

原田 達也

東京大学大学院情報理工学系研究科
教授

膨大なマルチメディアデータの理解・要約・検索基盤の構築

§ 1. 研究実施体制

(1) 原田グループ

- ① 研究代表者: 原田 達也 (東京大学 大学院情報理工学系研究科, 教授)
- ② 研究項目
 - ・研究全体の統合・統括
 - ・膨大なマルチメディアデータを認識する手法の開発
 - ・獲得したデータ間の関係性を発見し記述・要約する手法の開発

(2) 杉山グループ

- ① 主たる共同研究者: 杉山 将 (東京大学 大学院新領域創成科学研究科, 教授)
- ② 研究項目
 - ・マルチメディアデータの認識技術を支える数理的基盤の構築

(3) 大野グループ

- ① 主たる共同研究者: 大野 和則 (東北大学 未来科学技術共同研究センター, 准教授)
- ② 研究項目
 - ・能動的センシングにより獲得される実世界データの理解

(4) 塚田グループ

- ① 主たる共同研究者: 塚田 浩二 (公立はこだて未来大学 システム情報科学部情報アーキテクチャ学科, 准教授)
- ② 研究項目
 - ・パーソナルアプリケーションにより獲得される実世界データの理解

(5) 下坂グループ

- ① 主たる共同研究者: 下坂 正倫 (東京工業大学 大学院情報理工学研究科, 准教授)

② 研究項目

- ・人の行動データにより獲得される実世界データの理解と能動的知識獲得

§ 2. 研究実施の概要

膨大なマルチメディアデータを要約する基盤技術構築には、(1)データのコンテンツ認識アルゴリズム、(2)認識を支えるための知識獲得、(3)理解・要約・検索技術の有益さの検証、の3つの要素が必要不可欠である。これらの平成27年度の成果を述べる。

(1)データのコンテンツ認識アルゴリズム

画像に含まれるコンテンツを文章で記述すること(画像キャプション生成と呼ぶ)は、画像、映像を理解し要約する上で基盤となる技術である。これを実現するためには画像特徴と文章の構成要素である単語の組み合わせの関係性を学習する必要があるが、各特徴は高次元になり学習が困難となる。そこで今年度は、画像と文章の構成要素を共通の低次元空間に射影しつつ、その空間における識別則を同時に構成することで、高次元かつ少サンプルでも効率的に学習可能な手法を考案し、画像キャプション生成における有効性を確認した[1]。

昨年度に開発した生成確率分布を同定することなく統計的な依存関係の変化を直接検知するアルゴリズムの理論的な性質を精査し、高次元かつ複雑な依存関係をもつデータに対しても、依存関係の変化が限定的である場合は、少量のデータ標本から正確に変化を検知できる事を明らかにした。これは、これまでに得られている実験結果ともよく整合している。

要素間の関係性の理解と要約、では、個人差や多様な環境を考慮した多タスク学習法を継続して取り組んだ。また、人行動理解のためのセンシングシステムの環境に応じて再利用性が低下する問題に取り組み、環境毎に変化する最適なセンサ・パラメータを、貪欲法により効率よく取得する方法を構築した。

(2)認識を支えるための知識獲得

訓練時には正のクラスのデータしか与えられない分類問題(PU 学習と呼ぶ)に対して、凸最適化に基づく新しい機械学習の枠組みを構築した。これは、昨年度に開発した非凸最適化に基づくアプローチの大幅な改良になっており、大域的な最適解が高速かつ確実に求められるという優れた性質をもつ[2]。

人を活用した知識の能動的獲得では、得られたデータの処理方法に加え、どのようにして良質なデータを取得するかも重要であることから、システムの質の向上に寄与しうるデータの取得の決定法に取り組んだ。具体的にはベイズ最適化を組み込んだ能動的クラス選択法を開発した。屋内測位システム構築のための人行動データの取得に展開し、基礎的な性能評価を実施した。

(3)理解・要約・検索技術の有益さの検証

実世界で自律的に情報収集を行うロボットや動物を利用したデータの蓄積を継続して行った。具体的には、自動運転可能な車両、犬、飛行ロボットなどを利用して市街地の道路、模擬被災環境、老朽化した建造物の映像や形状データを収集した(図1参照)。サイバー救助犬や、自動運転の取り組みは、2015年度の仙台市主催の公開デモや、国内外のメディアを通じて一般公開した。また、収集したデータを利用して人間やロボットが利用する情報の抽出と可視化にも取り組んだ。収集した複数の情報限から、被災者の発見に繋がる情報の自動マッピングや[3]、整合性の取れたロボット用の地図を作る方法などを開発した。さらに、自動運転車両やUAVなどのロボットを利用した実験を実環境で行うにあたり、仙台市と東北大学が共同で特区申請を行い、2015年6月に近未来技術実証特区の認定を得た。



図1:ロボットや動物を利用した情報収集:自動運転車両(左),サイバー救助犬(右)

パーソナルファブ리케이션における実世界のモノづくり過程をマルチモーダルに蓄積・共有・再現する基盤システムを構築し,提案アルゴリズムを応用して認識検索や手順書作成の支援等を行う。本年度は, H26 年度に試作した実世界での部品の組立過程を写真等で記録し Web 上で共有するサービスについて, 動画データへの対応や本格運用に向けた足回りの強化を進めた。また, 実世界の様々な箇所に容易に固定でき, センサやクラウドと連携可能なカメラベースツールキットの実装を進めた。さらに, 学生主体での FabLab の運用に引き続き取り組み, ファブデータを収集／活用する土台となるコミュニティ作りを進めた。

ウェアラブル行動データの理解・要約・検索では, 市街地道路における人の運転行動の予測など, 複数のアプリケーションに応用し研究成果を発表した。

- [1] Yoshitaka Ushiku, Yusuke Mukuta, Masataka Yamaguchi, Tatsuya Harada. Common Subspace for Model and Similarity: Phrase Learning for Sentence Generation from Images. In the 15th International Conference on Computer Vision (ICCV 2015), pp.2668-2676, 2015.
- [2] Marthinus Christoffel du Plessis, Gang Niu, Masashi Sugiyama. Convex formulation for learning from positive and unlabeled data. In F. Bach and D. Blei (Eds.), Proceedings of 32nd International Conference on Machine Learning (ICML2015), JMLR Workshop and Conference Proceedings, vol.37, pp.1386-1394, Lille, France, Jul. 6-11, 2015.
- [3] Yuichi Komori, Kazuaki Ohno, Takuaki Fujieda, Takahiro Suzuki, Satoshi Tadokoro. Detection of Continuous Barking Actions from Search and Rescue Dogs' Activities Data. In 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp.630-635, 2015.