

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所、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

(4) 稲見グループ

- ① 主たる共同研究者: 稲見 昌彦 (慶應義塾大学 大学院メディアデザイン研究科、教授)
- ② 研究項目
 - ・検索に基づく Reading-Life Log

§2. 研究実施の概要

人間は、読むという行為を通して、知的活動に必要な多くの情報を得ている。我々の研究の主たるトピック Reading-Life Log は、人の読む行為を電子的に記録、解析することによって、人間の知的活動を補助するものである。具体的には、文書に対する Reading-Life Log(文書 RLL)、シーン中のテキストに対する Reading-Life Log(シーン RLL)の2つのメインテーマについて研究開発を進め、実証実験の準備をしているところである。以下では、各々の成果について概要を報告する。

(1) 文書 RLL

ユーザが読んだ語数を計測する万語計(wordometer)、読んだ文書のタイプ(漫画、教科書、雑誌、新聞など)を識別する手法、読んだ単語を時系列でログとして記録する手法、TOEIC テストの設問を読む際の視線データなどを解析し、英語能力を推定する方法を実現した。

上記のうち、英語能力の推定以外については、今年度は特に実証実験を踏まえた形で、安価なデバイス上での実現を中心に行った。具体的には、Tobii 社が1万円以下で販売しているアイトラッカ EyeX 上での wordometer 等の実装、JINS 社が発売を予定している、EOG センサを備えたインテリジェント眼鏡 JINS MEME を用いた実装を見据えて、EOG センサを使った万語計の実装[1]が成果としてあげられる。

また、成果の社会への浸透性を考慮して、英語能力の推定にも取り組んだ。具体的には、(1) fNIRS という計測機器を用いて脳の血流を計測し、そこから英語の問題に対する正答数を推定する手法、(2)アイトラッカを用いて英文を読む目の動きを計測、解析することにより、TOEIC の点数が上級(800 点以上)、中級(600 点以上 800 点未満)、初級(600 点未満)なのかを推定する手法[2]、(3) TOEIC の長文問題を解く様子を、アイトラッカを用いて計測することにより、TOEIC の点数そのものを推定する手法の3つを開発した。その結果、(1)については正答数の推定精度 80%、(2)については級の推定精度 91%、(3)については TOEIC 点数の推定誤差 36 点を得ることができた。

加えて、成果の実利用を促進する目的で、ノルウェーの Studix 社、インドの ZiZ Labs 社に対して、実時間文書画像検索技術を提供した。これらの会社は共に、紙に印刷された教科書と、電子的に作成された教育補助コンテンツをリンクするための事業を展開しており、そのリンクに我々の文書画像検索技術が利用される予定である。

(2) シーン RLL

文書中の文字だけではなく、シーン中の文字も、人間の行動を左右する重要な情報源である。シーン RLL では、このようなシーン中の文字を人間と同様に認識し、人間の知的活動を補助することを目的としている。

具体的には、次の2つのシナリオを設定しシステムを開発した。一つは、店舗内にある棚札を認識し、ユーザの購買行動を補助するものである。これについては、健常者はもとより、視覚に障害を持つユーザの助けとなることを意図している。もう一つは、薬局において処方箋や薬

に記載された薬の名称やロゴを読み取り, 正しく処方されているかどうかを検証するものである. こちらについては, 薬剤師の作業の妨げにならない形で, チェックを行うことを意図している.

さらに, 情景画像中のテキスト情報が我々の行動認識や環境理解にどの程度資するのかについて検討を行い, 肯定的な結論を得た. これは, 例えば「食べ物に関するテキストが多ければ, その環境はレストランである」という, 本質的ながら今まで全く見当されてこなかった, 文字と環境の関係解明に関する世界でも例を見ない研究である. この肯定的な結果は文献[3]にて公表された.

加えて, 成果の実利用を促進する目的で, 日本の Particular Design 社に, シーン中の文字認識の技術を提供した. この会社は失読症や弱視など視覚に障がいを持つ人の補助をするための眼鏡型デバイス OTON グラスの開発を進めている. OTON グラスでは, 我々の文字認識技術を用いてシーン中の文字を音に変換することを試みている.

- [1] Kai Kunze, Masai Katsutoshi, Yuji Uema, and Masahiko Inami, “How Much Do You Read? – Counting the Number of Words a User Reads Using Electrooculography”, Proceedings of Augmented Human, Singapore, 2015.3
- [2] 吉村 和代, Kai Kunze, 黄瀬 浩一, “読書時の眼球運動を利用した英語習熟度推定法”, パターン認識・メディア理解研究会 (PRMU)、仙台、2015年2月19日
- [3] Volkmar Frinken, Yutaro Iwakiri, Ryosuke Ishida, Kensho Fujisaki, Seiichi Uchida, “Improving Point of View Scene Recognition by Considering Textual Data,” Proceedings of the 22nd International Conference on Pattern Recognition, Stockholm, Aug. 2014.