

「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」

H26 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／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 大学院理学系研究科／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

### (2)「統計解析」グループ

- ① 主たる共同研究者: 池田 思朗 (情報システム研究機構 統計数理研究所、准教授)
- 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、上席特別研究員)
- 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 システム情報系(計算科学研究センター)、講師)

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

本研究の目的は、すばる望遠鏡に搭載された広視野撮像装置(Hyper Suprime-Cam)を用いた広域宇宙観測のデータを解析するアプリケーションを開発することである。画像ビッグデータ解析により宇宙の物質分布や遠方超新星の距離を測定し、宇宙進化の歴史を明らかにすることができる<sup>1)</sup>。観測計画は平成 26 年度に開始され、支障なく観測を続けている。

宇宙論グループは、宇宙の構造形成の大規模理論シミュレーションを開始し、HSC サーベイの解析に必要な模擬観測カタログを 10 例作成した。これは全天の観測を想定したものであり、1 つにつき HSC サーベイに対応する現実的な観測領域を 20 個以上切り出してくることができる。計 200 個の模擬カタログを用いて統計誤差を正確に計算した。次に、バリオン物質を取り入れた構造形成シミュレーションを行い、重力レンズ観測の統計解析の際に生じるパラメータバイアスについて定量的に明らかにした。

統計解析グループは統計学による超新星分類の検討を行った。超新星はいくつかのタイプに分類される。中でも Ia 型超新星は、地球からの距離の推定法が確立されており、本研究でも重要な天体である。観測により検出された突発天体が Ia 型かどうか判定するためには、一般に突発天体の色やその時間変動などの情報を集める必要がある。効率良く、必要最小限の情報によって精度良く超新星の分類を行うことは本研究の重要な課題である。我々はこの問題を統計的な変数選択問題と捉え、提唱されているいくつかの観測的指標を検討した。その結果、以前から認められている 2 つの変数、すなわち色指数と光度変動時間以外に、推定の精度を向上させる変数が見つからないことを示した。この結果は論文として投稿した。

画像解析機械学習グループは、画像解析の基盤として多数の画像の取得および操作環境の整備を行った。異なる時期と異なるスペクトルバンドで撮影された大量の画像から時間変動天体を検出するための基礎的検討を行い、画像特徴量、および検出手法に関する基礎検討を行った。

データ基盤グループはパイプラインシステムの高性能化、超新星爆発検出基盤の検討、近傍銀河検索処理の高性能化に関する研究を行った。超新星爆発検出基盤の検討について、配列データベースにおけるウィンドウ集約演算を高性能化する手法を提案した。高性能化のためにインクリメンタル計算スキーム(右上図)を適用した。提案手法を代表的な配列データベースである SciDB に実装して実験的に評価した。その結果、提案手法は従来手法に比べて最大 28 倍も高速であった。

