tracking. Proc. 12th International Conference on Document Analysis and Recognition (ICDAR 2013).
- [A-7] K. Kunze, H. Kawaichi, K. Yoshimura, K. Kise. The Wordometer – Estimating the Number of Words Read Using Document Image Retrieval and Mobile Eye Tracking. Proc. 12th International Conference on Document Analysis and Recognition (ICDAR 2013).
- [A-8] Hongxing Gao, Marçal Rusinol, Dimosthenis Karatzas, Josep Lladós, Tomokazu Sato, Masakazu Iwamura and Koichi Kise, “Key-region Detection for Document Images ---Application to Administrative Document Retrieval”, Proc. 12th International Conference on Document Analysis and Recognition (ICDAR 2013).
- [A-9] Sheraz Ahmed, Koichi Kise, Masakazu Iwamura, Marcus Liwicki, and Andreas Dengel, “Automatic Ground Truth Generation of Camera Captured Documents Using Document Image Retrieval”, Proc. 12th International Conference on Document Analysis and Recognition (ICDAR 2013).
- [A-10] Takuya Kobayashi, Masakazu Iwamura, Takahiro Matsuda and Koichi Kise, “An Anytime Algorithm for Camera-Based Character Recognition”, Proc. 12th International Conference on Document Analysis and Recognition (ICDAR 2013).
- [A-11] Masakazu Iwamura, Masaki Tsukada and Koichi Kise, “Automatic Labeling for Scene Text Database”, Proc. 12th International Conference on Document Analysis and Recognition (ICDAR 2013).
- [A-12] K. Kunze, S. Ishimaru, Y. Utsumi, K. Kise. My Reading Life – Towards Utilizing Eyetracking on Unmodified Tablets and Phones. Adjunct Proceedings of UBICOMP.
- [A-13] K. Kunze, K. Tanaka, M. Iwamura, K. Kise. Annotate Me – Supporting Active Reading using Real-Time Document Image Retrieval On Mobile Devices. Adjunct Proceedings of UBICOMP.
- [A-14] K. Kunze, A. Bulling, Y. Utsumi, S. Yuki, K. Kise. I know what you are reading – Recognition of document types using mobile eye tracking. International Symposium on Wearable Computers (ISWC).

- [A-15] Takumi Toyama, Wakana Suzuki, Andreas Dengel, Koichi Kise. User Attention Oriented Augmented Reality on Documents with Document Dependent Dynamic Overlay. Proc. International Symposium on Mixed and Augmented Reality (ISMAR 2013).
- [A-16] Masakazu Iwamura, Tomokazu Sato and Koichi Kise, "What Is the Most Efficient Way to Select Nearest Neighbor Candidates for Fast Approximate Nearest Neighbor Search?", Proc. 14th International Conference on Computer Vision (ICCV 2013).
- [A-17] Yuzuko Utsumi, Yuki Shiga, Masakazu Iwamura, Kai Kunze, Koichi Kise, "Document Type Classification Toward Understanding Reading Habits" Proceedings of 20th Korea—Japan Joint Workshop on Frontiers of Computer Vision, 3, pp.11-17 (2014-2)
- [A-18] Shoya Ishimaru, Jens Weppner, Kai Kunze, Andreas Bulling, Koichi Kise, Andreas Dengel and Paul Lukowicz, "In the Blink of an Eye — Combining Head Motion and Eye Blink Frequency for Activity Recognition with Google Glass." Proceedings of the 5th Augmented Human International Conference (2014)
- [A-19] Takahiro Matsuda, Masakazu Iwamura and Koichi Kise, "Performance Improvement in Local Feature Based Camera-Captured Character Recognition," Proceedings of the 11th Document Analysis Systems (Accepted) (2014)
- [B-1] Yutaka Katsuyama, Yoshinobu Hotta, Masako Omachi, and Shinichiro Omachi, "High Speed and High Accuracy Pre-Classification Method for OCR: Margin Added Hashing," IEICE Transactions on Information and Systems, vol.E96-D, no.9, pp.2087-2095, (2013-9)
- [C-3] 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).
- [C-4] 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).
- [C-5] 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.
- [C-6] Chihiro Nakamoto, Rong Huang, Sota Koizumi, Ryosuke Ishida, Yaokai Feng and Seiichi Uchida. Font Distribution Analysis by Network. The Fifth International Workshop on Camera-Based Document Analysis and Recognition (CBDAR 2013).
- [C-7] Renwu Gao, Faisal Shafait, Seiichi Uchida and Yaokai Feng. Saliency inside Saliency - A Hierarchical Usage of Visual Saliency for Scene Character Detection.

The Fifth International Workshop on Camera-Based Document Analysis and Recognition (CBDAR 2013).

- [C-8] Rong Huang, Palaiahnakote Shivakumara and Seiichi Uchida. Scene Character Detection by an Edge-Ray Filter. International Conference on Document Analysis and Recognition (ICDAR2013).
- [C-9] Takashi Kimura, Rong Huang, Seiichi Uchida, Masakazu Iwamura, Shinichiro Omachi and Koichi Kise, “The Reading-life Log --- Technologies to Recognize Texts That We Read”, Proc. 12th International Conference on Document Analysis and Recognition (ICDAR 2013).
- [C-10] Yugo Terada, Rong Huang, Yaokai Feng and Seiichi Uchida. On the Possibility of Structure Learning-Based Scene Character Detector. International Conference on Document Analysis and Recognition (ICDAR 2013).
- [C-11] Masanori Goto, Ryosuke Ishida, Yaokai Feng and Seiichi Uchida. Analyzing the Distribution of a Large-scale Character Pattern Set Using Relative Neighborhood Graph. International Conference on Document Analysis and Recognition (ICDAR 2013).
- [C-12] Song Wang, Seiichi Uchida and Marcus Liwicki. Part-Based Recognition of Arbitrary Fonts. International Conference on Document Analysis and Recognition (ICDAR 2013).
- [C-13] Rong Huang, Palaiahnakote Shivakumara, Yaokai Feng, Seiichi Uchida. Scene Character Detection and Recognition with Cooperative Multiple-Hypothesis Framework. IEICE Transactions on Information & Systems, vol.E96-D, no.10, pp.2235-2244, 2013.
- [C-14] Koichi Ogawara, Masahiro Fukutomi, Seiichi Uchida, Yaokai Feng. A Voting-Based Sequential Pattern Recognition Method. PLOS ONE, vol.8, issue 10, e76980, 2013.
- [C-15] Minoru Mori, Seiichi Uchida, Hitoshi Sakano, Global Feature for Online Character Recognition, Pattern Recognition Letters, vol.35, no.1 pp.142-148, 2014.
- [C-16] Marcus Liwicki, Seiichi Uchida, Akira Yoshida, Masakazu Iwamura, Shinichiro Omachi, Koichi Kise, More than Ink - Realization of a Data-Embedding Pen, Pattern Recognition Letters, vol.35, no.1 pp.246-255, 2014.
- [C-17] Cai Wenjie, Seiichi Uchida, Hiroaki Sakoe, Comparative Performance Analysis of Stroke Correspondence Search Methods for Stroke-Order Free Online Multi-Stroke Character Recognition, Frontiers of Computer Science (Accepted.)
- [C-18] Markus Weber, Christopher Scholzel, Marcus Liwicki, Seiichi Uchida, Didier Stricker, LSTM-Based Early Recognition of Motion Patterns, International Conference on Pattern Recognition (ICPR2014) (Accepted.)

[C-19] Volkmar Frinken, Yutaro Iwakiri, Ryosuke Ishida, Kensho Fujisaki, Seiichi Uchida, Improving Point of View Scene Recognition by Considering Textual Data, International Conference on Pattern Recognition (ICPR2014) (Accepted.)

[C-20] Kohei Inai, Marten Palsson, Volkmar Frinken, Yaokai Feng, Seiichi Uchida, Selective Concealment of Characters for Privacy Protection, International Conference on Pattern Recognition (ICPR2014) (Accepted.)

(3-2) 知財出願

① 平成 25 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)