

「科学的発見・社会的課題解決に向けた各分野のビッグデータ利活用推進
のための次世代アプリケーション技術の創出・高度化」

H29年度
実績報告書

平成 26 年度採択研究代表者

吉田 直紀

東京大学大学院理学系研究科／国際高等研究所カブリ数物連携宇宙研究機構
教授

広域撮像探査観測のビッグデータ分析による統計計算宇宙物理学

§ 1. 研究実施体制

(1)「宇宙論」グループ

- ① 研究代表者: 吉田 直紀 (東京大学／国際高等研究所カブリ数物連携宇宙研究機構 教授)
- ② 研究項目
 - ・撮像データ取得、可視化
 - ・理論シミュレーション
 - ・銀河形成シミュレーションコードの開発(SPPEXA)

(2)「統計解析」グループ

- ① 主たる共同研究者: 池田 思朗 (統計数理研究所 教授)
- ② 研究項目
 - ・宇宙論パラメータ推定法の開発

(3)「画像解析機械学習」グループ

- ① 主たる共同研究者: 上田 修功 (日本電信電話株式会社 NTT コミュニケーション科学基礎研究所 上田特別研究室長(NTT フェロー)／理化学研究所革新知能統合研究センター 副センター長)
- ② 研究項目
 - ・時間変動天体の自動検出法の開発

(4)「データ基盤」グループ

① 主たる共同研究者:川島 英之 (筑波大学システム情報工学研究科 准教授)

② 研究項目

- 解析パイプラインシステムの設計
- 天文データ処理システムの設計

§ 2. 研究実施の概要

すばる望遠鏡による広域宇宙観測計画 HSC サーベイは H26 年度に開始され、現在まで支障なく観測を続けてきた。H29 年度は、10-12 月にかけて主鏡の蒸着と望遠鏡のメンテナンス作業のために観測できない期間があったが、49 晩の観測を実施し、全体で 192 晩の観測が終了した。年度後半は天候に阻まれ、観測できない日々が続いたが、取得された宇宙画像データは当初予定どおりのクオリティが保たれており、本研究に使用することができる。

本研究の目的は、画像ビッグデータ解析により多数の超新星や銀河を検出し、宇宙の物質分布や遠方超新星までの距離を測定することである。その結果を大規模シミュレーションデータベースと統合して統計解析し、宇宙の進化を記述する基本量(宇宙論パラメータ)を推定することを目指している。H29 年度には、超新星探査を前半と後半に2回遂行した。地球から80億光年以上も離れた最遠方の Ia 型超新星を検出・分類し、米国 NASA のハッブル宇宙望遠鏡を用いて赤外波長域で観測すると共に、世界の大型望遠鏡(米国ケック、ジェミニ、欧州の VLT、スペインの GTC)による分光観測を遂行した。超新星の自動検出および分類には、本プロジェクトを通して開発した機械学習による分類が重要な指標となった。

宇宙論グループは、H28 年度までに構築された宇宙の構造形成のスーパーコンピュータシミュレーションデータベースを用いて、物質密度場に関わる各種統計量を高精度で推定する「エミュレータ」を開発した。また、画像解析機械学習グループと協働して、重力レンズ現象により歪められた銀河の画像から、レンズ天体の質量や形状を推定する深層学習アプリケーションを開発した。

画像解析機械学習グループは、昨年度までに開発した超新星の種類を分類する手法を発展させ、Ia 型以外の様々なタイプの超新星を自動的に分類可能な手法を開発した¹⁾。天文学分野で広く用いられる光度曲線モデルおよびそのモデルパラメータの操作方法を改良し、そのモデルに基づいて大量に生成した人工データを深層学習モデルの学習データとして利用することで、すばる HSC の実画像データに対して昨年度よりも優れた分類性能を得た。昨年度までと同様、本手法は実際の観測に適用され、より詳細な分光観測を適用する超新星候補の選定に貢献した。

統計解析グループは宇宙論グループと共同で宇宙論パラメータを推定するための統計量について評価する手法の検討を行った。この成果を日本天文学会など関連する学会や研究会で報告した。さきがけ研究「タイムドメイン宇宙観測用動画データの高速逐次処理法の開発」(PI 酒向重行)との共同研究を行い、地上中間赤外線観測に関して、大気放射を除去するための新たな方法であるスロースキャン観測の方法を提案した。

データ基盤グループは観測データ解析パイプラインにワークフローシステム Pwrake を組み込み、Pwrake を用いた複数ジョブの並列実行を評価した。Gfarm を高速化するために高速トランザクション処理技法を開発した。Gfarm のメタデータサーバのコード分析を行った。宇宙理論シミュレーションデータ分析に関して、空間索引の構築を劇的に高速化する必要があり、CPU に最適化した STR-R-R 木を提案した²⁾。さらに、相関関数計算のための厳密解法と近似解法の結合に取り組んだ。天体カタログデータ基盤の高性能化を検討すべく、データ偏りがある場合のシャッフル処理を高性能化した。

1) Akisato Kimura, Ichiro Takahashi, Masaomi Tanaka, Naoki Yasuda, Naonori Ueda and Naoki Yoshida, "Single-epoch supernova classification with deep convolutional neural networks", 2017 IEEE 37th International Conference on Distributed Computing Systems Workshops (ICDCSW), pp. 354-359, 2017

2) 三橋 龍也, 川島 英之, 西道 啓博, 建部 修見, 「CPU に最適化した STR-R 木による空間結合殻カウント」, 電子情報通信学会論文誌, 招待論文, Vol.J100-B, No.12, pp.961-973, 2017