

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
  - ・パーソナルファブ리케이션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 (東京大学 大学院情報理工学系研究科, 講師)

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26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

ビッグデータ活用場の増加に伴い必要となる計算能力の問題を、JavaScriptを用いた分散計算システムを構築して解決することを提案し、それを基幹システムである高速な行列計算ライブラリ、この高速なライブラリを活用した機械学習ライブラリ、可視化ライブラリを構築した。行列計算ライブラリはWebCLを通してGPUを利用することにより、JavaScriptでも実用的な速度で計算が行えることを示した。

カラー画像と距離画像を合わせた画像（RGB-D画像）から認識対象の物体を精度よく発見するアルゴリズムを開発した。このアルゴリズムでは、検出対象物体のRGB-D参照画像との対応点を求め、対応点ペア類似度行列の二次割当問題を解いて得られた対応点集合に対して、その正当性を評価する対応点集合類似度を導入し、パラメータを物体毎に学習する。提案手法は従来手法と比較して約15%の性能向上が確認できた[1]。

膨大なマルチモーダルデータを効率よく要約するためには、重要部分の切り出しが重要である。そこで、生成確率分布を明示的に同定することなく統計的な依存関係の変化を直接検出する密度比推定手法に関する理論研究に取り組んだ。従来の生成確率分布を推定するアプローチでは、変化の検知に必要なデータ数は、変化前と変化後のそれぞれの依存グラフのエッジ数に依存することが知られていた。そのため、依存グラフの変化分は疎であるが変化前と変化後のそれぞれ依存グラフは密であるという現実的な場面において、従来法で精度良く依存関係の変化検知を行うためには非常に多くのデータが必要であった。一方、密度比推定に基づく変化検知手法では、変化の検知に必要なデータ数は変化前と変化後のそれぞれ依存グラフのエッジ数には依存せず、依存グラフの変化エッジ数で決まることを理論的に明らかにし、数値実験により妥当性を検証した（図1参照）[2]。



図1

### (2) 認識を支えるための知識獲得

実世界情報処理では、得られたデータの処理方法に加え、認識技術に有益な良質な知識をどのようにして取得するかが本質的な課題である。この課題解決には人の活用が有益であり、そのためには人の行動データをスマートフォン、ウェアラブルセンサなどから絶え間なく収集し要約する技術に取り組む必要がある。本年度は、大量の行動データから行動パターンを抽出し要約・予測する技術への応用を念頭に、統計的な行動モデルの開発に取り組んだ。行動データそれ自体、もしくは行動パターンや行動予測モデルが、人の個性や行動データの取得状況に応じ変化することに注目し、この問題に対する枠組みとして多タスク学習法の開発を行った。具体的には、個性や状況

に応じてモデルを切り替える，非線形多タスク機械学習法を構築した．従来法では困難な，長期間にわたる行動の予測・シミュレーションといったタスクに対して優れた性能を有することを確認した．また，プロジェクト期間中に開発する「膨大なマルチメディアデータの理解・要約・検索技術」の適応対象である人間行動データを取得するためのプラットフォームを試作した．

### (3) 理解・要約・検索技術の有益さの検証

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，実世界でのモノづくり過程を手軽にマルチモーダルデータで記録・共有・再現するためのファブ環境の基盤システムの構築に着手した．まず，モノづくりの組立過程に焦点を当て，組立過程を記録／共有するプロトタイプを試作した．さらに，大学のプロジェクト学習の枠組みを用いて，学生主体での FabLab の運用に取り組み，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．さらに，ユーザの行動や判断をデータ収集時に活用する「日用品ライフログシステム」の実装を進めた[3]．

ロボット技術を利用して，実世界の日々変動する電子化されていないデータを収集し，要約する技術を開発する．本年度は，プロジェクトで開発する膨大なマルチメディアデータの理解・要約・検索技術を適用する対象となる実世界のデータを，地上移動体，飛行体，災害救助犬を利用して収集，蓄積を行った．着目すべきデータの種類と質について，防災やインフラ点検の関係者から聞き取りをおこなった．この知見に基づき，映像，形状，動作等のデータを収集した．また，異なるプラットフォームで収集したヘテロな大量データの統合方法についての検討を行った．また，自動運転による，能動的な情報収集の方法についても基盤整備を行った（図2参照）．



図 2

- [1] Asako Kanezaki, Emanuele Rodolà, Daniel Cremers, Tatsuya Harada. Learning Similarities for Rigid and Non-Rigid Object Detection. International Conference on 3D Vision (3DV 2014), pp.720-727, 2014.
- [2] Song Liu, Taiji Suzuki, Masashi Sugiyama. Support consistency of direct sparse-change learning in Markov networks. In Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence (AAAI2015), pp.2785-2791, Austin, Texas, USA, Jan. 25-29, 2015.
- [3] 門村亜珠沙, 李争原, 塚田 浩二, 朱浩華, 椎尾 一郎. 食行動改善を促すスマートフォン連動型センサ内蔵フォーク. 情報処理学会論文誌, Vol.56, No.1, pp.338-348, 2015.