

科学研究費助成事業 研究成果報告書

平成 29 年 6 月 12 日現在

機関番号：24403

研究種目：基盤研究(A) (一般)

研究期間：2013～2016

課題番号：25240028

研究課題名(和文) アクティビティ解析に基づく Knowledge Logの構築とその応用

研究課題名(英文) Construction of Knowledge Log Based on Activity Analysis and Its Applications

研究代表者

黄瀬 浩一 (Kise, Koichi)

大阪府立大学・工学(系)研究科(研究院)・教授

研究者番号：80224939

交付決定額(研究期間全体)：(直接経費) 36,300,000円

研究成果の概要(和文)：Knowledge Logとは、人が日々獲得している知識の記録である。従来から存在するライフログとの違いは、Knowledge Logが知識という高次記号情報をログの対象とするのに対して、ライフログは、信号情報あるいは記号であっても単純なもの(例えば人の座る、歩く、食べるなどの行動)を記述する点にある。人の知識の大半は、人の読むという行為によって獲得されていることに着目し、本研究では、読むことに関連した知識のログを中心に、その量や質について推定するための各種手法を構築した。また、その基礎となる特徴照合、文書画像検索、文字、顔、物体などの認識技術、大規模文字画像データベースについても開発した。

研究成果の概要(英文)：The Knowledge Log is a record of our knowledge acquired daily in our life. As compared to existing "life logs", which record signals or simple symbols such as sitting, walking and eating of our activities, we focus here on the record of knowledge as high-level symbols. Based on the fact that most of our knowledge is coming from our reading activities, we put special efforts on analyzing and acquiring knowledge logs through estimating quantity and quality of reading. In addition, we have also developed various fundamental technologies about feature matching, document image retrieval, recognition of characters, faces and objects, as well as a large-scale character image database.

研究分野：情報工学

キーワード：ライフログ Knowledge Log アクティビティ解析 アイトラッキング アイウェア ウエアラブルセンサ 機械学習 物体認識

1. 研究開始当初の背景

本研究を開始するまでに我々が行ってきた、文書、文字、三次元物体、ビデオ、顔、一般画像を対象とした大規模特定物体認識に関する成果の多くが世界のレベルにあった当時、これを用いて従来にないライフログを実現しようと考えた。従来のライフログの多くが、信号レベルの記録であったり、基本的な動作などの単純な記号レベルの記録に留まっていたりすることに着目し、本研究の目標を、「知識」という高次記号情報をログの対象とすることとし、研究テーマを「**Knowledge Log**」と名付けた。

研究の方法論としては、人のアクティビティのログを取得し、それを解析することである。ただし、この研究は、従来から行われているアクティビティ解析の研究とは以下の点で大きく異なる。従来のアクティビティ解析の研究では、解析対象となる人の行動が限定的であることに加えて、行動の記述に用いられる記号情報も、歩く、座る、食べるなど、基本的なものに限定されていた。また結果はあくまでも認識された行動であり、それが人にどのような価値や知識をもたらすのかを認識多少としていなかった。これに対して、本研究では、行動の記録ではなく、行動によって得られた知識の記録を得る点が大きく異なる。

2. 研究の目的

人が活動から得る知識をすべて記録することは究極の目標であるが、様々な技術的課題があることから、容易には達成できない。本研究では、その第一歩として、環境中にある記号情報に着目し、その記号情報と人の行動の関連で得られる知識を処理対象とすることとした。そして、人が日常生活でかかわる環境中の記号情報を認識し、それを他の認識結果(物体、顔、行動)と関連づけて記述することによって、得られる知識を把握することを、本研究の目標と定めた。特に、環境中の記号情報を認識する点については、文字認識や文書画像検索など、研究代表者らがこれまでに蓄積した研究成果があるため、他の追従を許さないものであると自負している。

3. 研究の方法

本研究の方法としては、文書画像検索、文字認識、物体認識、顔認識、行動認識などの各種認識技術、ならびにそれらの基盤となる技術の開発(基礎研究)と認識技術を用いた Knowledge Log の実現(応用研究)の2側面を考えた。基礎研究では、単に認識手法の開発、精度向上だけではなく、特に行動認識に対して、画像以外にどのようなセンサを用いると何が可能なのかについても検討した。具体的には、ウェアラブルセンサやアイトラッカなどのセンサを導入し、その利用技術を開発した。

応用研究については、研究計画当初は、一般的な知識記述枠組として RDF を導入し、そ

れに基づいて Knowledge Log を構築していく予定であったが、よりアプリケーションに適した形での記述が適しているであろうとの判断でアプリケーションごとの記述に変更した。具体的には、人が知識を獲得する際の入り口が読むことである点に注目し、「読み」に関連した知識の量や質について具体的なアプリケーションを作成していくこととした。

4. 研究成果

本研究の成果は、Knowledge Log に直接かかわる応用研究の成果と、それを支える基礎研究の成果に分類される。その各々について、基礎研究の成果から順に述べる。

(1) 基礎研究

① 近似最近傍探索

文字や物体などを認識するためには、特徴ベクトルを照合する必要がある。特に、データベース(辞書)中の特徴ベクトルが大規模である場合への対処が必要である。この問題に対して、本研究では BDH (Bucket Distance Hashing) と呼ぶ方法を開発し、世界最高レベルの性能が達成可能なことを示した。この手法は、本研究における様々な認識タスク(文字、顔など)で用いられている。

② 近似最近傍部分空間探索

認識手法が部分空間法に基づく場合、上記の近似最近傍探索によって高速化することはできない。そこで本研究では部分空間法に適した高速近似探索法を提案し、様々なタスクを用いて有効性を検証した。

③ 文字認識

図1に示すように、複雑な背景を伴い、かつレイアウトに規則性がない場合でも、高速かつ正確に文字を認識可能な手法を提案した。この手法は、我々がかつて物体認識のために提案した reference point と呼ぶ手法を文字認識に適用したものとなっており、研究代表者らの知る限り複雑背景下で文字認識を可能とした最初の手法である。



図1 複雑背景を伴う文字の認識

④ 顔認識

人の顔は、環境中の記号以外の認識対象として、きわめて重要なものの一つである。本研究では、図2に示すような、局所特徴量と投票処理を用いて、大規模データベースに対する高速顔認識を実現した。この手法は、10万顔画像に対して従来法より1000倍以上高速であり、300万顔画像データベースを対象とした認識では257ms/queryの速度で96.9%の1000位累積認識率を達成した。

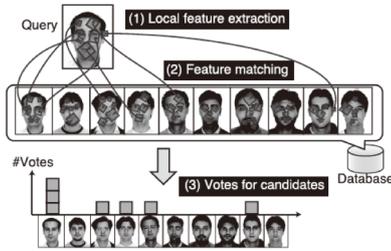


図2 局所特徴量と投票処理を用いた顔認識

⑤ 物体認識

物体認識の一分野である、一般物体認識（椅子や自動車などの一般的な物体カテゴリに対する認識）に取り組んだ。具体的には、PyramidSepDropと呼ぶ新しい構造のニューラルネットワークを用いることにより、CIFAR-10, CIFAR-100と呼ばれる標準的なベンチマークデータセットに対して、世界最高性能の認識率（2016年12月現在）を達成した。認識率の推移のグラフを図3に示す。

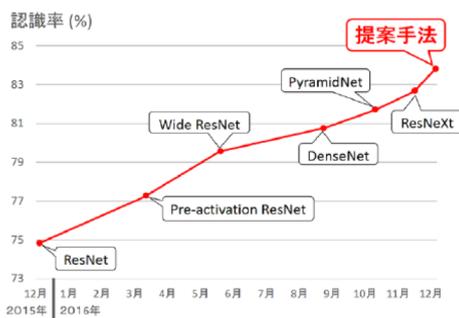


図3 CIFAR-100を対象とした一般物体認識の認識率の推移

⑥ 文字認識用データベース構築

研究用の新しいシーンテキストデータセットとして、Downtown Osaka Scene Text Dataset (DOST)を作成し、公開した。このデータセットは、280万文字を含むデータセットとなっており、シーンテキストのデータセットとしては世界最大クラスの

ものである。また、従来からあるシーンテキストのデータセットと比べて、より自然で現実的な撮影となっており、認識が格段に難しい(F値で8倍以上)ものである。

(2) 応用研究

① 万語計

人の知識の量は、読んだテキストの量に深く関連すると考えられている。これは、人の健康が歩いた歩数に関連することと類似している。そこで、我々は、知識の万歩計として、万語計(Wordometer)を開発した。万語計は、装着型アイトラッカ、据え置き型アイトラッカ、EOG眼鏡(JINS MEME)などのデバイスで動作する。誤差は使用デバイスや読む量によっても異なるが、概ね数パーセントから20パーセント程度であり、万歩計と同等のレベルにある。

② 読みの検出

いつ読み始めていつ読み終わったのか、は、万歩計と同様に獲得した知識の量を推定する基本的な特徴となる。また、体重計に毎日乗り続ければ、ダイエットが加速するように、読む時間や読んだ量を計測することは、獲得する知識の量を向上させる効果が期待できる。我々は、装着型アイトラッカやEOG眼鏡を用いて、高精度な読みの検出を実現した。

③ 文書タイプの認識

量だけではなくどの分野の知識を獲得したのかを知ることができれば、より詳細なKnowledge Logを実現できる。そこで本研究では、読んだ文書のタイプ(論文、新聞、小説、雑誌、教科書など)を認識する手法を構築した。この手法は、装着型アイトラッカ、あるいは簡易脳波計を用いるものであり、両者ともにユーザに応じた学習を行うことで、100%に近い識別精度を達成可能となっている。

④ 読んだ単語の記録

さらに詳細なKnowledge Logは、読んだ単語によって特徴付けられるものである。そこで本研究では、読んだ単語を推定し、その頻度付きリストを作成する手法を構築した。結果は、例えば、図4のような時系列で変化するタグクラウド形式(上側が読み始めて時間が経過すると下に向かう)で表示される。



図4 読んだ単語の記録

⑤ 未知単語推定

Knowledge Log の一つのアプリケーションとして、獲得し保有している知識を表現することではなく、獲得していない知識を表現することが考えられる。これは、次にどのような知識を獲得すべきかを考える上で、重要な情報源となる。

本研究では、そのような未習得知識の推定方法の最も基本的なものとして、未知英単語の推定法を提案した。この手法では、人が英文を読む際の視線データを解析し、その人にとって未知の単語を推定するものであり、Recall がほぼ 100% の際の Precision として 33.4% を実現した。

⑥ 英語能力推定

Knowledge Log のもう一つのアプリケーションとして、知識の質の推定がある。一般的に知識の質を推定することは大変困難であるが、分野を限れば可能性が見えてくる。本研究では、TOEIC の練習問題を解く際の視線データを解析することによって、TOEIC の点数を推定する手法を考案した。この手法には様々なバリエーションがあるが、例えば、推定に用いる問題に対して、他のユーザの視線データと TOEIC の点数の対が利用可能な場合（既知文書を用いた推定）、点数の推定平均誤差は 30 点程度であることを示した。

⑦ 多肢選択問題回答時の確信度推定

TOEIC パート 5 の四択問題を対象に、回答の確信度を推定する手法を提案した。この手法では、アイトラッカによって得られた、問題文と選択肢上の視線データを用いる。実験の結果 90% 以上の精度で確信の有無が推定できた。

⑧ AR sentence

直接的な Knowledge Log ではないが、我々が開発した手法を用いて、人が知識獲得する際の補助となるスマートフォンアプリケーションを、民間企業と連携して開発した。そのシステムは、図 5 に示す AR sentence と呼ぶものであり、学習参考書のページにスマートフォンをかざすと、

そのページに関連したビデオが再生されるというものである。これによって、参考書を読んだだけでは分からない場合でも、ビデオの説明を見ることによって、理解が促進される。

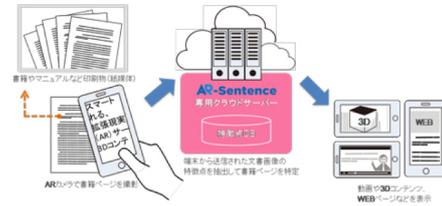


図5 AR sentence

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 17 件)

1. Yuzuko Utsumi, Tomoya Mizuno, Masakazu Iwamura and Koichi Kise, Fast Search Based on Generalized Similarity Measures, IPSJ Transactions on Computer Vision and Applications, Vol. 9, No. 1, 11 pages (2017-5), 査読有, DOI: [10.1186/s41074-017-0024-5](https://doi.org/10.1186/s41074-017-0024-5)
2. 志賀 優毅, 内海 ゆづ子, 岩村 雅一, Kai Kunze, 黄瀬 浩一, 視線情報と一人称視点画像を用いた文書カテゴリの自動識別, 電子情報通信学会論文誌, Vol. **J99-D**, No. 9 (2016-9), 査読有, http://search.ieice.org/bin/summary.php?id=j99-d_9_950
3. Charles Lima Sanches, Olivier Augereau, Koichi Kise, Vertical error correction of eye trackers in nonrestrictive reading condition, IPSJ Transactions on Computer Vision and Applications, **8**, 1, 7 pages, Springer Berlin Heidelberg (2016), 査読有, DOI: [10.1186/s41074-016-0008-x](https://doi.org/10.1186/s41074-016-0008-x)
4. 黄瀬浩一, 大町真一郎, 内田誠一, 岩村雅一, 文字・文書メディアの新しい利用基盤技術の開発への取組み, 電子情報通信学会誌, Vol. **98**, No. 4, pp.311-

- 327 (2015-4), 査読無,
<http://iss.ndl.go.jp/books/R100000002-I000000050067-00>
5. 内海 ゆづ子, 坂野 悠司, 前川 敬介, 岩村 雅一, 黄瀬 浩一, 局所特微量と投票処理を用いた大規模データベースに対する高速顔認識, 電子情報通信学会論文誌 D, Vol. **J97-D**, No. 8, pp.1263-1272 (2014-8), 査読有,
http://search.ieice.org/bin/summary.php?id=j97-d_8_1263
 6. Kai Kunze, Masakazu Iwamura, Koichi Kise, Seiichi Uchida, and Shinichiro Omachi, Activity Recognition for the Mind: Toward a Cognitive "Quantified Self", IEEE Computer, Vol. **46**, No. 10, pp.105-108 (2013-10), 査読無, DOI: [10.1109/MC.2013.339](https://doi.org/10.1109/MC.2013.339)
- [学会発表] (計 121 件)
1. Yuzuko Utsumi, Tomoya Mizuno, Masakazu Iwamura, Koichi Kise, Fast Search Based on Generalized Similarity Measure, IAPR International Conference on Machine Vision Applications, 2017.5.8-5.12, Nagoya (Japan).
 2. Olivier Augereau, Hiroki Fujiyoshi and Koichi Kise, Towards an Automated Estimation of English Skill via TOEIC Score Based on Reading Analysis, Proc. International Conference on Pattern Recognition, 2016.12.4-12.8, Cancun (Mexico).
 3. Masakazu Iwamura, Takahiro Matsuda, Naoyuki Morimoto, Hitomi Sato, Yuki Ikeda and Koichi Kise, Downtown Osaka Scene Text Dataset, 2nd International Workshop on Robust Reading, 2016.10.9, Amsterdam (The Netherlands).
 4. Wei Fan, Koichi Kise, and Masakazu Iwamura, Automatic Character Labeling for Camera Captured Document Images, Automatic Character Labeling for Camera Captured Document Images, 2016.9.25-9.28, Phoenix (USA).
 5. Shoya Ishimaru, Kai Kunze, Koichi Kise, and Andreas Dengel, The Wordometer 2.0 – Estimating the Number of Words You Read in Real Life Using Commercial EOG Glasses, 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, 2016.9.12-9.16, Heidelberg (Germany).
 6. Kai Kunze, Katsutoshi Masai, Masahiko Inami, Ömer Sacakli, Marcus Liwicki, Andreas Dengel, Shoya Ishimaru, Koichi Kise, Quantifying Reading Habits – Counting How Many Words You Read, 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing, 2015.9.7-9.10, Osaka (Japan).
 7. Masakazu Iwamura, Tomokazu Sato and Koichi Kise, What Is the Most Efficient Way to Select Nearest Neighbor Candidates for Fast Approximate Nearest Neighbor Search? 14th International Conference on Computer Vision, 2013.12.3-12.6, Sydney (Australia).
 8. Kai Kunze, Hitoshi Kawaichi, Koichi Kise and Kazuyo Yoshimura, The Wordometer - Estimating the Number of Words Read Using Document Image Retrieval and Mobile Eye Tracking, 12th International Conference on Document Analysis and Recognition, 2013.8.25-8.28, Washington DC (USA).
- [図書] (計 2 件)
1. Koichi Kise, Shinichiro Omachi, Seiichi Uchida, Masakazu Iwamura, Masahiko Inami, Reading-Life Log as a New

Paradigm of Utilizing Character and Document Media, Human-Harmonized Information Technology, Ed. By T. Nishida, Vol. 2, pp. 197-233, Springer (2017), DOI: [10.1007/978-4-431-56535-2_7](https://doi.org/10.1007/978-4-431-56535-2_7)

2. **Koichi Kise**, Page Segmentation Techniques in Document Analysis, Handbook of Document Image Processing and Recognition (Eds. by D.Doermann and K.Tombre), pp.135-175, Springer-Verlag, London (2014), DOI: 10.1007/978-0-85729-859-1_5

[その他]

(1) 国際会議開催

以下の国際ワークショップを開催し、研究成果を発表した。

- ① MANPU2016: 1st International Workshop on Comics Analysis, Processing and Understanding, 2016. 12. 6.
- ② WAHM2016: 3rd Workshop on Ubiquitous Technologies for Augmenting the Human Mind, 2016. 9. 12.
- ③ WAHM2015: 2nd Workshop on Ubiquitous Technologies for Augmenting the Human Mind, 2015. 9. 7.
- ④ WAHM2014: 1st Workshop on Ubiquitous Technologies for Augmenting the Human Mind, 2014. 9. 13.

(2) イベント出展

CEATEC JAPAN 2015 に研究成果を出展した。

(3) メディア掲載

- ① 日刊工業新聞：大阪府大，英語習熟度を推定する技術を開発-英文問題を読み解く眼球の動き解析，2016. 8. 3.
- ② 産経アプリスタ：[指先で文字を読む，腕の筋肉でゲーム操作 大学やベンチャーの最新技術が並ぶ](#)
- ③ EE Times Japan：[指先で文字を自動的に認識 CEATEC でデモ展示](#)

(4) 報道発表

- ① [一般物体認識分野で，府大生が世界一の認識精度を持つニューラルネットワークを開発](#)，2016. 12. 9.
- ② [紙媒体デジタルメディア化支援サービス AR-Sentence の開始について](#)，2016. 12. 1

(5) 受賞

- ① PRMU 研究奨励賞，山田良博，[PyramidNet における確率的な正則化の効果の検証](#)，2017
- ② PRMU 月間ベストプレゼンテーション賞，山田良博，2017.2
- ③ PRMU 研究奨励賞，小西将貴，[特徴空間の次元数と部分空間数にスケラブルな近似最近傍部分空間探索](#)，2015
- ④ Best Paper Award, 15th IAPR International Conference on Machine Vision Applications (学会発表 1)
- ⑤ Best Paper Award, IAPR International Conference on Document Analysis and Recognition (学会発表 8).

6. 研究組織

(1) 研究代表者

黄瀬 浩一 (KISE, Koichi)
大阪府立大学・工学研究科・教授
研究者番号：80224939

(2) 研究分担者

岩村 雅一 (IWAMURA, Masakazu)
大阪府立大学・工学研究科・准教授
研究者番号：80361129
岩田 基 (IWATA, Motoi)
大阪府立大学・工学研究科・准教授
研究者番号：70316008
内海ゆづ子 (UTSUMI, Yuzuko)
大阪府立大学・工学研究科・助教
研究者番号：80613489

(3) 連携研究者

クンツェ カイ (KUNZE, Kai)
慶應義塾大学・メディアデザイン研究科・准教授
研究者番号：00648040

(4) 研究協力者

デンゲル アンドレアス (DENGEL, Andreas)
ドイツ人工知能研究センター・Director・教授
外山 託海 (TOYAMA, Takumi)
ドイツ人工知能研究センター・研究員