



ISSN 1349-1113
JAXA-RR-12-006

宇宙航空研究開発機構研究開発報告

JAXA Research and Development Report

宇宙科学情報解析論文誌 第二号

2013年3月

宇宙航空研究開発機構

Japan Aerospace Exploration Agency

序 文

宇宙科学情報解析論文誌（Journal of Space Science Informatics Japan）第二号をお送りします。

JAXA 宇宙科学研究所・学際科学研究系では、その前身である宇宙科学情報解析研究系、宇宙科学情報解析センター（PLAIN センター）の時代より、主に科学衛星・探査機によって得られた観測データを収集し、広く国内外の研究者による利用を促進するためのサイエンスデータベースの開発、並びにデータ解析手法の研究を行ってきました。また、それに関連した研究発表および情報交換の場として、毎年「宇宙科学情報解析シンポジウム」を開催してきました。

近年、宇宙科学における様々な分野において、観測装置と計算機の大型化によって、以前に比べて飛躍的に大量のデータを取得することが可能になってきています。そのために、観測から科学的成果を産出するにあたって、大規模データをいかに効率良く解析し必要な情報を引き出すかが、新たな技術課題となりつつあります。また、デジタル技術の普及に伴い、デジタルプラネタリウムや様々なアプリケーションに代表されるように、宇宙観測データが科学研究以外の分野でも広く利用されるようになってきています。

そのような状況を鑑み、当研究系では、宇宙科学データ（シミュレーションデータを含む）に関する新しい処理、解析、利用の手法、データベース技術やそれを応用したシステムの構築・運用技術など、宇宙科学、情報科学、情報技術に関わる成果発表を集めた「宇宙科学情報解析シンポジウム」を引き続き開催するとともに、それに関連した研究論文を広く募集し、「宇宙科学情報解析論文誌」を発行して参ります。当論文誌が、周辺領域の研究者、技術者の間の情報交換や活発な議論につながることを期待しています。

2012 年 12 月

JAXA 宇宙科学研究所

学際科学研究系

海老沢 研、篠原 育、三浦 昭、山本 幸生

目 次

| | |
|---|----|
| 「大規模粒子シミュレーションデータ可視化ツール Zindaiji3 の開発について」 武田 隆顕 | 1 |
| 「宇宙科学データの可視化- モバイル環境に適した科学データの取り扱いについて -」 三浦 昭, 海老沢 研 | 9 |
| 「動的タイムワーピング距離を用いた X 線天文データの類似検索」 林 史尊, 天笠 俊之, 北川 博之, 海老沢 研, 中平 聡志 | 19 |
| 「MAXI/GSC による全天観測データ公開システムの開発」 中平 聡志, 海老沢 研, 根来 均, 三原 建弘, 杉崎 睦, 芹野 素子, 諏訪 文俊, 浅田 真, 富田 洋 | 29 |
| 「『あかり』中間赤外線全天 diffuse マップの作成」 石原 大助, 金田 英宏, 毛利 彰男, 近藤 徹, 鈴木 慧士, 山岸 光義, 大藪 進喜, 尾中敬 | 39 |
| 「トロムソ/ロングイアビンにおけるオーロラ観測用並列イメージャーの大量データ処理と 可視化」 小川 泰信, 門倉 昭, 元場 哲郎, 田中 良昌, 細川 敬祐 | 51 |
| 「IUGONET 解析ソフトウェアの現状と今後の発展」 田中 良昌, 新堀 淳樹, 梅村 宜生, 堀 智昭, 阿部 修司, 小山 幸伸, 林 寛生, 上野 悟, 佐藤 由佳, 谷田貝 亜紀代, 小川 泰信, 三好 由純, 関 華奈子, 宮下 幸長, 瀬川 朋紀 | 63 |
| 「IUGONET メタデータ登録・管理システムの処理性能評価」 堀 智昭, 梅村 宜生, 阿部 修司, 小山 幸伸, 田中 良昌, 林 寛生, 上野 悟, 新堀 淳樹, 佐藤 由佳, 八木 学 | 71 |
| 「宇宙天気情報サービスとサイエンスクラウド」 亘 慎一, 加藤 久雄, 村田 健史, 山本 和憲, 渡邊 英伸, 久保 康文, 國武 学 | 79 |

| | |
|---|-----|
| 「月・惑星探査データ表示・解析システム『WISE-CAPS』」 | |
| 寺藺 淳也, 中村 良介, 児玉 信介, 山本 直孝, 平田 成, 小川 佳子, 出村 裕英 | 89 |
| 「モデル駆動型システムズエンジニアリングに基づくモデル管理およびデータ解析 ーNano-JASMINE データ解析への応用ー」 | |
| 初鳥 陽一, 宮下 尚, 清水 淳也, 山田 良透 | 103 |
| 「xmlExcelHyper ¹ を用いた SIB2UI ² の開発 ¹ Excel 上で複雑な XML を扱う汎用ライブラリ ² 衛星情報ベース 2 作成ツール」 | |
| 松崎 恵一, 山下 美和子, 馬場 肇, 西村 佳代子, 宮野 喜和 | 113 |
| 「科学衛星のための衛星時刻校正システムの整備」 | |
| 岡田 尚基, 山本 幸生 | 131 |
| 「衛星自動運用システムの開発」 | |
| 永松 弘行, 齋藤 宏文 | 137 |

大規模粒子シミュレーションデータ可視化ツール Zindaiji3 の開発について

武田 隆顕*

About new particle data visualization tool ” Zindaiji3”

Takaaki TAKEDA*

Abstract

Visualizing 3-dimensional or 4-dimensional data is important in both aspects of tools for research and tools for public outreach. We have once developed 32-bit application “Zindaiji” to visualize particle simulation data and used it to make high quality movies. Recently, simulations with much large particle number are commonly performed. Thus, we are now developing 64-bit application “Zindaiji3” for visualization of large particle number data. Here, we report the current status of “Zindaiji3” .

Keyword: Visualization

概要

3次元, または4次元(3次元+時系列)のデータを可視化することは, 研究を推進する手段としても重要であり, パブリックアウトリーチの手段としても重要である. 国立天文台4D2Uプロジェクトでは, 粒子データを可視化するツール(Zindaiji)を開発し, それによる可視化映像の作成と公開を行なってきた. 近年より大規模な粒子シミュレーションが行われるようになり, そうしたデータに対応するべく新規に64-bit版として可視化ツール(Zindaiji3)の開発を開始した. Zindaiji3の開発の現状と今後について報告する.

1. はじめに

国立天文台4D2Uプロジェクトは, 最新の天文学の成果を可視化するためのプロジェクトである. 正式名称は4次元デジタル宇宙プロジェクトであり, 英語では4-Dimensional Digital Universeであり, アクロニム(頭文字)を使って4D2Uと読んでいる. 4D2Uにはまた, ”4-D to you”(4次元をあなたに), という意味も込められている. ここでの「4次元」は宇宙の次元(空間3次元と時間1次元)を意味している. プロジェクトの活動の一環として天文シミュレーションを元にした可視化映像の製作を行い, ウェブ公開やシアターにおける公開を行っている.

数値シミュレーションの可視化には, 一般的に大きく分けて2つの目的が存在する. 一つは研究者自身が直感的に把握しやすい形で計算結果を見るために行われる, 研究の補助ツールとしての可視化である. こうした可視化には, 見た目の綺麗な画像や映像を作ることよりも, 物理量が把握しやすく, 簡単に視点の移動などの操作ができるといった素早い応答性が最も重要な要素になる. もう一つは, 研究の結果を単なる点や線による図やグラフではなく, 専門外の人間に対しても理解しやすく印象的な画像, 映像を作成するための可視化, つまり「見せる」ための可視化である. 後者の可視化は, 研究成果を社会に還元するパブリックアウトリーチという意味でも重要である. 4D2Uプロジェクトでは主に後者に重きを置いた, 天文シミュレーションの可視化をおこなってきた.

粒子シミュレーションは, 天文学の分野では非常に多くみられるシミュレーションである. なぜなら, 星団内の星の運動や, 微惑星からの惑星の形成など, 天文学のトピックの多くで質点とみなせるような粒子同士の重力相互作用が重

* 国立天文台天文シミュレーションプロジェクト
(Center for Computational Astrophysics, National Astronomical Observatory of Japan)

要な役割を果たすためである。我々は、粒子系シミュレーションの可視化映像を製作するにあたり独自の可視化ツール (Zindaiji) の開発を行い、主にこのツールを用いて映像製作を行ってきた²⁾。Zindaiji は C++ で開発され、32-bit Windows 環境で動作するものとなっている。Zindaiji は、テクスチャの付加やカメラワークの編集機能の実装といった、CG 作成ソフトに近い機能を重視した粒子データ用の可視化ソフトウェアとして開発を行ったものである。Zindaiji は半径を持った球、もしくは光る点 (星) や簡易的なガス粒子を大量に表示する機能をもち、われわれはこれを用いて「月の形成」や「渦巻き銀河の形成」といった、粒子系の天文シミュレーションの映像化 (図 1) を行ってきた。

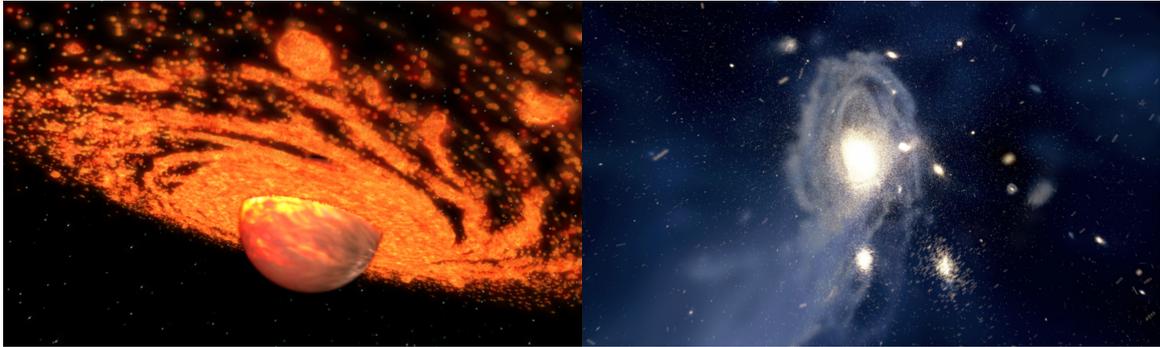


図 1. 左)「月の形成」のシーン。巨大衝突によって原始地球周りに撒き散らされた月の材料が集積する過程を示している³⁾。右)「渦巻銀河の形成」のシーン。星をあらわすスーパーパーティクルとガスをあらわす SPH 粒子を表示している。星とガスに 100 万粒子を用いたシミュレーションの可視化映像である。⁴⁾

初代 Zindaiji は 32-bit Windows アプリケーションとして開発されたため、主に利用できるメモリの制約から約 200 万粒子程度までが扱えるデータ量の上限となっていた。今日では、100 万粒子を用いる規模のシミュレーションは一般化し、より多くの粒子を用いた高分解能のシミュレーションも数多く行われている。

そこで、我々はより大規模なシミュレーションの可視化を行うために、64-bit アプリケーションとして Zindaiji3 の開発を開始した。Zindaiji3 では新たにコード開発を行い、大規模データに対応すると同時に、初代 Zindaiji の GUI における操作性上の不具合などを解消した、より強力で操作しやすいツールを目指して開発を行っている。セクション 2 では、Zindaiji3 の主な強化点について述べ、セクション 3 でまとめと今後について述べる。

2. Zindaiji3 における強化点

2.1 タイムライン GUI の強化

時間変化するデータを連続的に把握するためのもっとも素直な可視化方法は、時系列に沿ってスナップショット毎に連番画像を作成して、最後にそれをつなげた映像ファイルにすることである。しかし、本来は視点や時間、可視化用のパラメータなどを自由自在に操って、それをディスプレイ上に反映させることができることが望ましい。そうした、時間変化するデータを扱うためのアプリケーションで、もっともユーザーインターフェイスが進んでいるものとして映像編集用もしくは 3DCG アニメーション製作用のアプリケーションがあげられる。それらのソフトウェアでは、タイムライン上にカメラの位置やそのほか必要なパラメータのキーフレームを配置し、時間軸上でパラメータ値を自由に編集することで多種多様な映像の作成を行っている。初代 Zindaiji においてもタイムライン編集を軸にした可視化を行えるツールの開発を目指していたが、経験の不足から十分に目標を達成することができなかった。Zindaiji3 ではタイムライン編集機能を大幅に強化しパラメータのより自由な編集が可能となった (図 2)。

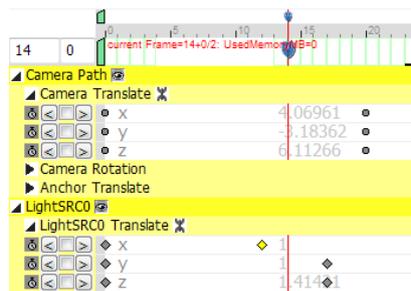


図2. Zindaiji3 になって強化されたタイムライン GUI. タイムライン上のキーフレームの配置やマウスによるドラッグといった機能や、表示非表示の選択を行える。タイムライン上部の緑のボックスは、メモリ上にデータが展開された時間ステップを表している。ユーザーの操作待ちの時間に、あらかじめ先の時間ステップのデータを読み込むプリフェッチ機能(セクション 2.3 参照)を実装することで、速い応答性と操作性の向上が図られている。

2.2 メモリ制限の軽減

32-bit アプリケーションでは、ソフトウェア上の制約から 2GB 程度のメモリしか使用することができないが、64-bit アプリケーションではメモリによる制限は大幅に軽減される。Zindaiji3 は 64-bit アプリケーションとして開発しているため、PC に積んでいる物理的なメモリ量、もしくは描画速度からくる制限が実質的に抜える粒子数の上限を決める。図 3 に、1 億体の粒子を用いたシミュレーションデータ内のフライスルー映像のスナップショットを示す。粒子の xyz 座標を単精度でメモリに保持していた場合、1 億体の粒子あたりに必要なメモリは 1.2GB になる。ただし、Zindaiji3 では一般的に時系列データを扱うため、速度、加速度や粒子の ID といった情報を保持している。そのため、それら全てのデータの保持に 6GB ほど使用する。また、元データとは別に補間済みのデータを作成しているため、それを保持するために 2GB メモリを使用しており、1 億体のデータを保持した状態で 8GB 程度のメモリ使用量となった。2012 年現在、16 ~ 32GB 程度のメモリを搭載した PC の調達是比较的容易であり、1 台の PC を用いて億の単位までは可視化を行うことができることになる。

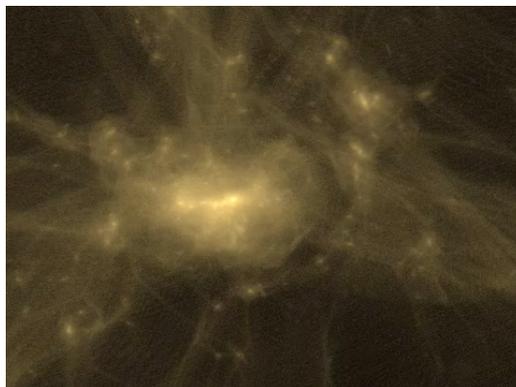


図3. 1億粒子を用いたシミュレーション内のフライスルーの一場面。(データ提供：石山智明) 5). 元のシミュレーションは約 170 億体の粒子を用いているが、そのうち特徴的な構造を含む領域から一億体分のデータを抜粋。映像の品質を上げるために、粒子は点 (GL_POINTS) ではなく、カメラの方向を向きテクスチャを張った小さな四角いポリゴンとして表示されている。

2.3 高速化と操作性の向上

64-bit 化に伴ってメモリに関する制約が少なくなったために、Zindaiji3 の実質的な有用性は、描画の速さや待ち時間の少なさに依存することになる。Zindaiji3 では実行速度の向上のためにアルゴリズムの見直しを行い、描画速度の向上とマルチスレッド化による待ち時間短縮が図られている。

OpenGL を用いた描画時間には、データの転送や描画そのものにかかる時間の他に、ステートの変更にかかる時間があげられる。ステートの変更とは、描画する色の変更や使用するテクスチャを変更する作業を指し、頻繁にステートの変更を行うと描画が遅くなることが知られている。ステートの変更をできるだけ少なくすることは 3DCG を用いたゲームなどで、高速化の基本的なテクニックとされている。

さまざまな物理量を持ち、色や粒子の種類（ガスと星など）が入り乱れたデータを描画する際には、最も単純に実装をすると図4の一番左の列のように、ループを用いて一つ一つの粒子の設定を行う描画ルーチンになる。

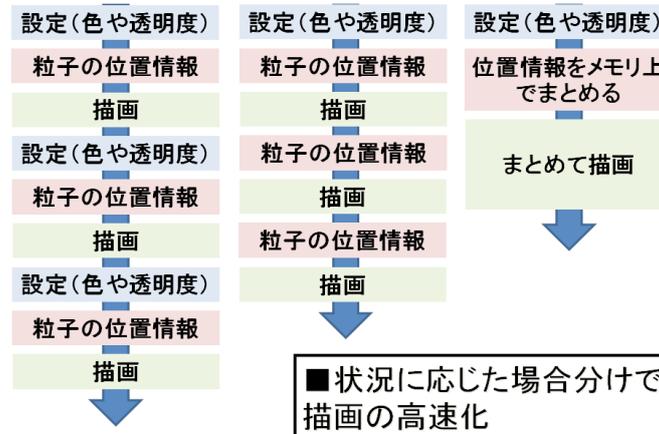


図4. Zindaiji3の描画に用いた場合分けを多用した効率化の模式図. 左が最も単純なループを用いた描画ルーチン.

しかし、描画の設定によっては、こうした毎回のステータスの変更は必要が無い場合がある。Zindaiji3においては、描画の設定に応じて場合分けを行い、ステータスの変更が必要のないデータに対してはステータス変更の無いルーチンへ分岐して描画をすることで、描画速度の向上を図っている。また、連続して同一種類の粒子を描画するような場合には、ステータス変更をスキップすることで描画時間の短縮が図れる。ガスと星からなる銀河衝突のデータを用いて確認してみたところ、Zindaiji3では初代Zindaijiと比べてほぼ半分の時間で描画することができ、描画アルゴリズムの変更によって高速化を行えることが確認できた。図3で用いた1億体の粒子は、すべて同一の粒子として描画されており、ステータスの変更はほとんど行われていない。このデータの場合では、グラフィックボードとしてNvidiaのGeForce 240を用いて、1枚の描画に約6秒の時間がかかっている。

上記の考えを拡張すると、ガスと星など描画方法の違う粒子が入り混じっている場合には、粒子の種類に応じてソートされた順番で描画をすることで描画速度の向上が期待できる。しかし、以下の理由により実際には適用できない場合が多い。SPHシミュレーションにおけるガス粒子のように不透明な粒子を描画する場合には、半透明なポリゴンを配置して表現を行う（図5）。この際、ポリゴンの相互の隠蔽の判定が行われるが、単純に粒子の配列順に表示を行うと、先に描画した手前のポリゴンの半透明部分によって奥のポリゴンが隠されるという現象が起きてしまう。そのため、半透明な物体の描画は奥からソートされた順番で行う必要があり、粒子の種類によってソートされた順番で描画を行うことができないことになる。そのため、現時点でZindaiji3では粒子の種類に応じたソートの実装はおこなっていない。

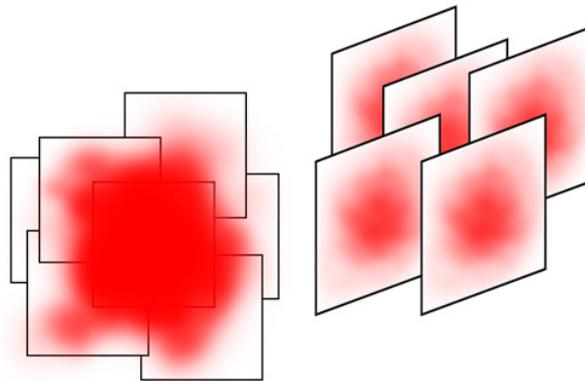


図5. SPH粒子などのガス粒子の表示方法の模式図. 半透明な板を透明度に応じて奥から表示することで、ガスの分布を表現する。

動画を作成するには描画以外にも、さまざまな処理による待ち時間が存在する。Zindaiji3では、それらの処理をレッド並列で処理することで待ち時間の短縮を図っている。具体的にはデータ読み込みと補間、描画後の画像出力を描

画とは別スレッドで行えるように作成し、描画を行っている間、もしくはユーザーによる操作待ちの間に次の時間フレームの読み込みなどを行うことができるようになってきている（図6）。また、ユーザーの操作待ちの時間には、次の時間フレームのデータに限らず、その先の時間フレームも予め読み込むプリフェッチ機能を実装している。これらの実装によって、操作時及びレンダリング時の待ち時間が減少し、操作性は大きく向上している。

動画作成時の速度を主に律速するのは、データ読み込み処理と描画処理のうち、より時間のかかる処理になる。粒子をすべて光の点として表示する場合には、前述したように1億体の描画に約6秒かかっている。一方1億体分のデータファイルは6GB程度のサイズになるために、HDDからの読み込みには1分ほどかかる。この場合には、時間フレーム間を10コマほどで補間するアニメーションを製作すると、描画時間が1分ほどでデータ読み込み処理にかかる時間と同程度になり、マルチスレッド処理がもっとも効率的となる。

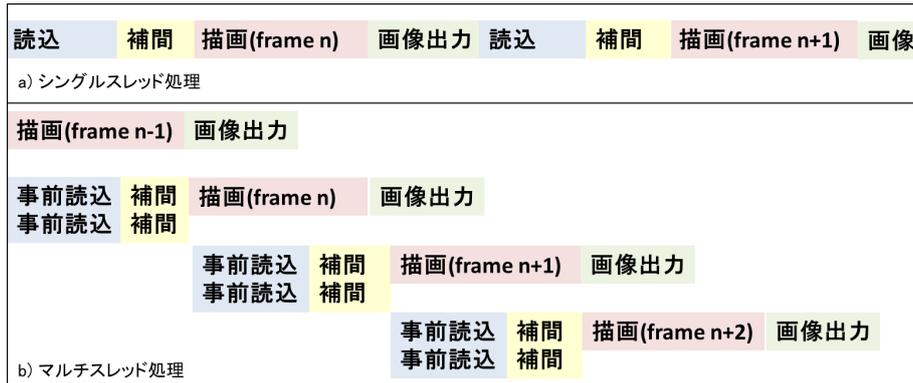


図6. Zindaiji3で実装したマルチスレッドによる並列化の模式図. a) シングルスレッドのみを使って、順番に処理を行う場合の模式図. b) 平行して処理を行うことで、高速化する場合の模式図. 事前読み込みと補間の処理が2行で書かれているのは、これらの処理をマルチスレッドで処理していることを表している。描画と画像出力はそれぞれ1スレッドを用いている。

2.4 描画品質の向上

大量の粒子データを描画するためには、描画品質の向上も重要である。たとえば単純に一粒子あたり一ピクセルで描画すると、大粒子数のデータでは簡単に画面を埋め尽くしてしまい、見たい構造が取り出せなくなる。また、図5のようなガス粒子を表示する場合、粒子数が多ければポリゴン一枚あたりの透明度が大きくなり、8ビット色深度のバッファ上で表示するには精度が足りなくなり、カラーバンディングが発生する。これらの問題から、Zindaiji3では16ビット色深度の2048×2048又は4096×4096ピクセルのバッファFrame Buffer Object (FBO)上に描画を行い、その結果を最後に画面上に表示をしている。これによりカラーバンディングの発生を抑制し、見かけ上の1ピクセルよりも小さいサイズの粒子の表示も扱うことができるようになった。

また、Zindaiji3は専門家以外の人に魅力的に感じる可視化映像を作ることが大きな柱になっている。そのため、追加機能としてより綺麗な出力をする機能を複数実装した。以下はそうした機能をいくつか列挙する。

大きさを持った球として粒子を表示する場合、影を実装することによって立体構造の表現力は大きく向上する。初代Zindaijiにおいても1つの光源からの影をシャドウボリューム法⁹⁾によって実装していたが、Zindaiji3ではFBO上で表示された画像を複数合成することによって、複数の光源からの影を表現する拡張を行った。光源を複数配置することで凹部に光が侵入しづらいという物理が再現され、複雑な形状でもその立体感が強調される（図7）。

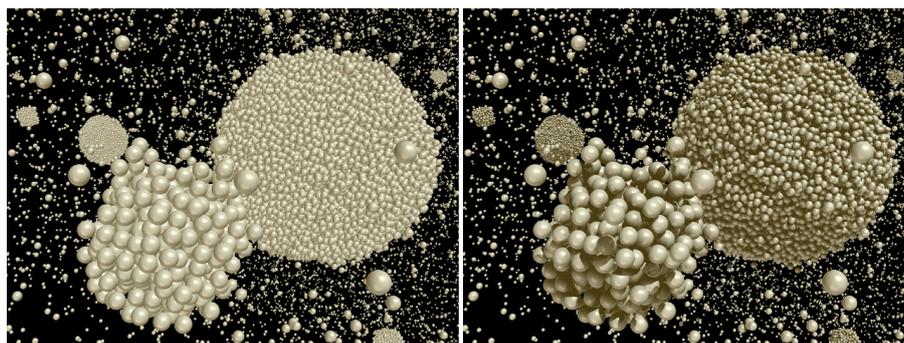


図7. 分子動力学計算による核生成過程のスナップショット. 左) 影のない表示. 右) 複数の光源からの影を入れた表示. 影によって凹凸がはっきりするほか、粒子の密集度に関しても受ける印象が変化する。(データ提供: 田中今日子)⁷⁾

コンピュータグラフィックにおいては、フォトリアルな表現を目指すほか、トゥーンレンダリングのように、より象徴度を高めたレンダリングを行うことがある。Zindaiji3 では、GLSL を利用することでトゥーンレンダリングの実装を行った(図8)。トゥーンレンダリングには、イラスト的でスタイリッシュな出力が得られるという単なる見かけの目新しさの他に、輪郭線が強調されることにより影に隠れた部分も認識しやすいなどの利点がある。

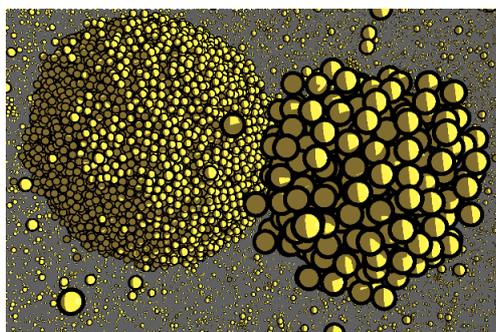


図8. トゥーンレンダリングを用いた表示. 用いたデータは図7と同じもの。

3. まとめ

Zindaiji3 によって、今まで用いていた初代 Zindaiji と比べ 100 倍程度、億の単位の粒子シミュレーションの可視化映像の作成を行うことができるようになった。演出を含めた映像制作においては、カメラワークや見た目の調整のためにトライアンドエラーを繰り返すことになり、できるだけ小回りの利く可視化プロセスであることが望ましい。そのため Zindaiji3 では、まず 1 台の PC でできる限り大規模なデータを綺麗に出力できる可視化ツールを目指し開発を行った。現時点では実装されていない機能はまだ多く残っているが、ある程度その目標を達成することができた。しかし、最新のスーパーコンピュータをフル稼働するようなシミュレーションでは、数百億といった粒子数を扱うことがある。そうした本当に大規模なデータの可視化は、もはや 1 台の PC での実行は難しく、少なくとも複数の PC を用いた作業が必要になると考えられる。次の大きな目標は、複数の PC を用いてより大規模な粒子データの可視化を行うことのできる、ネットワークレンダリングの実装となる。Zindaiji3 は、GUI 部分にマルチプラットフォームなライブラリである wxWidgets⁸⁾ を利用し、Windows/LINUX/Mac 上で実行することができる(ただし現在 Mac 版のみ 32-bit となっている)。現在ソースコードも含めて以下の URL で公開を行っているの、興味のある方はぜひダウンロードして試してみたい。

<http://www.cfca.nao.ac.jp/~takedatk/COMPUTER/ZINDAIJI3/Zindaiji3Top.html>

参考文献

- 1) 小久保英一郎ほか: 4次元デジタル宇宙プロジェクト, 情報処理, Vol. 45, No.12 (2004) pp. 1229-1233
- 2) 武田隆顕: 大規模天文シミュレーションの可視化 - 国立天文台 4D2U プロジェクトでの試み, 可視化情報学会誌 Vol. 28, No. 110 (2008) pp. 186-192

- 3) Takeda, T. and Ida, S.: Angular Momentum Transfer in the Protolunar Disk, *The Astrophysical Journal*, Vol. 560, (2001), pp. 514-533
- 4) Saitoh, T. R., Koda, J., Okamoto, T., Wada, K., and Habe, A.: Tidal Disruption of Dark Matter Halos around Proto-Globular Clusters, *The Astrophysical Journal*, Vol. 640 (2006) pp. 22-30.
- 5) Ishiyama, T. (private communication)
- 6) Crow F.C.: Shadow algorithms for computer graphics, *Proc. SIGGRAPH' 77*(1977), pp242-248.
- 7) Tanaka, K., Tanaka, H., Yamamoto, T., and Kawamura, K.: Molecular dynamics simulations of nucleation from vapor to solid composed of Lennard-Jones molecules, *The Journal of Chemical Physics*, Vol. 134 (2011), 204313
- 8) <http://www.wxwidgets.org/>

宇宙科学データの可視化

ーモバイル環境に適した科学データの取り扱いについてー

三浦 昭^{*1} 海老沢 研^{*1}

Visualization of Space Science Data ~Science Data Handling Suitable for Mobile Devices~

Akira MIURA^{*1}, Ken EBISAWA^{*1}

Abstract

We have been working on methods of visualizing space science data. For the purpose of education and public outreach, devices built on mobile computing platforms (iOS and Android, etc.) are fascinating ones to represent space science data. While mobile devices require a large amount of computing resources to visualize science data, wireless data services for mobile devices are not yet fast enough to instantly transfer the sufficient amount of data. This paper introduces a method to progressively download science data in order to reduce the latency. The proposed method here is also expected to reduce memory usage and CPU usage in comparison with conventional space science data I/O libraries.

Keywords: visualization, space science data, mobile platform

概 要

かねてより筆者らは宇宙科学データの可視化について検討してきた。最近の iOS や Android 等に代表されるモバイルデバイスは、教育・広報の観点からも無視できない存在になっている。科学データの可視化の1つの手段として、モバイルデバイスの性能は年々向上しているが、モバイル通信環境は未だ発展途上にあり、科学データを即時に転送できるまでには至っていない。本稿では、利用者を待たせることなく科学データを可視化するために、科学データを逐次取込みながら処理するための手法について提案する。本手法は従来手法と比較して、データ転送に伴う待ち時間を短縮するのみならず、デバイス内のメモリやCPUの使用率低減にも寄与することが期待される。

*1 宇宙航空研究開発機構宇宙科学研究所
(Institute of Space and Astronautical Science, Japan Aerospace Exploration Agency)

1. はじめに

筆者らは、従来から JAXA が保有する宇宙科学データを広報や教育等の目的に供すべく、可視化や可聴化についての技術検討を行ってきた^{[1][2][3]}。かような目的に適ったデバイスは、かつては PC であったが、昨今ではスマートフォンやタブレット等、モバイル型のデバイスが急速に普及しており、その勢力は無視できない規模となっている。本稿では、まず研究室や家庭の PC と比較して低速の CPU やモバイルデータ通信サービス、低容量のメモリ等、制約の多いモバイルデバイスにおいて科学データを取り扱う際の考察を行う。続いて、かような環境において、利用者の待ち時間低減やデバイスのメモリ効率向上を目的とした、科学データの逐次取込みと可視化の併用について提案する。本稿で提案する手法により、モバイルデバイスを利用した科学データの教育・広報が一層拡大することを期待するものである。

1.1 対象とするデータ

本稿では、「科学データ」の一例として、X 線天文衛星「あすか」や「すざく」の観測データを取り上げる。X 線天文衛星の観測データは、衛星に搭載された X 線望遠鏡が観測した個々の光子の飛来時刻や飛来方向等を時系列に保存して FITS^[4]形式にしたものである。以後本稿では、特に言及しない限り、「科学データ」は X 線天文衛星の観測データ (FITS 形式) を指すものとする。応用として、類似のフォーマットを有するデータに対しては、本稿で提案する手法が同様に適用できるものと考えられる。例えば筆者らが可視化・可聴化を試みた^[3]PWS データもこの類である。

1.2 モバイルデバイスに纏わる動向

1.2.1 ハードウェア性能

iPod touch 登場以降のモバイルデバイスの高度化・高機能化は著しく、iPod touch から iPhone、iPad に連なる iOS デバイスや、各社が参入する Android デバイス等がしのぎを削っている。これらのモバイルデバイスは急速に普及しており、2011 年末時点の iOS デバイス累計販売台数は 3 億台以上¹、Android デバイスが 2 億 5 千万台以上²とされている。昨今の販売台数は、もはや PC を超えており³、携行用のデバイスとして主流になっていると言っても過言ではない。

最近の典型的なデバイスのスペック比較を、表 1 に示す。表中、CPU、RAM、解像度は主要家電メーカー及び携帯電話各社の公表値による。モバイルデバイスに搭載される CPU は、PC (デスクトップ機等) 用とアーキテクチャが異なるため、同一 CPU クロックの最新 PC より性能は数段劣る。通信環境はモバイルデータ通信サービス (iOS、Android) 及び有線のデータ通信サービス (家庭用 PC) について、関連する情報サイトの実測値^{4,5,6,7}を参考にした。なお無線アクセスポイントを用いる等、通信環境に配慮することによって、モバイルデバイスであっても、これ以上の通信速度を享受することは可能である。

表 1 各種デバイスのスペック比較 (2012 年夏の例)

| 系統 | 形態 | CPU [Hz] x コア数 | RAM [Bytes] | 画面解像度 | 通信環境 |
|---------|---------|-------------------|-------------|--------------|-----------|
| iOS | スマートフォン | 800M x 2 | 512M | 960 x 640 | 数 Mbps ~ |
| | タブレット | 1G x 2 | 1G | 2048 x 1536 | 数 Mbps ~ |
| Android | スマートフォン | ~ 1.5G x 2 ~ | 1G ~ | ~ 720 x 1280 | 数 Mbps ~ |
| | タブレット | ~ 1.5G x 2 ~ | 1G ~ | ~ 1280 x 800 | 数 Mbps ~ |
| 家庭用 PC | 据置 | 2G ~ x 2 ~ | 2G ~ | 1600 x 900 ~ | 数十 Mbps ~ |
| | ラップトップ | 1.5G ~ x 2 ~ | 2G ~ | 1366 x 768 ~ | 数十 Mbps ~ |

1 http://www.excite.co.jp/News/pc/20120217/Cobs_ip_201202_2011iosmac28.html

2 <https://plus.google.com/+LarryPage/posts/jcyvVa5K4JW#+LarryPage/posts/jcyvVa5K4JW>

3 <http://japan.cnet.com/mobile/35013776/>

4 <http://www.musen-lan.com/speed/htmldata/>

5 http://www.bsptest.jp/stat1_1.html

6 <http://wimaxspeedmap.com/>

7 http://mmd.up-date.ne.jp/news/detail.php?news_id=1014

PCの上位機種はこれらに比べると圧倒的な性能を実現しており、研究室等での解析・閲覧用途としてはモバイルデバイスと比べてPCが有利であることは言うまでもないが、家庭用等の低価格帯の製品を比較した場合、携帯会社の販売促進等の効果もあり、モバイルデバイスのコストパフォーマンスは決して悪くない。実際、ネットブックと呼ばれる低価格ノートPCに匹敵するスペックのデバイスも少なくない。とは言え、モバイルデバイスに搭載されるCPUの性能やメモリ容量、バッテリー容量等はPCと比較して決して潤沢とは言えないので、大きな科学データを扱うアプリケーション開発においては、例えば高負荷の解析処理等、電力使用量が高くなる処理を避けつつ、同時にメモリ使用量も考慮する必要がある。

1.2.2 アプリケーション実行環境

モバイルデバイスとPCとでアプリケーションの実行環境を比較する。

PCにおいては、ブラウザの多機能化・高性能化に伴い、FlashやJava、JavaScript等を用いたWebベースのサービスが隆盛を極めている。これらのサービスは、ネイティブコードのアプリケーションに比べると、中間言語やスクリプティング言語を用いる分、性能面での不利益はあるが、PCの性能向上がそれを補っている。一方でPC用のアプリケーションをインストールすることは、利用者がインストーラを能動的に取得した上で起動し、さらに幾つかの手間をかけることであり、一般の利用者にとっては敷居の高いものである。

これに対してモバイルデバイスでは、iPodの楽曲ダウンロード(iTunes)に始まる諸々のダウンロードサービスがアプリケーション配布にも拡大されており、様々なアプリケーションの取得やインストールは、とても簡易なものになっている。iOS向けアプリケーションはネイティブコードであるため、ブラウザベースのサービスと比べて、ハードウェア性能を十分に発揮できる。Androidではハードウェア環境の互換性に重きを置いたJavaアプリケーションと、性能に重きを置いたネイティブコードのアプリケーションと2種類の開発環境が提供されている。

1.2.3 通信環境

端末のハードウェア環境は着実に進化しているモバイルデバイスであるが、通信環境は必ずしも劇的な進化は遂げていない。家庭(建家)内の通信環境がADSLからFTTHに進化し、インタフェース速度が100M~1Gbpsになりつつあるのに対して、モバイルデータ通信サービスも高速の新商品が次々と発表されており、カタログスペックは数十Mbpsに至っているが、実測値は測定箇所によって様々である。1.2.1節で参照した情報サイトに記された各地点の通信速度から推察するに、場所を限れば10Mbpsを超えるサービスを受けられる状況になりつつあるが、通勤や行楽等、広範囲のモバイル環境でサービスを受けられる速度としては数Mbps程度からであると考えられる。

すなわち、現在のモバイルデバイスは、処理能力は高くなっているが、それに供するデータを高速に取得することは必ずしも容易ではない状況にある。

そこで本稿では、低速のネットワーク環境においても利用者を待たせることなく科学データを取得し、可視化する手法の検討について述べる。

2. モバイル環境における科学データの可視化

2.1 配信方法の比較(配信と可視化の順序)

モバイル環境で科学データの可視化を実現するにあたっては、大きく2つの方法が考えられる。ひとつは科学データを予め映像化して配信するものである。もうひとつは科学データそのものを配信してデバイス上で可視化するものである。執筆時点でiOS向けのアプリケーションを“FITS”というキーワードで検索した結果、FITS形式に対応していることを謳っているものは、フランスの線形加速器研究所(Laboratoire de l'Accélérateur Linéaire)が提供するioda⁸と呼ばれるソフト1件のみであった。このiodaは、研究者向けも含めた種々のフォーマットのデータを表示できるソフトであるが、「あすか」や「すざく」のデータ形式には対応していないものであった。また表示できるデータはローカルのデータに限られるため、インターネット上のデータを表示するためには、事前にデータをダウンロード(もしくはPCからデバ

8 <http://ioda.lal.in2p3.fr/>

スにデータをコピー) する必要がある。これに対して、YouTube 等の映像配信サービスはモバイルデバイスにも対応しており、科学映像も潤沢に提供されている。

2.1.1 映像を配信するメリット・デメリット

インターネット上の映像配信サービスは、視聴の待ち時間を短縮できるメリットがある。昨今の映像配信の手法は、大きく2種類に分けられる。一つはストリーミングであり、もう一つはプログレッシブダウンロードである。ストリーミングは映像配信を制御するセッション (RTSP として標準化されている) と映像のデータグラム (RTP として標準化されているが、これに限定されるものではない) で構成される。映像の再生・停止等の制御や、再生に必要なパラメータ等は RTSP で受け渡される。映像のデータグラムは、予めファイルに保存された映像でもリアルタイムに生成される映像でも配信可能である。映像が途切れることなく再生できるよう、端末側には幾ばくかのバッファが設けられている。端末側のアプリケーションによっては長時間の映像を保持できる場合もあるが、一般には映像全体のサイズに比べてバッファのサイズは僅かである。プログレッシブダウンロードは、予めファイルに保存された映像を配信する手法であるが、映像ファイルの冒頭に再生に必要なパラメータ類を集約することにより、ファイルの受信途中に映像を再生できるものである。受信したファイルは、再利用のため端末側に一時保存されることが多い。

比較的低速の通信環境でも再生できるよう、低解像度の映像から高解像度の映像まで選択できるようになっていれば、通信環境によらず映像視聴の待ち時間は低減できる。また異なるパラメータ設定で映像化したファイルを複数用意しておけば、様々な切り口での映像を視聴することも可能になる。

ただしこの方法は低ビットレートにしても、長時間の視聴に際しては通信量が大きくなる恐れがある。パラメータを変更して視聴する場合は、その都度異なるファイルを受信することとなるため、トータルの通信量は、さらに大きくなる。また映像を提供する側としても、視聴が想定される全科学データについて、予めシナリオを作成して映像化する必要があるため、映像化作業の負担は無視できない。

2.1.2 科学データを配信するするメリット・デメリット

科学データそのものを配信することにより、受信した端末上でデータを操作できるようになる。映像配信に比べて、一度受信したデータを用いて様々な切り口の可視化ができるため、複数の側面から可視化を試みる場合、トータルの通信量が低減できるメリットがある。最近のモバイルデータ通信サービスはパケット定額制を謳うサービスが主流となっているが、料金を定額にする範囲を限定するケースや、定額の代りに高速のサービスを受けられる通信量に上限を設けるケースが多く、トータルの通信量を低減することは利用者にとってもメリットがある。

ここで問題となるのが、科学データの受信に要する時間である。例えば 10MB 程度の科学データの可視化を考えた場合、研究室の LAN 環境や家庭の高速インターネットサービスであれば数十 Mbps 程度の通信環境が期待でき、受信にかかる待ち時間は数秒程度となる。これに対して、モバイルデバイスで想定されるような数 Mbps の通信環境では、受信に数十秒かかることとなる。一般ユーザとして数秒程度ならば待てるかも知れないが、数十秒の待ち時間は許容範囲を超えていると思われる。

2.2 科学データの速やかな可視化にとって望ましい手法とは

以上の点に鑑み、科学データをモバイルデバイスで可視化するにあたって望ましい手法を検討する。端末上でデータを操作でき、総通信量を低減できることを考慮すると、科学データそのものを配信して利用に供することが妥当と考えられる。しかしながら科学データを配信する方法では、長い待ち時間がデメリットとなることは前述の通りである。このような状況下で利用者を退屈させない方策、すなわち科学データ取得の間に「待っている」と認識させないような策を講じなければならない。

良く見られる対策としては、簡単なものは時計マークやプログレスバーで進捗状況を表示するものが多い。これは明らかに待ち時間であると宣言しているものではあるが、利用者が残り時間を把握できるという効果はある。その他、関連する情報を表示して、利用者が読んでいる間に処理が完了するもの等も見受けられる。

これに対して本稿では、映像配信の手法に倣い、受信途中にも可視化処理が可能な、「プログレッシブダウンロード」を科学データに適用することを検討する。検討項目は2点ある。ひとつは科学データを格納したファイルが適用条件を満たしているか、もうひとつは科学データを読み込むための既存ライブラリが条件を満たしているか、である。

2.2.1 「プログレッシブダウンロード」が適用できる条件

ここでは、一般に映像配信で用いられるプログレッシブダウンロードが本稿で取り上げる科学データにも応用可能であるか検討する。ここで実現を目指すのは、科学データを受信しながらの可視化処理である。映像配信サービスにおけるプログレッシブダウンロードを参考に、科学データを格納したファイルの適合条件を挙げると以下のようになる。

- (a) ファイルの中で、可視化に必要なパラメータが冒頭にある。
- (b) ファイルの中で、可視化に必要なデータが時系列に並んでいる。
すなわち、データ取込みの順に可視化できる。

X線天文データを格納する FITS ファイルのデータ構造（「あすか」や「すざく」の例）を表2に示す。可視化に必要なパラメータ類は、ファイルの冒頭にあるプライマリ HDU とエクステンションのヘッダに記されている。それに続くデータ部には、観測された X 線の個々の光子についての情報（飛来時刻、飛来方向、エネルギーに関する情報等）が、観測された時刻順に格納されている。

表2 X線天文データ (FITS ファイル) のデータ構造

| Header & Data Unit の種類 | 構成要素 | 備考 |
|-----------------------------|------|-----------------|
| プライマリ HDU | ヘッダ | ファイル全般の情報 |
| エクステンション (観測データがここに置かれる) | ヘッダ | 観測データに関するパラメータ類 |
| | データ | 時系列の観測データ |
| エクステンション | ヘッダ | 補助的なデータ類 |
| | データ | |

このフォーマットで作成されたファイルのヘッダ情報（可視化に必要なパラメータ類）を読み取った後、データが格納された順に個々の光子に関する情報を可視化すれば、天文衛星が観測した時の状況を時刻に沿って再現できる。すなわち、このフォーマットは前述の、プログレッシブダウンロードが適用できる条件を満たしている。また全データが揃わなくても、個々の光子の飛来方向 (X-Y 方向) 等に基づいて、ある期間の観測データを可視化でき、また、その輝度変化を再現する事もできる。極限すれば、この形式の FITS ファイルは、特殊なフォーマットの動画ファイルであるとも言える。

2.2.2 FITS ファイルの取込み

科学データを「プログレッシブダウンロード」するためには、それに適合した FITS ファイルの読み込み処理が必要である。従来から提供されている FITS ライブラリ (FITS ファイルの読み書きのためのライブラリ) ⁽⁴⁾⁽⁵⁾ はデータの完全性を保証する構造となっており、ファイル全体を読み込んで初めてデータの取扱いが可能となる。従来型 FITS ライブラリのデータ取込み動作を図1に示す。図中ではデータ長を N としている。左端が観測データを表しており、従来型 FITS ライブラリは、先ずこのデータを一括解釈する。その後、個々のデータにアクセスして、可視化することが可能となる。

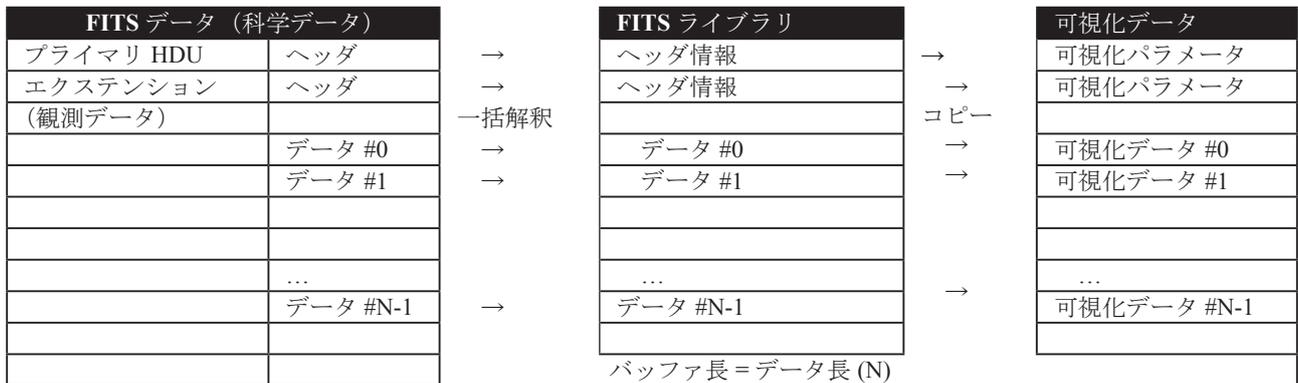


図1 従来型 FITS ライブラリのデータ取込み動作

これは研究者等に対してデータの完全性を保証する上では妥当な設計であるが、FITS ファイルを読み込み次第可視化するには適さない。また FITS データ取込み用にライブラリが作成するバッファのデータ長は、オリジナルの FITS ファイルのデータ長と同等となり、さらには、読み取ったバッファの中には、可視化に必要な要素も含まれている。一時的とは言え、可視化処理中には、このような大きなバッファと、それに基づいて作成される可視化データと、二重にメモリを消費することとなる。限られたリソース上で動作するモバイルデバイスへの適用を考えた場合、従来型 FITS ライブラリには、メモリの利用効率と、FITS データ読み込み完了までの待ち時間という 2 つのデメリットがある。このように従来型 FITS ライブラリでは「プログレッシブダウンロード」は実現できず、新たにライブラリを開発する必要がある。またメモリの利用効率の観点では、小規模のバッファを用いて映像を逐次再生するストリーミングの手法にも見習うべきものがあると考えられる。

3. FITS 逐次取込み型データ処理

以上の点に鑑み本稿では、メモリの効率的な利用と、科学データ読み込み完了までの待ち時間短縮の観点から、FITS 逐次取込み型のデータ処理を提案する。提案手法の要点は、次の 2 つである。一つは、科学データを到着順に取込み、逐次可視化処理に供することができることである。もう一つは、ライブラリとしてのバッファ領域を必要最小限にし、メモリの利用効率を高めることである。これにより、モバイルデバイスでの低速ネットワーク環境における待ち時間短縮と、限られたメモリの有効利用が期待できる。

3.1 処理の流れ

データ処理の過程を図 2 に示す。図中、リングバッファ長 n はデータ長 N より十分小さな値としている。この処理過程においては、まず FITS データの冒頭からヘッダ情報を読み取り、可視化に必要なパラメータと、後続のエクステンション (binary table extension) に格納された観測データの構造 (レコード毎の値の並び及びそれらの型・サイズ) を抽出する。次に、観測データの構造に基づいて、各データを読み取るためのリングバッファ、及び、リングバッファから観測データ構造に即して値を解釈するためのメソッド群を構築する。続いて、エクステンション中の観測データを逐次読み取る。リングバッファの入力はエクステンションから読み取った各レコードのバイナリ値である。これらの値は、バイトオーダーの変換以外の解釈はせずにリングバッファに格納される。リングバッファに短期的に取込まれている間に、可視化に必要な要素のみを可視化に適した形式に変換して保持する。すなわち、可視化に不要な観測データは、リングバッファにバイナリデータとして転記されるのみであり、値を解釈するオーバーヘッドが低減される。リングバッファは元の観測データ長に比べて十分短く、バッファ中の情報は取込みの過程で新しい観測データに上書きされるため、従来型の FITS ライブラリと比べて、読み取り時に占有するメモリ量を控えることができる。

可視化データへの変換処理は、取込まれた情報がリングバッファ中に維持されている期間内に実行するため、可視化データは観測データ取込み開始早々から順次利用可能となり、取込み完了までの待ち時間の間にも、可視化データをディスプレイに表示することが可能となる。

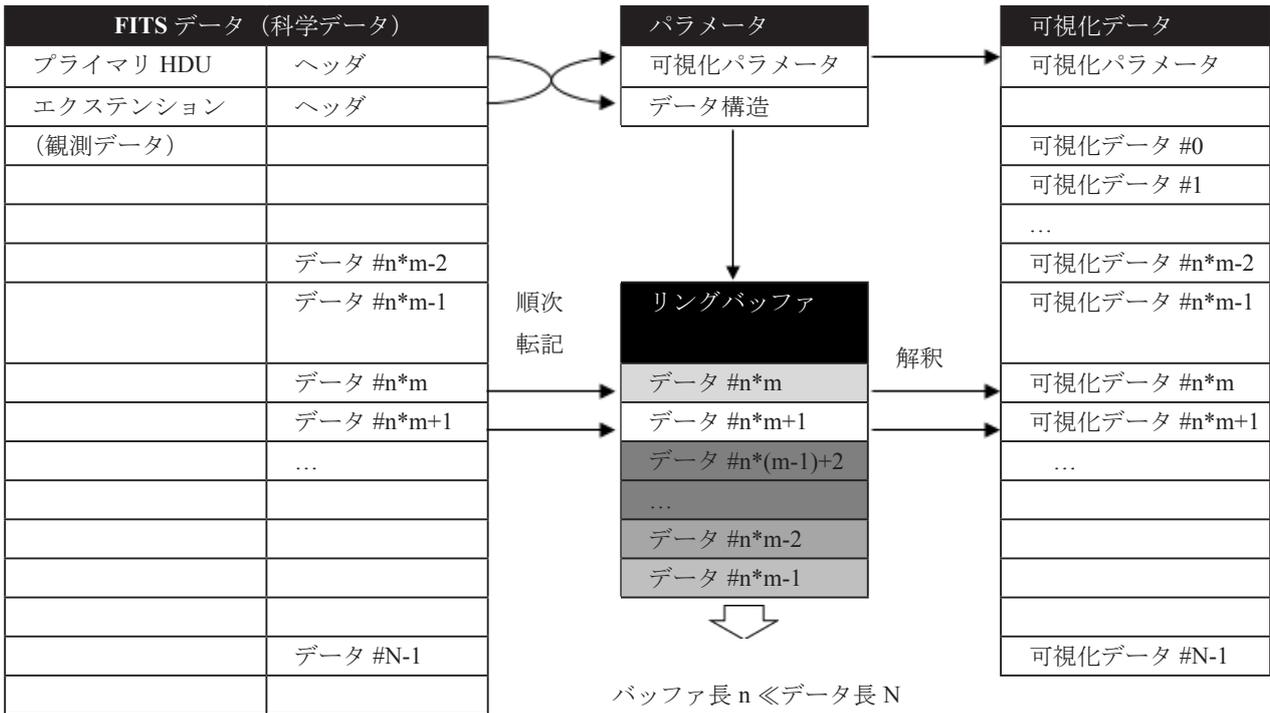


図 2 FITS 逐次取込み概要

3.2 対象となるデータ型

FITS データを逐次取込む都合上、FITS フォーマットで定義されているデータ型の内、一部は提案手法の適用対象外とする。本手法において、binary table extension 中の対応可能なデータ型を表 3 に示す。

表 3 FITS 逐次取込み型データ処理の適用対象

| データ型 | 適用対象 | 内部表現 |
|---------------------------------|------|-------------------|
| Logical | ○ | char |
| Bit | ○ | char |
| Unsigned byte | ○ | uint8_t, int8_t |
| 16-bit integer | ○ | int16_t, uint16_t |
| 32-bit integer | ○ | int32_t, uint32_t |
| 64-bit integer | ○ | int64_t, uint64_t |
| Character | ○ | char |
| Single precision floating point | ○ | float |
| Double precision floating point | ○ | double |
| Array Descriptor (32-bit) | × | 適用外 |
| Array Descriptor (64-bit) | × | 適用外 |

Binary table extension に定義されている型の内、Array Descriptor は後続のデータ領域中に格納された値を指し示すポインタであり、短いリングバッファでは取込んだ順に値を解釈することが困難である。長いバッファを使用することは、本稿で課題としている待ち時間の短縮やメモリ使用効率の向上とは相反するものであり、このような型を使用する FITS データは、提案手法の効果が薄いと判断される。幸いにして今回検討対象とした X 線天文衛星「あすか」や「すざく」の FITS データでは使用されていない型でもあり、Array Descriptor (32-bit)、Array Descriptor (64-bit) 共に、提案手法の適用対象外とした。

4. 模擬環境による比較

4.1 模擬環境の条件

本稿では、提案手法と従来手法との比較を以下の模擬環境下で行った。

- (a) 開発環境 : Mac Pro (Early 2008), Mac OS X, Xcode
iOS ベースのモバイルデバイスとの開発環境共有が可能であるため、Mac OS X と Xcode を基本としたプログラミングを行っている。Android は複数の CPU アーキテクチャが混在するため、互換性を確保するためには Java に基づいたアプリケーション開発が必要となるが、CPU 毎のネイティブコードで開発できる環境も整いつつあり、Xcode で開発したソースの共用も可能になると考えられる。
- (b) 通信環境模擬 : USB 1.1 カードリーダー + SD カード
低速のデータ通信を、USB 1.1 経由のファイル読み込みで模擬した。USB 1.1 の通信速度は 12Mbps であるが、模擬環境下でのファイル読み込みの実測値は約 3Mbps 程度であった。この速度は、筆者の執務室におけるモバイルデータ通信サービス (UQ WiMAX) の利用環境と比較すると、実際の通信速度に近いものであるが、さまざまなモバイルデータ通信サービス全般を近似できるものではない。
- (c) 従来手法で用いたライブラリ : sfitsio⁽⁴⁾
C++ でコーディングされた FITS ライブラリである。C++ の機能を用いて分かりやすいインタフェースが提供されている。C++ の高度な知識は要求せず、伝統的な CFITSIO と比べて利便性の高いライブラリとなっている。
- (d) 可視化ライブラリ : Open GL と C++ を用いた独自開発のライブラリ
Android や iOS との共用ができるよう、Open GL ES (Android や iOS で採用されているバージョン) との互換性を考慮している。
- (e) 取込み対象の FITS データ : 「あすか」が観測した、かにパルサーのデータ
DARTS で提供されている FITS データの中から、「あすか」衛星搭載の GIS 観測装置が 1995 年 9 月 15 日に取得した ad10405000g300170h.evt (10.6MB) を用いた¹⁰。
- (f) FITS データ取込みと可視化手順 : プログラム起動直後に軌道要素と FITS データの取込みを開始する。取込んだ FITS データは、従来手法・提案手法それぞれにおいて、ライブラリの FITS データ取得関数が戻り次第可視化を開始し、全データ取得が完了するまで、観測データの可視化データへの変換と、変換結果の描画を繰り返すものとする。従来手法においては、FITS データ取得開始から完了までが 1 回の処理となるため、実際に繰り返しの処理が発生するのは、提案手法のみとなる。
- (g) 可視化する項目 : 対象天体 (かにパルサー) 観測時の「あすか」の軌道
観測期間中に刻々と変わる「あすか」の軌道上の位置をプロットする。位置計算に必要となる軌道要素は、上記 FITS データと共に DARTS で提供されている
<http://darts.jaxa.jp/pub/asca/data/10405000/aux/frf.orbit.204.gz>
を用いた。なお FITS データ取込みと並行して可視化できる項目は、これに限定されるものではなく、一般に時系列に可視化できるものは、補助的なデータ (キャリブレーション等) が必要となる場合があるものの、原則として FITS データ取込みと並行して可視化可能である。
閏秒については、現時点ではプログラムのソースに補正項目を埋め込んでいる。

4.2 模擬環境における実行結果

実行画面を図 3 に示す。上段が従来手法の実行結果、下段が提案手法の実行結果である。それぞれ左から右へ時系列に示す。各シーン下の数字は、プログラム起動時からの経過時間 (分:秒;フレーム) である。画面が空白となっているシーンは、可視化に要する最初の観測データが取込みルーチンから戻ってくるのを待っている時間帯である。6 つの球体が描かれているシーンは、取込んだ観測データに基づいて、その時刻の衛星の位置を描画したものである。球体が地球を表し、6 方向から見た衛星の位置を白丸でプロットしている。

¹⁰ <ftp://ftp.darts.isas.jaxa.jp/pub/asca/data/10405000/screened/ad10405000g300170h.evt.gz>

従来手法



提案手法

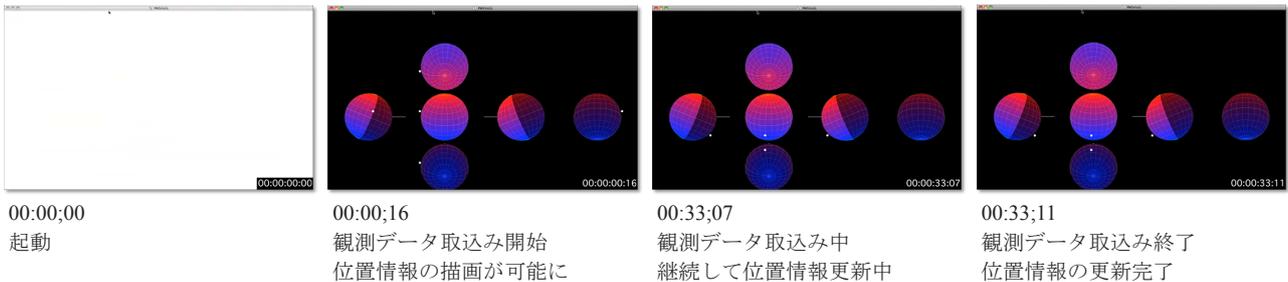


図3 従来手法と提案手法の実行結果

プログラム起動からデータ取込み終了までの時間は、従来手法・提案手法共に33秒程度となり、殆ど差がなかった。従来手法ではこの33秒余りの間は取得途中の観測データを参照できず、衛星の位置情報も描画できない。提案手法においては、0秒の16フレーム目（約0.5秒目）から衛星位置の描画が可能となっており、この時点から取込みが完了する33秒の11フレーム目までの間は、取得途中の観測データから抽出した観測時刻情報に基づいた衛星の位置をアニメーションで描画できている。これが実際の科学データ可視化アプリケーションであれば、利用者にとっては、30秒余りの観測データ取得時間は待ち時間と認識されなくなると期待される。

5. おわりに

モバイルデバイスを前提としたデータ通信サービスにおいても利用者を待たせることなく科学データを可視化する手段として、FITSデータの逐次取込みについて提案した。この手法の利点は、低速の通信環境においても利用者の待ち時間を低減できることや、デバイス側のメモリ消費量を低減できること等が挙げられる。また高速ネットワークを利用している場合であっても、ネットワークの輻輳などに起因する通信速度の低下に対しては同様の効果が期待できる。今回比較に用いた科学データは10MB余りと、昨今のモバイルデバイスの記憶容量と比較すると十分小さいとみなすこともできるが、観測によっては100MBを超えるデータもあり、通信環境やデバイスの処理能力等が限られる中では、従来手法と比べて提案手法を用いた可視化は適用できる科学データの範囲が広がるものと期待される。

一方で従来のFITSライブラリが、一旦取込んだデータは、明示的に消さない限り、継続してアクセスできるのに対して、提案する手法の場合、デバイス上にバッファされているデータは限られており、開発者はバッファされているデータの範囲を常に意識しながらプログラミングしなくてはならない。また提案手法は科学データのフォーマットに即した取込み処理を新規に作り上げることとなり、旧来から実績のあるライブラリに比べて、データの再現性等の信頼性を確立するには、尚一層の検証が必要であると考えられる。ただし信頼性の検証については、従来型のFITSライブラリを用いた場合も、ライブラリ開発者が想定していなかったモバイルデバイスに移植する際には、適切なコンパイルがなされているか逐一検証する必要がある。動作検証の労力は提案手法のみに発生するものではない。

今後は提案手法の動作検証を重ねると共に、同手法を用いたモバイルデバイス上での可視化について応用例を積み重ねて行く必要がある。また教育・アウトリーチの他に、研究者向けの応用も考えられる。詳細な解析に用いるためには

デバイスの性能や容量等とのトレードオフが必要であるが、モバイルデバイスを用いた衛星運用状況のモニターや観測データの Quick Look 等、軽快なレスポンスを必要とする用途には適用可能であると思われる。

参考文献

1. 大規模宇宙科学データの特徴量抽出と映像・音声化 - 教育利用として -. 三浦昭, 海老澤研. 2011 年, 日本天文学会 2011 年春季年会予稿集.
2. 宇宙科学データの時系列可視化と携帯型デバイスへの応用検討. 三浦昭, 海老澤研. 2012 年, 日本天文学会 2012 年春季年会.
3. 宇宙科学データの可視化・可聴化: 教育・広報利用. 三浦昭, ほか. 2012 年, 宇宙科学情報解析論文誌, ページ: 13-22.
4. *SFITSIO - A Next-Generation FITS I/O Library for C/C++ Users*. **Yamauchi, C.** 2010, *Astronomical Data Analysis Software and Systems XIX*, Vol. 434, pp. 469-472.
5. *Definition of the Flexible Image Transport System (FITS), version 3.0*. **W. D. Pence, L. Chiappetti, C. G. Page, R. A. Shaw and E. Stobie.** A42, 2010, *Astronomy and Astrophysics*, Vol. 524.
6. *CFITSIO, v2.0: A New Full-Featured Data Interface*. **Pence, William.** 1999, *Astronomical Data Analysis Software and Systems VIII*, Vol. 172.

動的タイムワーピング距離を用いた X線天文データの類似検索

林史尊^{*1} 天笠俊之^{*2,*3} 北川博之^{*2} 海老沢研^{*3} 中平聡志^{*3}

Similarity Search of Astronomical X-ray Data using Dynamic Time Warping Distance

Fumitaka HAYASHI^{*1}, Toshiyuki AMAGASA^{*2,*3}, Hiroyuki KITAGAWA^{*2},
Ken EBISAWA^{*3} and Satoshi NAKAHIRA^{*3}

Abstract

Explosive growth of data volume in scientific domains has caused demand of machine processing of the massive scientific data. In this paper we propose several schemes of similarity search over astronomical X-ray data for X-ray outbursts. Specifically, we first detect outbursts from the original data, followed by smoothing for reducing noise and normalization. Having detected outburst patterns, we apply dynamic time warping (DTW), which is known to be robust against time scaling to evaluate similarities between two outburst patterns. We propose several variations based on DTW by taking features of the outburst patterns and requirements for the similarity search; we apply derivative DTW, which is a variant of DTW, and apply sliding windows to evaluate partial similarities. We evaluate feasibility of the proposed schemes by experiments using real X-ray astronomy data.

Keywords: Astronomical X-ray Data, Similarity Search, Dynamic Time Warping.

概要

科学分野で扱われるデータ量は爆発的に増加しており、膨大なデータに対する機械的処理への要求は極めて高い。論文では、天体物理学におけるX線天体のX線アウトバースト現象を対象に、その観測データの類似検索を行う手法を提案する。具体的には、観測データからアウトバースト部分の自動検出を行なう。得られたデータに対し、ノイズ除去を目的とした平滑化と正規化を施した上で、動的タイムワーピング(DTW; Dynamic Time Warping)法を適用する。DTW法は、長さが異なる時系列データに対しても適用可能であるだけでなく、時間軸方向のスケーリングに対しても頑健なマッチングを行うことが距離である。さらに、X線アウトバーストの持つ特徴や、類似検索に対する要求を考慮し、DTW法の改良手法であるDerivative DTW法や、DTW法に滑り窓を適用した手法など、いくつかの新たな手法を提案する。さらに、実データとの比較によってその有効性を評価する。

キーワード: X線天文データ, 類似検索, 動的タイムワーピング。

*1 筑波大学大学院システム情報工学研究科 (Graduate School of Systems and Information Engineering, University of Tsukuba)

*2 筑波大学システム情報系 (Faculty of Engineering, Information and Systems, University of Tsukuba)

*3 宇宙航空研究開発機構宇宙科学研究所 (Institute of Space and Astronautical Science, Japan Aerospace Exploration Agency)

1 はじめに

近年、科学分野で扱われる観測データやシミュレーションデータは膨大なものとなっている。そのため、膨大なデータに対して高速な検索や分析の手段を提供することは、科学分野の進展のために不可欠なものとなっている。天文分野においても、日々蓄積される観測データへの対応は重要な課題となっている。本研究では X 線天体の観測データを取り上げる。X 線天体とは強力な X 線を放出する天体であり、その例にはブラックホール、中性子星がある。これらの天体が放出する X 線の強度を観測すると、短期間に強度が大きく上昇する現象が観測される。この現象は **X 線アウトバースト** と呼ばれ、このとき、天体は重力エネルギーを X 線として大量に放出している。

X 線アウトバースト現象の物理過程は、完全には解明されていない。例えば、観測される X 線強度の変化の様子は、天体や観測されるアウトバースト毎に異なるが、まれに類似した波形を示すことが明らかにされている¹⁾。これは、アウトバースト現象の物理過程に何らかの共通性があるという可能性を示唆しており、興味深い。このため、波形の類似するアウトバーストを検出することは、アウトバースト現象の起源を解き明かす上で重要である。例えば、図 1 は、全天 X 線観測装置 MAXI²⁾ によって観測された異なる天体のアウトバーストである。X 線強度に関する正規化を行うと、X 線強度がなだらかに上昇した後、急激に落ちるところなど、共通点があることが分かる。

X 線天体の観測を行っているセンサーは MAXI の他にもあり、数多くの観測データが蓄積されつつある。このため観測されるデータは膨大であり、類似する波形を人手で発見するのは極めて困難である。

一方、時系列データに対する検索やマイニングには

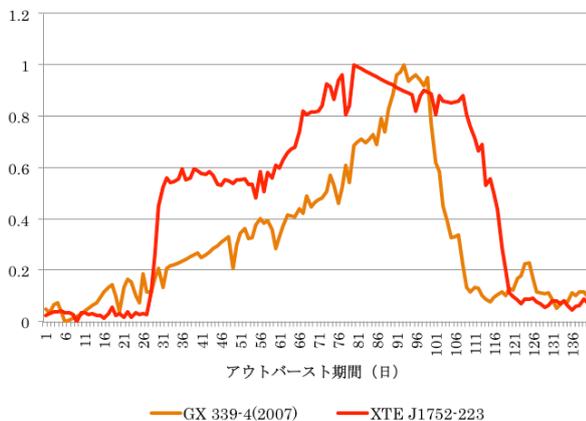


図 1 波形の類似した X 線アウトバースト。

多数の研究が存在する。特に、類似検索については、フーリエ変換³⁾、ウェーブレット変換⁴⁾、主成分分析⁵⁾など多数の方法が提案されており、分子生物学、経済学、音声認識、移動体分析などに広く応用されている。時系列データの類似度を検討する際には、対象となるデータの特徴や検索に対する要求などを考慮し、適切な手法を選択することが重要である。

本研究では、X 線アウトバーストの時系列観測データを対象に、類似検索を行なう手法を提案する。具体的には、オリジナルの観測データに対し、アウトバースト部分の自動検出を行なう。次に、抽出された時系列データに対して、ノイズ除去を目的とした平滑化と正規化を施す。得られた時系列データに対して、時系列データに対する距離関数を適用し、両者の類似度を評価する。本研究では、動的タイムワーピング (Dynamic Time Warping) 法 (以下、DTW 法)⁶⁾ を適用する。これは、動的計画法に基づく時系列データ上の距離関数であり、上で述べた関連手法に比べて、長さの異なる時系列データに対しても適用可能である、時間軸方向のスケーリングに対して頑健もマッチングを行うことが可能である等の特徴を持つ。X 線アウトバーストの時系列観測データは、長さが一定でなく、時間軸方向のスケールも一定していないという特徴があるため、DTW 法を用いることとした。さらに、DTW 法の改良手法である DerivativeDTW (DDTW) 法⁷⁾ の適用も検討する。

一方、アウトバーストの波形の類似検索には、例えば立ち上がり (下がり) の勾配が類似しているなど、波形のうち部分的な類似性に着目した検索に対する要求が存在すると考えられる。これに対応するため、DTW 法 (DDTW 法) に滑り窓を適用することで、時系列データのうち部分的な類似性を考慮した検索を可能にすることを提案する。また、時系列データを X 線強度が最大になる点の前後で二分割し、それぞれについて DTW 法 (DDTW 法) で類似度を評価する手法についても検討する。最後に、以上の手法を実データに適用し、その有効性を評価する。

本論文の構成は以下の通りである。2 節では、データの前処理とアウトバースト部分の検出方法を説明する。3 節で提案手法の詳細を述べ、4 節で提案手法を実験により評価する。5 節はまとめである。

2 アウトバーストの自動検出と前処理

X 線観測データにおいて、アウトバーストが検出されている部分は全体のうちごく一部である。また、一つの天体から複数のアウトバーストが検出されるため、アウ

トバーストが検出されている部分だけを自動的に抽出する処理が必要である。また、生データにはノイズが乗っており、また X 線強度も天体によって異なるため、ノイズ除去や正規化など適切な前処理を施す必要がある。以下では、アウトバースト部分の自動検出と前処理について説明する。

2.1 アウトバーストの自動検出

ある観測データ $x = \langle x_0, x_1, \dots, x_{n-1} \rangle$ におけるアウトバーストの自動検出は、次の手順で行う。

1. **欠損値の補完** さまざまな理由により、観測値が欠損している場合がある。その場合は、前後の観測値に基づく線形補完を行い、欠損値を補う。
2. **基準 X 線強度の設定** 観測値である X 線強度の最大値 $x_{max} = \max(x_0, x_1, \dots, x_{n-1})$ と最小値 $x_{min} = \min(x_0, x_1, \dots, x_{n-1})$ を求める。次に、 x_{min} と x_{max} の間を k 等分（本論文では $k = 10$ ）し、各区間における観測値出現頻度を計算する。最も出現頻度の大きい区間について、その区間に該当する観測値の平均を**基準 X 線強度** x_{base} とする。
3. **m 点平均の計算** m 点毎の観測値の平均 $y_j = \frac{\sum_{i=mj}^{m(j+1)-1} x_i}{m}$ ($j = 0, m, 2m, \dots$) を算出する。本研究では $m = 20$ とした。
4. **アウトバーストの検出** m 点平均が二つ以上連続して $x_{base} + \frac{x_{max} - x_{base}}{k} \leq y_j$ となる区間をアウトバーストとして検出し、 y_j に対応する観測データ x_i を抽出する。なお、抽出する観測値には、上記区間に加えて、前後の m 点（合計 $2m$ 点）も含めることとする。すなわち、 $y_j, y_{j+1}, \dots, y_{j+l}$ が連続して上記の条件を満たしていたとすると、 y_{j-1} と y_{j+l+1} に対応する観測値も抽出する。

2.2 平滑化

観測値にはノイズが含まれているため、平滑化によってその影響を低減する。第 2.1 節では、観測値の強度の推移からアウトバーストを検出することを目的としてい

たため単純な m 点平均を計算していたが、ここでは株価のチャート分析などにも利用される n 点線形加重移動平均を用いる。 **n 点線形加重移動平均**は、以下のように計算される。データ列 $q = \langle q_0, q_1, \dots, q_{n-1} \rangle$ が与えられたとき、平滑カゴのデータ列 $Sq = \langle Sq_0, Sq_2, \dots, Sq_{n-1} \rangle$ は図 2 に示した式に従って求められる。

2.3 X 線強度の正規化

観測される X 線の強度は天体毎に大きく異なるため、最小値が 0、最大値が 1 となるよう、 $[0;1]$ 区間に正規化する。図 3 は、GX339-4, XTEJ1752-223, 4U1608-52 の三つの天体のデータから提案手法で抽出したアウトバーストのデータに平滑化を施したデータをプロットしたものである。なお、GX339-4 については、2007 年のデータと 2010 年のデータがあり、それぞれ GX339-4(2007) と GX339-4(2010) としている。正規化前（上）では、4U1608-52 の X 線強度が他に比べて強く類似性が判別しづらいが、正規化によって、パターンの特徴による類似性ははっきりすることが分かる（下）

3 アウトバーストの類似検索

本節では、アウトバーストの類似検索手法を述べる。前処理を施した二つのアウトバーストの時系列データに対して類似検索を行なう場合、一般的には両者の（非）類似度を距離によって評価する。距離が小さいなら、両者はより類似していることになる。このとき、最も適切な距離尺度を選択することが重要である。例えば、最も馴染み深いユークリッド距離は、比較対象となる時系列の次元数（要素数）が等しくなければ適用することができない。このため、長さの異なる時系列データの間の距離尺度が多数提案されている。本研究では、その中でも広く用いられているものの一つである動的タイムワーピング（DynamicTimeWarping）法（以下、DTW 法）⁶⁾、およびその改良手法である DerivativeDTW（DDDTW）法⁷⁾を採用する。

$$Sq_i = \begin{cases} q_0 & (i = 0) \\ \frac{1}{weightsum} \left\{ \frac{n+1}{2} q_i + \sum_{k=1}^N \left(\frac{n+1}{2} - k \right) (q_{i-k} + q_{i+k}) \right\} & (0 < i < m) \\ q_m & (i = m) \end{cases}$$

$$N = \begin{cases} i & (i < (n+1)/2 - 1) \\ n - i & (i > m + 1 - (n+1)/2) \\ \frac{n+1}{2} - 1 & \text{otherwise} \end{cases}$$

$weightsum$: 重みの合計。

図 2 n 点線形加重移動平均。

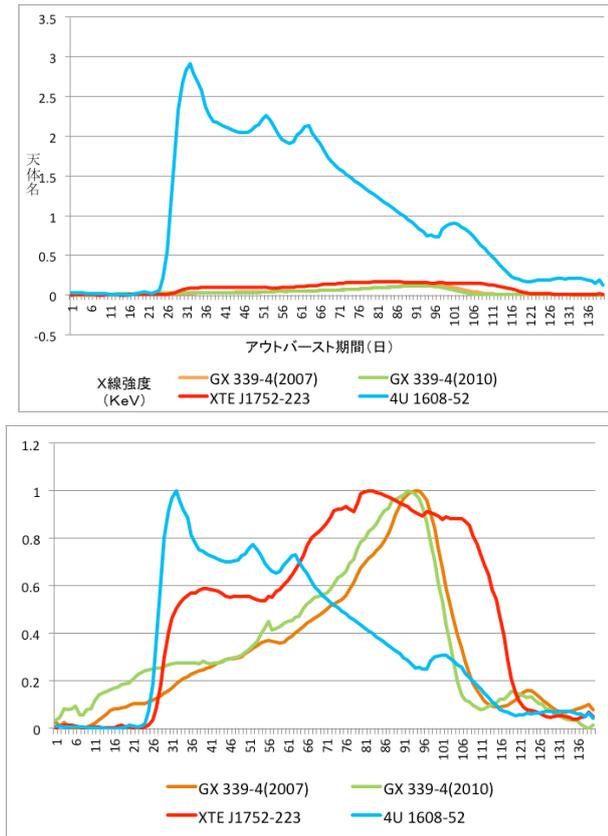


図3 正規化前(上)／正規化後(下)。

以下では、まずその概要を説明する。続いて、時系列データの部分的な類似性に着目した類似検索を可能にするため、滑り窓を用いる方法と、時系列データをX線強度が最大となる点の前後で二分割する方法を説明する。

3.1 基本事項

動的タイムワーピング (DTW) 法

動的タイムワーピング (Dynamic Time Warping; DTP) 法⁶⁾とは、二つの時系列データの最適なマッチングを動的計画法⁸⁾によって求め、そのマッチングに基づいて距離を計算する手法である。その特徴として、長さの異なるデータの比較にも利用でき、また時間軸方向のスケールリングに対しても頑健という性質を持っている。このため、音声認識など幅広い分野で用いられている^{9,10)}。

$$D(x, y) = \gamma(m, n)$$

$$\gamma(i, j) = \begin{cases} d(x_0, y_0) & (i = j = 0) \\ d(x_0, y_j) + \gamma(0, j - 1) & (i = 0, j > 0) \\ d(x_i, y_0) + \gamma(i - 1, 0) & (i > 0, j = 0) \\ d(x_i, y_j) + \min\{\gamma(i - 1, j - 1), \gamma(i - 1, j), \gamma(i, j - 1)\} & (i > 0, j > 0) \end{cases}$$

図4 DTW 距離。

二つの時系列データ：

$$x = \langle x_0, x_1, \dots, x_i, \dots, x_{n-1} \rangle$$

$$y = \langle y_0, y_1, \dots, y_j, \dots, y_{m-1} \rangle$$

に対して、DTW 距離 $D(x, y)$ は図4に示した漸化式によって求められる。ここで、 $d(x_i, y_j)$ は x_i と y_j の距離であり、差の絶対値に対して単調増加性があれば任意の関数を利用して良い¹¹⁾。本研究では、 $d(x_i, y_j) = |x_i - y_j|^2$ とした。また、DTW 距離の計算は動的計画法によって行なうことができる。

図5に、図3(下)で示した四つのアウトバースト (GX339-4(2007), GX339-4(2010), XTEJ1752-223, 4U1608-52) にDTW法を適用して距離を計算した結果を示す。元の波形を見ると、GX339-4(2007)とGX339-4(2010)が最も類似している。次に、この二つとXTEJ1752-223が、最大X線強度までならかに上昇し、その後急激に落ちているという点で類似性が見られる。DTW法による距離もそれを反映していることが分かる。

DerivativeDTW 法

既に述べたように、DTW法は時系列データの距離を計算する際、極めて有用であるが、いくつか欠点も指摘されており、それに対応するための改良手法が提案されている^{7,9,12-14)}。その中でも最も大きな問題の一つが、スパイクのような急激な変化があったときに、その点の周辺で二つの時系列間の点同士の対応関係がいびつになってしまう、正しく距離が評価されないという問題である。この問題に対応するため、KeoghらはDerivativeDTW法を提案した⁷⁾。DerivativeDTW法の基本的なアイデアは、各点の距離ではなく、変化量を比較しようとするものである。このため、時系列データに対して以下の前処理を行なう。変換対象のデータを $q = \langle q_0, q_1, \dots, q_{n-1} \rangle$ とすると、以下の式によって得られる Dq_i が変換後のデータとなる。

$$Dq_i = \begin{cases} \frac{(q_i - q_{i-1}) + (q_{i+1} - q_i)/2}{2} & (0 < i < n) \\ Dq_1 & (i = 0) \\ Dq_{n-1} & (i = n) \end{cases}$$



図 5 DTW 法による距離計算の例.

注目する点自身を含む周囲 3 点における平均変化量が、新たな時系列データとして生成される。例として、図 3 (下) に示した GX339-4(2007) に対して上記の変換を施した例を図 6 に示す。変換されたデータ同士に対して、通常の DTW 法を適用することで、DDTW 法による距離計算ができる。図 7 に、図 3 (下) で示した四つのアウトバースト (GX339-4(2007), GX339-4(2010), XTEJ1752-223, 4U1608-52) に DTW 法を適用して距離を計算した結果を示す。DTW 法の結果に比べて、4U1608-52 との距離が相対的に上がっていることがわかる。これは、通常の DTW 法では、4U1608-52 の最大 X 線強度の前後において、他の点とのいびつな対応付けが見られるのに対し、DDTW 法ではそれが解消されるためである。

3.2 アウトバーストの部分的な類似性に着目した類似検索

DTW 法あるいは DDTW 法をそのまま用いることによって、アウトバーストの波形全体による類似度の計算

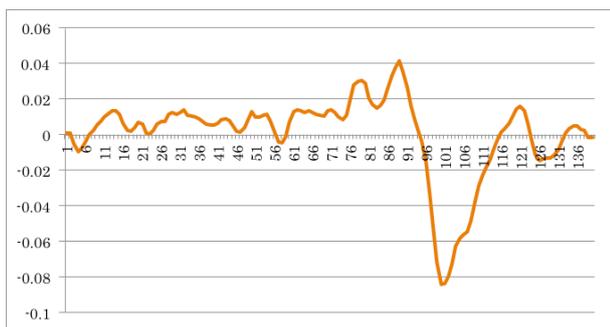


図 6 DDTW 法によるデータ変換の例.

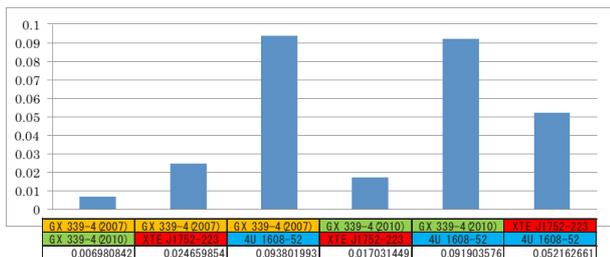


図 7 DDTW 法による距離計算の例.

が可能である。しかしながら、時系列データ全体ではなく、部分的な類似度を考慮した検索を行ないたい場合も考えられる。このため以下では、滑り窓 (slidingwindow) の適用と、時系列データを X 線強度が最大になる点で前後に二分割して、前半 (後半) の類似度を評価する方法を提案する。これにより、波形全体のうち 50% 以上が類似しているアウトバースト、X 線強度の上昇 (下降) のしかたが似ているアウトバーストといった検索が可能となる。

滑り窓の利用

滑り窓 (slidingwindow) とは、窓 (window) と呼ばれる固定幅の部分系列に関して距離計算を行ない、その後、窓をずらし幅 s だけずらして、同様の計算を繰り返す行なう計算手法のことである。これに DTW (DDTW) 法を適用する。その結果、部分系列 (窓) 間の距離が、窓をずらした回数だけ得られる。これは、部分系列間の距離が波形の先頭から末尾に向かってどのように推移するかを示している。波形全体の距離は、得られた窓内の距離の総和で評価できる。また、ある閾値以下の距離に対応する窓の位置から、波形が部分的に類似している区間を見つけることも可能である。通常の滑り窓法は、窓幅 w とずらし幅 s は比較対象の二つの時系列データに対して共通である。しかし、時系列データ全体の長さが極端に異なる場合、考慮が必要である。本研究では、 w と s を、比較対象の波形の長さに対する相対値で与えることでこれに対応した。

最大点の前後による二分割

アウトバーストの X 線強度の変化は、基本的に最大強度まで上昇し、その後減少することから、最大値の前後によって二分割することが可能である。このとき、最大強度まで (以降) の X 線強度の上昇 (下降) のしかたに着目した類似度の評価を行ないたい場合などが考えられる。これには、波形を最大値の前後で二分割し、それぞれについて DTW (DDTW) 法を適用すれば良い。

4 評価実験

提案手法の有効性を検証するため、実データによる実験を行なった。

4.1 実験環境およびデータセット

実験は、CPU に 4 コア Intel(R) Xeon(R) CPU E5310 (1.60GHz)、メモリ容量 5GB の PC を使用した。言語は、DTW 法以外の実装と計算には JVMversion1.6.030 上で

動作する Scalaver.2.9.1 を利用した。DTW 法の計算には、FastDTW^{15,16)} に含まれる DtwTest.java を使用した。データセットとしては、人工衛星 Swift に搭載された硬 X 線モニター BAT によって観測された公開データを使用した¹⁷⁾。BAT の観測対象となっている X 線天体数は 951 あるが、その中で X 線強度が強い 155 の天体を対象とした。これらのデータに対して、2 節で説明したアウトバーストの検出を行ない、平滑化および正規化を行なった。なお、一つの天体から複数のアウトバーストが検出された場合は、その中から期間の長い二つを選択した。結果として 153 件のアウトバーストを検出した。図 8 に、各天体について何件のアウトバーストが検出されたかをまとめた図を示す。抽出されたアウトバースト全ての組合せ 11,268 通り (=153*152/2) について、以下に説明する手法で距離を計算し、その上位について結果を評価した。

4.2 実験結果

DTW 法, DDTW 法による類似検索

DTW 法および DDTW 法によって、類似したアウトバーストが検索できるかどうかを評価した。距離としては DTW 法, DDTW 法を用い、上位のデータを比較した。表 1, 表 2 に、それぞれ DTW 法, DDTW 法による類似検索結果の上位 20 件を示す。表において、1, 3 列目はアウトバーストが観測された天体名、2, 4 列目は当該天体の中で観測された何番目のアウトバーストであるか、5 列目は類似度 (DTW 距離) を表している。なお、ここでの類似度は、値が小さければ小さいほど二つの時系列データが類似していること示している。

また、上位 6 件の実際の時系列データのプロットを図 9 に示す。なお、左上から右の順に、1 位、2 位の順に並べてある。これから分かるように、どちらの手法とも、類似した時系列データが検索できていることが分かる、DTW 法の 2 位には、類似していない時系列データ

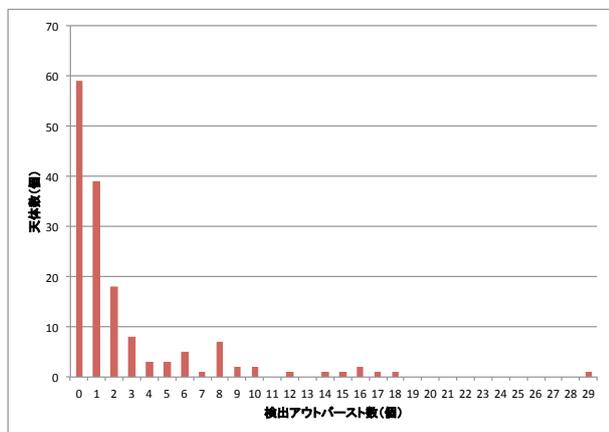


図 8 天体毎のアウトバースト検出数の分布。検出バースト数が突出して多い (29 個) 天体は SMCX-1 である。

表 1 DTW 法による類似検索 (類似度上位 20 件)

| 天体1 | OB No. | 天体2 | OB No. | 類似度 |
|-----------------|--------|------------------|--------|-------------|
| 4U0115p634 | 1 | V0332p53 | 1 | 0.018236731 |
| NGC5506 | 1 | SWFTJ1756.9-2508 | 1 | 0.018513513 |
| 4U0115p634 | 2 | MXB0656-072 | 3 | 0.018897482 |
| SAXJ1747.0-2853 | 1 | SAXJ1808.4-3658 | 1 | 0.022549738 |
| 4U0115p634 | 2 | V0332p53 | 1 | 0.02857227 |
| 1A0535p262 | 3 | GRJ1655-40 | 1 | 0.029279561 |
| AXJ1749.1-2639 | 1 | H1417-624 | 1 | 0.030407902 |
| 1A0535p262 | 2 | V0332p53 | 1 | 0.030461565 |
| 1A0535p262 | 2 | 4U0115p634 | 1 | 0.03081885 |
| 1A0535p262 | 3 | 4U0115p634 | 1 | 0.032547949 |
| EXO 2030p375 | 1 | V0332p53 | 1 | 0.032651616 |
| 4U0115p634 | 1 | EXO 2030p375 | 1 | 0.033370397 |
| EXO 2030p375 | 1 | H1417-624 | 1 | 0.035118511 |
| 1A0535p262 | 3 | V0332p53 | 1 | 0.035500701 |
| 4U0115p634 | 1 | GX304-1 | 1 | 0.035695783 |
| 1A0535p262 | 3 | GX304-1 | 1 | 0.035700165 |
| GX304-1 | 1 | V0332p53 | 1 | 0.036771114 |
| 1A0535p262 | 2 | 4U0115p634 | 2 | 0.0392177 |
| GX339-4 | 2 | H1417-624 | 1 | 0.041358905 |
| GheA1843n00 | 1 | MXB0656-072 | 3 | 0.045164234 |

表 2 DDTW 法による類似検索 (類似度上位 20 件)

| 天体1 | OB No. | 天体2 | OB No. | 類似度 |
|----------------|--------|-----------------|--------|-------------|
| 1A0535p262 | 3 | 4U0115p634 | 2 | 0.002832915 |
| 4U0115p634 | 2 | MXB0656-072 | 3 | 0.003044645 |
| 1A0535p262 | 3 | MXB0656-072 | 3 | 0.003189259 |
| 1A0535p262 | 3 | GRJ17586-2129 | 2 | 0.00457833 |
| 4U0115p634 | 1 | V0332p53 | 1 | 0.004600994 |
| 4U0115p634 | 2 | GRJ17586-2129 | 2 | 0.005378599 |
| 1A0535p262 | 3 | SAXJ1747.0-2853 | 1 | 0.00576885 |
| 1A0535p262 | 2 | 4U0115p634 | 2 | 0.006898757 |
| 4U0115p634 | 2 | Mrk509 | 1 | 0.007081451 |
| 4U0115p634 | 2 | SAXJ1747.0-2853 | 1 | 0.007212816 |
| GRJ17586-2129 | 2 | MXB0656-072 | 3 | 0.007630367 |
| AXJ1749.1-2639 | 1 | EXO 2030p375 | 1 | 0.007875901 |
| 1E1743.1-2843 | 2 | GRJ17586-2129 | 2 | 0.008599417 |
| GRJ1655-40 | 1 | MXB0656-072 | 3 | 0.009489001 |
| 1A0535p262 | 2 | MXB0656-072 | 3 | 0.009878791 |
| GX304-1 | 1 | MXB0656-072 | 3 | 0.010134271 |
| GX339-4 | 2 | XTEJ1752-223 | 1 | 0.010366353 |
| 4U0115p634 | 2 | GRJ17497-2821 | 1 | 0.012344659 |
| 4U0115p634 | 1 | MXB0656-072 | 3 | 0.012853684 |
| 1A0535p262 | 3 | GX304-1 | 1 | 0.01286813 |

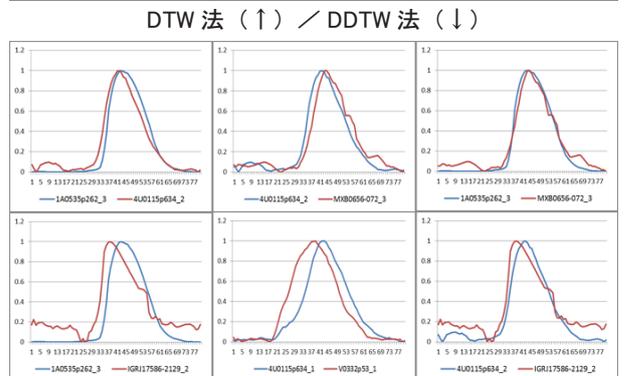
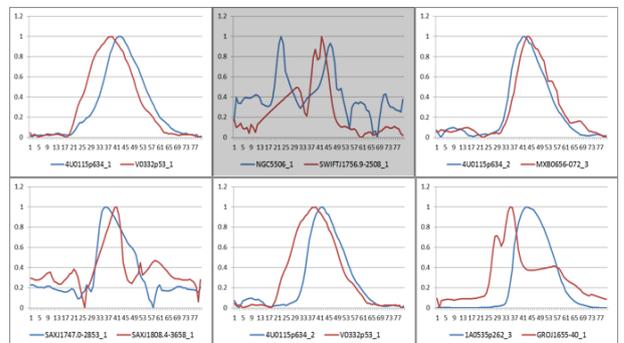


図 9 DTW 法, DDTW 法による類似検索 (左上から右に 1 位から 6 位まで)。

がランキングされている。これは、ある天体に対する観測値が時間的に一定ではなかったり、天体が暗くなることなどによる統計誤差である。このため DTW 法における特異点の不自然な対応付けによって、距離が低く評価されてしまったものと思われる。精度向上のためには、このようなノイズの乗ったデータを事前にフィルタリングする、観測値に付随している観測値に関するメタデータを利用したランキングの補正を行うことなどが考えられる。一方、DDTW 法では、DTW 法のような問題に対応しているため、当該データの順位は低く押さえられている。

滑り窓による類似検索

滑り窓法による類似検索を行ない、結果を評価した。具体的には、それぞれ DTW 法、DDTW 法による距離の閾値を 0.5, 0.025 とし、これより小さい距離を示した区間が連続して三つ以上出現したものについて、その平均距離の順に上位 20 件を抽出した (表 3, 表 4)。ここで、表における類似部平均は、DTW 距離が閾値を下回った部分区間の DTW 距離の平均である。図 10 は上位 6 件のプロットである。DTW 法の結果については、時系列データ全体の類似度では上位にランクされなかった、部

表 3 滑り窓による類似検索 (DTW 法, 類似度上位 20 件)。

| 天体1 | OB No. | 天体2 | OB No. | 類似部平均 |
|----------------|--------|-------------------|--------|-------------|
| 1A0535p262 | 3 | XTEJ1752-223 | 1 | 0.009636776 |
| EXO 2030p375 | 1 | IGRJ17091-3624 | 1 | 0.029473847 |
| 1A0535p262 | 2 | XTEJ1752-223 | 1 | 0.032223213 |
| 1A0535p262 | 2 | MXB0656-072 | 3 | 0.032564217 |
| 1A0535p262 | 3 | MXB0656-072 | 3 | 0.035630179 |
| 1A0535p262 | 3 | 4U0115p634 | 2 | 0.036466685 |
| 1A0535p262 | 2 | 4U0115p634 | 2 | 0.037292779 |
| 4U1702-429 | 3 | XTEJ1701-407 | 1 | 0.037755554 |
| 1E1743.1-2843 | 2 | 4U0115p634 | 2 | 0.03797757 |
| EXO 2030p375 | 1 | MXB0656-072 | 3 | 0.038530666 |
| 1E1743.1-2843 | 2 | MXB0656-072 | 3 | 0.039156731 |
| Ginga1843p00 | 2 | SW FTJ1539.2-6227 | 1 | 0.04002407 |
| 4U0115p634 | 2 | AXJ1744.8-2921 | 1 | 0.040474071 |
| 4U0115p634 | 2 | MXB0656-072 | 3 | 0.044306802 |
| GR0 J1655-40 | 1 | SW FTJ1539.2-6227 | 1 | 0.04590276 |
| IGRJ17473-2721 | 1 | IGRJ17586-2129 | 2 | 0.04634014 |
| IGRJ17497-2821 | 1 | XTEJ1810-189 | 1 | 0.052331623 |
| IGRJ17473-2721 | 1 | XTEJ1810-189 | 1 | 0.052364995 |
| EXO 2030p375 | 2 | V0332p53 | 1 | 0.05412503 |
| H1417-624 | 1 | SW FTJ1539.2-6227 | 1 | 0.05486193 |

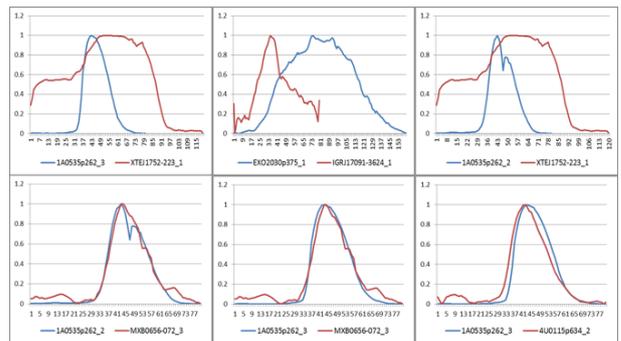
表 4 滑り窓による類似検索 (DDTW 法, 類似順上位 20 件)。

| 天体1 | OB No. | 天体2 | OB No. | 類似部平均 |
|----------------|--------|----------------------|--------|-------------|
| 1A0535p262 | 3 | MXB0656-072 | 3 | 0.001423682 |
| 4U0115p634 | 2 | MXB0656-072 | 3 | 0.001616227 |
| 1A0535p262 | 3 | 4U0115p634 | 2 | 0.001620353 |
| 1E1740.7-2942 | 4 | GX3p1 | 4 | 0.002083352 |
| 1A1118-61 | 1 | SW FTJ1357.2-0933 | 1 | 0.002356777 |
| 4U1630-472 | 1 | IGRJ17497-2821 | 1 | 0.002480686 |
| 1A0535p262 | 2 | 4U0115p634 | 2 | 0.002637977 |
| GR0 J1655-40 | 1 | XMMUJ054134.7-682550 | 1 | 0.003065809 |
| 1A1118-61 | 1 | GR0 J1655-40 | 1 | 0.003101865 |
| 1E1740.7-2942 | 3 | IGRJ17473-2721 | 1 | 0.003120295 |
| 4U0115p634 | 2 | Mrk509 | 1 | 0.003164238 |
| 4U1630-472 | 1 | XMMUJ054134.7-682550 | 1 | 0.003225372 |
| AXJ1749.1-2639 | 1 | SW FTJ1753.5-0127 | 1 | 0.003267231 |
| GR0 J1655-40 | 1 | H1417-624 | 1 | 0.003394996 |
| 4U0115p634 | 2 | IGRJ17586-2129 | 2 | 0.003484066 |
| 1E1740.7-2942 | 4 | GX9p9 | 2 | 0.003490676 |
| GX3p1 | 4 | XTEJ1701-462 | 2 | 0.00352943 |
| 1E1740.7-2942 | 4 | GRS1724-308 | 1 | 0.003600989 |
| 1A0535p262 | 3 | 1E1743.1-2843 | 2 | 0.003705902 |
| 1A1118-61 | 1 | XMMUJ054134.7-682550 | 1 | 0.003845263 |

分的な類似区間を含むデータが上位にランクされていることが分かる。これにより、滑り窓による部分的な類似度による類似検索によって、単純な DTW 法 (DDTW 法) では検索できないデータを検索することが可能であることが示された。一方、DDTW 法の 4 位, 5 位では、一見すると類似していないデータが上位にランキングされていることが分かる。これは、ノイズ等何らかの理由により一定の強度が持続的に維持されているデータ同士に対する結果である。DDTW 法は信号強度の絶対値ではなくその変化率に着目しているため、一定の値を維持している区間について、滑り窓内の類似度が極めて高くなってしまふ。DDTW 法では一定の強度が維持されるようなデータを含む場合に、不適切な結果が出力される可能性があることが分かった。これに対しては、滑り窓を適用する区間や幅などを調整すること、一定のレベルが維持されるデータについてはアウトバーストではないと判断して、事前にフィルタリングするなどの対応が必要であると考えられる。

二分割による類似検索

二分割による類似検索を行ない、結果を評価した。具体的には、X 線強度が最大になる点で時系列データを二分割し、前半 (後半) のみの部分系列に対して、DTW 法、DDTW 法による距離計算を行なった。ここでは、後半部分のみの結果を示す。表 5, 表 6 は上位 20 件の結果、



DTW 法 (↑) / DDTW 法 (↓)

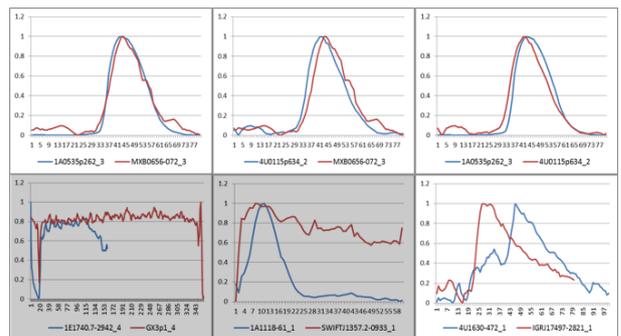


図 10 滑り窓による類似検索 (左上から右に 1 位から 6 位まで)。

表 5 二分割による類似検索 (DTW 法, 類似度上位 20 件)

| 天体1 | OB No. | 天体2 | OB No. | 類似度 |
|----------------|--------|-----------------|--------|-------------|
| 4U0115p634 | 1 | V0332p53 | 1 | 0.005059078 |
| 4U0115p634 | 2 | V0332p53 | 1 | 0.005499665 |
| 1A0535p262 | 3 | MXB0656-072 | 3 | 0.006590148 |
| 4U0115p634 | 2 | MXB0656-072 | 3 | 0.00678463 |
| 4U0115p634 | 1 | MXB0656-072 | 3 | 0.007124647 |
| MXB0656-072 | 3 | XTEJ1752-223 | 1 | 0.00728251 |
| 1A0535p262 | 2 | MXB0656-072 | 3 | 0.007335766 |
| 1A0535p262 | 3 | V0332p53 | 1 | 0.007446669 |
| MXB0656-072 | 3 | V0332p53 | 1 | 0.007546729 |
| AXJ1749.1-2639 | 1 | EXO2030p375 | 1 | 0.00769272 |
| 1A0535p262 | 3 | XTEJ1752-223 | 1 | 0.008206249 |
| 1A0535p262 | 3 | 1A1118-61 | 1 | 0.008858237 |
| Ginga1843p00 | 2 | SAXJ1750.8-2900 | 2 | 0.008911676 |
| 1A1118-61 | 1 | MXB0656-072 | 3 | 0.009128038 |
| 1A0535p262 | 3 | AqlX-1 | 3 | 0.009276455 |
| 1A1118-61 | 1 | 4U0115p634 | 2 | 0.009444722 |
| GX304-1 | 1 | MXB0656-072 | 3 | 0.009608333 |
| 1A0535p262 | 3 | GROJ1655-40 | 1 | 0.010601421 |
| 1A0535p262 | 3 | MAXIJ1409-619 | 1 | 0.011048155 |
| 1E1743.1-2843 | 2 | GX339-4 | 2 | 0.01155077 |

表 6 二分割による類似検索 (DDTW 法, 類似度上位 20 件)

| 天体1 | OB No. | 天体2 | OB No. | 類似度 |
|----------------|--------|----------------|--------|-------------|
| 1A0535p262 | 3 | 4U0115p634 | 2 | 0.000445267 |
| 1A0535p262 | 3 | IGRJ17586-2129 | 2 | 0.000762409 |
| V0332p53 | 1 | XTEJ1752-223 | 1 | 0.000885903 |
| 1A0535p262 | 3 | MXB0656-072 | 3 | 0.000950769 |
| 4U0115p634 | 2 | MXB0656-072 | 3 | 0.000954245 |
| 4U0115p634 | 2 | V0332p53 | 1 | 0.001163465 |
| 1A0535p262 | 3 | 1E1743.1-2843 | 2 | 0.001166002 |
| 1E1740.7-2942 | 4 | MXB0656-072 | 3 | 0.001173052 |
| 1A0535p262 | 3 | 4U0115p634 | 1 | 0.001268383 |
| 4U0115p634 | 2 | IGRJ17586-2129 | 2 | 0.001271939 |
| 1A0535p262 | 3 | V0332p53 | 1 | 0.001277466 |
| 1A0535p262 | 3 | XTEJ1752-223 | 1 | 0.001315496 |
| 1E1743.1-2843 | 2 | IGRJ17586-2129 | 2 | 0.001364595 |
| 4U0115p634 | 1 | V0332p53 | 1 | 0.001400574 |
| 4U0115p634 | 1 | XTEJ1752-223 | 1 | 0.001469094 |
| 1E1740.7-2942 | 4 | 4U0115p634 | 2 | 0.001555429 |
| 4U0115p634 | 2 | GX1p4 | 2 | 0.001619554 |
| 4U1702-429 | 4 | EXO2030p375 | 2 | 0.001665785 |
| IGRJ17586-2129 | 2 | XTEJ1946p274 | 3 | 0.001814627 |
| 4U0115p634 | 2 | XTEJ1946p274 | 3 | 0.001837680 |

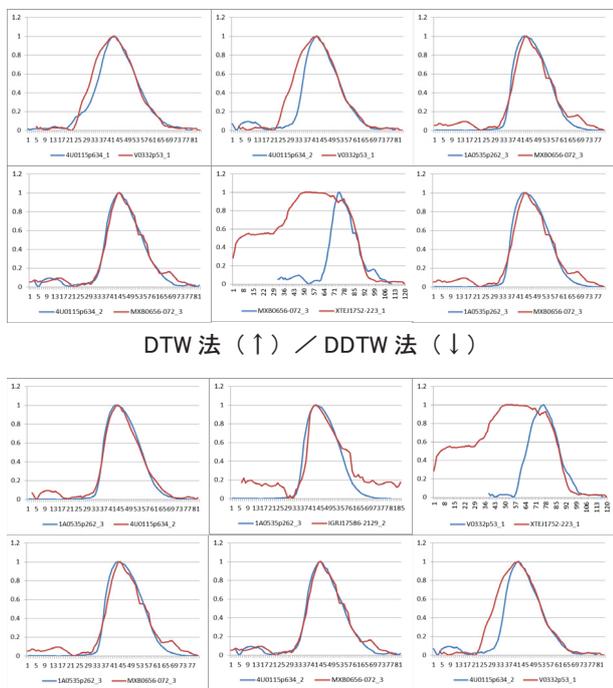


図 11 二分割による類似検索 (左上から右に 1 位から 6 位まで)

図 11 は上位 6 件のプロットである。図から分かるように、DTW 法, DDTW 法とも、最大値からの値の下がり方が類似しているデータを検索できていることが分かる。

5 まとめ

本研究では、X 線アウトバーストの時系列データの類似検索手法を提案した。時系列データの検索手法である DTW 法, DDTW 法をベースとして、時系列データの部分的な類似性を考慮に入れた検索を行なうために、滑り窓を用いた方法と時系列データを X 線強度が最大になる点で前後に二分割し、部分データに対して距離計算を行なう手法を提案した。また、実験による評価によって、提案手法の有効性を検証した。今後は、アウトバースト部分の検出精度を向上し、ノイズの乗ったデータに対する適切な処理を行うことによって、類似検索自体の精度向上を目指す。また、既存のデータに対し網羅的に類似検索を行ない、新たな知見の獲得に取り組む。さらに、提案した手法を Web サービス化して公開することも検討する予定である。

参考文献

- 1) Satoshi Nakahira and MAXI Team. Black hole novae observed with MAXI. *The Astronomical Herald*, Vol. 103, No. 3, pp. 166–175, 2012.
- 2) M.Matsuoka. et al. The MAXI mission on the ISS: Sciences and instruments for monitoring all sky Xray images. *PASJ*, Vol. 61, No. 999, 2009.
- 3) Davood Rafiei and Alberto Mendelzon. Similaritybased queries for time series data. In *Proc. 1997 ACM SIGMOD international conference on Management of data (SIGMOD'97)*, 1997.
- 4) I. Popivanov. Similarity search over time-series data using wavelets. In *Proc. 18th International Conference on Data Engineering (ICDE2002)*, pp. 212–221, 2002.
- 5) Kiyoung Yang and Cyrus Shahab. A PCA-based similarity measure for multivariate time series. In *Proc. 2nd ACM international workshop on Multimedia databases (MMDB' 04)*, pp. 65–74, 2004.
- 6) H. Sakoe and S. Chiba. A dynamic programming approach to continuous speech recognition. In *Proc. 7th International Congress on Acoustics*, pp. 65–69, 1971.
- 7) Eamonn J. Keogh and Michael J. Pazzani. Derivative

- dynamic time warping. In *Proc. 1st SIAM International Conference on Data Mining (SDM2001)*, 2001.
- 8) Jon Kleinberg and Éva Tardos. *Algorithm Design*. Addison Wesley, 2005.
 - 9) Donald J. Berndt and James Clifford. Using dynamic time warping to find patterns in time series. In *Proc. KDD Workshop*, pp. 359–370, 1994.
 - 10) E. G. Caiani, A. Porta, G. Turiel, M. Muzzupappa, S. Pieruzzi, F. Grema, C. Malliani, A. Cerutti, and S. Cerutti. Warped-average template technique to track on a cycle-by-cycle basis the cardiac filling phases on left ventricular volume. In *IEEE Computers in Cardiology*, Vol. 25, 1998.
 - 11) 石川雅弘, 吉川昂伯, 陳漢雄, 古瀬一隆, 大保信夫. 類似部分区間検索のためのタイムワーピング距離の下限值計算. *DBSJ Letters*, Vol. 6, No. 1, pp. 25–28, 2007.
 - 12) F. Itakura. Minimum prediction residual principle applied to speech recognition. *Audio*, Vol. 23, No. 1, pp. 67–72, 1975.
 - 13) H. Sakoe and S. Chiba. Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics Speech and Signal Processing*, Vol. 26, No. 1, pp. 43–49, 1978.
 - 14) J. B. Kruskal and M. Liberman. *Time Warps, String Edits, and Macromolecules: The Theory and Practice of Sequence Comparison*, chapter The Symmetric Time-Warping Problem: from Continuous to Discrete, pp. 125–161. Addison-Wesley, 1983.
 - 15) Stan Salvador and Philip Chan. FastDTW: Toward accurate dynamic time warping in linear time and space. *Time*, Vol. 11, No. 5, pp. 70–80, 2004.
 - 16) fastdtw: Dynamic time warping (DTW) with a linear time and memory complexity. <http://code.google.com/p/fastdtw/>.
 - 17) NASA. Swift/BAT hard X-ray transient monitor. <http://swift.gsfc.nasa.gov/docs/swift/results/transients/>.

MAXI/GSC による全天観測データ公開システムの開発

中平聡志^{*1}, 海老沢研^{*2}, 根来均^{*3}, 三原建弘^{*4}, 杉崎睦^{*4}, 芹野素子^{*4},
諏訪文俊^{*3}, 浅田真^{*3}, 富田洋^{*1}

Development of the MAXI/GSC all-sky data archive system

Satoshi NAKAHIRA^{*1}, Ken EBISAWA^{*2}, Hitoshi NEGORO^{*3}, Tatehiro MIHARA^{*4},
Mutsumi SUGIZAKI^{*4}, Motoko SERINO^{*4}, Fumitoshi SUWA^{*3}, Makoto ASADA^{*3} and Hiroshi TOMIDA^{*1}

Abstract

MAXI, the only all-sky X-ray monitor on the International Space Station (ISS), started observation in 2009 August. MAXI has two major objectives, 1) to issue rapid alerts of new X-ray novae/transients into astronomers worldwide and 2) to acquire long-term all-sky X-ray data. Thanks to the good network connection between the ISS and the ground station, we can obtain 70% of MAXI data through the real-time communication path. We built a ground system (referred as “MAXI-DB”) to handle real-time data efficiently; with the MAXI-DB and the alert system we issued 29150 rapid alerts in the three years since MAXI started observation. It is crucial to provide MAXI data to scientists all over the world, because MAXI data includes rich information. The MAXI data has been released to public step by step. Daily data products including images, light curves and energy spectra for 29300 pre-registered X-ray sources have been archived on the MAXI web site (<http://maxi.riken.jp/top>). Since the MAXI data taken by the slit-scan observation includes complex information on all-sky X-ray sources, it is not easy to collect data for a given source quickly on demand. We thus developed the MAXI on-demand archive system which sort all event data in orders of the sky region and the time. The new archive system is confirmed to process data for any celestial source 29100 times faster than the former system.

Keywords: X-ray astronomy, MAXI, data archive, real-time data

概要

MAXI(Monitor of All-sky X-ray Image)は2009年8月から「きぼう」船外実験プラットフォーム上で観測を続けているX線全天モニタである。MAXIには2つの主要な役割があり、1つはX線突発変動天体の速報、もう1つは長期にわたる全天のX線による定常的な観測である。MAXIはISSの通信環境により、観測時間の70%はリアルタイムにデータを取得できる。MAXIは、リアルタイムデータを効率良く扱うため、MAXI-DBという地上システムを構築し、それらを利用してこれまでの3年間の観測で150件以上ものX線天体の変動現象に関する観測速報を行った。MAXIの観測データには多くの天体に関する様々な科学情報が含まれるため、データを全世界の科学者に利用しやすい形式で提供することは重要である。我々は段階的にデータ公開を実現してきており、現在までに300程度のカatalog天体の日々の画像、ライトカーブ、エネルギースペクトルデータや、ユーザーが任意に入力した位置/時刻に対応したデータ(オンデマンドデータ)を公開している。MAXIのデータは時間/空間方向で複雑に入り組んでいるため、1日毎に全データを1ファイルにアーカイブする方式では非効率で、オンデマンド公開で天体毎の解析処理の際、要求される処理時間を達成できなかった。そこで、X線データを時刻や到来方向によって整理し、幾つかのファイルに分割し、結合する方式を使ったシステムを構築した。その結果、100倍程度の効率化を達成することができた。

キーワード: X線天文学, MAXI, データアーカイブ, リアルタイム

*1 宇宙航空研究開発機構宇宙科学研究所 ISS 科学プロジェクト室 (ISS Science Project Ofce, ISAS, JAXA)

*2 宇宙航空研究開発機構宇宙科学研究所宇宙科学情報解析研究系
(Department of Space Science Information Analysis, ISAS, JAXA)

*3 日本大学理工学部物理学科 (Department of Physics, Nihon University)

*4 理化学研究所 MAXI チーム (MAXI team, RIKEN)

1 はじめに

全天 X 線監視装置 MAXI(Monitor of All-sky X-ray Image)[1] は、国際宇宙ステーション (ISS) 日本実験棟船外実験プラットフォーム上で、現在観測を継続中である。JAXA と理化学研究所が主体となり、東京工業大学、青山学院大学、大阪大学、日本大学、京都大学、宮崎大学、中央大学の研究者や大学院生によって運用および宇宙物理学の観測的研究が行われている。MAXI は ISS の周回に伴い、進行方向と垂直に細長い視野で走査観測を行うことで周回ごとに全天の大部分を観測する。検出感度は過去に存在した同種の観測装置と比較して 1 桁優れており (瞬間的な検出感度、長期観測による限界値共に)、これまで 2 年半の観測により数百の X 線源を検出している (図 1)。この高い検出感度と全天カバー率、観測頻度に加え ISS の恵まれた通信環境を活用することで、MAXI は

1. 突発的な天体の出現、既知の天体の変動現象について、即時に全世界に通報する
2. あらゆる天空方向に対する、長期 X 線観測データを恒久的にアーカイブし全世界に対して提供する

という使命を与えられている。

この使命を達成するためには、過去の観測装置または現在運用中の指向型観測装置と全く異なるデータの取り扱いが必要であり、我々はその開発を行ってきた。本論文では、これまでに実現している天体変動現象の速報と、オンデマンドデータ公開を実現するための開発について説明する。まず 2 章では MAXI の装置と、**データ公開**および**データアーカイブ**の意義、これまでに行なってきたデータ公開の履歴について述べる、3 章では MAXI-DB と呼ばれる地上系を中心としたリアルタイムデータの扱いと突発天体変動現象の速報、4 章ではユーザーの入力した時刻と位置に対応するデータを提供する「オンデマンドデータ公開」の概要および、短時間でデータを提供するために構築した、短期アーカイブシステムについて



図 1 運用開始以降、約 2 年半の観測データを積分した全天画像。2-4, 4-8, 8-16keV の X 線強度をそれぞれ赤、緑、青に対応させた。銀河面上の明るい天体から、0.5mCrab 程度の暗い天体まで数百の X 線源が検出されている。

説明する。最後に、MAXI のデータを永久的に保存・公開するための長期データアーカイブシステムについて述べる。

2 MAXI とデータ公開

2.1 観測装置と観測のメカニズム

X 線天文学の黎明期より後に登場したほとんど全ての全天モニターは、主検出器としての指向型観測装置に対する、副検出器という位置づけであり、限られた性能しか与えられなかった。一方、MAXI は全天観測に特化した装置として考案され、電源、冷却、通信などのリソースが用意されている ISS という環境において、X 線検出器に十分な重量を割くことができ、それによって全天モニターとしては過去最高の性能を実現することができた。具体的には、以下の 2 種類の科学観測装置を搭載している。

■ **GSC(Gas Slit Camera)** GSC([2], [3]) は 12 台のガス比例計数管 (10m の炭素芯線と Xe99%+CO₂3% ガスを使用) によって構成されていて、2-30keV の X 線に感度を持ち、12 台合計で 5000cm² と非常に大きな検出面積を持つ。MAXI は ISS の進行方向と垂直方向に視野を持ち、GSC はの各視野の瞬間的な視野は 3° × 160° である。

■ **SSC(Solid State-slit Camera)** SSC([4], [5]) は 32 枚の CCD カメラチップを持ち、0.5-12keV の観測帯域と、3 × 90° の視野を、GSC 同様 2 つ持っている。また、その高エネルギー分解能を生かして、より詳細な輝線スペクトルの測定が可能である。

各カメラは、スリットとコリメーターによって制限/分割された細長い視野を持ち、検出器の 1 次元位置決定性能とあわせて X 線光子の到来方向を決定する事ができる。約 92 分毎の ISS の周回に伴って、この進行方向に対して垂直に細長い視野による走査観測を行うことで、全天の広い領域を見ることができる。MAXI の観測効率、検出感度などについて、表 1 にまとめた。X 線検出器は地球上の高緯度帯など、粒子バックグラウンドの高い領域では観測を停止するが、1 周回あたり全天の 5-70% 程度、1 日では 95% 程度を観測することができる。そのことで、ある天空領域に対しては、40-100 秒程度の観測を 1 日に数回から最大で ~ 20 回行うことができる。

2.2 データ公開

高感度で全天を走査する MAXI の観測データには多様な天体からの情報が含まれ、思わぬ発見につながる

表 1 MAXI による典型的な観測効率と、GSC, SSC それぞれの検出感度.

| | |
|---------------------|---|
| 全天カバー率 (GSC+SSC) | 50-70% / orbit ~ 95% / day (ある位置に対しては ~ 十数 scan/day) |
| リアルタイム ・データリンク | 60-70% |
| 検出感度 (5σ) | GSC : 50 mCrab/scan, 20 mCrab/day SSC : 100 mCrab/day |

現象が潜んでいる可能性があり得る. よって, 全世界に対して使いやすい形でそのデータを提供し, より多くの科学者が利用できるようにすることが本質的に重要である. 我々は, 解析やデータ公開の準備が整い次第, 理化学研究所に設置した MAXI 公式 Web ページ¹を通じて, 科学的利用が可能な処理済みデータを以下に示す様に段階的に提供してきた.

■ 2009 年 12 月~光度曲線データ MAXI によって検出された天体, もしくは過去の観測装置や現在運用中に他の観測衛星によって検出されて監視が必要な天体の位置の X 線光度変化で, 図 2 にはその一例を示した. 最新の数日間の簡易プロットや運用開始以降にわたる数値データが入手可能である.

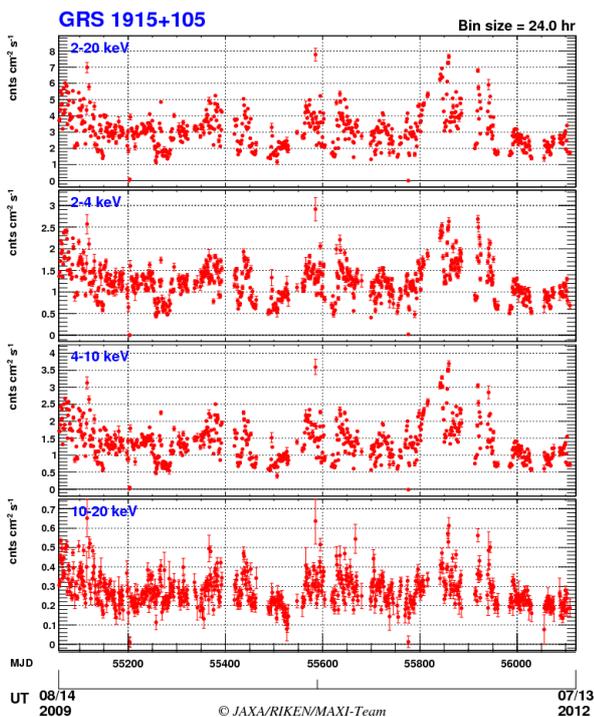


図 2 MAXI 公式 Web ページから取得した光度曲線の図. ここでは, マイクロクェーサー GRS 1915+105 の全観測データを示した.

■ 2010 年 09 月~1 日毎のスペクトルデータ公開

MAXI によって有意に検出されている天体のうち比較的明るいものに限り, 1 日毎に積算したエネルギースペクトルと応答関数データを提供.

■ 2011 年 11 月~オンデマンドデータ公開 任意の天体・時刻に対する画像, 光度曲線, スペクトルを, ユーザーの入力に応じて提供する. 詳しくは 4 章で説明する.

3 リアルタイムデータと地上システム

2009 年 8 月の観測開始以来, 現在にまでに, MAXI はブラックホール連星, 中性子星連星, X 線パルサー, 活動銀河核といった様々な高エネルギー天体を連続的にモニターし, 状態遷移などの変動現象や, ガンマ線バーストや星フレアなど, 数多くの突発現象を捉えてきた. 検出された変動現象のうち価値が高いものは, MAX チームの提供するメーリングリスト², The Astronomer's Telegram(ATel)³ および The Gamma-ray Coordinates Network(GCN)⁴ などの天文学観測コミュニティに対して速報され, 現在までに 100 件以上となっている. この速報を実現するためには, ISS の通信環境と地上系が大きな役割を果たしている. 本研究は, ISS から地上系のデータの流れと MAXI-DB を起点とする MAXI 地上系の概要について説明する. データフロー及び地上系のより詳細と, MAXI-DB および速報系の技術的な面については出版準備中の別論文 [6] で述べられる.

3.1 観測データのフロー

図 3(a) に簡略化した機上から地上へのデータの流れを示した. X 線観測データや House Keeping(HK) データ, ISS 補助データからなる MAXI データは, オンボードのデータ処理系によってテレメトリデータとして処理され, データ中継衛星を経由してつくば宇宙センターに転送される. 転送系には低速系と呼ばれる 1553b と, 中速系と呼ばれる Ethernet の両方が用いられ, 低速系は転送容量が小さいため中速系と比較してデータの品質を低下させているが, それでも科学的目的が達成できるように設計されている. 60 ~ 70% 程度の時間は ISS と地上とのリンクが確立されていて, 取得から数秒で地上へ届くリアルデータとして取得でき, リアルタイムに転送されなかったデータは, Communication Outage Recorder に蓄積され, リンク再開後に再生され取得可能である.

1 <http://maxi.riken.jp/top/>

2 <http://maxi.riken.jp/mailman/listinfo>

3 <http://www.astronomerstelegram.org/>

4 <http://gcnc.gsfc.nasa.gov/>

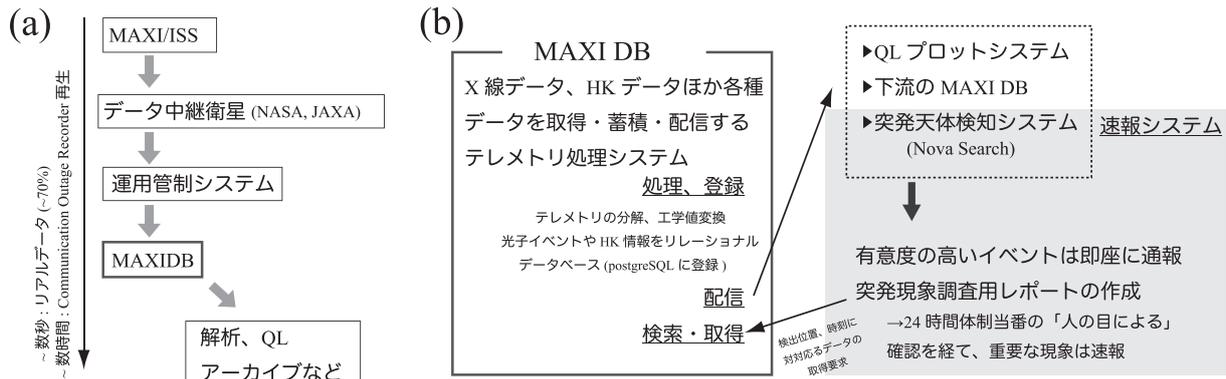


図 3 (a)ISS から MAXI 地上系までのデータの流れの概略 (b)MAXI-DB と MAXI 地上系

3.2 地上システム

運用管制システムを通過したテレメトリデータは、MAXI チーム独自のシステムである MAXI-DB に転送され、MAXI-DB を起点として様々な処理がなされる (図 3(a)).

■ **MAXI-DB** MAXI-DB は MAXI チームと SEC 社によって開発されたシステムで、テレメトリを受信し、データを分解・整理、検出器のキャリブレーション情報に基づいて加工し、PostgreSQL[7] によるリレーショナルデータベースへの登録を行なうという機能を持つ。また、ユーザーは Quick Look(QL) 用インターフェイスに接続することでリアルタイムのデータ配信を受け、光子イベントや補助ファイルのデータベースに接続し、検索とデータ取得を行なうことができる。MAXI-DB のデータ配信機能を利用して我々は、観測装置の状態を確認するための QL 表示や、Nova Search と呼ばれる変動現象探索・通報システムを運用している。突発天体や変動天体の候補が発見されると、さらなる解析のため、データベースから注目する時刻と位置に対応する光子イベントを取得し、調査用のレポート Web ページを生成し、24 時間体制の人の目による確認を経て、重要性の高いイベントについては速報を行なっている。また、データベースに蓄積された光子イベントや、ISS 姿勢、太陽パネル角などの科学解析に必要な情報は、日毎に FITS ファイルとして抽出され、理化学研究所を通じて共同研究機関に配布されている。

4 オンデマンドデータ公開と空域イベントファイルシステム

MAXI チームは、2011 年 11 月より、任意の時刻/位置に対する処理済みデータを提供するオンデマンドデー

タ公開を提供している。オンデマンドデータ公開では既知天体のデータ公開とは異なり、ユーザーの入力に対応してリアルタイムにデータ取得を行ない、イメージ、スペクトルおよび光度曲線データと表示用の Web ページを生成する。このように、MAXI の膨大な観測データから必要なデータを、実用的な処理時間で抽出するには観測の特性を考慮に入れたデータアーカイブが必要である。以下では、まずオンデマンド公開で行なっている処理の概要について説明し、その後 MAXI のような観測装置のデータから効率的にデータを抽出するために考案したデータアーカイブシステムについて説明する。

4.1 オンデマンド公開の処理

オンデマンドデータ公開では、ユーザーからの入力を受けると同時に以下の様な手順で処理を行い、データを提供している。

1. ジョブの受付

(a) Web インターフェイス

Web インターフェイス⁵を通じてユーザーは、興味のある天体の座標、時刻範囲、ファイル名として使う名前、イメージおよび光度曲線のエネルギー範囲、光度曲線の積分時間範囲を入力する。詳細なオプションとして、自動作成リージョン(後述)の半径や、ds9[8]形式のリージョンファイルを入力する事もできる。

(b) データ処理サーバーへジョブの投入

入力パラメータに対して簡単なチェックを行い、ソケット通信を使って解析サーバーにジョブを投入する。現在のところ処理時間に対して利用ユーザーは多くないので、解析サーバーは 1 台のみである。

⁵ <http://maxi.riken.jp/mxondem/>

2. 初期処理

(a) 結果表示用 Web ページの作成

データプロダクトと情報を表示するための html ファイル(図4)を作成・転送し、ユーザーは中間ページを介して、表示用ページへリダイレクトされる。表示用ページに15秒毎にリロードを繰り返すタグを埋め込みことで、処理が進むごとに情報を更新する。

(b) 周辺天体の抽出と解析リージョンの決定

入力した位置の周辺の既知天体を表示するために、カタログを検索する。また、解析に用いる ds9 リージョンを生成する。現時点での解析はすべてアパ

チャ測光を用い、ソース領域とバックグラウンド領域にはそれぞれ入力位置の周囲を円状/円環状に利用する。同時に、周辺天体の情報から2.2で説明したカタログ天体に対する光度曲線データを調べ、明るさに応じて付近の既知天体からの寄与を解析から取り除く。

(c) X線イベントの抽出

後述のアーカイブシステムを利用し、解析に必要なX線イベントデータを抽出する。

(d) 補助データリストの作成

1日ごとにアーカイブされた、補助データ(ISSの姿勢情報など)のうち解析時間範囲に対するリストファイルを作成。

(e) Good-Time Interval(GTI)情報の作成

MAXIデータを用いた、アパチャ測光方式による精度の良い光度曲線とエネルギースペクトルの作成には、その天空領域内での各スキャン観測が不連続にならないようなデータ選択が必要であることがわかっている。ここではHK情報とISSの姿勢情報を利用することで、指定した期間内に存在する多数のスキャン観測のうち、解析領域内で観測が開始/停止することなく連続的に観測したスキャンのみを選択し、残り(1-2割程度に相当)を破棄するGTIを作成している。

3. データプロダクトの作成

以下のようなデータを作成する。データセットを作成すると共に、表示用htmlを更新し、転送する。

(a) 天体周辺のイメージ

X線イベントを天空座標に投影し、FITS形式の画像ファイルを作成する。全データを用いた、観測量の補正やバックグラウンド差し引きのない生のイメージが生成される。

(b) エネルギースペクトル

上記で作成したGTIに含まれる観測を利用し、FITS形式のエネルギースペクトルファイルと、応答関数ファイルを作成する。スペクトル解析に関する詳細や、精度については、同様の手順を用いて解析した論文[9],[10]で詳しく述べられている。

(c) 光度曲線の作成

入力エネルギー範囲に対応するX線強度変化を入力された積分時間範囲ごとに計算し、ascii形式で作成する。全データ処理が終了すると同時に、自動更新を中止するためhtmlファイルの値を変更し、更新して終了。

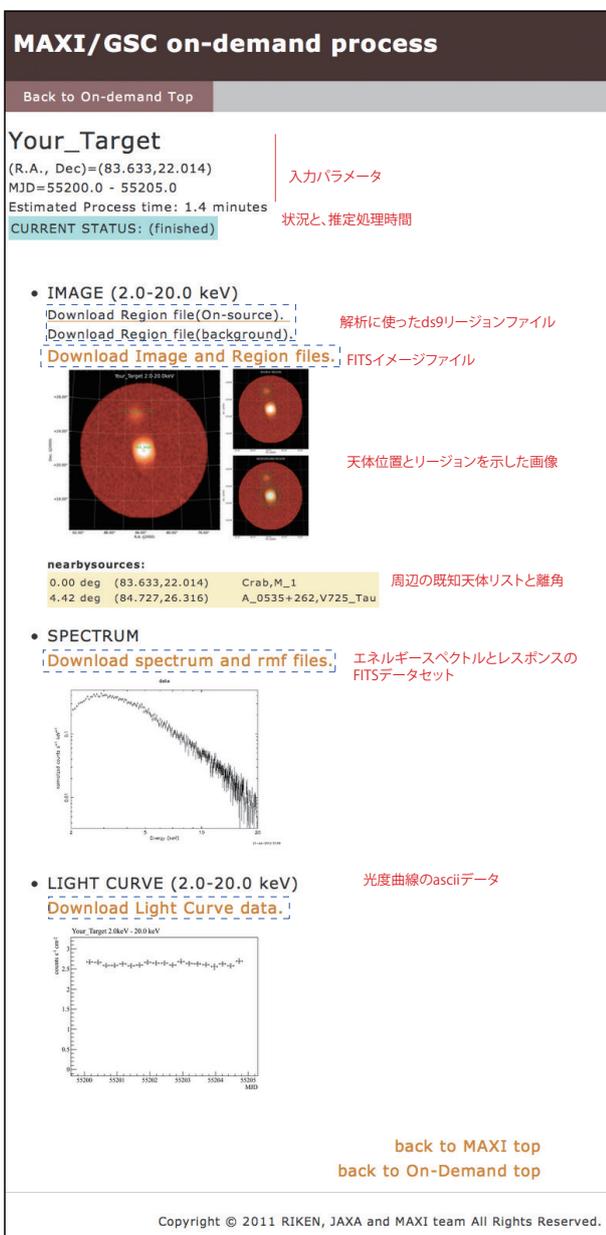


図4 オンデマンド公開、結果表示 Web ページの例。赤字で項目の説明を行い、青い点線の枠でデータプロダクトへのリンクの位置を示す。

4.2 空域イベントファイルシステム

まず、MAXIの観測の特性について指向型観測装置との比較という点から再度整理する(表2)。すぎく衛星を始めとする、望遠鏡を備えた観測衛星はターゲット天体への1指向観測を単位とし、ターゲット位置を使って必要とする観測データを取り出すことが出来る。一方でMAXIは、指向型観測装置の典型的な視野(全天の100万分の1)と全く異なり、全天の~95%の領域を1日に何度も観測することで、その観測データは、時間・空間方向に複雑に入り組んでいる。また、原理的に、一度各イベントを読み込んで必要な領域に含まれるか判定するまでは、それが必要なデータか判断できない。つまり一天体に注目した解析では、データ抽出クエリ毎に全データを読むことになり非効率である。

4.2.1 概要

データ抽出の効率化のため、我々は観測データをアーカイブする際にHEALPIXライブラリ[11]を利用して均等に分割した領域によって予め整理しておく手法を試みた。ここでは、HEALPIXのパラメータを $n_{\text{side}}=16$ に設定することにより、全天をMAXIのPSF全体が含まれる大きさ:約 $3 \times 3 \text{ deg}^2$ より少し大きい $3.6 \times 3.6 \text{ deg}^2$ のピクセルからなる3072の領域に分割した。図5に分割の様子が表されていてそれぞれ、標準的な解析の時に利用する半径 8° の円(マゼンタの領域)、HEALPIX上の1ピクセル(灰色)、半径 8° の円に含まれるピクセルの集合(青色)を表している。今回構築したシステムの上

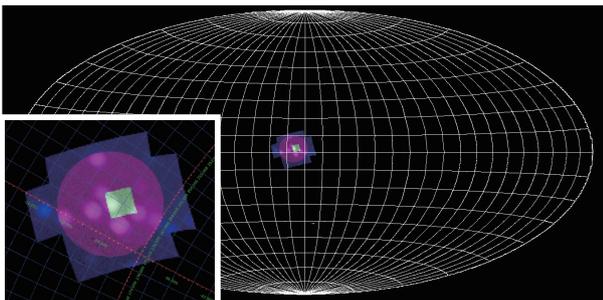


図5 全天を表す領域中に示した、(白)healpixの1ピクセル、(赤)ある位置の周囲 8° と(青)それを含むピクセルの集合

では各MJD(Modified Julian Date⁶)で1ピクセル毎に1つのFITSファイルを作成し(空域イベントファイル)、以下に例を示すようなディレクトリ構造でファイルを格納した。

```
EVENT_TOPDIR/
|-- MJD55060
| |-- gsc_MJD55060_pixel0000.fits
| |-- gsc_MJD55060_pixel0001.fits
| |-- ... ..
| +-- gsc_MJD55060_pixel3071.fits
|-- MJD55061
| |-- gsc_MJD55061_pixel0000.fits
| |-- ... ..
|-- ... ..
```

その後、必要な時間範囲、中心位置の座標と抽出半径を与えると、1日/1ピクセルの時間/空間幅で必要な空域イベントファイルをリストするソフトウェアを作成し、データ取得に利用する。

4.2.2 効率化の度合い

本システム構築以前に利用されていた、1日毎に全イベントデータを1ファイルをアーカイブする方式(全天イベントファイル)と比較して、イベントデータを天空領域に分割した「空域イベントファイル」を利用したシステムがどの程度効率的なのかを調査した。

■**検証方法** ここではまず、MAXIの全カメラが稼働を開始した2009年8月15日から2012年7月17日までの実観測データを使い、現在までに確立している解析では使わないエネルギー帯域や検出器位置のデータをあらかじめ除く一次処理を施した上で、以下で説明する各方式にしたがってアーカイブを作成した。全天イベントファイルでは、X線イベントは1日毎に1ファイルに、空域ファイルシステムは4.2.1で説明したとおりに、1日毎、HEALPIXピクセルごとにFITSファイルとしてアーカイブされている。

X線イベントの計数は、FITSファイルのイベント情

表2 MAXIと指向型観測装置の観測と観測データの比較

| | 指向型観測装置 | MAXI |
|--------------------------|---------------------------------|--------------------------------------|
| 1日あたりの観測領域 (全天に対する割合) | $\text{several} \times 10^{-6}$ | 0.95 (同じ位置を周回ごとに繰り返し観測) |
| 観測データ | ある位置に対する 固定観測の集合 | 観測する空域が連続的に 変わり時間・空間方向 入り組んでいる |

⁶修正ユリウス日、1MJDは1日に対応する

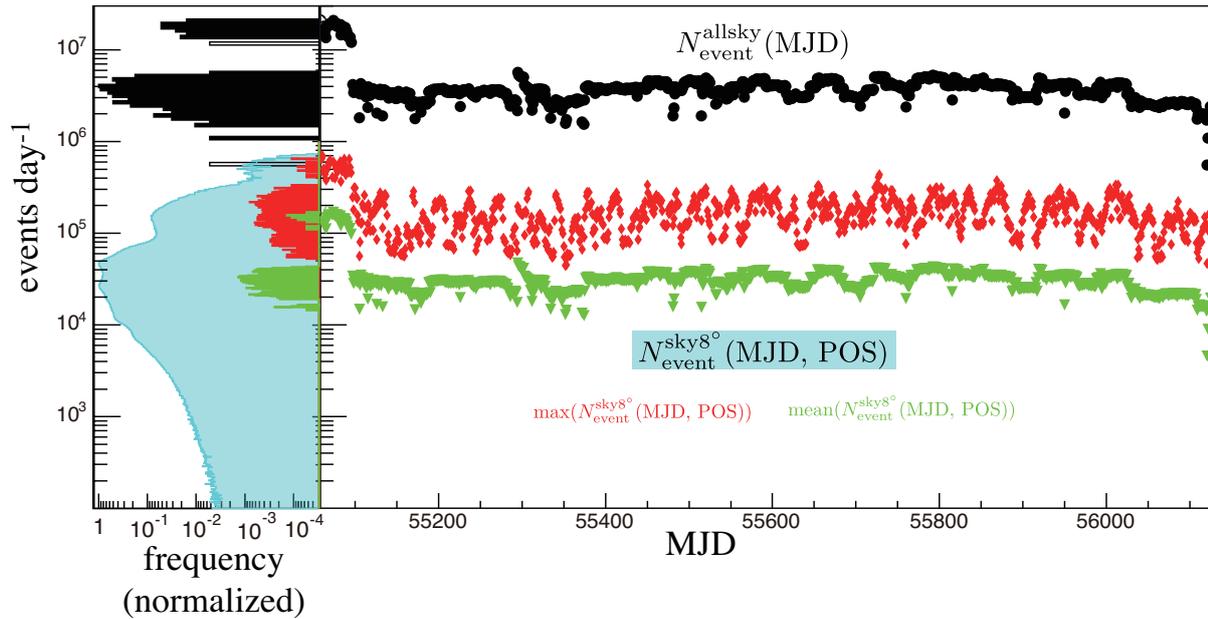


図6 2009年8月15日から2012年7月17日の期間の一次処理済み全天イベントファイルに含まれるX線イベントの一日ごとの数 $N_{\text{event}}^{\text{allsky}}(\text{MJD})$ (黒)と、各ピクセルを中心に8度に相当するピクセルのイベント数を積算した $N_{\text{event}}^{\text{skys}8^\circ}(\text{MJD}, \text{POS})$ の最大値(赤)と、平均値(緑)。左図にはその度数分布を作成し(最大値を1に規格化)、すべての例に対する度数分布をシアンで示した。MJD55100以前は、地球位置に依存する観測停止条件が異なり、運用時間効率が高いためイベント数が多くなっている。

報を含むテーブルのヘッダーから行数の情報を読みとることで行なった⁷。全天イベントファイルシステムでは、指定する天体位置に関係なくデータ抽出時に読み込むイベント数が一定であるので、1日ごとのX線イベント数 $N_{\text{event}}^{\text{allsky}}(\text{MJD})$ を求めた。一方で、空域イベントファイルシステムでは位置によって異なるため、各1日分のデータに対してHEALPIXピクセルの周囲 8° に含まれるピクセルのイベント数を積算する操作を、各3072ピクセルに対して繰り返すことで、1日毎に各領域のX線イベント数 $N_{\text{event}}^{\text{skys}8^\circ}(\text{MJD}, \text{POS})$ を求める。

■検証結果 上記の手法で調べた、各アーカイブ方式によるX線イベント数の1067日に渡る時間的な変化とその度数分布を図6に示した。 $N_{\text{event}}^{\text{allsky}}(\text{MJD})$ は、初期を除いて $(2-4) \times 10^6 \text{ events day}^{-1}$ に分布している。その変化は、運用するカメラの台数の変化、検出器保護のため観測を停止する地球上の領域の変更、ISSの放射線環境の変化によるバックグラウンド計数率の変動(太陽活動の変化や、着陸用の高度計に ^{137}Cs 線源を利用しているソユーズ宇宙船の発着などの影響が大きい)など、種々の条件で決まる各天空領域に対する観測量に依存して変化している。 $N_{\text{event}}^{\text{skys}8^\circ}(\text{MJD}, \text{POS})$ は1日ごとに3072点あるの

で時間変化の図からは省略し、代わりにその最大値と平均値を示した。それらの度数分布は左図に示してあり、各方式ごとに最大の度数が1になるように規格化した。 $N_{\text{event}}^{\text{allsky}}(\text{MJD})$ に対して、 $N_{\text{event}}^{\text{skys}8^\circ}(\text{MJD}, \text{POS})$ の最大値は典型的に数十分の1、平均値は約1/200であった。

次に、 $N_{\text{event}}^{\text{skys}8^\circ}(\text{MJD}, \text{POS})$ の抽出位置に対する関係を調査した。ここでは、指定した中心位置ごとにX線イベント数を積算するため、各抽出領域の観測効率⁸の高さと含まれる天体の明るさに依存する値になる。図7をみてわかるように、X線帯域で最も明るく見えるさそり座X-1が抽出領域に含まれる場合にイベント数が多く、次いで銀河中心近傍、かにパルサーやGRS 1915+105など明るい天体の付近と、図の右下に見られる観測効率の高い領域のイベント数が多く、その他の領域の、位置による違いはfactor2~3であった。

最後に効率化の度合いを調べた。図8には、 $N_{\text{event}}^{\text{skys}8^\circ}(\text{MJD}, \text{POS})/N_{\text{event}}^{\text{allsky}}(\text{MJD})$ の度数分布を示している。観測期間×ピクセル数だけ存在する全パターンのうち95%は、 1.6×10^{-1} 以下であり、67.5%は 1×10^{-2} より小さいことから、典型的には2桁程度高速化したと言える。

7 厳密に言えば、全天イベントファイルに対しては各イベントの位置と抽出位置の離角を計算するコスト、領域イベントファイルに対してはFITSファイルのオープン/クローズの回数が増えるコストが考慮されていない。

8 ISSは51.6度の軌道傾斜角で周回し、多くの時間を進行方向に対して同じ姿勢で費やす。通常姿勢では、検出器上のある位置が観測する天空領域が決まっている。結果的として有効面積の方向依存性やカメラの稼働台数により天空領域上の観測量が一様にはならない。

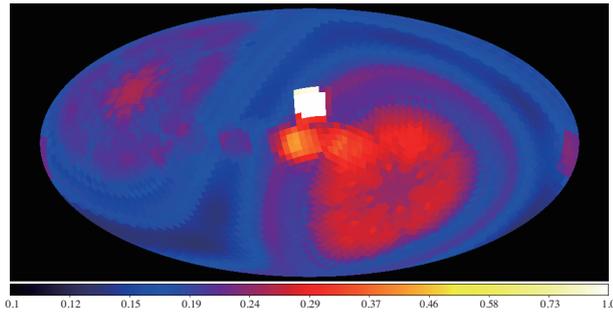


図7 $N_{\text{event}}^{\text{sky}8^\circ}$ (MJD, POS) の抽出の中心位置による依存性. 各中心ピクセルに対しては全期間のデータを積算した後, 最大のピクセルが1になるよう規格化した上で銀河座標系を使って投影した. 画像中心からやや上に位置するさそり座 X-1 を始めとする明るい天体の付近や, 観測効率の良い天空領域において値が高くなっている.

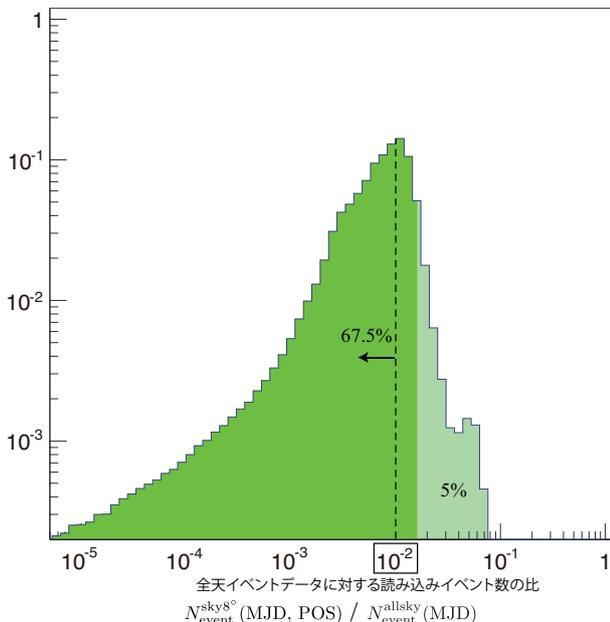


図8 空域イベントファイルシステムを使った, 効率化の割合を示す図. $N_{\text{event}}^{\text{sky}8^\circ}$ (MJD, POS) / $N_{\text{event}}^{\text{allsky}}$ (MJD) の度数分布と作成し, 積分値が1になるよう規格化した. 薄い色で示した領域が全体の5%を示していて, 0.01以下のサンプルの割合は全体の67.5%である.

4.2.3 議論と今後の展望

旧方式である全天イベントファイルシステムでは, ある天体の周囲のデータを抽出する際に1日程度を擁していたが, 新方式である空域イベントファイルシステムでは100倍程度の効率化により, 長期間のクエリに対しても現実的な時間でデータの提供ができるようになり, そのことでオンデマンド公開が実現できた.

現在は周囲画像のチェックを容易にするため, X線イベントを抽出する半径を 8° (3072分割した場合26-27

ピクセル)に設定しているが, ある天体のデータに限れば, 現在の解析方法でも半分の半径 4° のデータがあれば十分であり, ピクセル数と明るい天体が含まれる割合が減少することにより更に効率化する. また, バックグラウンドの再現性の調査を進め, 疑似イベントを発生するバックグラウンドジェネレータが利用できるようになればバックグラウンド領域が不要となり, 天体のPSFが入る1-4ピクセルで科学解析を行える.

5 まとめ

ISSに搭載された全天X線監視装置MAXIは2009年8月以降, 3年近くをわたって観測を継続しており, 段階的にデータ公開を実現してきた. 現在, 理化学研究所からMAXIチームの研究者に, 1日ごとに1ファイルにアーカイブされた全天データファイルが配布されているが, オンデマンドデータ公開を実現するためには, 効率的なアーカイブシステムの開発が必要であった. 今回, 我々はX線イベントデータを時間と天空領域を使って整理し, アーカイブする空域イベントデータシステムを開発することで, 10^2 倍程度の効率化を実現し, オンデマンドデータ公開において実際に利用している.

オンデマンドデータ公開で提供しているデータ, アーカイブシステムの項目で説明した内容はGSCについてのみであるが, 我々はGSCと同じ観測原理を利用したSSCのデータについても, 本論文で示した方式によるアーカイブ化とオンデマンド解析の実装を完了しており, ごく近い将来にはその提供が開始される予定である.

全MAXIデータは将来的に, 宇宙科学研究本部の科学衛星運用・データ利用センターに配置, 恒久的にアーカイブ化され, ftoolsに準拠した解析ソフトウェアと共に未来の科学者の利用に供される. その際に利用されるアーカイブシステムの方式として, 今回開発したシステムの採用が検討されている.

参考文献

- [1] Matsuoka, M., et al., PASJ, 61, 999 (2009)
- [2] Mihara, T., Nakajima, M., Sugizaki, M., et al., PASJ, 63, 623 (2011)
- [3] Sugizaki, M., Mihara, T., et al., PASJ, 63, 635 (2011)
- [4] Tomida, H., et al., PASJ, 63, 397 (2011).
- [5] Tsunemi, H., et al., pasj, 62, 1371 (2010)
- [6] Negoro et al., in prep.
- [7] <http://www.postgresql.org/>
- [8] Payne, H. E. et al., ADASS XII ASP Conference Series,

Vol. 295, 489

- [9] Nakahira, S. et al., PASJ, 64, 13 (2012)
- [10] Shidatsu, M. et al., PASJ, 63, 803 (2011)
- [11] Górski, K. M. et al., 622, 759 (2005)

「あかり」 中間赤外線全天 diffuse マップの作成

石原大助^{*1}, 金田英宏^{*1}, 毛利彰男^{*1}, 近藤徹^{*1}, 鈴木慧士^{*1}, 山岸光義^{*1}, 大藪進喜^{*1},
尾中敬^{*2}

Development of AKARI Mid-Infrared All-Sky Survey Diffuse Map

Daisuke ISHIHARA^{*1}, Hidehiro KANEDA^{*1}, Akio MOURI^{*1}, Toru KONDO^{*1}, Satoshi SUZUKI^{*1},
Mitsuyoshi YAMAGISHI^{*1}, Shinki OYABU^{*1} and Takashi ONAKA^{*2}

Abstract

AKARI, a Japanese infrared astronomical satellite, has surveyed all the sky with the 6 photometric bands centered at from 9 to 160 μ m. The mid-infrared all-sky survey was performed by one of the on-board instruments, the Infrared Camera (IRC), with two photometric bands centered at 9 μ m and 18 μ m. The AKARI 9 μ m all-sky map is unique as a tracer of the inter-stellar polycyclic aromatic hydrocarbon (PAH) emission. We have developed further new special reduction processes for the previous maps employed in generating the point source catalog; for example, we have corrected the ionizing radiation effects in the South Atlantic Anomaly (SAA) and the stray light component of the bright moon. Using the improved diffuse maps, we can investigate faint diffuse emission, which leads us to new science.

Keyword: Data reduction, infrared observation

概要

日本初の赤外線天文衛星「あかり」は、波長 9 ~ 160 μ m の 6 つの赤外線波長帯で全天を観測した。この内、搭載装置 IRC によって行われた波長 9 μ m 帯および 18 μ m 帯での全天観測は、過去の IRAS 衛星の成果を一桁凌ぐ感度と空間分解能で塗り替えた。波長 9 μ m 帯のマップは、星間空間の芳香族炭化水素 (PAH) 輝線放射のトレーサーとしてもユニークである。我々は、点源カタログを作成する際に作成したマップに対し、南大西洋磁気異常帯 (SAA) での放射線ヒットの間接的影響の補正、月の迷光の補正、明るい天体の迷光の補正、黄道光の季節間差の補正など、詳細で地道な解析をさらに追加し、より信頼性の高いマップを作成した。これにより、淡く広がった放射成分の議論が可能になり、様々なサイエンスに繋がっている。

1. はじめに

日本初の赤外線天文衛星「あかり」¹⁾は、2006年2月に打ち上げられ、波長 9 ~ 160 μ m の 6 つの赤外線波長帯で全天を観測した。この内、中間赤外線の 2 波長帯 (9 μ m 帯と 18 μ m 帯) の観測は、搭載装置の一つ近中間赤外線カメラ (IRC)²⁾ の検出器のユニークな運用方法により実現した³⁾。IRAS の 12 μ m 帯、25 μ m 帯を一桁凌ぐ感度 (9 μ m 帯で 90 mJy) と空間分解能 ($\sim 4''$) で全天をカバーし、観測結果は既に点源カタログとして C-SODA から世界へ公開している⁴⁾。このカタログは、星形成・晩期型星から系外銀河に至るまで、天文学の様々な分野の研究に活用されている^{5,6,7)}。

星間空間からの放射の観測において、「あかり」9 μ m 帯マップは、IRAS を感度と空間分解能で凌ぐだけでなく、芳

*1 名古屋大学 理学研究科 物理学専攻

*2 東京大学 大学院 理学研究科 天文学専攻

香族炭化水素 (Polycyclic Aromatic Hydrocarbon; PAH) 輝線放射のトレーサーとしてのユニークさを持つ (図 1, 図 2). PAH は星間空間における有機物の最小単位として注目されている. また, 一般的なダストとは加熱や破壊の特性の異なる物質として, 星間現象の解明に利用することができる^{8,9)}.

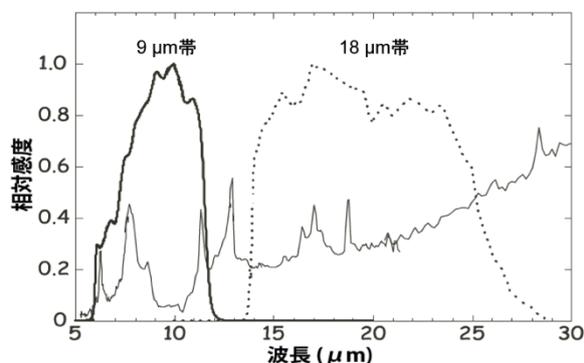


図 1 「あかり」9 μm 帯, 18 μm 帯の波長感度曲線と, 典型的な星間空間のスペクトルの比較. 「あかり」9 μm 帯は星間空間の PAH の 6.2, 7.7, 8.6, 11.3 μm 放射をカバーし, 18 μm 帯は数百 K に加熱された暖かいダストからの熱放射をカバーする.

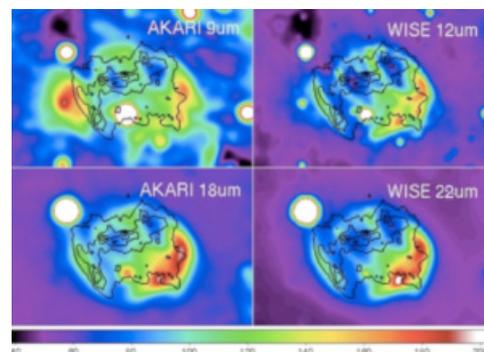


図 2 超新星残骸 W49B の「あかり」9 μm 帯, 18 μm 帯と, WISE 12 μm 帯, 22 μm 帯の画像の比較. Contour は「すざく」X線 (0.2–12 keV). 「あかり」9 μm 帯は特にユニークな情報を提供する.

観測データには, 銀河系の星間空間からの放射 (太陽系外からの放射) 以外に, 黄道光 (太陽系内の惑星間空間塵の熱放射), artifacts (検出器や光学系に由来する偽のシグナル), ショットノイズが乗っている. 太陽系外の星間空間からの放射を定量評価したい場合には, 観測データから黄道光と artifacts を除去しなければならない. Artifacts には, 様々な空間スケールを持つものがあり, ゲインとして乗ってくる成分 (感度の劣化など真の出力に対するかけ算効果になる成分) も, オフセットとして乗ってくる成分 (迷光や散乱光など真の出力に加算される成分) もある. これまでの, 点源カタログの作成⁴⁾や, 同じ観測装置を使った IRC 指向観測のデータ解析²⁾の際にも, artifacts の解析・処理はある程度行われてきた. しかし, 点源カタログを作成する際は, 点源周囲のローカルなスカイ (~数十秒角) を差し引いてから測光をするので, それより大きな空間スケールを持つオフセット成分に関しては補正する必要がなかった. また, 点源カタログ作成の際には, 検出天体数よりも信頼性を重視するコンセプトで処理していたので, artifacts の影響を比較的強く受けたデータについては補正せずに捨てていた. また, 指向観測のデータ解析でも, 視野 (~10 分角) より大きな空間スケールを持つオフセット成分は, 補正する必要がなかった. ところが広域マップを作成し, 広がった放射成分の議論をする場合には, 空間スケールの大きな (数十秒角 ~ 数度角) artifacts の影響も受ける. しかも, 複数スキャンのデータを重ね合わせて深いマップを作ることで, artifacts も強調されてしまうので, より精密な処理が必要となる. また, 評価したい対象の広がった天体を構成するデータの中に artifacts の影響を強く受けたデータがあり, それを取り除くと穴が空いてしまう場合は, 捨ててしまわずに可能な限り補正を施して利用することで, 不定性の範囲内で議論が可能になる.

そこで我々は, 点源カタログを作成する際に点源の抜き出しのために作成したマップに対し, 南大西洋磁気異常帯 (South Atlantic Anomaly; SAA) での放射線ヒットの間接的影響の補正, 月の迷光の補正, 明るい天体の迷光の補正, 黄道光の季節間差の補正など, 他の赤外線衛星プロジェクトでは手を入れていない高度な処理をさらに幾つか追加し, より信頼性の高いマップ・よりカバー率の高いマップの作成を試みた. 検出器の光検出の原理や光学設計など, 観測装置の物理に基づいてデータを解析することで, artifacts に埋もれた真の信号を引き出し, 新しい科学的発見に繋がり, 次の衛星の設計にも重要な情報を得ることができる.

2. データ解析の概要

2.1 観測方法とデータ解析の概要

「あかり」の軌道は, 半年で全天をカバーできる, 上空 700km の太陽同期極軌道である. 半年毎に同じ領域を観測するが, 半年後の観測ではスキャン方向が 180° 回転し, 1年後の観測では再び同じスキャン方向になる. 冷媒の液体ヘリウムは約 16 ヶ月保ったので, 殆どの領域を 3つの季節で, 残り領域も 2季節で観測している. スキャン幅は $10'$ であ

り、その軌道は黄道で4'ずつずれるので、連続的にサーベイしていれば、1回の季節の観測のみでオーバーラップを生じながら空を埋め尽くすことができる。マップの作成では、これら重複したデータをフルに活用し、領域あたり1つのマップを作る。ただし、「あかり」は時々、天文台型運用(指向観測)も行っており、その間に通りすぎた空はデータ欠損となる(gap)。Gapを除いても、全天の90%以上を観測できている⁴⁾。

マップ作成のプロセスの概略を図3に示す。生データはまず、点源カタログ作成時と同様、スキャンパス毎に基本パイプラインで処理する⁴⁾。ここでreset anomaly, linearity, flat, カメラ内散乱光, 宇宙放射線ヒットのスパイクノイズの除去, など基本的な処理を行う。この基本的な処理を終えた段階では、マップのピクセルスケール(4.68")でのノイズレベル(~1MJy/sr)に比べて、最大で数十倍(例えばSAAや月の迷光中)、平均で2~3倍程度のartifactsの影響が残っている。次にこの後述べる、マップの作成のための特殊処理, SAAにおける放射線ヒットの間接的影響の補正, 月の迷光の差し引き, 明るい天体周囲に生じる迷光の補正等, を追加した。各スキャン毎に処理したデータは、重ね合わせ・繋ぎ合わせて1枚の画像にする。

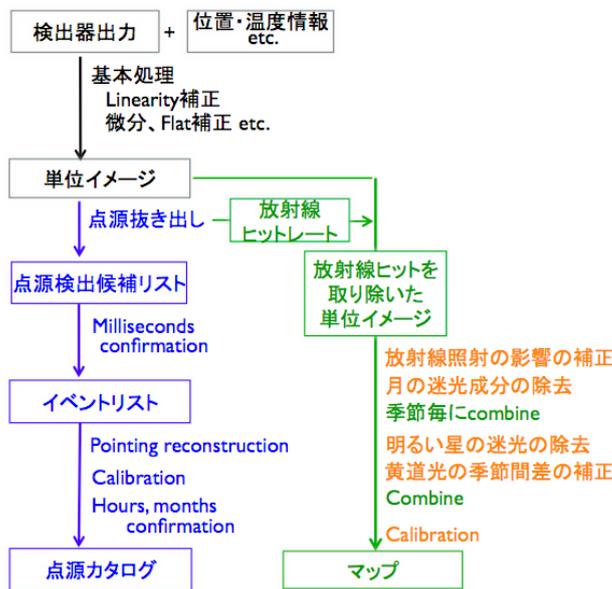


図3 マップ作成プロセスの概要。黒字：共通プロセス部分。青字：点源カタログ作成のためのプロセス。緑字：マップ作成のためのプロセス。赤字：マップ作成のために今回新たに追加したプロセス。

2.2 SAA における放射線照射の間接的影響の補正

南大西洋磁気異常帯(SAA)とは、地球の南大西洋にあって磁場が顕著に弱い領域である¹⁰⁾。「あかり」がそのSAA上空を横切る間、宇宙放射線は地球磁場に妨げられないので、他の天域よりも10倍から最大で200倍も多く観測視野に降り注ぐ。その宇宙放射線は、直接的には観測データ上にスパイクノイズを添加するが、間接的にも出力全体にわずかな増加をもたらすという現象が見られた(図4左)。「あかり」は1日に地球を14.4周するが、この内3~4周がSAAに重なるため、影響を受ける領域は多い。詳細な解析により¹¹⁾、(1)この増加成分がゲインではなくオフセットであること、(2)放射線ヒットレートの関数で表されること(図5)、(3)原因は、除去しきれていない小さなスパイクノイズの集合ではなく、読み出し回路のオフセット変動でもなく、暗電流のわずかな増加であること、をつきとめ、放射線ヒットレートの関数で、系統的に補正することに成功した(図4右)。

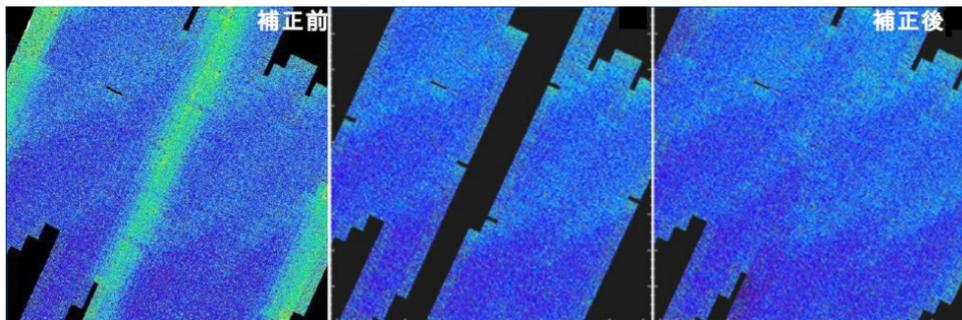


図 4 SAAにおける宇宙放射線照射の間接的影響とその補正。例として、(R.A., Dec.)=(23h50m48s, -19° 42' 00") 付近の 1.7° 四方の領域に対して、1 季節目の観測機会で取得したデータのみを使用し、マップを構成したものを示す。(左) SAA の影響を受けているデータも使って作成したマップ。スキャン方向を矢印で示す。衛星は、SAA の中心部を 14.4 周に 1 度の頻度で通過し、軌道(スキャンパス)は周回ごとに 4' ずつずれるので、スキャン方向に沿って約 1° おきに SAA の影響のピークが現れている。(中) SAA 通過中に取得したデータを除外したマップ。(右) この章で述べる方法で、放射線照射の影響を補正したマップ。

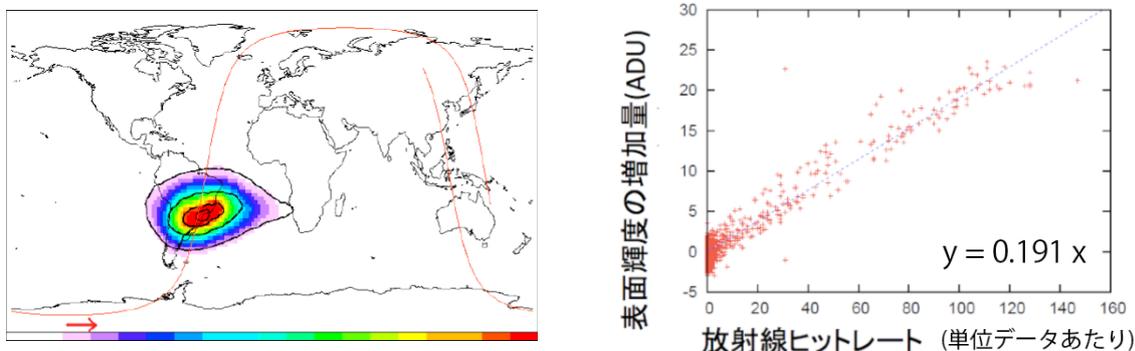


図 5 (左) 宇宙放射線照射に起因すると思われる出力の増加量(カラーマップ)と放射線ヒットレート(contour)の分布を地球座標で示す。ヒットレートの高い場所ほど出力の増加量も多い。赤線と矢印は、とある周回での衛星の軌道を示す。(右) 宇宙放射線のヒットレートと検出器出力の増加量の関係。ヒットレートは、データ処理上の「単位データ」あたり、検出器が受けている放射線ヒットの数(1 素子 30 μ m 角の 256 \times 2 素子の検出器が 13.5 秒の間に受ける放射線ヒットの合計数)を表している。このプロットから、極端に外れている点は除外し、一次式(青点線)でフィットし、この関係を補正式として使うことにした。

2.3 月の迷光成分の除去

観測視野の近くに月があると、直入射を防いでも迷光の影響を受ける(図 6a)。月齢によって迷光の形状や強度が異なり、トラス形状を反映した特徴的な構造も示すため、使えるデータと使えないデータを、単純に月離隔だけで区別するのは難しい。

そこで、月の迷光のパターンを定量的に評価した。まず月入射の前後のデータを集め、月を中心とした検出器座標系(縦軸がスキャン方向、横軸がクロススキャン方向)で、強度マップを作成した(図 7 左)。この時、銀河面が写っているデータは除外し、点源は差し引いた。さらに、月の影響を受けていない部分(クロススキャン方向に離隔 $\sim 50^\circ$ 付近)のデータから黄道光の成分を推測し、この強度マップからは差し引いた。銀河面成分である淡い小構造が含まれる場合もあるが、各月で作成した強度マップの median をとることで、月起因で無い成分は落とされる。月との相対ベクトル空間を全て埋められる訳ではないので、データが無い部分(図 7 左の黒い筋に相当、全体の約 50%)は内挿した。

月入射の機会ごとに作成した月の迷光強度マップの median 画像を、図 7 中に示す。迷光は、(1) 離隔 40° に渡って広がっており、(2) 非対称な構造で、(3) トラス起因と思われる 3 対の円弧構造が見られる。また、このパターンは、季節間や上向き下向きスキャンの間で、比較的安定であることが分かった。図 7 右に、各季節での迷光パターンの分散画像を示す。よって、この迷光強度マップを使ってマップから迷光成分を差し引き、差し引き誤差の大きな部分(迷光パターンの季節間分散が大きい部分)はマスクすることで、使えないデータは明確に排除し、月の影響を受けたデータもある程度(離隔 17° 以下を捨てた場合に失っていた領域の 50% 以上を)再利用できることになった。

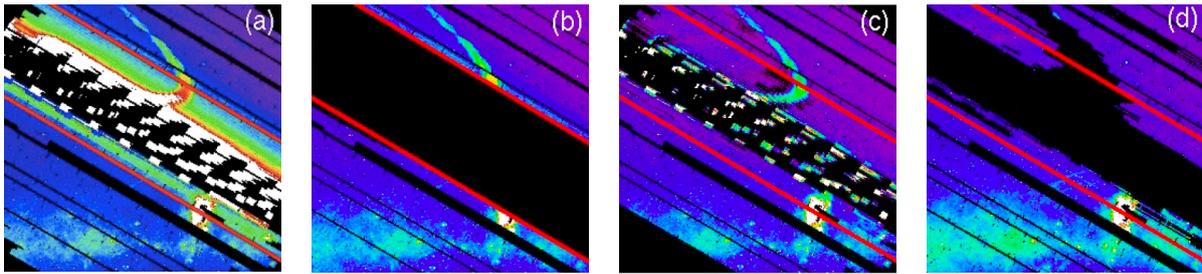


図 6 月の迷光の影響とその補正の様子. 例として, $(l, b)=(354.5^\circ, 2.5^\circ)$ 付近の 8° 四方の領域のマップを示す. (a) 月が近くにいる 1 季節目の観測機会で, 月の迷光を受けたマップを示す. 多数回のスキャンの間に観測視野と月との相対位置が変わっていくので, 月の迷光もまたクロススキャン方向に少しずつずれていく様子を表している. 図の中央上部分に見られる釣り針状の構造は, トラス起因の迷光である. 2本の赤線に囲まれた範囲が月からの離隔 17° 以内の領域に相当する. 左下の赤線付近を観測している時は, 月本体は視野の右上の離隔 17° 付近に存在している. 月の影響を一番受けている真ん中付近を観測している頃には, 月も視野付近に存在している. 右上の赤線付近を観測している頃には, 月本体は視野の左下の離隔 17° 付近に存在している. 中心部で白黒の斑模様になっているのは, 飽和ぎりぎりの高いレベルと, 飽和したデータ(マスクされてゼロになっている)が入れ子になっているため. (b) 月離隔 17° 以下のデータを除外して作成したマップ. (c) この章で述べる方法で迷光成分を差し引いたマップ. (d) この章で述べる方法で迷光成分を差し引いた上, 差し引き誤差の大きい部分をマスクしたマップを示す. このマスクされた部分は, 半年後もしくは1年後に正常な観測ができれば, 埋めることができる.

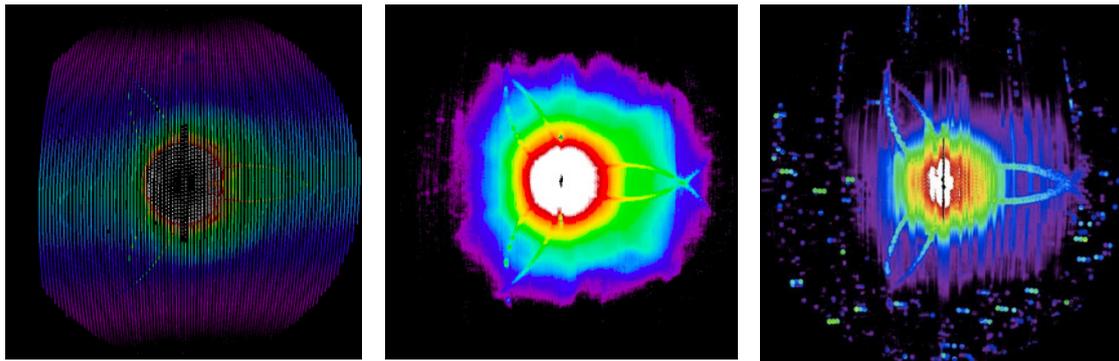


図 7 月を中心とした検出器座標で描いた迷光強度マップ ($9 \mu m$ 帯). 縦軸がスキャン方向, 横軸がクロススキャン方向に相当する. サイズは $80^\circ \times 80^\circ$. (左)2006/6/1 ~ 6/8 のデータから構成したデータ. (中)左図に対し, データが無い領域は内挿し, 黄道光を差し引いたものを, 各月入射の前後で作成し, それらの median 画像を作成したもの(月の迷光のテンプレート). (右)中図に対する σ 画像. 強度が強い部分は, 各月入射の度に再現性が低いことを示す.

2.4 明るい天体の迷光の除去

観測視野からある方向に離隔 1° くらいのところに明るい天体があると, 視野内に強い迷光が現れる. 2.3 章の月の迷光とは, 強度・形状ともに異なる現象である. マップ上でこの迷光は, 明るい天体を中心とし楕円弧形状に現れる(図 8a,8b). 幾つかの事象の解析結果から, 迷光の表面強度は迷光源の数 % 程度であり, 円弧の半径はだいたい 1° であることが分かった. しかし, 多くのケースで迷光源が飽和していたり構造をもっていたりするため, 迷光パターンを迷光源の強度の関数で定量的に表現するのは難しい.

そこで個々のケースに対して, データの性質に基づいて個別に対応することにした. 2.1 章で述べたように, 「あかり」は半年後に同じ領域を反対向きでスキャンする. 迷光の円弧形状がスキャン方向に対して非対称であるため, 半年前と後では天球座標に対する迷光パターンの形状が異なる(図 8a,8b). このため, 同じ領域の半年前と後での観測データの差分をとると, リアルな成分は消え, 迷光成分だけが残る(図 8c). この迷光成分を, それぞれのマップからマスクすることにした(図 8d). 迷光成分を同定する際, 図 8c において, 単純にあるレベル以上の値を示す部分を迷光と定義してしまうと, ノイズも拾ってしまうので, smooth をかけて空間周波数の高いノイズは落として, 迷光形状を同定するようにした. ただし, 2つの円弧の交点では両方のデータに迷光が載っているため, 両方のデータに対してマスクが必要で

あり、結果的にデータが残らない。マップ利用の際には、明るい天体のスキャン方向上下約 1° 付近にある2点の黒い gap に注意する必要がある。このような処理がなされた明るい天体は、全天で計20天体存在する。

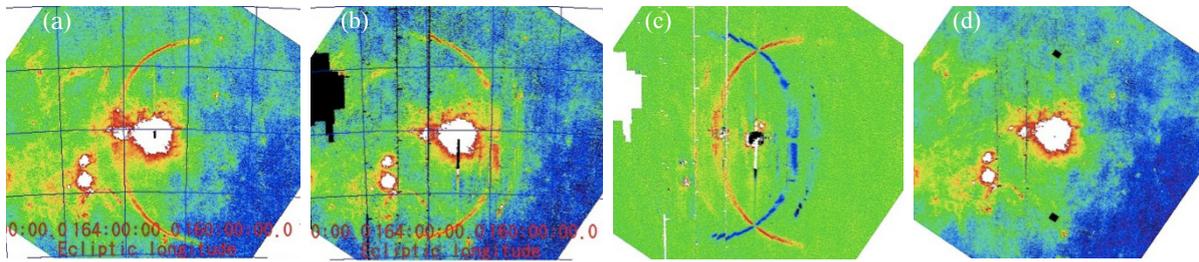


図8 明るい天体周囲に現れる迷光成分とその除去方法。RCW38 周囲 3° 四方のマップの例。(a) 季節1の観測で得られた画像。(b) 季節2の観測で得られた画像。矢印は衛星のスキャン方向を示す。(c) (a) - (b) の差分マップ。リアルな構造は両方のマップで等しく移っているため差分画像上には残らず、(a) のマップ上にだけ乗っている迷光パターン(左側の円弧)がプラス成分として、(b) のマップ上にだけ乗っている迷光パターン(右側の円弧)がマイナス成分として現れている。(d) この章で述べる方法により、迷光成分を除去した画像。

2.5 黄道光の季節間差の補正

「あかり」は、ある天域を観測した半年後に、地球の公転軌道上の反対側の位置から同じ天域を観測する。しかし、衛星の惑星間空間の中での位置は、半年前と後で異なるので、前景放射成分(オフセット)としての黄道光(惑星空間塵放射)のレベルは異なる。一番効いているのは、earth trailing dust cloud 成分¹²⁾と考えられる。このオフセット成分の空間周波数は低い(～数十分角)ので、数 \times 数 $^\circ$ 程度のマップ上では、季節間差を2次曲面でfittingし、片方の季節のマップから差し引くことでレベルを揃えることができる(図9)。現在は、最初に観測した季節(2006/5/8-2006/11/8)でのオフセットレベルに揃えている。

なお、黄道光の差し引きは行っていない。現在知られている黄道光モデル¹²⁾は比較的単純であるため、とくに黄道面周辺のダストバンド付近で、実際のデータと表面輝度のずれが大きい。とくに中間赤外線長の波長では黄道光放射のピークにあたるため、この不定性の影響を他波長よりも大きく受けてしまう。IRAS の配布マップ(ISSA)でも、 $12\ \mu\text{m}$ および $25\ \mu\text{m}$ では、黄道光差し引きにより混入したartifactsが強く残っている。このため現時点では、黄道光をモデルで差し引かないデータをプロダクトとして保持する方針にしている。

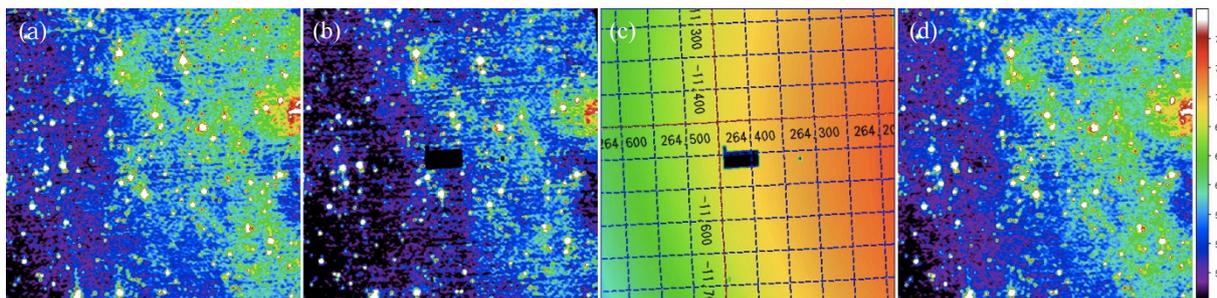


図9 背景光(黄道光)の季節間差の補正例。(a) ある領域($l, b=(353^\circ, -0.7^\circ)$)付近の 0.4° 四方)の季節1の画像。(b) 同じ領域の季節2(季節1から半年後の観測機会)の画像。(c) (a) - (b)。空間周波数の低いオフセット成分が乗っている。グリッドは黄道座標を示す。(d) オフセットを合わせてcombineした画像。(a),(b),(d)のスケールは統一している。(c)は、わずかなオフセット成分を強調するため、(a),(b),(d)に比べ20倍スケールを拡大している。

2.6 面輝度較正

全天サーベイ観測で取得した画像のピクセル値(ADU pix^{-1})を物理単位(MJy sr^{-1})に変換する、キャリブレーションのための換算係数を求めた。指向観測で、同じ装置(IRC)の同じフィルタで取得した画像は、標準星の観測結果を元にキャリブレーションすることができる¹³⁾。そこで、指向観測モード(584 ms 露光)で観測・キャリブレーションしたdiffuse

天体 (反射星雲 NGC7023) を、全天サーベイ観測モード (44 ms 露光) で観測し、この換算係数を求めた (図 10)。また、これと並行して、MSX 衛星や Spitzer 衛星で確立されつつある、赤外線標準面光源¹⁴⁾ によるキャリブレーションも試し、クロスチェックを行った。

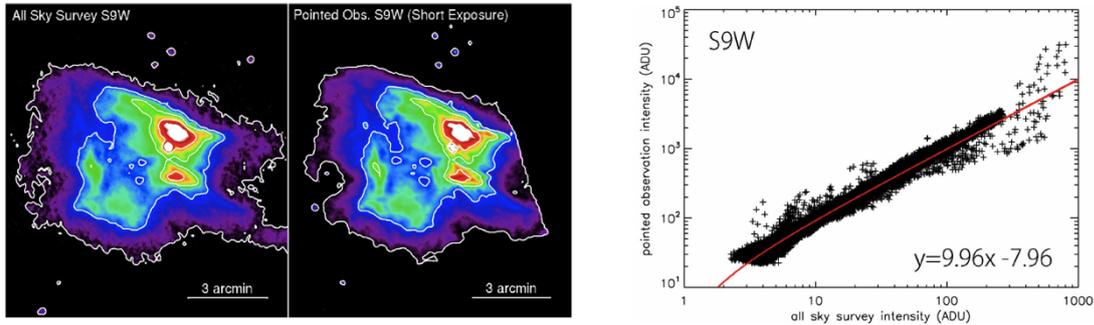


図 10 (左) NGC7023 の全天サーベイモードでの取得画像 (44 ms 露光)。(中) 同じ天体を撮像モードで取得した画像 (600 ms 露光)。図右端は視野外のためデータが切れている。(右) 撮像モードでの出力値 vs. 全天サーベイモードでの出力値。フィッティングで求めた傾きと、撮像モードでのキャリブレーション結果から、全天サーベイモードでの換算係数を決めた。

3. 結果

図 11 に、2 章で述べた様々な補正を施し、全観測機会のデータを使って作成した $9 \mu\text{m}$ 帯マップの例を示す。様々な artifacts の影響を、典型的な銀河面でのマップのオリジナルのピクセルスケール ($4.68''$) でのショットノイズレベル ($\sim 1 \text{ MJy/sr}$) 以下に抑えることができた。このデータは、C-SODA を通じて全世界への公開を計画している。

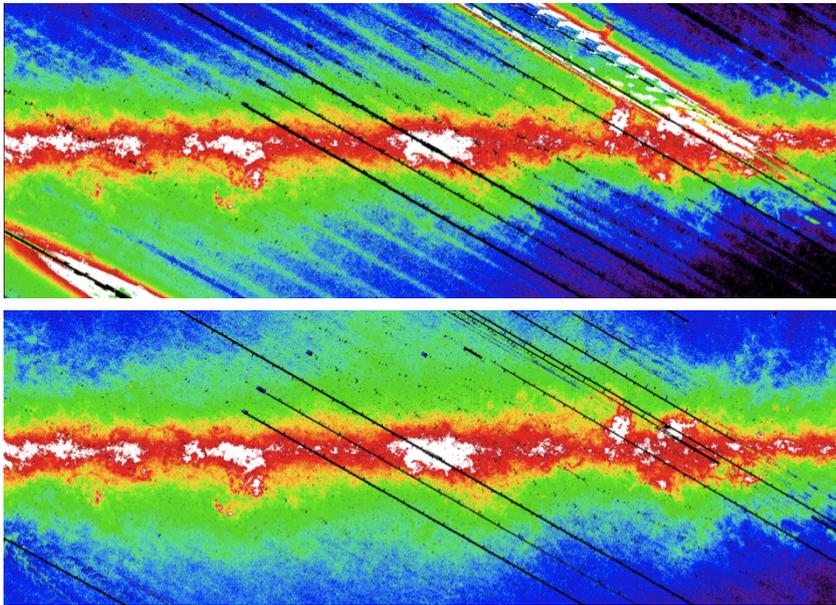


図 11 銀河中心付近 $40^\circ \times 20^\circ$ の $9 \mu\text{m}$ 帯マップの例。(上) 2 章で述べた様々な補正を行う前。とくに目立つのは、SAA 中の放射線照射に起因する縞々構造と、マップ右上と左下の月の迷光の影響。(下) 補正を施した後。画面上に残る黒い筋は観測データが無い領域であり、データ処理上の問題では無い。

4. 議論

4.1 マップ作成プロセスの改善により可能になったサイエンス例

2 章で述べた様々な補正により、マップの精度や信頼性を上げることで、可能になるサイエンスがある。これまでの報告例の中から幾つかを紹介する。

磁気浮上ループの物質の物理状態の解明⁹⁾

太陽表面のコロナのように、銀河磁場のパーカー不安定性によりループ状に物質が持ち上げられた構造が、「なんてん」望遠鏡での電波観測により銀河面上で見つかった(磁気浮上ループ)¹⁵⁾。持ち上げられた物質が落下してきた根元では、激しい衝突が推定される。2章で述べた様々な補正によって精度が改善された「あかり」中間赤外線マップを用いて、この磁気浮上ループの周辺の放射構造を解析した。その結果、まずこのループの一部に沿った赤外線放射構造が明らかになった(図12左)。また、星間空間において、PAH放射と遠赤外線のcold dustの放射強度は、(個数密度比がほぼ一定であることと、必要な放射場硬度が同程度であるため)一般に相関が良いと言われているが¹⁶⁾¹⁷⁾¹⁸⁾、「あかり」9 μm 帯とIRAS 100 μm 帯の強度の比のマップから、一般的な星間空間のダストに対するPAH放射の割合が、ループの根元で小さくなっていることが分かった(図12右)。(なお、「あかり」の遠赤外線画像はこの時点で較正が終わっていなかったため、IRAS 100 μm をcold dust放射として利用した)。「あかり」2~5 μm 帯の分光観測でも、ループ根元付近では、PAH放射が一般的な星間空間よりも弱く、そのかわり炭素系ダストの破壊により生じる鎖状炭化水素の放射が増えていることが分かり、小さな粒子であるPAHが選択的に破壊されている状況が明らかになった。これらの解析により、分子雲観測とは独立に、星間空間のダストの物理の視点から、磁気浮上ループの存在とループ根元でのショック状態を、検証することができた。

Type Ia 超新星残骸でのダスト生成の発見⁸⁾

Tychoの超新星残骸に於いて、「あかり」18 μm 帯でトレースする暖かい(~100 K)星間ダストの分布と、電波観測で調べた分子雲の分布と、X線観測で調べた放出物質の分布・衝撃波面の位置・不連続面(放出物質の先端)の位置を比較することで、超新星の衝撃波との相互作用で加熱された星間物質と、放出物質の中で新しく生成されたダストを区別することに成功した(図13)。Type Iaの超新星残骸でダスト生成の兆候が見つかったのは世界で初めてであり、宇宙の化学進化の研究に大きなインパクトをもたらした。

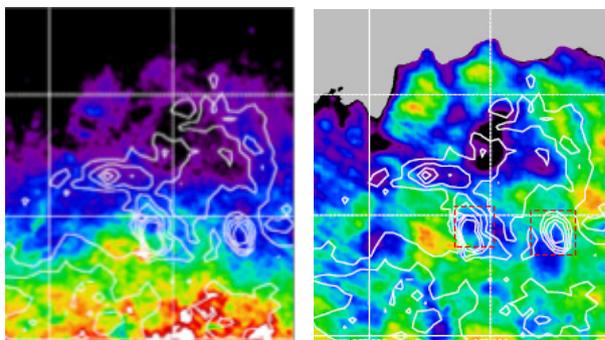


図12 (左) $(l, b) = (-3.8^\circ, +1.5^\circ)$ 付近の $2.5^\circ \times 2.7^\circ$ の「あかり」9 μm 帯画像。「なんてん」電波観測により見つかった、磁気浮上ループの一つ、分子雲ループ215を白のcontourで示す。(右)同領域の「あかり」9 μm 帯 / IRAS 100 μm 帯放射の比の画像。赤点線で囲った部分が、ループの根元に相当する。

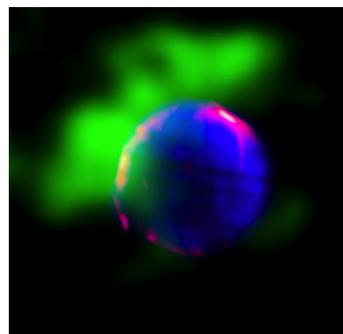


図13 Tychoの超新星残骸、 $(l, b) = (120.1^\circ, 1.415^\circ)$ を中心とした $20' \times 20'$ の3色合成画像。青がSuzaku X線(0.2-12 keV)、緑が電波観測による一酸化炭素輝線の分布、赤がAKARI 18 μm 帯放射を表す。

4.2 今後の課題

これまでに、様々なartifactsの補正によりマップの精度と信頼性が向上し、新しいサイエンスが可能になった。今後も様々な補正を追加することで、さらに淡い構造や大きな空間スケールの構造の議論が可能になるはずである。「あかり」中間赤外線全天サーベイデータ、とくに波長9 μm 帯のデータは、天文学的にも貴重なリソースであり、労力を費やして精度良いマップを作る価値は高い。今後3年のタイムスケールで取り組むべき課題としては、以下の項目が挙げられる。

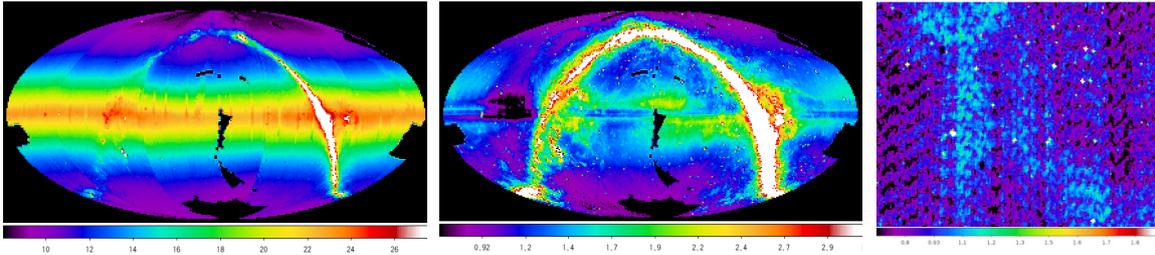


図 14 (左)黄道光込みの「あかり」 $9 \mu\text{m}$ 全天マップ (中)モデル 12)に基づいて黄道光を差し引いた「あかり」 $9 \mu\text{m}$ 全天マップ。差し引き残渣が見える。(右)中図の $(l, b) = (90^\circ, -38.5^\circ)$ 付近を拡大し、 $3'$ のガウシアンカーネルで smooth した図。1 MJy/sr 以下のレベルではあるが、linearity と reset anomaly 補正及び、フラット補正とカメラ内散乱光補正の不十分が原因と思われる縞模様が見られる。左図及び中図で共通して見える黒い部分は、全天のマスプロセスの際に生じたバグによる物であり、将来的には埋まる予定。

黄道光の定量評価と除去

波長 $9 \mu\text{m}$ 帯および $18 \mu\text{m}$ 帯では、天域によっては、黄道光が銀河面放射よりも強い (図 14 左)。黄道光の分布は空間周波数が低いので、小さな (< 数度) 放射構造の議論では、「スカイ」を差し引けば問題ないが、銀河面の大きな (数度～数百度) 空間スケールの構造を議論する時には、この黄道光の差し引き精度が問題になる。この黄道光成分を差し引き、太陽系外からの赤外線放射 (銀河面からの放射、銀河系外からの放射) のマップを作成する。IRAS マップ (IRSA) に於いても、波長 $12 \mu\text{m}$ 帯および $25 \mu\text{m}$ 帯のマップでは、黄道光の差し引き誤差が大きく、「あかり」のデータに含まれる黄道光もこれまでの標準的なモデル¹²⁾では説明できないことが分かっており¹⁹⁾、難しい課題である (図 14 中)。そのため、既存のモデルに依存するのではなく、全天の各画素での SED フィッティングを行い、地球軌道付近のダストの熱放射 (~270K) 相当の成分を分離するなどの、独自のアプローチを試している。

地球の高層大気の洩れ光の定量評価と除去

「あかり」の軌道面は、地球公転面の法線に対して夏至の頃に最大 35° 傾く。この前後に、極付近で地球の高層大気の放射の洩れ込みが原因と思われる迷光が現れる。さらに、望遠鏡の遮光面は太陽方向に対して高くなっているため、この時期、南極よりも北極で地球に対する遮蔽が低くなり、迷光も強く出る (図 15)。この影響を正確に評価し、差し引いたマップを作成する。

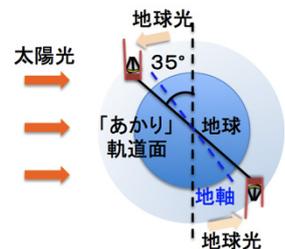


図 15 夏至の頃の「あかり」軌道面を横からみた模式図。

明るい天体を見た後の感度劣化の定量評価と補正

月を見た後 100 分以上のタイムスケールで、検出器の感度の低下が見られる。この影響を定量的に評価し、マップを補正するとともに、その現象を物理的に理解し、次の衛星の検出器試験計画に情報を反映する。

近軌道デブリが視野を通過したデータ等の除外

機械的な処理で除外できない異常なデータの除去を、一つ一つ手作業で行う。例えば、全天サーベイ中に時々、数秒間程度の間、視野全体が明るくなる現象があるが、この原因は、近軌道のデブリが視野を横切ったためと考えられている。

Linearity と reset anomaly の補正式の改善

電荷蓄積型の検出器を使い、非破壊読み出しにより観測を行っているので、光電荷が蓄積すると検出器にかかる実効バイアスが低下し、観測中にゲインが低下してしまう (linearity 劣化)⁴⁾。また、蓄積電荷をクリアするために定期的リセットをかけているが、このリセット時に流れる電流によって極低温部の読み出し回路が一瞬温度上昇し、その後ゆっくり冷えてゆくため、出力のオフセット変動を招く (reset anomaly)⁴⁾。Linearity 劣化や reset anomaly によって、出力信号は最大 8 割変化する (それ以上補正が必要なデータはマスクしている)。この 2 つの現象は、「検出器が一定の光量を受けていても、リセットから時間が経過すると出力が低下してくる」という同じ外見を示すので、分離して評価するのが難しい。点源カタログ作成時は、この影響を、実験室で定常光を当てた時のデータを元に、電荷蓄積開始直後は linearity 劣化をゼロと見做す、全ピクセルの素性が同じと見做すなど、様々な仮定を入れて評価していたが、とくに点源フラックス評価時には影響が少ないがマップ作成の際には影響の大きい、reset anomaly の評価が不十分であった (図 14 右)。全天サーベイ時の多くの時間帯は、星などの天体が見えず、空全体に広がる黄道光成分を見ている。これらを材料にすれば、様々

な空の強度・様々な電荷蓄積量に対する応答を調べるための大量のサンプルが得られるので、1ピクセルずつ linearity と reset anomaly を同時に考慮して fitting を行い、より現実的な補正関数を作成することが可能である。

フラットとカメラ内散乱光補正式の改善

背景光を光源とし、カメラ内で生ずる散乱光が、検出器上の全ピクセルにまたがって淡く広がっていることが分かっている。検出器全体に散乱光(ピクセル毎に違うオフセット成分)が乗っている状態では、「一様な空(黄道面や銀河面から離れた、一見何も写っていないような天域の多数枚の画像のモードを取るなどして作成したスカイフラット)」の観測から flat(ゲインのピクセル間ばらつき)を求めることができない。そのためこれまでは、同じ星を様々なピクセルで観測することで、flatの滑らかな成分(空間周波数の低い成分)だけを評価していた。しかし、「一様な空」の観測から高空間周波成分も評価し、より精度の高い flat を作成したい。また、点源のカメラ内散乱光については多数の具体的なサンプルが得られているため、点源の強度と視野中の位置の関数で、カメラ内散乱光の強度・パターンを表現し、convolve することで、カメラ内に一様に入ってきた光や広がった光源に対する散乱光も精度良く記述できると期待される。

5. まとめ

「あかり」の波長 $9 \mu\text{m}$ 帯と $18 \mu\text{m}$ 帯での中間赤外線全天マップは、感度と空間分解能で過去の IRAS サーベイマップを上回り、とくに波長 $9 \mu\text{m}$ 帯は PAH のトレーサーとしての恒久的価値も持つ。我々は、点源カタログを作成する際に作成したマップに対し、幾つかの新しい処理を追加し、マップの精度と信頼性を向上させた。マップの artifacts のレベルは、オリジナルピクセルスケールでのノイズレベル以下に抑えることができた。新しく追加した処理は、SAA での放射線ヒットの間接的影響の補正、月の迷光の補正、明るい天体の迷光の補正、黄道光の季節間差の補正など、他の赤外線衛星プロジェクトでは稀に見る地道な解析を含む。改善したマップにより、淡く広がった放射成分の定量評価に基づく議論が可能になり、銀河面の磁気浮上ループの物理状態の解明や、超新星残骸中で熱放射が見えているダストの起源の解明など、様々なサイエンスに繋がっている。マップの改善を通じて、装置の特性が明らかになり、次の赤外線天文衛星の開発に反映できる新しい技術情報も得られた。今後もデータ解析方法の開発を継続することにより、さらに淡い放射成分や広がった放射成分の議論を可能にし、「あかり」全天サーベイデータでしかできない新しいサイエンスに繋がりたい。

謝辞

This research is based on observations with AKARI, a JAXA project with the participation of ESA. This work was supported by a Grant-in-Aid for Scientific Research (24740122) from Japan Society for the Promotion of Science (JSPS), and the Nagoya University Global Center of Excellence Program, “Quest for Fundamental Principles in the Universe” (QFPU) from the Japan Society for the Promotion of Science and the Ministry of Education, Culture, Sports, Science and Technology of Japan. We thank all the AKARI/IRC team members, for beneficial suggestions on the processing of the in-flight data. We also express many thanks to the anonymous referee for a careful reading and constructive comments.

参考文献

- 1) Murakami, H., et al. : The Infrared Astronomical Mission AKARI, Publications of the Astronomical Society of Japan, Vol.59, (2007), pp.369-376
- 2) Onaka, T., Matsuhara, H., Wada, T., et al. : The Infrared Camera (IRC) for AKARI - Design and Imaging Performance, Publications of the Astronomical Society of Japan, Vol.59, (2007), pp.401-410
- 3) Ishihara, D., Wada, T., Onaka, T., et al. : Mid-Infrared All-Sky Survey with the Infrared Camera (IRC) on Board the ASTRO-F Satellite, The Publications of the Astronomical Society of the Pacific, (2006), Vol. 118, pp.324-343
- 4) Ishihara, D., Onaka, T., Katata, H., et al. : The AKARI/IRC mid-infrared all-sky survey, Astronomy & Astrophysics, Vol. 514, (2010), id. 1

- 5) Ita, Y., Matsuura, M., Ishihara, D., et al. : AKARI's infrared view on nearby stars. Using AKARI infrared camera all-sky survey, 2MASS, and Hipparcos catalogs, *Astronomy & Astrophysics*, Vol. 514, (2010), id. A2
- 6) Ishihara, D., Kaneda, H., Onaka, T., et al. : Galactic distributions of carbon- and oxygen-rich AGB stars revealed by the AKARI mid-infrared all-sky survey, *Astronomy & Astrophysics*, Vol. 534, (2011), id. 79
- 7) Oyabu, S., Ishihara, D., Malkan, M., et al. : AKARI detections of hot dust in luminous infrared galaxies. Search for dusty active galactic nuclei, *Astronomy & Astrophysics*, Vol. 529, (2011), id. A122
- 8) Ishihara, D., Kaneda, H., Furuzawa, A., et al. : Origin of the dust emission from Tycho's SNR, *Astronomy & Astrophysics*, Vol. 521, (2010), id. L61
- 9) Kaneda, H., Ishihara, D., Mouri, A., et al. : Processing of Polycyclic Aromatic Hydrocarbons in Molecular-Loop Regions near the Galactic Center Revealed by AKARI, *Publications of the Astronomical Society of Japan*, Vol.64, No.25
- 10) Adriani, O., Barbarino, G. C., Bazilevskaya, G. A., et al. : The Discovery of Geomagnetically Trapped Cosmic-Ray Antiprotons, *The Astrophysical Journal Letters*, Vol. 737, (2011), id. L29
- 11) Mouri, A., Kaneda, H., Ishihara, D., et al. : Effects of High-Energy Ionizing Particles on the Si:As Mid-Infrared Detector Array on Board the AKARI Satellite, *The Astronomical Society of the Pacific*, Vol. 123, (2011), pp.561-567
- 12) Kelsall, T., Weiland, J. L., Franz, B. A. et al. : The COBE Diffuse Infrared Background Experiment Search for the Cosmic Infrared Background. II. Model of the Interplanetary Dust Cloud, *The Astrophysical Journal*, Vol. 508, (1998), pp.44-73
- 13) Tanabé, T., Sakon, I., Cohen, M., et al. : Absolute Photometric Calibration of the Infrared Camera (IRC) aboard AKARI, *Publications of the Astronomical Society of Japan*, Vol.60, (2008), pp.375-388
- 14) Cohen, M., Green, A. J., Meade, M. R., et al. : Absolute diffuse calibration of IRAC through mid-infrared and radio study of HII regions, *Monthly Notices of the Royal Astronomical Society*, Vol. 374, (2007), pp.979-998
- 15) Fukui, Y., Yamamoto, H., Fujishita, M., et al. : Molecular Loops in the Galactic Center: Evidence for Magnetic Flotation, *Science*, Vol. 314, (2006), pp.106-109
- 16) Onaka, T., Yamamura, I., Tanabe, T., Roellig, T. L., & Yuen, L. : Detection of the Mid-Infrared Unidentified Bands in the Diffuse Galactic Emission by IRTS, *Publications of the Astronomical Society of Japan*, Vol. 48, (1996), pp. L59-L63
- 17) Draine, B. T., Dale, D. A., Bendo, G., et al. : Dust Masses, PAH Abundances, and Starlight Intensities in the SINGS Galaxy Sample, *The Astrophysical Journal*, Vol. 663, (2007), pp.866-894
- 18) Bendo, G. J., Draine, B. T., Engelbracht, C. W., et al. : The relations among 8, 24 and 160 μ m dust emission within nearby spiral galaxies *Monthly Notices of the Royal Astronomical Society*, Vol. 389, (2008), pp.629-650
- 19) Pyo, J., Ueno, M., Kwon, S. M., et al. : Brightness map of the zodiacal emission from the AKARI IRC All-Sky Survey, *Astronomy & Astrophysics*, Vol.523, (2010), id. A53

トロムソ/ロングイアビンにおける オーロラ観測用並列イメージャーの大量データ処理と可視化

小川泰信^{*1} 門倉昭^{*1} 元場哲郎^{*1,2} 田中良昌^{*1} 細川敬祐^{*3}

Processing and visualization of large amounts of auroral data obtained with All-sky/Narrow field-of-view parallel imagers in Tromsø and Longyearbyen

Y. OGAWA^{*1} A. KADOKURA^{*1} T. MOTOKI^{*1,2} Y. TANAKA^{*1} and K. HOSOKAWA^{*3}

Abstract

We report current status and future plans of processing, analysis methods, and visualization of large amounts of auroral data obtained with All-sky/Narrow field-of-view parallel imagers in Tromsø Norway and Longyearbyen Svalbard. Observations of the parallel imagers started in October, 2010, and the optical data provide important information for various studies on auroral physics, as well as other in-situ and ground-based data such as European incoherent scatter (EISCAT) radars. We have developed several methods of data processing and visualization of the auroral database in order to find valuable auroral data among the abundant data. This experience can be applied to development of a global network aurora observation system using automated imagers with high sensitivity in the future.

Keywords: Polar ionosphere, aurora, imager, EISCAT, IUGONET

概要

本論文では、トロムソ及びロングイアビンで国立極地研究所が展開している全天/狭視野並列イメージャー観測による、高時間分解能のオーロラカラー画像の大量データベース化や、データ解析手法、データ可視化手法を中心に現状と今後の予定を紹介する。このデータベースは、様々なオーロラの研究を進める上で重要なオーロラの基本情報を含んでおり、欧州非干渉散乱 (EISCAT) レーダー等の関連観測データを組み合わせた総合的な研究を全国共同利用で実施する上でも不可欠であると考えている。得られた大量のオーロラ画像データから、オーロラ研究に利用可能な (ごく一部の) 質の高いオーロラ画像データを抽出する方法や、その抽出された画像データを更に詳細かつ総合的に解析する手法の開発・発展が求められており、我々は様々な画像処理プログラムや複数のデータを同時に可視化するプログラムの開発に取り組んでいる。このイメージャー観測システム構築と可視化を含むデータ処理の経験を基に、将来的には多地点のオーロラ画像データを統合して扱うことが可能なシステム開発や、省電力・無人オーロラ観測システムの多点展開等への発展が期待される。

*1 国立極地研究所 (National Institute of Polar Research, Japan)

*2 Now at APL/JHU (The Johns Hopkins University Applied Physics Laboratory, USA)

*3 電気通信大学 (The University of Electro-Communications, Japan)

1. 背景

極域におけるオーロラ研究は長い歴史を持っているが、近年では、高時間・高空間分解能の人工衛星観測や地上レーダー観測などと組み合わせた総合的なデータ解析が行われることが多く、従来に比べてより時間分解能の高い（数秒以内程度の）データを連続的に取得することの重要性が増してきている。また、解析に使用出来るオーロラ観測データは晴天夜のものに限られるため、そうした連続的に取得された大量のデータから晴天夜のデータをまず抽出し、さらに解析対象とするオーロラ活動が観測されたデータを抽出する、という作業も必要となってきた。

微弱なオーロラ発光の高速サンプリング観測には、高価かつ慎重な扱いが必要な機材（ICCDカメラなど）が従来用いられてきた。そのような機器を用いた主にキャンペーンベースの観測に加え、最近普及が進んでいる簡易型の超高感度CCDカメラを用いた定常観測が、例えば南極昭和基地やアイスランドのオーロラ観測拠点で2009年から開始されている。この簡易型カメラ観測の利点としては、月が多少出ている場合でも CCD の劣化をあまり気にせずにオーロラ観測が可能であるため、定常観測に適していることが挙げられる。

このような先行観測の経験を基に、ノルウェー・トロムソの EISCAT サイト敷地内にある分反射実験（PRE）用ビルディング別棟の光学観測施設（図1参照。地理緯度：北緯 69.6°、東経 19.2°、不変磁気緯度：66.2°）と、スバルバル・ロングイアビンの Kjell Henriksen Observatory 光学観測所（図1参照。地理緯度：北緯 78.2°、東経 16.0°、不変磁気緯度：75.2°）の2地点において、全天及び狭視野カメラによる日陰期間（10-3月）のオーロラ定常観測を2010年10月より本格的に開始した。その後様々な試行錯誤を経て、現在は複数の簡易型高速イメージャーの組み合わせを用いたオーロラの高速度カラー画像データの定常取得を実現している。オーロラの定常観測は極域の様々な地点で行われているが、カラー画像データを1秒以内の時間分解能で取得する全天オーロラ定常観測は他に類を見ない。オーロラ画像を白黒画像ではなく、多波長のカラー画像として取得するメリットについては、(1) 画像内の明るい領域がオーロラか雲かを区別しやすいこと、(2) オーロラの色の違いから様々な情報（発光高度や降下粒子エネルギー等）を引き出し得ること、(3) オーロラ以外の特定の光（ナトリウム光など）を光学フィルターで取り除くことができること、等が挙げられる。

高時間分解能によるオーロラ定常観測では、膨大な量のデータが蓄積されるが、最近では高い演算処理能力をもつ CPU や速いデータ転送速度を持つネットワーク技術が普及し、また大量のデータを扱うことが可能なストレージシステムも容易に構築可能となってきた。また、自動運用や遠隔運用を可能にする技術も進んできている。本論文で扱っているオーロラ観測用並列イメージャーも、そのようなデータ転送及びストレージ容量に関する最近の劇的な発展の上に成り立っている。1つのイメージャーで得られる1画像当たりのファイルサイズは約 80 kBytes とそれほど大きくないが、高いサンプリングレートで複数のカメラを同時に運用しているため、全体のデータ量は莫大となり、2010年10月から2012年3月までの全データ容量は、約 22 TBytes にも及んでいる（その内訳は、観測日数（約 250 日）×1日当たりの平均観測時間（約 40,000 秒）×サンプリング頻度（約 2 枚/秒）×画像サイズ（70-140 kBytes）×イメージャーの数（5 台/8 台/12 台）の計約 16 TBytes に加え、3色合成処理後のデータなど約 6 TBytes）。これらの膨大な画像データを、研究目的に合わせて効率よく処理し可視化するシステムが必要とされる。本論文では、全天/狭視野並列イメージャーシステムとそのデータ取得の概要を説明すると共に、そのデータ処理や可視化の現状と、今後の予定や方向性を紹介する。



図1 (左図) トロムソ大学が運用する分反射実験 (PRE) 用ビルディング別棟の光学観測施設。(右図) ロングイアビン・スバルバル大学センター (UNIS) が運用する光学観測所 Kjell Henriksen Observatory (KHO)

2. 全天/狭視野並列イメージャーシステム及びデータ取得について

全天/狭視野並列イメージャーシステムは、同じ種類の複数のカメラ及びレンズ、透過波長の異なるフィルター、カメラからのビデオ信号を画像ファイルに変換するビデオエンコーダ、ネットワークを介してデータを保存する PC 及びハードディスクで構成される。全天観測と狭視野観測では、用いているレンズ(魚眼レンズもしくは狭視野レンズ)と光学フィルターを変えており(図2参照)、それに応じてカメラの観測設定を調整している。このオーロラ観測用並列イメージャー観測に用いている機材と仕様を、表1にまとめて記載する。カメラについては、超高感度 CCD を用いたフレーム積分が可能な WAT-120N+ (ワテック社製) を複数台用いている。レンズは光学的に明るく CS マウントが可能な製品を利用し、光学フィルターは(オーロラ入射光の角度を考慮して)半値幅が約 100 nm と広く、かつ透過率の高い(約 90%)製品を加工して用いている。並列イメージャーシステムで用いている光学フィルターの波長透過特性を図3に示す。これらの光学系により、微弱なオーロラ発光を1秒以内の高時間分解能で取得することに成功している。

これまでの観測状況の概要を表2にまとめて記載した。2010年3月のトロムソでの試験観測を経て、2010年10月から2011年3月の観測シーズンに定常観測をトロムソ及びロングイアビンで開始した。その後、本格的な並列イメージャーシステムの構築を2011年4月から開始し、2011年10月から2012年3月の観測シーズンに現地におけるシステム運用を行った。これまでに得られたデータのデータベース化や解析処理を進めると共に、2012年9-10月からはトロムソ及びロングイアビンにおいて統一された運用システムでの同時観測を実施する。

観測地点における観測システムの概要を図4の左側(a)に示す。その特徴として、(1)データ量が莫大となるため、データ欠損が発生しないように転送速度を最大限考慮したシステムを構築していること、(2)PCやハードディスク等の記録装置に不具合が生じた場合のバックアップ体制を用意していること、の2点が主に挙げられる。(1)に関して、ビデオエンコーダには1Gbpsでのネットワーク転送が可能な AXIS Q7404 を使用し、PC 転送されたデータは、SATA 接続された内蔵ハードディスク或いは USB3.0 接続の外付けハードディスクにカメラ毎に別々に分けて記録される。(2)に関して、並列イメージャーシステム毎に2台の PC を用意しており、もし1台の PC に不具合が生じた場合には、ビデオエンコーダがその状況を察知し、もう1台の PC に画像データを記録するように自動で切り替える設定となっている。

これらの観測データを用いた研究対象としては、(1)数秒で大きく変化するオーロラ(例:周期が数秒程度の脈動オーロラ¹⁾やオーロラブレイクアップ前に生じるオーロラビーズ構造²⁾³⁾)、(2)ミクロスケール(数 km 以内)のオーロラの微細構造⁴⁾、(3)ミクロスケールとメソスケール(数 10 km - 数 100 km)のオーロラ構造の間のスケール間結合、(4)オーロラサブストームの発達過程⁵⁾、等が挙げられる。

トロムソとロングイアビンでは、EISCAT レーダー⁶⁾が稼働しており、オーロラ発光が生じている電離圏大気の温度や密度などの物理量を定量的に取得することが可能である。また、名古屋大学太陽地球環境研究所がトロムソに展開しているモノクロ全天イメージャーや中間圏-下部熱圏温度観測を行うナトリウムライダー⁷⁾などの複数の地上光学機器観測(参照ウェブページ: http://www.stelab.nagoya-u.ac.jp/~eiscat/prephoto_data.html)との相補的な同時観測もこれまでに実施してきている。

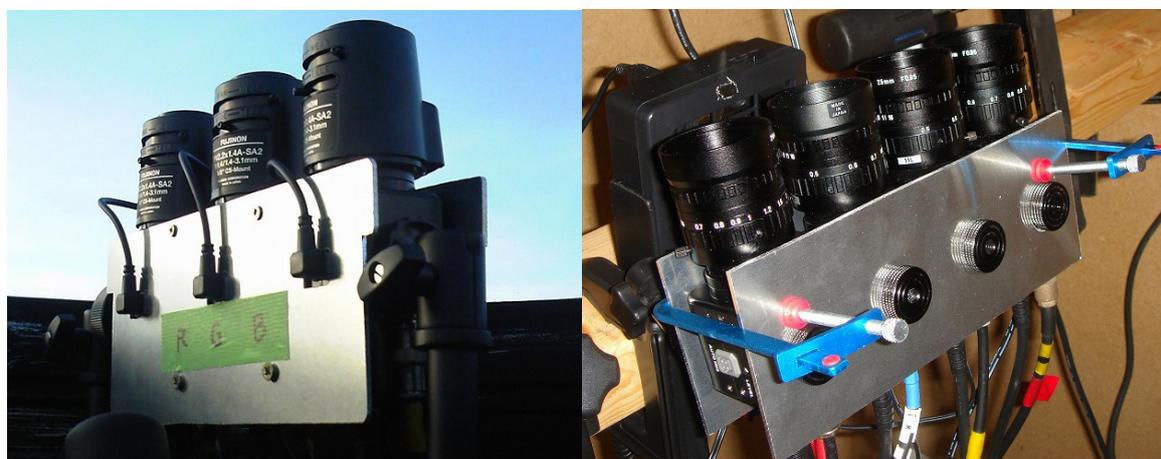


図2 (左図) 全天ワテック並列イメージャー(AWI)の写真。(右図) 狭視野ワテック並列イメージャー(NWI)の写真。NWIの背後には、イメージャー冷却用のファンが設置されている。

表 1 オーロラ観測用並列イメージャー観測に用いる機材とその仕様

| | |
|-------------|--|
| ビデオカメラ | ワテック社製 WAT-120N+ (CCD : Sony 社製 ICX418ALL) 最低被写体照度 0.00002 ルクス . フレーム積分可能 (最大 8.5 秒間) |
| 全天観測用レンズ | フジノン社製魚眼レンズ YV2.2x1.4A-SA2 (絞り DC 駆動) (トロムソ) , YV2.2X1.4A-2 (絞り手動駆動) (ロングイアビン) 焦点距離 : 1.4 mm, 絞り : F1.4, 画角 : 185°x 185° (水平 x 垂直) |
| 狭視野観測用レンズ | 八雲社製狭視野レンズ YMV2595N 焦点距離 : 25 mm, 絞り : F0.95, 画角 : 14.6°x 11.0° (水平 x 垂直) |
| 光学フィルター | バーダー社製 LRGB 撮影用フィルター (型番 2458475) |
| ビデオエンコーダ | アキシス社製 AXIS 241Q ビデオサーバー (2010 年 10 月 ~ 2011 年 3 月) アキシス社製 AXIS Q7404 ビデオエンコーダ (トロムソ : 2011 年 11 月 ~ , ロングイアビン : 2011 年 10 月 ~) 画像分解能 : 640 ピクセル x 480 ピクセル 保存ファイルフォーマット : jpeg 形式 (8 ビット) |
| ビデオタイムインサータ | Alexander Meier Elektronik 社製 TIM-10 (2010 年 10 月 ~ 2011 年 3 月) (注 : 2011 年 10 月からは NTP サーバー PC を用いた時刻同期を実施) |

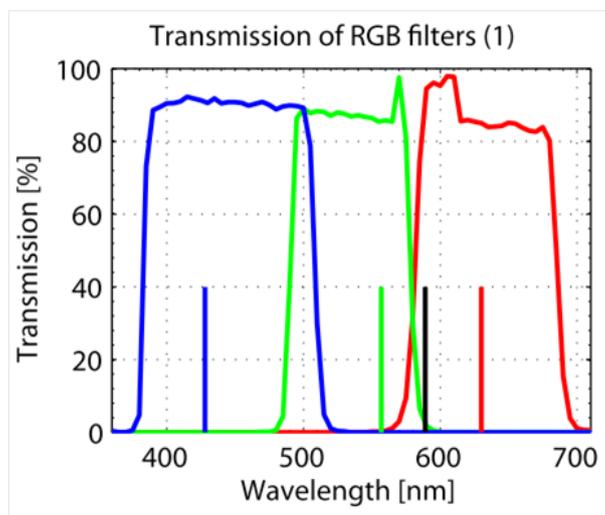


図 3 イメージャー観測に用いた光学フィルターの波長透過率特性 (国立極地研究所の光学校正機器を用いて測定) . 青色, 緑色, 赤色の直線で描かれた箇所は典型的なオーロラの発光波長 (それぞれ 428, 558, 630 nm) の位置を表す . 黒い直線で描かれた波長はナトリウム発光 (589 nm) を表す .

表 2 トロムソとロングイアビンにおける観測状況の概要

| | トロムソ AWI | トロムソ NWI | ロングイアビン AWI | ロングイアビン NWI |
|----------------------|--|--|--|--|
| 2010年3月 | テスト観測実施 | — | — | — |
| 2010年10月～ 2011年3月 | BW (2 Hz) [11月～] | BW (2 Hz)[10月] G (2 Hz)[11月～] | BW (2 Hz) [1月～] | G&R (2 Hz) [1月～] |
| 2011年10月～ 2012年3月 | RGB (2 or 1 Hz) & BW (2 Hz) | BW (2 Hz) [10月～1月] G2 台 & R2 台 (4 Hz)[1月～] | RGB (1 Hz) | BW (2 Hz) |
| 2012年9月～ | RGB (1 Hz)& BW (2 Hz) | RGB (2 Hz)& BW (4 Hz) | RGB (1 Hz)& BW (2 Hz) [10月～] | RGB (2 Hz)& BW (4 Hz) [10月～] |

注 AWI : 全天ワテック並列イメージャー
NWI : 狭視野ワテック並列イメージャー

BW : カメラ1台によるパングロ (白黒) 画像観測
RGB : カメラ3台による RGB カラー画像観測
(**G** : 緑フィルター, **R** : 赤フィルター, **B** : 青フィルターによる観測.
括弧内はサンプリング周波数を表す.)

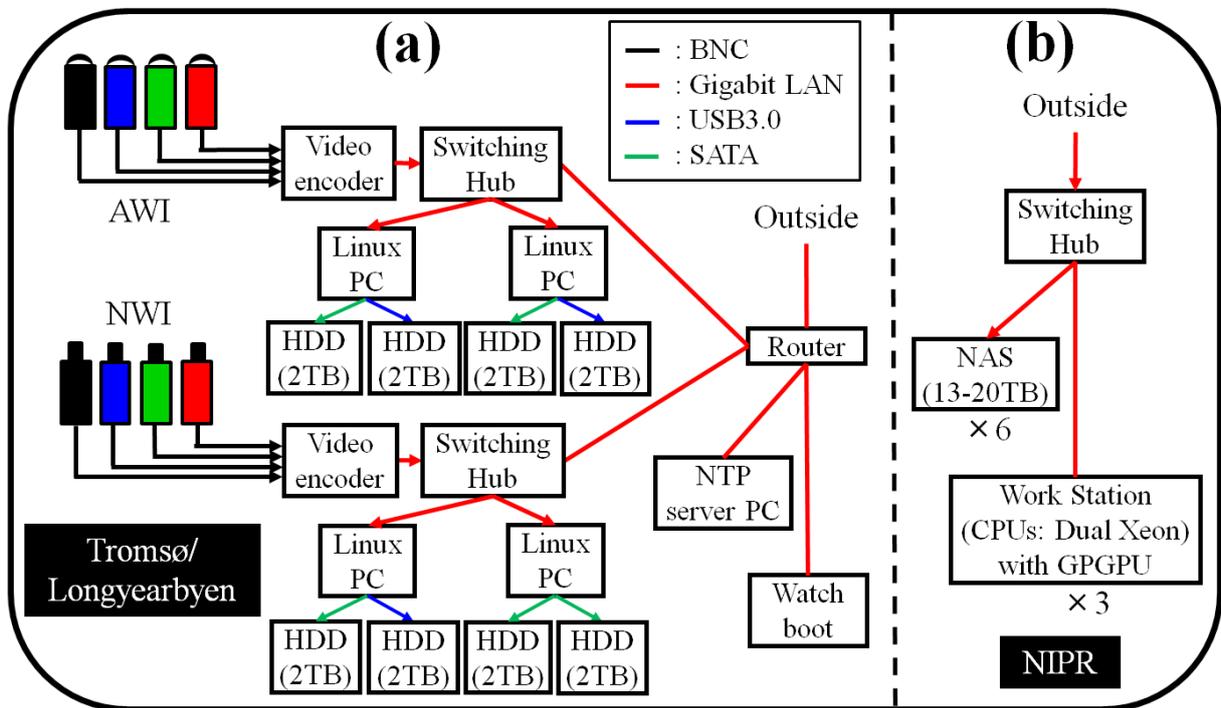


図 4 観測システム概要 (左側(a). 2012年9月以降のトロムソでの観測)と極地研究所内のデータ処理システム概要 (右側(b)). 矢印は画像データの流れを表す.

3. 全天/狭視野並列イメージャーデータ処理と画像解析について

3.1. ルーチン的なデータ処理の流れ

トロムソ/ロングイアピンの両観測点で取得された画像データは極地研に転送され、極地研内に構築された画像データベースに収納される（そのデータ処理システム構成図については図4の右側（b）に示す）。表3にそのデータ処理の流れを示す。観測シーズン（10-3月）の日陰時に取得されたリアルタイムのデータから3色合成画像を30秒毎に自動作成し、極地研ウェブサーバーに転送している。その画像をホームページに掲載することにより、リアルタイムのオーロラ観測状況やシステムの不具合をチェックできるようにしている。現地PCに蓄えられた画像データについては、1時間毎にディレクトリに分けて保管し、晴れてオーロラの見られた時間帯のデータのみ優先して極地研に転送している。現在はその判断を観測サマリープロットから目視で行っているが、画像から雲やオーロラを認識することによる自動化を進めている。極地研に優先して転送された画像データについては、3色合成画像（図5参照）や1時間毎の動画ファイル、発光輝度の南北子午面ケオグラムやアスキーファイルなどの作成を速やかに行い、ホームページ（<http://polaris.nipr.ac.jp/~eiscat/optical/>）上で公開している。また、トロムソとロングイアピン共に晴天でオーロラが見られるイベントについては、（オーロラの発生高度を高度120kmと仮定して）両方の画像を地図上に投影した図を作成し、広い緯度幅（約1400km）のオーロラ分布と変動を把握可能なプロット（図6参照）を作成している。また、観測シーズン終了時（毎年3月）には、すべての画像データを現地から極地研究所に持ち帰っている。

表3 データ処理の流れ

| | |
|-----|---|
| (1) | リアルタイムデータを30秒毎に極地研ウェブサーバーに転送し、ホームページ内（ http://polaris.nipr.ac.jp/~eiscat/optical/realtime.html ）に公開。 （RGB3色画像データについては3色合成画像を自動作成しホームページに公開。） |
| (2) | 0.25-1秒毎に得られる複数の画像データを1時間毎にディレクトリに分けて保管。 |
| (3) | 晴れてオーロラの見られた時間帯のピックアップ。 （現在は目視で、画像から雲/オーロラを認識することによる自動化を試行中。） |
| (4) | RGB3色画像データについては、色調整と3色合成した画像の作成。 さらに全天画像と狭視野画像を並べた画像（図5参照）及び動画の作成。 |
| (5) | 動画やケオグラム、数値ファイル（10秒平均値と1.4Hzオリジナルデータ）の作成と、ホームページ（ http://polaris.nipr.ac.jp/~eiscat/optical/ ）での公開。 |
| (6) | トロムソ及びロングイアピン画像を地図上に投影した画像（図6参照）の作成。 |

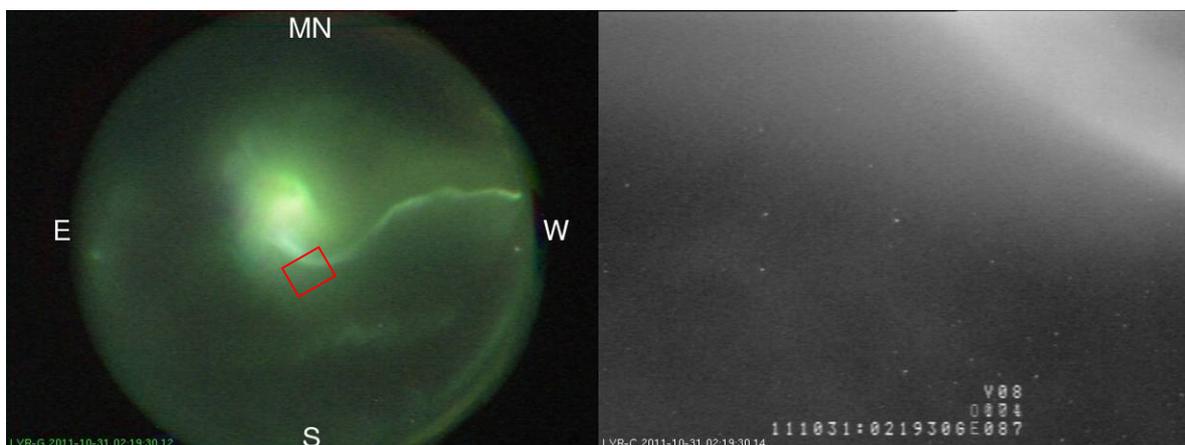


図5 ロングイアピン KHO で取得した同時刻の全天及び狭視野画像の一例。左側の全天イメージャーの視野内に赤色の四角で描かれている領域は、右側の沿磁力線方向の狭視野イメージャーの視野の大きさを表す。

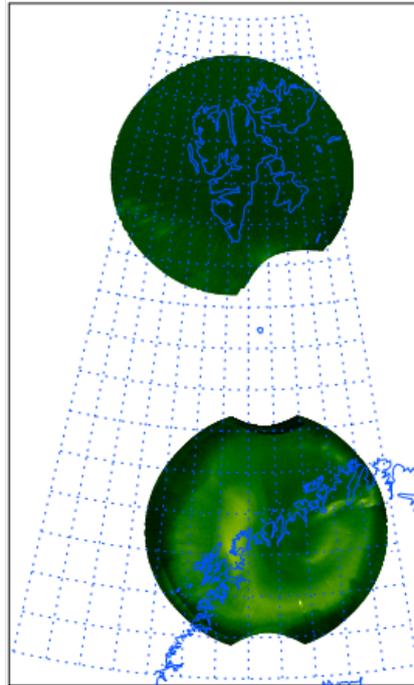


図 6 トロムソ画像とロングイアビン画像を地図上に投影した画像例 (2012年01月22日2107:00 UT 取得データを利用. オーロラの発生高度を120 kmと仮定して地図上に投影.)

3.2. 高度画像処理

これらの画像データベース構築及びその公開作業と平行して進めている, 高度な画像処理の現状と今後想定している展開を以下に記載する.

まず, 画像から雲やオーロラを認識する方法として, 均等量子化法による減色処理を行い, RGB画像を(4³)の64ビット化して定量的に扱うことを行っている. その例を図7に示す. 図7の左図のような558 nmの緑色のオーロラ発光の場合, 緑色のみ強度が高いピクセル(赤色の矢印の部分)が多く存在するようになる. それに対し, 曇りの場合には町明かりを雲が反射して, RGB全色が同等な色分布を示す. このような違いを利用して, 雲やオーロラの観測を自動認識する手法の開発を現在行っている. その手法を観測現地PCに組み込み, 優先してデータ処理すべき時間帯を自動でピックアップするシステム構築を検討している. さらに今後, オーロラを(オーロラアークなどの)形状からも認識できるように, アルゴリズムを改良することが期待される.

2つ目にコンポジット積分によるランダムノイズの軽減が挙げられる. その結果の例を図8に示す. 図8の右図のように, 4つの並列イメージャーで同時刻に得られた4枚の画像を重ね合わせた合成画像(注: この画像取得時には4台のイメージャー共に光学フィルターを不使用)では, 1枚の画像に比べランダムノイズが1/2に軽減し, 薄いオーロラや小さな星がより明確になっている様子が見て取れる. このような手法は, ディフューズオーロラ中の脈動オーロラなどの暗いオーロラ発光(すなわちシグナルがノイズレベルに近い場合)に対して効果的と考えている.

その他に, 本データから特徴的なオーロラの波数を取り出すためのスペクトル解析についても現在進めているところである. オーロラアークの崩壊時などには, オーロラビーズやカールなどが特徴的な波数を持つ構造が発生する²⁾. これらの現象はオーロラ微細構造の生成機構やサブストーム発達機構を理解する上での鍵となると考えられる. また, データベース構築や各種画像処理にはかなりの時間を要する状況(具体的には, 1時間分のデータに対して約30分間の処理時間)であるが, 現在は複数のワークステーション(図4の右側(b)参照)を用いて, できる限り効率よく行っている. 今後はMatlabのParallel Computing ToolboxとGPUによる汎目的計算(GPGPU)を組み合わせ用いた並列処理によるデータ処理速度の改善を図ることを予定している.

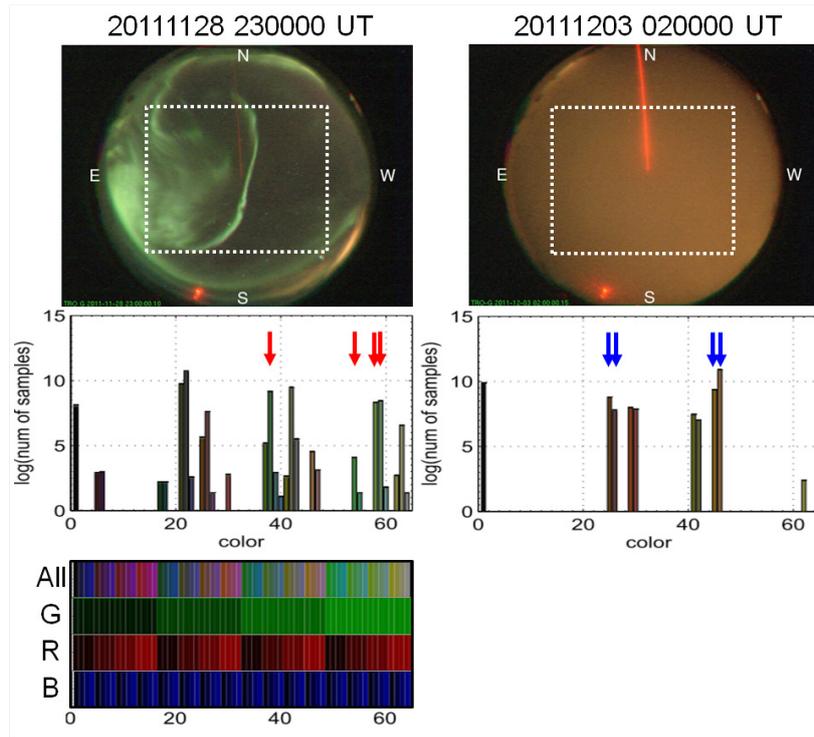


図7 均等量子化（64ビット）を用いた画像解析例．左側に2011年11月28日2300:00 UTに得られたオーロラ全天画像とその色分布を、右側に2011年12月03日0200:00 UTに得られた曇り画像とその色分布を示す．一番上の画像に白い点線で描かれた四角の領域のデータを中央のグラフで用いている．オーロラに起因する典型的な色の位置を赤色の矢印で、雲に起因する色の位置を青色の矢印で示している．なお、全天カメラ画像の北から中央に伸びている赤い直線はナトリウムライダーによる波長589 nmの発光である．一番下の図は色の配列（上から順に、3色合成、緑色、赤色、青色）を表す．

(Before) only 1 picture



(After) a composite picture

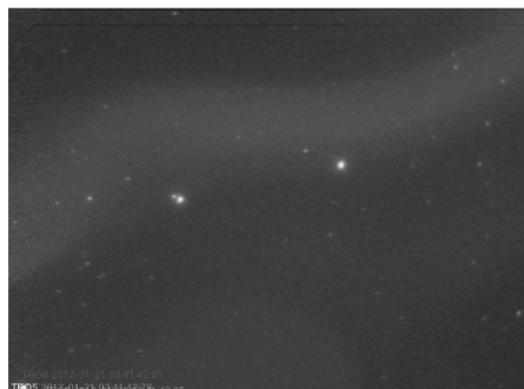


図8 狭視野イメージャーによるオーロラ画像のコンポジット合成処理を行った結果の例．左側は処理を行う前の元画像．右側は4つの並列イメージャーで同時刻に得られた4枚の画像を重ね合わせた合成画像．ランダムノイズが軽減し、薄いオーロラや小さな星がより明確になっている様子が見られる．

4. 全天 / 狭視野並列イメージャーデータの可視化と検索手法について

背景で述べたようにオーロラの総合観測・研究を進める上では、本並列イメージャーで得られたオーロラデータを、電離圏プラズマ物理量等と同時に表示することが求められる。そこで、大学間連携プロジェクト (IUGONET, <http://www.iugonet.org/>) で開発が進められている、超高層大気の地上観測データを表示・解析するためのソフトウェア TDAS/UDAS⁸⁾ 上で、本並列イメージャーデータを表示するプログラム開発を進めてきた。現在はウェブ上に公開している数値ファイル (10 秒平均値) を読み込んで、同時刻の関連観測データとの比較プロットを容易に作成することが可能である。図 9 にその同時観測データのプロット例を示す。この例では、15 UT 頃に太陽風動圧 (図 9 (e)) の急激な増加に伴い、オーロラがロングイアビン上空 (図 9 (c)) からトロムソ上空 (図 9 (d)) に移動し、その後 18 UT からオーロラが極方向に急激に広がる様子 (ポールワードエクспанション) が見て取れる。また、トロムソ VHF レーダーデータ (図 9 (b)) からは、19 UT 付近のオーロラの増光に伴い、高度 120 km 付近の電子密度が増加していることが分かる。

この TDAS/UDAS を用いて表示可能なプロットは、現在のところ南北子午面方向の発光強度ケオグラムのみである。今後は数値ファイルを CDF フォーマット化することにより、画像データのファイル容量を減らし、任意の方向のケオグラムと特定の時刻のオーロラの投影画像 (図 6 参照) を TDAS/UDAS 上で表示できるようにする予定である。さらに、学术论文に使えるように質を高めたプロットを使えるように、データの精度を高めたデータベースと画像データの投影などのデータ処理プログラムの向上を図っている状況である。

その他に、複数のイメージャーによる定常観測を行うようになり、利用可能なデータの情報は莫大となったため、様々な研究者がデータを効率よく検索するシステムの構築が不可欠と言える。この並列イメージャーの観測機器の情報に関しては、国立極地研究所の学術データベース (<http://scidbase.nipr.ac.jp/>) にメタ情報を登録し随時更新すると共に、IUGONET プロジェクトのメタデータ検索ホームページ (<http://search.iugonet.org/iugonet/>) でも検索可能な体制を整えている。観測データのメタ情報 (データの有無や観測の設定情報等) についても、IUGONET メタデータ検索と連携を取り、日時を指定することでイメージャーデータの有無と詳細情報を検索可能なシステム構築を進めているところである。現状では光学観測用ホームページ内に観測日時とデータ処理の進行状況をリスト化してホームページ内に公開している。

データ公開と利用のポリシーに関しては、全国共同利用によりデータが多目的で広く用いられるために、得られたデータをリアルタイムで公開すると共に、速やかにデータ処理をしてウェブページに公開することを基本としている。非営利の研究及び教育目的の場合には、ウェブに公開したオーロラ画像データを (引用元を記載した上で) 自由に使用できる方針としている。また、大量の画像データを扱う場合や 2 次処理を必要とする場合には、データ責任者に事前にコンタクトをとることを推奨している。このようなデータ利用ポリシーは、日本を含めた 6 ヶ国による国際共同で運営されている EISCAT レーダーで得られたデータの利用ポリシーと同じ方針と言える。

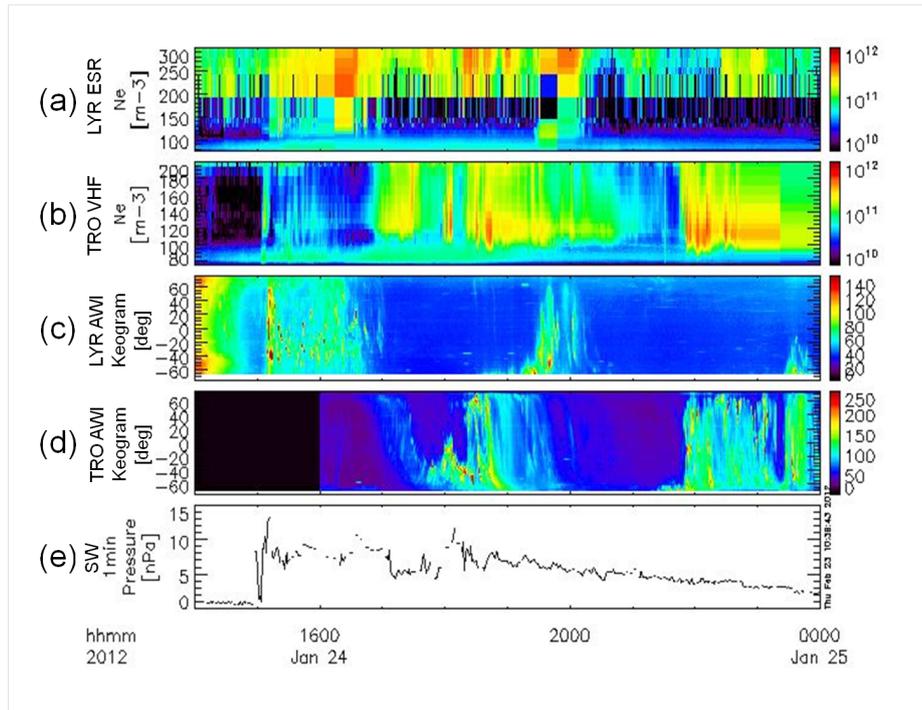


図9 TDAS/UDASを用いて作成した時系列プロット(2012年1月24日14-24 UTの各データを使用)。上から順に、(a) ロングイアビン ESR32m 電子密度分布(高度80-300 km)、(b) トロムソ VHF 電子密度分布(高度80-250 km)、(c) ロングイアビン AWI オーロラ発光分布(磁気南北子午面)、(d) トロムソ AWI オーロラ発光分布(地理南北子午面)、(e) 太陽風動圧の時間変化。

5. まとめと今後の方向性について

本論文では、トロムソ及びロングイアビンで国立極地研究所が展開している全天/狭視野並列イメージャー観測による、高時間分解能のオーロラカラー画像の大量データベース化や、データ解析手法、データ可視化手法を中心に現状と今後の予定を紹介した。このデータベースは、様々なオーロラの研究を進める上で重要なオーロラの基本情報を含んでおり、EISCAT レーダー等の関連観測データを組み合わせた総合的な研究を全国共同利用で実施する上でも不可欠であると考えている。得られた大量のオーロラ画像データから、オーロラ研究に利用可能な(ごく一部の)質の高いオーロラ画像データを抽出する方法や、その抽出された画像データを更に詳細かつ総合的に解析する手法の開発・発展が求められており、我々はオーロラ/雲画像認識プログラムや複数のデータを同時に可視化するプログラム等の各種開発に取り組んでいる。

この並列イメージャー観測システム構築と可視化を含むデータ処理の経験を基に、将来的には(1) アイスランドや南極昭和基地、南極点などに多点展開をしているオーロラ画像データを統合して扱うことが可能なシステム開発、(2) フィンランド気象研究所(FMI)が中心に推進しているGlobal Auroral Imaging Access(GAIA)プロジェクト(<http://gaia-vxo.fmi.fi/viewer/>)との連携、(3) 省電力・無人オーロラ観測システムの多点展開への発展等が期待される。

謝辞

本論文で述べられているオーロラ観測用並列イメージャー観測及びデータ処理の一部は、国立極地研究所研究プロジェクト KP-9 (「EISCAT レーダーならびに地上拠点観測に基づく北極圏超高層・中層大気の国際共同研究」) の支援を受けて実施している。KP-9プロジェクト及び光学観測施設の維持に責任を持つ宮岡宏博士と、並列イメージャー観測データの整備に尽力を費やして下さっている杉田郁美氏に感謝の意を表す。また、 Tromsø 及びロングイアビンにおける光学観測を継続して行うに当たり、観測施設を運用する Tromsø 大学 Tromsø Geophysical Observatory (TGO) 及び The University Centre in Svalbard (UNIS) のスタッフに感謝する。

参考文献

- 1) Hosokawa, K., Y. Ogawa, A. Kadokura, H. Miyaoka, and N. Sato, Modulation of ionospheric conductance and electric field associated with pulsating aurora, *J. Geophys. Res.*, Vol. 115 (2010), A03201, doi:10.1029/2009JA014683.
- 2) Sakaguchi, K., K. Shiokawa, and E. Donovan, Azimuthal structures of ray auroras at the beginning of auroral substorms, *Geophys. Res. Lett.*, Vol. 36 (2009), L23106, doi:10.1029/2009GL041252.
- 3) Motoba, T., K. Hosokawa, A. Kadokura, and N. Sato, Magnetic conjugacy of northern and southern auroral beads, *Geophys. Res. Lett.*, Vol. 39 (2012), L08108, doi:10.1029/2012GL051599.
- 4) Blixt, E. M., and M. J. Kosch, Coordinated optical and EISCAT observations of black aurora, *Geophys. Res. Lett.*, Vol. 31 (2004), L06813, doi:10.1029/2003GL019244.
- 5) Akasofu S-I., The development of the auroral substorm. *Planet. Space Sci.*, Vol. 12 (1964), pp 273-282.
- 6) 小川泰信, 野澤悟徳, Ingemar Häggström, 大山伸一郎, 元場哲郎, 津田卓雄, 齋藤昭則, 宮下幸長, 田中良昌, 堀智昭, 上野玄太, 宮岡宏, 藤井良一, 欧州非干渉散乱 (EISCAT) レーダーの大規模データ処理と可視化, *宇宙科学情報解析論文誌*, 第 1 号 (2012), pp 83-89.
- 7) Tsuda, T. T., S. Nozawa, T. D. Kawahara, T. Kawabata, N. Saito, S. Wada, C. M. Hall, S. Oyama, Y. Ogawa, S. Suzuki, T. Ogawa, T. Takahashi, H. Fujiwara, R. Fujii, N. Matuura, and A. Brekke, Fine structure of sporadic sodium layer observed with a sodium lidar at Tromsø, Norway, *Geophys. Res. Lett.*, Vol. 38 (2011), L18102, doi:10.1029/2011GL048685.
- 8) 田中良昌, 新堀淳樹, 鍵谷将人, 堀智昭, 阿部修司, 小山幸伸, 林寛生, 吉田大紀, 河野貴久, 上野悟, 金田直樹, 米田瑞生, 田所裕康, 元場哲郎, 三好由純, 関華奈子, 宮下幸長, 瀬川朋紀, 小川泰信, IUGONET 解析ソフトウェアの開発, *宇宙科学情報解析論文誌*, 第 1 号 (2012), pp 91-98.

IUGONET 解析ソフトウェアの現状と今後の発展

田中 良昌^{*1} 新堀 淳樹^{*2} 梅村宜生^{*3} 堀 智昭^{*3} 阿部 修司^{*4} 小山 幸伸^{*5} 林 寛生^{*2,*6} 上野 悟^{*7}
佐藤 由佳^{*1} 谷田貝 亜紀代^{*2} 小川 泰信^{*1} 三好 由純^{*4} 関 華奈子^{*4} 宮下 幸長^{*4} 瀬川 朋紀^{*4}

Current status and future development of IUGONET data analysis software

Yoshimasa TANAKA^{*1} Atsuki SHINBORI^{*2} Norio UMEMURA^{*3} Tomoaki HORI^{*3} Shuji ABE^{*4}
Yukinobu KOYAMA^{*5} Hiroo HAYASHI^{*2,*6} Satoru UENO^{*7} Yuka SATO^{*1} Akiyo YATAGAI^{*2} Yasunobu OGAWA^{*1}
Yoshizumi MIYOSHI^{*4} Kanako SEKI^{*4} Yukinaga MIYASHITA^{*4} and Tomonori SEGAWA^{*4}

Abstract

In the IUGONET (Inter-university Upper atmosphere Global Observation NETWORK) project, we have developed UDAS (iUgonet Data Analysis Software), which is a software to visualize and analyze various kinds of upper atmospheric data distributed by five universities/institutes (Tohoku Univ., Nagoya Univ., Kyoto Univ., Kyushu Univ., and NIPR). UDAS is a plug-in software of TDAS (THEMIS Data Analysis Software suite) written in IDL (Interactive Data Language), and thus can call useful routines to visualize and analyze time series data and GUI included in TDAS. We released a beta version of UDAS at the IUGONET website in May, 2011, and a formal version in February, 2012. We are planning to incorporate UDAS into TDAS and release it from the THEMIS software website after September, 2012. Automatic test tools for the UDAS programs were developed to reduce the workload for the development team. Furthermore, we built an executable file of TDAS that can run on the IDL Virtual Machine environment without any IDL licenses and released it for public testing. These developments will contribute to the promotion of the use of the UDAS and the ground-based observational data distributed by the IUGONET institutions.

Keyword: IUGONET, THEMIS mission, upper atmosphere, data analysis software, IDL

概要

大学間連携プロジェクト「超高層大気長期変動の全球地上ネットワーク観測・研究」(IUGONET)では、IUGONET参加機関である東北大学、名古屋大学、京都大学、九州大学および国立極地研究所の5機関が所有する多種多様な超高層大気データを可視化・解析できるデータ解析ソフトウェアUDAS (iUgonet Data Analysis Software)を開発している。UDASは、IDL言語で書かれたTHEMISミッションのデータ解析ソフトウェアTDAS (THEMIS Data Analysis Software suite)のプラグインソフトであり、TDASが持つ様々な時系列データの可視化・解析ツールやGUI機能を利用することができる。我々は、2011年5月よりIUGONETウェブページにおいてUDASのβバージョンの公開を開始し、2012年2月に本公開を開始した。2012年10月以降には、UDASをTDASに組み込み、THEMISソフトウェアウェブサイトから公開することを計画している。UDASを提供するにあたり、プログラムの動作確認作業を軽減するためのテストツールの開発も行った。さらに、IDLライセンスを持たない研究者のために、IDL Virtual Machine環境で走るTDASの実行ファイルを開発し、テスト公開を開始した。これらの新しい展開により、UDAS及びIUGONET所属機関が所有する地上観測データの利用が促進されることが期待される。

*1 国立極地研究所 (National Institute of Polar Research)

*2 京大生圏研究所 (Research Institute for Sustainable Humanosphere, Kyoto University)

*3 名古屋大学太陽地球環境研究所 (Solar-Terrestrial Environment Laboratory, Nagoya University)

*4 九州大学国際宇宙天気科学・教育センター (International Center for Space Weather Science and Education, Kyushu University)

*5 京都大学理学研究科附属地磁気世界資料解析センター (Data Analysis Center for Geomagnetism and Space Magnetism, Graduate School of Science, Kyoto University)

*6 Now at 富士通エフ・アイ・ピー株式会社 (FUJITSU FIP CORPORATION)

*7 京都大学理学研究科附属天文台 (Kwasan and Hida Observatories, Graduate School of Science, Kyoto University)

1. はじめに

平成 21 年度から 6 ヶ年計画でスタートした大学間連携プロジェクト「超高層大気長期変動の全球地上ネットワーク観測・研究」(IUGONET: Inter-university Upper atmosphere Global Observation NETwork)¹⁾は、東北大学、名古屋大学、京都大学、九州大学、国立極地研究所の 5 機関が所有する全球規模の地上観測ネットワークで取得された多種多様な超高層大気データを利用して、複数の要因が複雑に絡み合った超高層大気長期変動のメカニズムを解明することを目的としている。そのために、本プロジェクトでは、複数の機関が分散管理している様々な観測データを検索、取得、解析するための”e-infrastructure”として、主にメタデータデータベース^{2),3)}とデータ解析ソフトウェア⁴⁾の開発を行っている。本論文では、後者の解析ソフトウェア (UDAS : iUgonet Data Analysis Software) の概要と現状を報告し、今後の展望について述べる。

2. UDAS の概要と現状

UDAS は、THEMIS ミッション⁵⁾ で得られた衛星・地上観測データを描写・解析するための総合解析ツール TDAS (THEMIS Data Analysis Software suite) のプラグインソフトである。THEMIS ミッションは、2007 年 2 月に打ち上げられた 5 機の磁気圏探査衛星と北米の約 20 ヶ所の全天カメラ、磁場観測装置で取得されたデータから、オーロラが爆発的に発達する現象「サブストーム」の発生機構を解明することを目的としたミッションである。IUGONET プロジェクトにおける解析ソフトウェアを作成するにあたって TDAS をベースに採用した理由は、以下に述べるような利点が活用できるためである。(a) IDL (Interactive Data Language) 言語で書かれたオープンソースのライブラリであり、自由にプログラムを書き変えることができる。(b) 遠隔地で分散管理されているデータファイルを、個々の所在情報を意識することなく、インターネット越しにダウンロードできる。(c) 複数データを並べて表示する可視化ルーチンや、フィルター、フーリエ変換、ウェーブレット変換等の解析ルーチンが多数用意されている。(d) GUI (Graphical User Interface) が用意されており、IDL 初心者でも比較的簡単に利用できる。

UDAS を利用するには、UDAS を IUGONET ウェブサイトからダウンロードし、UDAS を TDAS よりも先に読み込むように IDL のパスを設定する。UDAS のインストールが成功すれば、TDAS の GUI のロードデータウィンドウに新たに「IUGONET Data」タブが現れる。

UDAS は、2011 年 5 月に β バージョンが IUGONET ウェブサイト (URL: <http://www.iugonet.org/software/install.html>) から公開され、数度のバージョンアップを経た後、2012 年 2 月に正式バージョン (UDAS v1.00.1) が公開された。表 1 に 2012 年 6 月 30 日時点での最新バージョン (UDAS v2.00.1, TDAS v7.00 に対応) に含まれている 14 種類のデータのロードプログラムを、図 1 に UDAS を使って複数の IUGONET 関連データをプロットした例を、それぞれ示す。なお、UDAS は、内部磁気圏探査を目的とした ERG (Energization and Radiation in Geospace) ミッション⁶⁾のサイエンスセンター (ERG-SC) との協力の下、開発が行なわれている。

表 1 UDAS v2.00.1 に含まれているロードプログラム。(*)印は、ERG 関連データのロードプログラム (erg_load_xxx) のエイリアス。

| 観測データ | プログラム名 |
|------------------------|----------------------|
| 飯館電波望遠鏡データ | iug_load_iprt |
| 境界層レーダーデータ | iug_load_blr_rish |
| L バンド下部対流圏レーダーデータ | iug_load_ltr_rish |
| 赤道大気レーダーデータ | iug_load_ear |
| MU レーダーデータ | iug_load_mu |
| 流星レーダーデータ | iug_load_meteor_rish |
| MF レーダーデータ | iug_load_mf_rish |
| ウィンドプロファイラデータ | iug_load_wpr_rish |
| SuperDARN レーダーデータ (*) | iug_load_sdfit |
| EISCAT レーダーデータ | iug_load_eiscat |
| 地磁気指数, WDC 地磁気データ | iug_load_gmag_wdc |
| 昭和基地, アイスランド地磁気データ (*) | iug_load_gmag_nipr |
| 210° 地磁気観測網データ (*) | iug_load_gmag_mm210 |
| MAGDAS 地磁気観測網データ | iug_load_gmag_serc |

3. UDAS と TDAS の統合

UDAS のインストールには、TDAS とは別に IUGONET ウェブサイトから UDAS パッケージをダウンロードする方式が採用されているが、UDAS を TDAS の標準機能として組み込み UCB (University of California, Berkeley) の THEMIS ソフトウェアウェブサイトから公開することが決まっており、現在その準備が進められている。これには、UDAS をインストールする手間を減らすことができるということ以外に、既に TDAS を利用しているユーザーが IUGONET プロジェクト関連データを容易に使用できるようになり、データ利用促進に繋がるというメリットがある。現時点では、2012 年 10 月以降に TDAS v7.1 として公開することを計画している。

TDAS v7.1 において、UDAS のプログラム群は、TDAS のディレクトリ構造の最上位にある themis, erg 等のミッションのディレクトリと並列に iugonet というディレクトリ名で配置される。また、TDAS v7.1 では、GUI で IUGONET プロジェクトに関連するデータをロードする際にデータポリシーをポップアップウィンドウで表示する機能を新たに追加する。これは、GUI にはコンソール画面が無くデータポリシーが表示できないという理由から追加した機能であり、表示されたデータポリシーに対して OK あるいは Cancel を選択し、OK を選択した場合のみデータがロードされる。また、OK をクリックしたという情報は IDL 環境変数として保存され、一度 OK をクリックした後は IDL を終了するまで同種のデータについてデータポリシーは表示しない仕組みになっている。

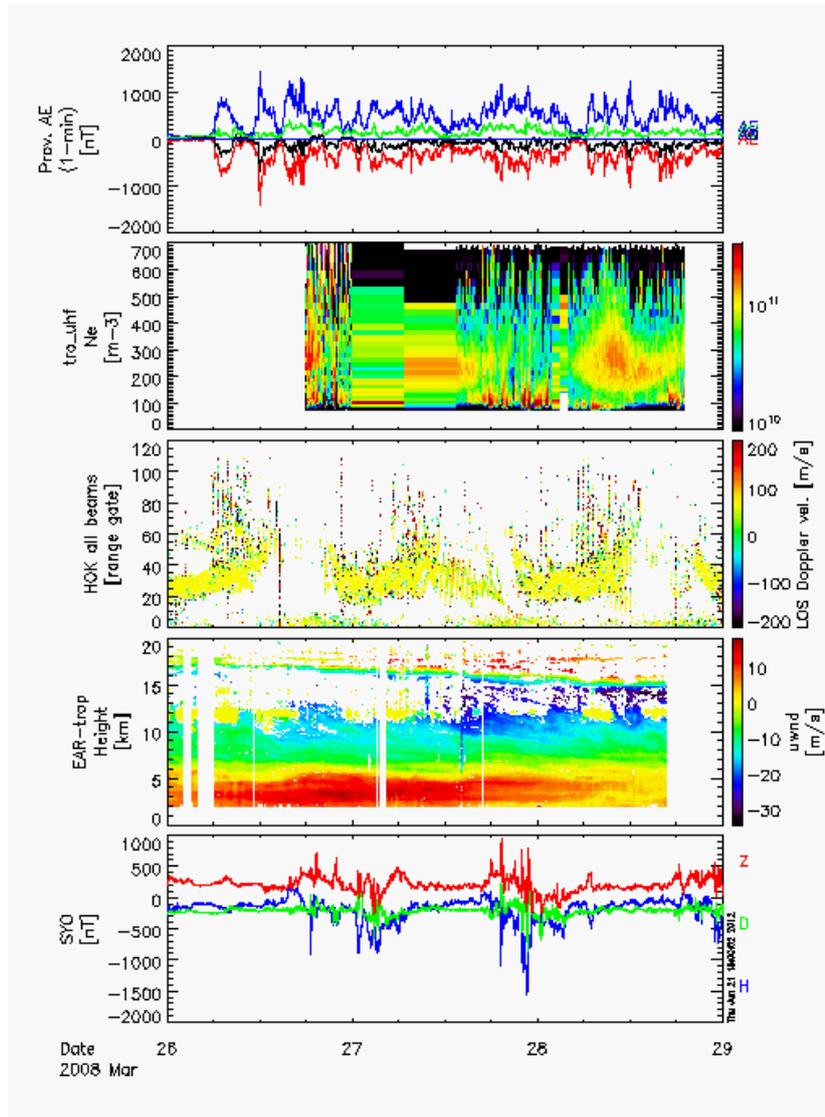


図 1 UDAS のロードプログラムを使って作成したスタックプロット。上から、2008 年 3 月 28 日の AE 指数、EISCAT トロムソ UHF レーダーデータ、SuperDARN 北海道レーダーデータ、赤道大気レーダーデータ、南極昭和基地地磁気データ。

4. UDAS プログラムのテストツールの開発

THEMIS ミッションは、TDAS の動作に関して一定の基準を満たすことを確認するために、独自のテストケース (TEST SUITE) を策定している。第3章で述べたように UDAS を TDAS に組み込むためには、IUGONET も、THEMIS ミッションが策定する TEST SUITE に従って、予め UDAS のプログラムの動作を確認する必要がある。TEST SUITE には、(1) プログラムの読み込みとデータロード、(2) パラメータの投入、(3) データのプロット、(4) 解析、(5) プロット画像の出力、などの機能を確認するための、約 100 項目にわたる詳細な実行手順とその期待値が記載されている。実行手順のうち、プログラムの名称およびプログラムに依存するパラメータの名称の部分については、テストするプログラムごとに読み替える必要があるが、実行手順および期待値は THEMIS ミッションが定めるものと同じである。このすべてのテスト項目をクリアすることにより、UDAS を THEMIS ミッションに提供できる状態となる。

IUGONET プロジェクトは、(1)1 つのプログラムあたり約 100 項目もの基準をすべて満たすまでテストを何度も繰り返す必要がある、(2)IUGONET が独自に開発したプログラムは本稿執筆時点(2012年6月時点)で14と多い、(3)プログラムの機能を修正または追加するごとにテストする必要がある、(4)IDL または TDAS のバージョンが上がるごとにすべてのプログラムをテストする必要がある、つまり、テストの頻度が多く、その実施には多大な人的コストを必要とすることから、テストを自動で実行するツールを開発することとした。THEMIS ミッションは、TEST SUITE の実施にあたり、GUI を使うことを条件としている。これに対し、IUGONET は、前段階として、まず CUI(Command User Interface) を使って自動でテストを実行するツールを開発した。CUI を使ったテストは、THEMIS ミッションの基準を満たすものではないが、IUGONET プロジェクトにおいて独自に UDAS の動作を確認するためには有用である。また、GUI を用いたテストと CUI を用いたテストでは、操作手法が異なるだけであり、TDAS および UDAS が内部で実行する処理をテストするという点では等価である。

CUI を使ったテストでは、UDAS のプログラムが出力するメッセージを確認することと、予め作成したプロット画像とテスト実行時に作成したプロット画像を比較することを、判定要素とした。プログラムの動作状況については、前者の手法を用いて容易に確認することができる。一方で、プログラムの内部においてパラメータとプログラムコードが想定どおりに動作したことを確認するのは目視では困難であることから、後者の手法も用いることとした。プログラム内部におけるパラメータとプログラムコードの振る舞いは、最終的に、プロットを実行する tplot コマンドの引数に入力する値 (tplot 変数) に影響する。従って、実際に tplot コマンドを実行して得たプロット画像と、予め作成したプロット画像を比較すれば、それらの振る舞いの妥当性を確認できるという仕組みである。

テストツールの論理構成を、図2に示す。テストツールは、テストを実行するモジュール、設定ファイル、プロット画像を保存する領域(ディレクトリ)の3つで構成する。モジュールは IDL と TDAS, UDAS の実行と、UDAS から出力した画像の比較・判定を行う。前者の部分は、IDL と TDAS, UDAS をインストールしている環境であれば容易に動作することを目的としてシェルスクリプトで実装した。後者の部分は、実行結果を、CUI のコンソール画面のみでなく、IUGONET プロジェクトのメタデータデータベースの開発において様々なテストの結果を可視化するために採用している Cruise Control へ組み込むことを視野に入れ、Java のテストフレームワークである JUnit で実装した。なお、JUnit の実行に必要なライブラリは、予めテストツールに同梱している。また、前者のみ、または後者のみの実行も可能としているが、1つのコマンドで両者の処理を実行してテスト結果を得ることができるように、各モジュールを1つのシェルスクリプトでラッピングした。設定ファイルは、実行する crib sheet(詳細については後述する)を選択するためのテキスト形式のファイルである。プロット画像を保存するディレクトリは、予め作成したプロット画像を格納する expected ディレクトリと、TDAS の実行ごとに作成するプロット画像を格納する actual ディレクトリの2つで構成する。

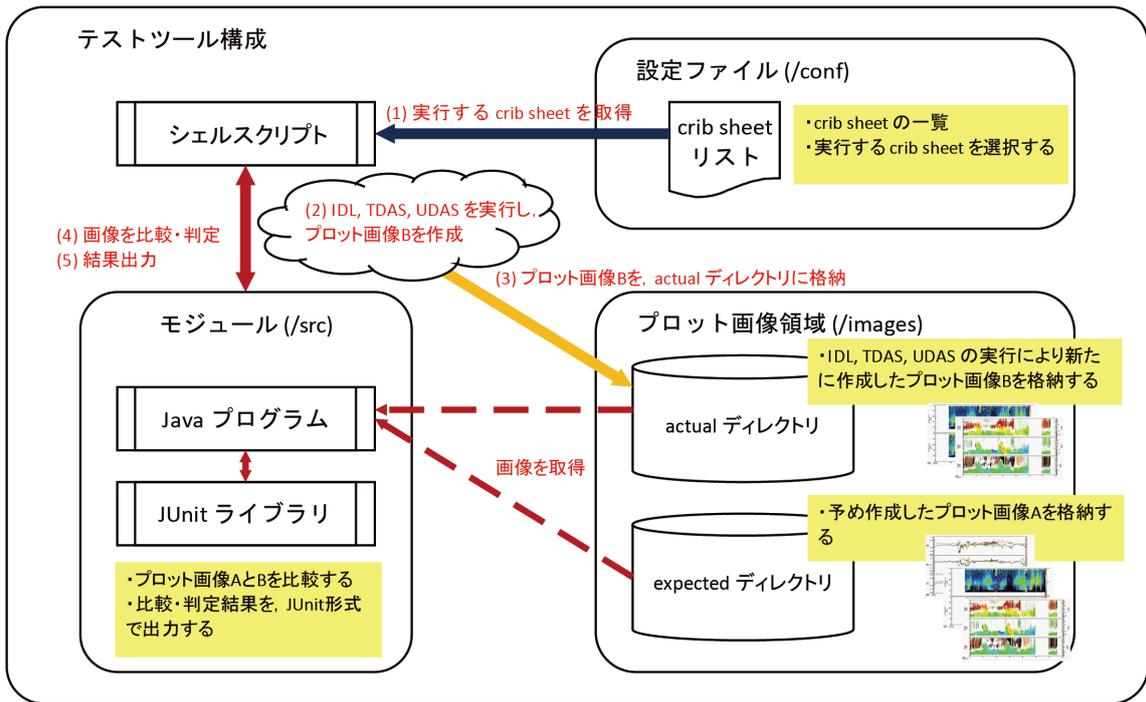


図 2 テストツールは、モジュール、設定ファイル、プロット画像を保存する領域（ディレクトリ）の 3 つで構成され、モジュールをラッピングしたシェルスクリプトの実行のみで、プログラムの動作結果を得ることができる。

次に、UDAS テストのフレームワークを、図 3 に示す。まず、テスト実施者は、改版前のプログラムと crib sheet を使い、プロット画像 A を得る。ここで、改版前のプログラムとは、IUGONET プロジェクトが正式バージョンとして既に公開しているものである。また、crib sheet とは、例えば TDAS の起動 (thm_init)、日時指定 (timespan)、プログラムの読み込みとデータロード (iug_load_xxx; ここで xxx はプログラムに依存)、プロット (tplot)、画像出力 (makepng あるいは makejpg など)、終了 (end) などの、一連のコマンドを記述したものである。crib sheet は、基本的には 1 つのプログラムに対して 1 つであるが、複数のテストケースを別々に策定する場合は 2 つ以上となる。

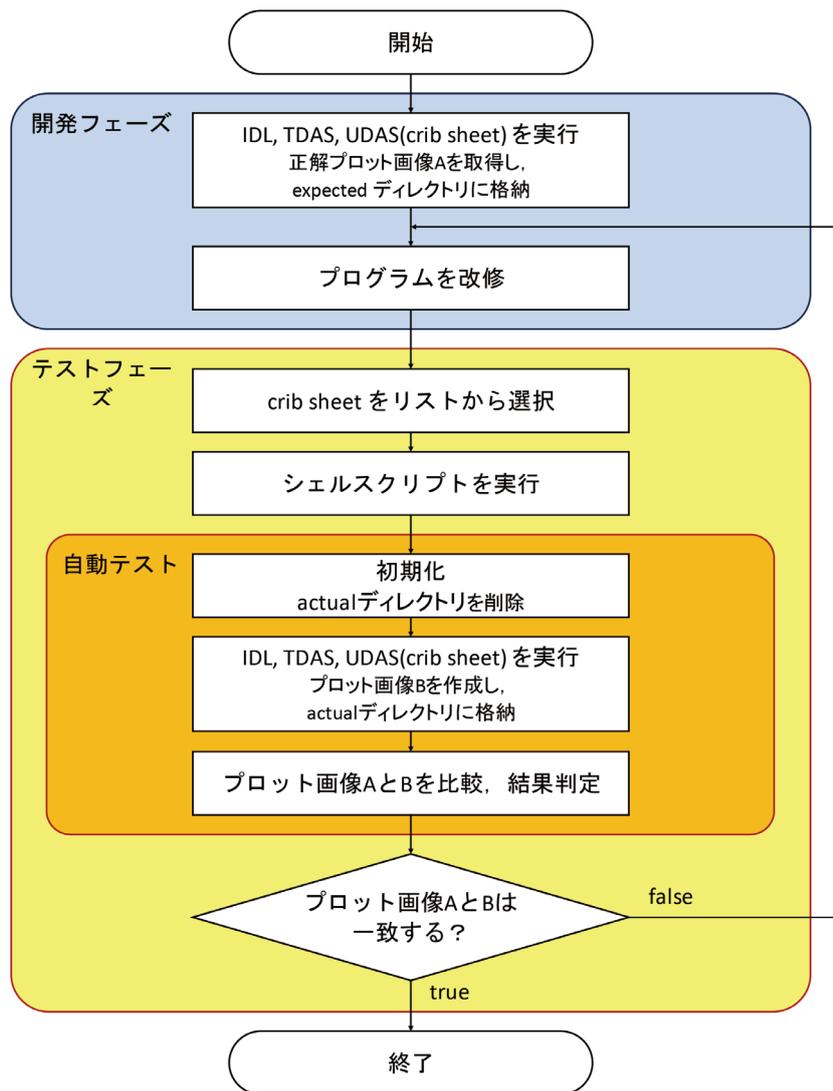


図3 UDAS テストのフレームワーク。まず、改版前のプログラムからプロット画像 A を作成し、次に、自動テストを使って、改版後のプログラムからプロット画像 B を作成し、かつ、プロット画像 A と B を比較・判定することによって、プログラムの動作を確認することができる。

次に、改版中のプログラムをテストするフェーズに移行する。まず、テスト実施者は、設定ファイルを編集し、テストしたいプログラムに対する crib sheet を選択する。なお、設定ファイルにおいては、複数の crib sheet を選択することも可能である。次に、シェルスクリプトを実行する。シェルスクリプトは、まず、初期化処理として、actual ディレクトリに含まれるプロット画像を削除する。次に、IDL を起動し、設定ファイルで選択された crib sheet を順に実行する。crib sheet に記述した TDAS, UDAS のコマンドを順に実行することで、最終的に、改版したプログラムのパラメータとプログラムコードに則したプロット画像 B を作成し、actual ディレクトリへ出力する。すべての crib sheet に対して処理を終えると、次に、シェルスクリプトは Java プログラムを呼び出す。Java プログラムは、expected ディレクトリに保存したテスト対象のプロット画像 A と、actual ディレクトリに新たに作成されたプロット画像 B を比較する。2つの画像からバイナリコードを取得し、互いに一致した場合は判定結果として真(true)を、一致しない、または actual ディレクトリに画像が出力されていないなど異常の場合は偽(false)を返し、コンソール画面に出力する。テストにより作成されたすべての画像に対して同様の判定と出力を行った後、JUnit の assertTrue 関数にその結果を渡し、総合判定する。すべての判定結果が真(true)である場合は総合結果として真(true)を、1つでも偽(false)が存在する場合は偽(false)を返し、コンソール画面にメッセージを出力する(図4)。テスト実施者は、このテスト結果をプログラムの開発者に報告するとともに、仮に、判定結果に偽(false)が含まれている場合は、プログラムの再修正を依頼する。

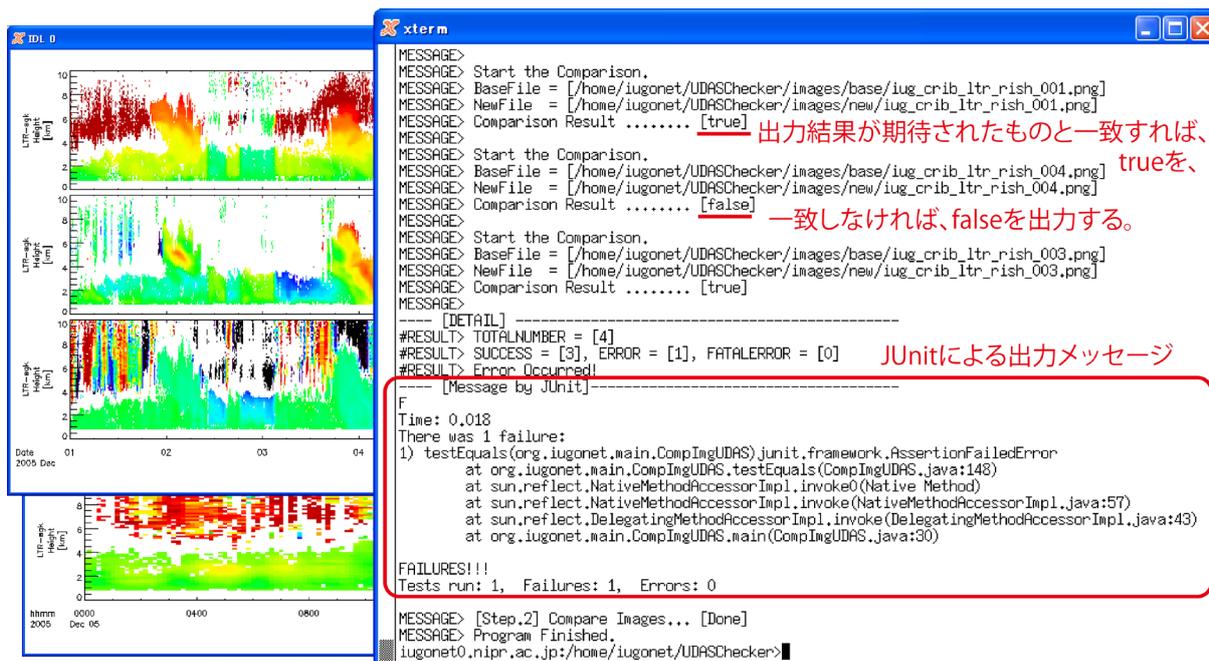


図 4 CUI を使った自動テストの実行例. シェルスクリプトを実行することで, 自動的に, crib sheet の実行とプロット画像の出力, 新旧プロット画像の比較・判定を行う. (左側のプロット画像は, テスト実施中は表示と非表示を繰り返すが, この図においては, 事例として示すために, 別途, 手動で表示した).

1つのプログラム(1つの crib sheet)に対するテストは, 実行するコマンドの数にもよるが, 10秒から20秒程度で終了する. 従来の手動テストでは, プログラムの動作異常が開発者にフィードバックされるまでに時間を要したり, 開発スケジュールにおいてテストが多大な時間を占めていたりしたのに対し, 自動テストでは, 開発者は, 即時にプログラムのテスト結果を得ることができ, その結果, プログラムのソースコードを見直すことに専念できるようになった. また, 前述のとおり, テスト実施者は, シェルスクリプトの実行のみで結果を得ることができるため, IDL, TDAS, UDAS の操作に不慣れであっても, 統一的な手法でテストを実施できるようにもなった. IUGONET は, 現在, CUI のテストツールのみを使っているが, THEMIS ミッションと継続的に連携していくためにも, TDAS v7.1 のリリース後には GUI のテストツールも整備する予定である.

5. TDAS on IDL-VM の開発

IDL は有料のソフトウェアであり, IDL ライセンスを持たない研究者は, TDAS を使用することはできない. このような IDL ライセンスを持たないユーザーにも TDAS 及び UDAS を利用してもらうため, IDL Virtual Machine (IDL-VM) 環境で走る TDAS の開発を行った. IDL-VM とは, Exelis Visual Information Solutions により無償提供される IDL の実行環境であり, IDL でコンパイルしてできた実行ファイルをライセンス無しで走らせることができる. IDL-VM 環境では, プログラムを修正・コンパイルすることはできないが, GUI で行う決まった命令を実行することができる. 既に述べたように TDAS には高機能の GUI が用意されており, GUI 上でフィルターやフーリエ変換, ウェーブレット変換等の解析ツールが利用できる. また, Mini Language と呼ばれる (CUI では calc コマンドを使うことで実行できる) スクリプトを GUI から実行可能であり, GUI でも複数のデータ変数に対して簡単な算術計算を行うことができる.

TDAS の実行ファイルの作成には, IDL の標準ライブラリと TDAS のライブラリを全てコンパイルする方法を取った. このとき, TDAS ライブラリの中でプログラム・関数名が同じでファイル名が異なるファイルが複数ある場合は必要なものを残して削除する, コンパイルオプション (compile_opt) の記述があるプログラム・関数は最後にコンパイルする, 等のことに注意する必要がある.

開発された実行ファイルは, 一通りの動作確認の後, β バージョンとして UCB の FTP サイト (ftp://apollo.ssl.berkeley.

edu/pub/THEMIS/SCI/Soft/TDAS%20VM/idlvm_tdas_7_00_b1.zip) でテスト公開された。将来的に UDAS のプログラムを含んだ TDAS v7.1 が公開された後には、TDAS v7.1 についても実行ファイルを作成し、THEMIS ソフトウェアウェブサイト (<http://themis.ssl.berkeley.edu/software.shtml>) からダウンロードできるようにする予定である。

6. まとめ

大学間連携プロジェクト「超高層大気長期変動の全球地上ネットワーク観測・研究」(IUGONET) で開発しているソフトウェア UDAS は、THEMIS ミッションのデータ解析ツール TDAS のプラグインソフトである。2011 年 5 月に IUGONET ウェブサイトで UDAS を β 公開して以降、数度のバージョンアップを経て、2012 年 2 月から正式公開を行なっている。2012 年 10 月以降に UDAS を TDAS v7.1 の標準機能として組み込んで THEMIS ソフトウェアウェブサイトから公開できるよう準備を進めており、IUGONET プロジェクト関連データの一層の利用促進が期待される。また、UDAS を TDAS に統合するにあたって、UDAS プログラムの動作テストにかかる人的コストを軽減するために、自動テストツールを開発した。さらに、IDL-VM 環境で走る TDAS の実行ファイルを開発し、THEMIS ソフトウェアウェブサイトですべて公開を開始した。これにより、IDL ライセンスを持たない研究者への TDAS の普及が見込まれる。

謝辞

大学間連携プロジェクト「超高層大気長期変動の全球地上ネットワーク観測・研究」は、文部科学省特別教育研究経費(研究推進)[平成 21 年度]および特別経費(プロジェクト分)[平成 22 年度~]の交付を受けて、平成 21 年度より 6 カ年計画で実施している事業である。IUGONET プロジェクトで使用される観測データのデータベース構築に携わった全ての方々に感謝する。特に、国立極地研究所宙空圏研究グループのデータ整備に尽力下さった余川真純氏、京都大学生存圏研究所のデータ整備に尽力下さった橋口典子氏、京都大学理学研究科附属天文台のデータ整備に尽力下さった金田直樹氏に感謝の意を表す。また、UDAS の開発、及び、TDAS on IDL-VM の開発に多大な協力をいただいた THEMIS サイエンスサポートチームに感謝する。

参考文献

- 1) 林 寛生, 小山幸伸, 堀 智昭, 田中良昌, 新堀淳樹, 鍵谷将人, 阿部修司, 河野貴久, 吉田大紀, 上野 悟, 金田直樹, 米田瑞生, 田所裕康, 元場哲郎, 大学間連携プロジェクト「超高層大気長期変動の全球地上ネットワーク観測・研究」, 宇宙科学情報解析論文誌, 第 1 号, 113-120, 2012.
- 2) 小山幸伸, 河野貴久, 堀 智昭, 阿部修司, 吉田大紀, 林寛生, 田中良昌, 新堀淳樹, 上野悟, 金田直樹, 米田瑞生, 元場哲郎, 鍵谷将人, 田所裕康, 超高層物理学分野の為のメタデータ・データベースの開発, 宇宙科学情報解析論文誌, 第 1 号, 99-104, 2012.
- 3) 堀 智昭, 鍵谷将人, 田中良昌, 林 寛生, 上野 悟, 吉田大紀, 阿部修司, 小山幸伸, 河野貴久, 金田直樹, 新堀淳樹, 田所裕康, 米田瑞生, IUGONET 共通メタデータフォーマットの策定とメタデータ登録管理システムの開発, 宇宙科学情報解析論文誌, 第 1 号, 105-111, 2012.
- 4) 田中良昌, 新堀淳樹, 鍵谷将人, 堀智昭, 阿部修司, 小山幸伸, 林寛生, 吉田大紀, 河野貴久, 上野悟, 金田直樹, 米田瑞生, 田所裕康, 元場哲郎, 三好由純, 関華奈子, 宮下幸長, 瀨川朋紀, 小川 泰信, IUGONET 解析ソフトウェアの開発, 宇宙科学情報解析論文誌, 第 1 号, 91-98, 2012.
- 5) Angelopoulos, V., The THEMIS mission, Space Sci. Rev., Vol.141 (2008), pp.5-34, doi:10.1007/s11214-008-9336-1.
- 6) Miyoshi, Y., K. Seki, K. Shiokawa, T. Ono, Y. Kasaba, A. Kumamoto, M. Hirahara, T. Takashima, K. Asamura, A. Matsuoka, T. Nagatsuma and ERG Working Group, Geospace Exploration Mission: ERG Project, Transactions of the Japanese Society for Artificial Intelligence, Aerospace Technology Japan, Vol.8 (2010), pp.Tm_1-Tm_6.

IUGONET メタデータ登録・管理システムの処理性能評価

堀 智昭^{*1} 梅村宜生^{*1} 阿部 修司^{*2} 小山 幸伸^{*3} 田中 良昌^{*4} 林 寛生^{*5,*6}
上野 悟^{*7} 新堀 淳樹^{*5} 佐藤 由佳^{*4} 八木 学^{*8}

Performance study of IUGONET metadata management system

Tomoaki HORI^{*1}, Norio UMEMURA^{*1}, Shuji ABE^{*2}, Yukinobu KOYAMA^{*3}, Yoshimasa TANAKA^{*4},
Hiroo HAYASHI^{*5,*6}, Satoru UENO^{*7}, Atsuki SHINBORI^{*5}, Yuka SATO^{*4} and Manabu YAGI^{*8}

Abstract

This paper reports on the benchmark test of the metadata management system developed by the Inter-university Upper atmosphere Global Observation NETwork (IUGONET) project. The IUGONET metadata management system consists of the metadata reception module based on file repositories of Git, which is a widely-used version control software, and the metadata import module to register metadata to the metadata database using the DSpace commands. We made the performance assessments for these two modules in terms of the processing speed for registering/importing metadata both on a native Linux platform and a virtualized Linux platform. As a result, the metadata reception module processes properly even millions of registered metadata within several minutes at most, showing roughly the similar performance for the native and virtualized platform. On the other hand, the metadata import module needs much longer processing times of several to a few tens of hours for the same amount of metadata. As compared with the virtualized platform, the native platform gives a four times faster speed for updating indices for the built-in search engine based on Lucene. From a perspective of the regular operation of the metadata management system, the metadata import module runs more efficiently on the native Linux platform and it is estimated by our performance test that it takes about 5 hours to newly import 10^5 metadata files and subsequently takes about 21 hours to update the Lucene indices for the registered metadata of 5×10^6 in total.

Keyword : IUGONET, metadata, metadata management, Git, DSpace

概要

本論文では、大学間連携プロジェクト「超高層大気長期変動の全球地上ネットワーク観測・研究」(IUGONET)において、作成されたメタデータの受付やメタデータ・データベースへの登録を行っている「メタデータ登録・管理システム」のメタデータ登録処理に関する性能試験の結果について報告する。メタデータ登録・管理システムは、フリーのバージョン管理ソフトウェアである Git を利用したメタデータ受付部と、そこからメタデータ・データベースへメタデータを登録するメタデータインポート部の、2つの処理部で構成される。この2つの処理部の性能を評価するために、実際に作成・登録されたメタデータを用いて、物理マシンおよび仮想マシン上に構築した Linux 環境のそれぞれについて、メタデータの処理速度を計測した。その結果、Gitによるメタデータ受付部は物理マシンか仮想マシンかに依らず、メタデータ総数が500万件レベルまで増えても数分以内で処理が完了するのに対して、メタデータインポート部は処理時間が数時間から数十時間とかなり長くなり、特に検索エンジンである Lucene のインデックス更新については、物理マシンにおける実行速度が仮想マシンのそれより約4倍高速であることがわかった。メタデータインポート部を物理マシン上に実装した場合、メタデータ総数500万件では、新規登録メタデータ数10万件につき5時間弱プラス Lucene インデックス更新で21時間ほど処理時間がかかると見積もられる。

*1 名古屋大学太陽地球環境研究所 (Solar-Terrestrial Environment Laboratory, Nagoya University)

*2 九州大学国際宇宙天気科学・教育センター (International Center for Space Weather Science and Education)

*3 京都大学理学研究科附属地磁気世界資料解析センター

(Data Analysis Center for Geomagnetism and Space Magnetism, Graduate School of Science, Kyoto University)

*4 国立極地研究所 (National Institute of Polar Research)

*5 京大学生存圏研究所 (Research Institute for Sustainable Humanosphere, Kyoto University)

*6 富士通エフ・アイ・ピー株式会社 (FUJITSU FIP CORPORATION)

*7 京都大学理学研究科附属天文台 (Kwasan and Hida Observatories, Graduate School of Science, Kyoto University)

*8 東北大学理学研究科附属惑星プラズマ・大気研究センター

(Planetary Plasma and Atmospheric Research Center, Tohoku University)

1 はじめに

超高層大気長期変動の全球地上ネットワーク観測・研究 (IUGONET: Inter-university Upper atmosphere Global Observation NETwork)¹⁾ は、2009年度より6ヵ年計画で実施されているプロジェクトであり、国立極地研究所宙空圏研究グループ、東北大学大学院理学研究科地球物理専攻太陽惑星空間物理学講座並びに東北大学惑星プラズマ・大気研究センター、名古屋大学太陽地球環境研究所、京大大学生存圏研究所、京都大学理学研究科附属地磁気世界資料解析センター、京都大学理学研究科附属天文台、および九州大学国際宇宙天気科学・教育センターの5機関7組織が参加している。各研究機関が得意とする観測手段、データを持ち寄ることで、幾つかの領域が互いに様々な相互作用をしている超高層大気という領域での、各種物理量の長期変動について研究を行い、地球温暖化など地球規模での大気循環・変動への理解を深めていくことが、このプロジェクトの目的である。

IUGONET 参加研究機関は、これまでレーダー、磁力計、光学観測器、可視光・赤外・紫外線・電波望遠鏡などを用いた超高層大気環境の観測を数十年に渡って行っている。蓄積された大量のデータは各機関が独自に構築したデータベースにアーカイブされているが、データ自体やデータアクセスに関する情報の整備が不足していたり、また内容が非常に専門化されているため、同じ超高層物理分野であっても、別の領域を専門とする研究者にとって、必ずしも利用しやすい状態にはなっていない。その点を改善し、超高層大気の異なる領域に跨るような総合解析に基づく研究を促進するため、IUGONET では各機関が保有するデータに関する様々な情報(メタ情報、またはメタデータ)の抽出・収集を行い、それらをアーカイブして各種検索ができるようなデータベース、すなわちメタデータ・データベースの構築を行っている。またIUGONETでは、このメタデータ・データベースとして、オープンソースの学術リポジトリ用ソフトウェアであるDSpaceⁱをベースにしたシステムの開発を行っている²⁾。

各研究機関から提出されるメタデータ(実際にはSPASEデータモデル³⁾をベースにしたIUGONET共通メタデータフォーマットに沿って作成されたXMLファイル)の登録・管理のために、フリーの版管理ソフトウェアであるGit^jを利用し、Gitによるメタデータリポジトリを核としたメタデータ登録・管理システムを構築して運用している⁴⁾。このシステムによって、新規メタデータの受付や既登録メタデータの改変・修正の履歴管理を行っているほか、IUGONETメタデータ・データベースへの登録もこのメタデータ登録・管理システムを通じて行われる。

このメタデータ登録・管理システムは2010年より本格的に運用を開始し、各部の改良を行いつつ2年余りに渡って運用されている。本論文では、このメタデータ登録・管理システムの運用の現状と、これからも増えていくメタデータ処理に対して運用を最適化するために、その基礎データとなるシステムの性能試験の結果について報告する。またIUGONETでの運用のみならず、フリーソフトであるGitによるメタデータ管理やDSpaceベースのメタデータ・データベースシステムは、モジュール単位で他プロジェクトでも利用可能である。IUGONETでは、プロジェクトで開発したこれらのシステムを他プロジェクト・研究機関に積極的に提供することも行っており、実際に2011年度末より開始した情報通信研究機構からのメタデータ受け入れに伴い、同研究所にメタデータ登録・管理システムおよびメタデータ・データベース一式を提供し、実際に運用も行われている。このように外部プロジェクト・研究機関に提供するという観点からも、システム運用に関する基礎データをまとめておくことは非常に重要である。

2 IUGONETメタデータ登録・管理システムの運用

図1に、メタデータ登録・管理システムを通じて最終的にIUGONETメタデータ・データベースへ登録されたメタデータの総数の推移を示す。ここでは、メタデータ登録・管理システムとメタデータ・データベースとが連動した、定常運用開始(2010年1月)以降のメタデータ数をグラフ化しており、2012年6月末の時点で、メタデータ総数は340万件近くに達している。また図からわかるように、メタデータの増加の仕方は一様ではない。これは新規データセットのメタデータが不定期で登録されるため、その際に、データセットの規模にも依るが、メタデータ数が数千~数十万件程度急増することになる。またグラフがほぼ横ばいになっているように見える期間(例えば2011年7月辺り)でも実はメタデータ数は1日数百件程度ずつ増加しており、これはリアルタイムで観測データが取得されるのに伴って、メタデータが自動生成・登録されるためである。

メタデータ登録・管理システムを通じたメタデータの登録は、実際には大きくわけて4つの作業・処理を経て行われる。図2は、メタデータ作成担当者からメタデータ・データベースに登録されるまでのメタデータの流れを概略的に示している。図の右側がメタデータ登録・管理システムであり、左側が各研究機関でメタデータを作成する担当者のメタデータ作成環境を表している。現状

i <http://www.dspace.org/>

j <http://git-scm.com/>

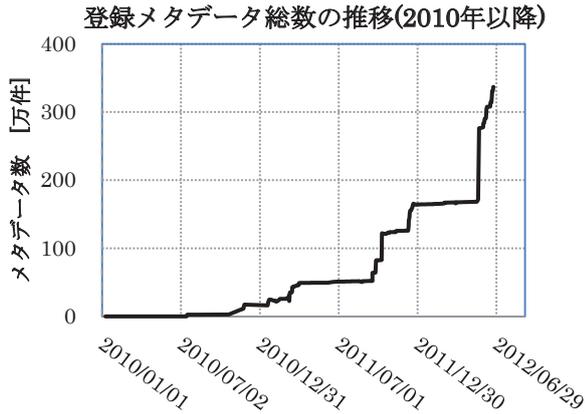


図1 IUGONET メタデータ・データベースに登録された総メタデータ数の推移(2010年1月～2012年6月)

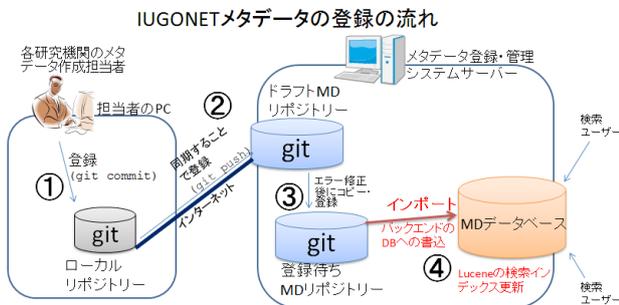


図2 IUGONET メタデータ登録・管理システムを通じたメタデータ提出・登録の流れ

のシステム構成では、メタデータ登録・管理システムとメタデータ・データベース本体は1台のLinuxサーバー上で運用されている。

登録・管理システムのうち、ドラフト/登録待ちメタデータ(MD)リポジトリで構成される部分をメタデータ受付部と呼び、また登録待ちMDリポジトリに入っているメタデータをメタデータ・データベースに登録する機能を担う部分(図中の赤字の部分)をメタデータインポート部と呼ぶ。登録・管理システム側のメタデータ受付部でGitを用いているため、作成担当者側でのメタデータの編集・履歴管理でもGitリポジトリが用いられる。なお、Gitリポジトリ運用の詳細については堀ほか⁴⁾を参照されたい。

メタデータが作成者の手元からメタデータ・データベースに登録されるまでに、主に4つのステップを経ることになる。まず、作成者のローカルな計算機上で作成されたメタデータを、ローカルなGitリポジトリに登録する(図中の①)。具体的には、作成者の計算機上でgit add/commitコマンドが実行されることになる。次にそのローカルリポジトリへの変更分を、メタデータ

登録・管理システムサーバー上のドラフトMDリポジトリとインターネット越しに同期させることで、サーバーに転送する(②)。このローカル-サーバー間の同期はgit pushコマンドによって行われる。このドラフトMDリポジトリ上で簡易的な文法・内容チェックを通過したメタデータは、同サーバー上にある登録待ちMDリポジトリへコピー・登録される(③)。最後に、この登録待ちリポジトリに入ったメタデータは、定期的に自動実行されるDSpaceのimportコマンドでメタデータ・データベースに登録され、また同時に、検索エンジンであるLuceneのインデックス更新が自動的に行われる(④)。このimportコマンドはJavaで実装されており、具体的にはメタデータ・データベースのバックエンドになっているPostgreSQLのDBテーブルへのメタデータの書き込みを行っている。このDBへの書き込みの後、Luceneインデックス更新が完了すると、メタデータ・データベースでの検索でメタデータがヒットようになる。

3 メタデータ登録・管理システムの処理性能評価

現状の登録・管理システムの運用では、図2の①～④が十数時間以内で完了するだろうという見通しのもとで、サーバー上でのメタデータ・データベースへの登録処理を1日1回の頻度で自動起動している。しかしながら、処理プロセスの多重起動を防いだり、作成者がサーバーに転送(図2の②)したメタデータがなるべく早くメタデータ・データベース上で検索可能となるよう登録処理ジョブスケジュールを最適化するためにも、上記の①～④の処理が実際の計算機上でどれくらいの時間を要するかを、あらかじめ見積もっておくことが重要である。また、処理能力の観点で、日々増え続けるメタデータ数に対して、この登録・管理システムがどれくらいのメタデータ総数および1回で登録するメタデータ量まで現実的な運用ができるのかについて、あらかじめ試験しておく、システムの限界などを把握しておくことは不可欠である。

この点については、既に河野ほか⁵⁾でメタデータ・データベースへの登録の部分についてのみベンチマークの結果が報告されている。しかしながら、河野ほか⁵⁾ではメタデータを1000件ずつ計100万件まで追加していった際の登録速度について調査を行ったのみであった。この調査では、Gitリポジトリへの登録、サーバー上のGitリポジトリとの同期に関する性能評価は実施されていない。

また現在では既にメタデータ総数が300万件を超えて

おり、また通常1回に登録されるメタデータ数は数万～数十万件におよぶこともあり、この点では前回の調査で前提としたメタデータ登録数と、現状が既にかげ離れつつあると言える。さらに、前回の調査時から、登録・管理システムサーバーのOSが新しいものに置き換わっており、かつメタデータ・データベース自体にも幾つかの機能追加・強化が施されており、その意味で現状の環境で再度ベンチマーク試験を行う必要があると考えられる。

また河野ほか⁵⁾では行われなかった、今回の新しい試みとして、従来の物理マシンでのシステムだけでなく、仮想マシン上にも同様のシステムを実装して試験を行った。これは、ここ数年でサーバー管理作業を容易にするため仮想化サーバーで各種サーバープログラムを運用する事例が増えてきており、IUGONETでも将来のサーバー運用形態の有力候補として、仮想サーバー検討しているからである。このサーバーの仮想化により、バックアップ作業や、異なる計算機環境への移行が容易になり可搬性が向上するというメリットが得られる。また、期限付きプロジェクトであるIUGONETが終了した後は、限られた人的リソースで、メタデータ登録・管理システムやメタデータ・データベースの運用・メンテナンスを行っていく必要があり、その意味で、サーバー管理の作業コスト削減を模索していくことは、IUGONETプロジェクト後期での重要な課題である。

表1 性能評価を行った計算機環境

| | |
|---------------|--|
| CPU | Intel Xeon 2.93GHz(物理 4core) × 2CPU |
| メモリー | 24GB |
| OS | Scientific Linux 6.2 64bit (Redhat Linux6 互換) |
| ベースとなる DSpace | ver. 1.7.0 |
| PostgreSQL | ver. 8.4.11 |
| Apache Lucene | ver. 2.9.3 |
| HDD | Hardware RAID 1 by Intel Workstation board S5520SC |
| 仮想化アプリケーション | Oracle Virtualbox ver. 4.1.14 |
| 仮想マシンのスペック | Science Linux 6.2 64bit, 2コア, メモリー 12GB, HDD200GB |

表1は、今回のベンチマーク試験を行ったLinux環境についてまとめたものである。メタデータ・デー

タベースのベースとなる DSpace や、バックエンドの PostgreSQL 及び Apache Lucene のバージョンは河野ほか⁵⁾での試験環境と同じであるが、河野ら⁵⁾が使用した Redhat Enterprise Linux 5 互換である CentOS 5.3 64bit 版から、Redhat Enterprise Linux 6 互換である Scientific Linux 6.2 64bit 版に置き換わっている。また、仮想マシンと物理マシンとの性能差を評価するために、今回の試験では Git 及びメタデータ・データベースシステムを物理マシン上と仮想マシン上に構築し、それぞれの環境で同じベンチマーク試験を行った。仮想マシンの構築には、Oracle 社の Virtualbox⁴⁾を用いて、その上にホスト OS と同じ Scientific Linux 6.2 およびアプリケーションをインストールすることで、ソフトウェア的に同一のシステムを準備した。幾つか存在する仮想化ソフトの中で Virtualbox を採用している理由は、フリーであり、かつ他の仮想化ソフトではサポートされないことが多い Mac や Solaris をホスト OS として使うことができ、仮想マシンの可搬性というメリットを最大限享受できるからである。また現時点では Virtualbox を採用して試験を行ったが、他の仮想化ソフトも開発が進められており、将来より良いソフトが現れれば、代わってそちらを使用して運用することも視野に入れるべきである。

3.1 Git リポジトリへのメタデータ登録に関する性能評価

まず、Git リポジトリへのメタデータ登録処理にかかる時間を計測することで、Git リポジトリの性能評価を行った。具体的な試験の方法は、表1のスペックの物理マシン上に空の Git リポジトリを作成してローカルリポジトリとして、さらに上記の物理マシンと同一の物理サブネットワーク内にある別の Linux 機上の空ディレクトリを Git リポジトリとして初期化することでリモートリポジトリとした。次にローカルリポジトリに 10000 件の XML 形式のメタデータをコピーして、git add/commit コマンドで登録、さらに git push コマンドを ssh 経由で使うことでリモートリポジトリと同期、という工程を繰り返し実行し、この3種の処理にかかった時間を計測した。また同様の計測を、物理マシンと仮想マシン (Virtualbox) 上の両方でを行い、実行時間の比較を行った。

物理マシンで行った試験の結果を図3に示す。計測は、空のローカルリポジトリに 10000 件ずつメタデータを追加登録していき、総登録数が 500 万件に達するまで行った。この結果から明らかなのは、メタデータファイルのコピー及びリモートリポジトリとの同期は、その

⁴⁾ <https://www.virtualbox.org/>

時の登録メタデータ総数に依らず、常に数秒以内で完了していることである。一方、ローカルリポジトリへの登録は、その時のメタデータ総数に比例して実行時間が長くなっていることがわかる。これは、Gitの動作として、ローカルリポジトリへの登録時にリポジトリ内の全ファイルをスキャンして更新の有無を確認するために、リポジトリ内のメタデータが増えてスキャンするファイル数が増えれば増えるほど実行に時間がかかる、と考えると理解できる。

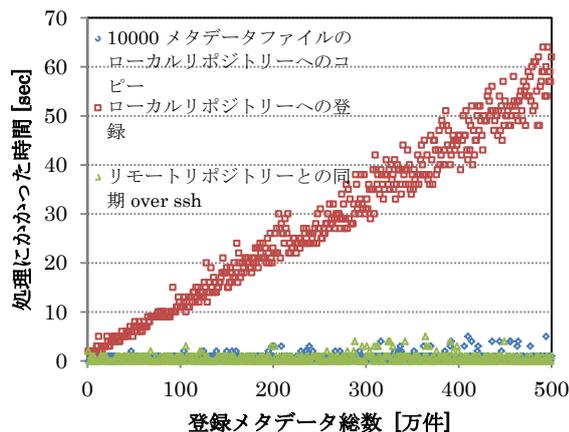


図3 物理マシン上において、10000個のメタデータファイルのコピー、ローカルGitリポジトリへの登録、及びリモートリポジトリとの同期を繰り返した際の、各処理の実行時間。

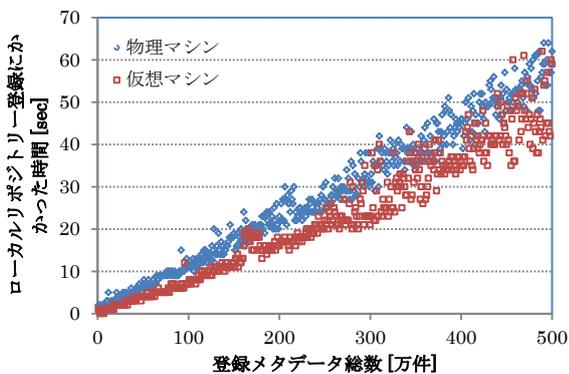


図4 物理マシンおよび仮想マシン (Virtualbox) 上のLinux環境での、ローカルリポジトリへの登録にかかる時間の比較

次にローカルリポジトリへの登録にかかる時間について、物理マシンと仮想マシン (Virtualbox) とで比較した結果を図4に示す。仮想マシンについても物理マシンで行ったのと全く同じ処理を行った。実際にはファイルコピーとリモートリポジトリとの同期についても計測を行ったが、これらは物理マシンと同様に数秒以内であったので、違いが顕著であったローカルリポジトリ

登録の結果のみ図示している。

得られた結果は予想に反し、物理マシンよりも仮想マシンでの処理の方が20-30%ほど高速であった。これには2つ理由が考えられる。1つは、仮想マシン (Virtualbox 上のLinux) では必要最低限のプロセスしか動いておらず、またネットワークを通じた外部からの通信もないため、余計な負荷が全く無い状態で試験を行ったため、今回の試験単体は速く実行できた、という可能性が考えられる。またもう1つ原因として考えられるのは、仮想マシンでのファイルアクセスは、ホストOSのメモリなどを使った読込/書込処理の高速化が図られているかもしれないということである。今回の試験で用いた仮想マシンには12ギガバイトという比較的大きなメモリが割り当てられているのに対して、10000件のメタデータXMLファイルの総量はわずか20メガバイト程度であるので、新規に書き込まれたメタデータファイルがキャッシュ機能の範囲内で扱われ、それゆえに全体としてGitの処理が高速化しているというのは十分考えられることである。これはVirtualboxという仮想化アプリケーションの内部仕様に関わることなので、残念ながらこのことを裏付ける資料をOracle社サイト等で見つけることはできなかった。しかし、もしメモリ等を利用したキャッシュ機能でファイルアクセスを見かけ上高速化しているとすれば、そのキャッシュ量の限界以上のアクセスをしようとすると、恐らく本来のnative環境より遅い速度になると考えられる。

実際の運用という観点では、図2でのメタデータの登録処理の中で、Gitに関する処理が伴うのは①～③である。今回の性能試験より、メタデータ総数が500万件でもGitリポジトリは正常に機能し、また処理時間は①～③合計で、せいぜい数分程度と見積もることができる。ただし、②の処理については、ローカルリポジトリとリモートリポジトリとの間のネットワーク速度にも依存すると考えられる。しかしながら、後述のように、④の処理時間の方がずっと長いので、メタデータ処理にかかる全体の時間の中で①～③が占める割合は非常に小さいと言える。

3.2 メタデータ・データベースへの登録に関する性能評価

次に、図2の④の部分に当たる、メタデータ・データベースへの登録に関する性能評価の結果について述べる。この処理は2つの要素とそのため準備作業にわけることができる。それぞれ、メタデータ・データベースのバックエンドであるPostgreSQLのDBテーブルへのメタデータの書込、検索エンジンであるLucene用の内部インデックスの更新、及びこの2つの処理を行うための

Java 実装のコマンドの実行スクリプト作成 (Ruby¹ スクリプト及び Unix shell スクリプトで実装) などの前後処理がある。ただし、メタデータ・データベースへの登録は非常に時間がかかるため、前述の Git リポジトリの性能試験とは異なり、10 万件ずつメタデータを登録していった時の、各処理の実行時間を計測している。それでも、現在のところ物理マシンについては 130 万件、仮想マシンについては 150 万件までしかデータが得られていない。この試験は現在も継続中であり、少なくとも Git リポジトリと同じ 500 万件までは試験を続ける予定である。以下では、それぞれ 130 万件および 150 万件までのデータに基づいて議論を行う。

図 5 は、物理マシンで行ったメタデータ・データベースへの登録に関する処理の実行時間の内訳と、登録メタデータ総数への依存性を示したものである。青丸が PostgreSQL の DB テーブルへの書き込みにかかった時間、赤の四角がその後に実行される Lucene インデックス更新にかかった時間、緑の三角が登録の全処理にかかった時間を表す。それぞれのメタデータ総数において、青丸と赤い四角の和は緑の三角が表す時間の約 70-80% 程度となっており、このことから、Ruby+shell スクリプトによる前後処理にかかる時間は、登録にかかる全時間のうちの 20-30% であることがわかる。

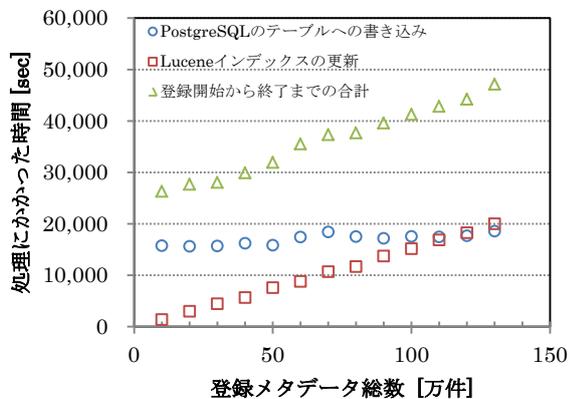


図 5 物理マシンでのメタデータ・データベースへの登録に関する処理の実行時間の内訳と、メタデータ総数への依存性

また図 5 より、PostgreSQL テーブルへの書き込みはメタデータ総数に依らずほぼ一定なのに対して、Lucene インデックス更新にかかる時間はメタデータ総数にほぼ比例して長くなっていることがわかる。この違いは、定性的には両者の処理の性質の違いとして理解できる。つまり、PostgreSQL テーブルへの書き込みはまさにその時に登録するメタデータの量で決まるので、毎回 10 万件とい

う同じ件数を処理しているがゆえに実行時間は毎回同じとなる。逆に Lucene インデックス更新は新しいメタデータの追加によりインデックス構成自体が変わることがあり、その場合 DB テーブル中の全メタデータを読みなおしてインデックスを再構築する必要があるため、実行時間が登録済みのメタデータ総数に比例することになる、と考えられる。また重要なのは、総数 110 万件的の辺りで、PostgreSQL テーブル書き込みにかかる時間と、Lucene インデックス更新にかかる時間の大小が逆転している。つまり、システムとしての運用を考えた場合、メタデータ数が多くなってくると、処理時間としてネックになるのは Lucene インデックス更新になるということである。

次に同じ性能試験を Virtualbox 上の仮想マシンで行った場合の結果を図 6 に示す。結果として、定性的には同じようなメタデータ総数への依存性を示しているが、Lucene インデックス更新にかかる時間が、物理マシンのそれと比べて約 4 倍に伸びている。この結果は、Virtualbox 上でのファイルアクセス、特に仮想ディスクへのファイル書き込みに何らかのキャッシュ機能が介在しているという前述の推測と整合する。もちろんこれは Virtualbox の内部仕様が不明であるがゆえに想像の域を出ないのだが、PostgreSQL テーブルへの書き込み量はメタデータ 10 万件ということから約 200 メガバイト程度とそれほど大きな量ではないが、登録メタデータ総数が 100 万件を超えると Lucene インデックスは数十ギガバイトのオーダーとなり、またインデックス作成に伴い大量の新規ファイルを作成し実際に仮想ディスクに書き込むことになるため、キャッシュ機能で吸収できなくなるのではないかと考えることができる。その結果、仮想マシンではメタデータ総数が 30 万件を超えると Lucene インデックス更新の方が処理時間が長くなり、150 万件的段階ではメタデータ登録の全時間のうちの 8 割強がこのインデックス更新に費やされることになる。

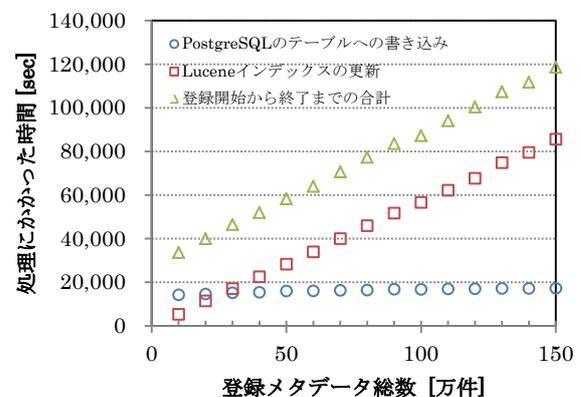


図 6 仮想マシン (Virtualbox) でのメタデータ・データベース登録処理にかかった時間の内訳

¹ <http://www.ruby-lang.org>

4 考察

以上の性能試験から、現状のメタデータ登録・管理システムの運用方針について、幾つかの見通しを得ることができる。

まず、図2に示す①～④の処理のうち、処理時間の観点で一番のネックになるのは④のメタデータ・データベースへの登録の部分であり、かつその中でも Lucene インデックス更新の処理が支配的である。また通常はメタデータの新規追加量はせいぜい数十～数百件/1日であることを考えると、このインデックス更新の処理時間が、登録・管理システムの実行可能な最短サイクル、つまりメタデータ・データベースの更新可能頻度の上限を決めることになる。

例えば今回の物理マシンを用いて、かつメタデータ総数が130万件の場合、書込メタデータ数に比例すると予想される PostgreSQL テーブル書込速度は約17000秒(5時間弱)/10万件、Lucene インデックス更新は総数130万件に対して約20000秒(6時間弱)かかり、前後処理を含めると計13時間かかることになる。つまり、この場合に Git リポジトリ経由で新規メタデータを10万件登録しても数分程度しかかからないが、それが実際にメタデータ・データベースで検索可能となるまでに、理論的には最短で13時間かかるという計算である。

また今回の性能試験は総数150万件までしかデータが取得できていないが、2012年5月下旬に実際に IUGONET で運用中のメタデータ・データベース(現状は物理マシンで運用)に一度に約105万件の新規メタデータが登録され、結果として登録総数は約270万件となった。この時の記録によれば、PostgreSQL テーブルの書込と Lucene インデックス更新に、それぞれ約144,000秒(約40時間)、47000秒(約13時間)かかっており、今回の試験の結果からそれぞれ線形外挿して得られる値におおよそ近いものになっている。この線形外挿からメタデータ総数が500万件に達した場合を予想すると、Lucene インデックス更新にかかる時間は約21時間となる。1つ注意すべきことは、追加メタデータ数が少なくともインデックス更新の時間は変わらない(登録総数に比例)ので、例えば追加メタデータが1件であろうが1万件であろうが、結果として総数500万件であればインデックス更新にはやはり約21時間かかってしまうということである。

この21時間という値は、システムの性能的にメタデータ総数500万件あたりからデータベースの毎日更新が間に合わなくなるということを暗示している。ここで、実際にメタデータ・データベースの機能として要求される更新頻度との兼ね合いで評価する必要があるが、

IUGONET で取り扱う観測データは毎日生産されるものが多く、可能であればメタデータの方も毎日更新され新規観測データのメタデータが即座に検索可能となるのが望ましい。もちろんデータベース更新の頻度は高ければ高いほどよいのであるが、現実的には、例えば毎日1回の観測データのデータベース更新にメタデータ・データベースの方が遅延なく同期できるかどうか、1つの目安と言えよう。この観点で今回の結果を評価すると、現状のシステム性能では、仮想マシンに比べて高速な物理マシン上のシステムを用いたとしても、メタデータ総数が500万件というレベルになると、更新ジョブを走らせる頻度を1日1回以下の頻度に下げざるをえなくなり、つまり毎日更新には追従できないことを意味している。

このように、メタデータ総数が増えてきた時に Lucene インデックス更新にかかる時間が益々ネックになるという予想を踏まえると、この部分の処理を速く行うためには、メタデータインポート部を含むメタデータ・データベース本体については、仮想マシンではなく、物理マシンで運用する方がよいと思われる。ただし仮想化によって得られるシステム保守等の利便性も大きく、また将来的にハードウェア性能が向上すれば実際の処理時間は現実の運用では問題にならない程度に短縮されるかもしれない。このように、実際のデータベース運用でどちらを選択するかは、将来のハードウェア性能向上に合わせて継続的な評価を行いつつ、要求されるデータベースの更新性能と、仮想化による利便性向上との兼ね合いで判断されるべきである。

またメタデータ・データベース更新処理の高速化について、IUGONET 開発チームでも幾つかの施策を検討している。各部の処理を少しずつ改善していくなどの地道な手段の他に、メタデータ・データベースを新バージョンの DSpace ベースに移行することも検討しており、それによって PostgreSQL などのバックエンドが新バージョンとなり、より効率的なデータベース更新処理ができるようになると期待される。

5 結論

IUGONET プロジェクトで作成したメタデータを管理しメタデータ・データベースへの登録を行うメタデータ登録・管理システムに関して、その中枢機能を構成している Git リポジトリの処理とメタデータ・データベースへの登録処理に関する性能評価を行った。Git リポジトリでの処理はメタデータ総数が500万件レベルになってもほぼ数分以内で完了し、処理速度は物理マシンでも仮想マシンでもおおよそ同じくらいであると推測さ

れる。一方メタデータ・データベース登録部分の処理は千倍以上の時間がかかり、特にメタデータ総数が数百万件レベルになってきた場合、処理時間上ネックになるのは Lucene インデックス更新処理である。この Lucene インデックス更新については、物理マシンの方が仮想マシンより約4倍高速に処理できることがわかった。

寛生, 新堀淳樹, 田中良昌, 鍵谷将人, 上野悟, 金田直樹, 田所裕康, DSpace を用いた超高層物理学のためのメタデータ・データベースの構築, *Proc. of the 3rd Forum on Data Engineering and Information Management*, 2011, C8-5, <http://db-event.jpn.org/deim2011/proceedings/pdf/c8-5.pdf>.

謝辞

大学間連携プロジェクト「超高層大気長期変動の全球地上ネットワーク観測・研究」は、文部科学省特別教育研究経費(研究推進)[平成21年度]および特別経費(プロジェクト分)[平成22年度~]の交付を受けて、平成21年度より6ヵ年計画で実施している事業である。IUGONET プロジェクトで使用される観測データのデータベース構築に携わった全ての方々に感謝する。特に、国立極地研究所宙空圏研究グループのデータ整備に尽力下さった余川 真純氏, 京大大学生存圏研究所のデータ整備に尽力下さった橋口 典子氏, 京都大学理学研究科附属天文台のデータ整備に尽力下さった金田 直樹氏に感謝の意を表す。

参考文献

- 1) 林 寛生, 小山幸伸, 堀 智昭, 田中良昌, 新堀淳樹, 鍵谷将人, 阿部修司, 河野貴久, 吉田大紀, 上野 悟, 金田直樹, 米田瑞生, 田所裕康, 元場哲郎, 大学間連携プロジェクト「超高層大気長期変動の全球地上ネットワーク観測・研究」, 宇宙科学情報解析論文誌, 第1号, 2012年3月, 113-120.
- 2) 小山幸伸, 河野貴久, 堀 智昭, 阿部修司, 吉田大紀, 林 寛生, 田中良昌, 新堀淳樹, 上野悟, 金田直樹, 米田瑞生, 元場哲郎, 鍵谷将人, 田所裕康, 超高層物理学分野の為のメタデータ・データベースの開発, 宇宙科学情報解析論文誌, 第1号, 2012年3月, 99-104.
- 3) King, T., J. Thieman and D. A. Roberts, SPASE 2.0: a standard data model for space physics, *Earth Sci. Inform.*, vol. 3, 2010, 67-73.
- 4) 堀 智昭, 鍵谷将人, 田中良昌, 林 寛生, 上野 悟, 吉田大紀, 阿部修司, 小山幸伸, 河野貴久, 金田直樹, 新堀淳樹, 田所裕康, 米田瑞生, IUGONET 共通メタデータフォーマットの策定とメタデータ登録管理システムの開発, 宇宙科学情報解析論文誌, 第1号, 2012年3月, 105-111.
- 5) 河野貴久, 小山幸伸, 堀 智昭, 阿部修司, 吉田大紀, 林

宇宙天気情報サービスとサイエンスクラウド

亘 慎一^{*1} 加藤 久雄^{*1} 村田 健史^{*1} 山本 和憲^{*1} 渡邊 英伸^{*1}
久保田 康文^{*1} 國武 学^{*1}

Space Weather Information Services and Science Cloud

Shinichi WATARI^{*1}, Hisao KATO^{*1}, Ken. T. MURATA^{*1}, Kazunori YAMAMOTO^{*1},
Hidenobu WATANABE^{*1}, Yasubumi KUBOTA^{*1} and Manabu KUNITAKE^{*1}

Abstract

Variation of space environment, which causes failure of manmade infrastructure, such as artificial satellites, is called space weather. Research of space weather covers vast space. Observation by spacecraft is important for space weather. However, it is difficult to cover whole area of the space only by the observation. We need a new computing environment, which enables to analyze observation data and simulation data together.

In this paper, we report services of space weather based on “Space Weather Cloud”, which is composed by distributed mass storage system using the Grid Datafarm (Gfarm) architecture technology, high performance computer, servers for visualization using AVS and IDL programs, RCM (R&D Chain Management) System for job service, servers for Solar-Terrestrial data Analysis and Reference System (STARS), and so on.

Keyword: space weather, science cloud, cloud computing, informatics

概 要

人工衛星など社会的なインフラに障害を発生させるような宇宙環境の変動を扱う宇宙天気では、太陽から地球周辺の宇宙空間までの広大な領域を扱う必要がある。宇宙機による観測は重要であるが、この広大な領域を観測データだけでカバーするのは困難である。そこで、観測データと数値シミュレーションデータを統合的に処理してサービスを提供できる情報プラットフォームの構築が必要となる。情報通信研究機構が構築している Gfarm による大容量分散ディスクシステム、スーパーコンピュータ、AVS や IDL などインストールした可視化サーバ群、ジョブサービスを行う RCM (R&D Chain Management) System, 様々な観測データをダウンロードしてプロットや解析を行う STARS (Solar-Terrestrial data Analysis and Reference System) のサーバなどからなる「宇宙天気クラウド」を利用した宇宙天気の情報サービスについて報告する。

*1 情報通信研究機構
(National Institute of Information and Communications Technology)

1. はじめに

宇宙環境の変動により宇宙機の障害、衛星測位システムへの影響、短波通信障害、送配電システムの障害などが発生することがある。このような人工衛星などの社会システムに影響を与えるような宇宙環境の変動は「宇宙天気」と呼ばれている¹⁾。情報通信研究機構では、国際宇宙環境サービス (ISES: International Space Environment Service) と呼ばれる国際的な組織の日本の宇宙天気予報センターとして宇宙天気情報のサービスを行っている。ISESの予報センターは、オーストラリア (シドニー)、ベルギー (ブリュッセル)、ブラジル (サンジョゼ・ドカンポス)、カナダ (オタワ)、チェコ共和国 (プラハ)、中国 (北京)、インド (ニューデリー)、日本 (東京)、韓国 (チェジュ)、ポーランド (ワルシャワ)、ロシア (モスクワ)、スウェーデン (ルンド)、南アフリカ (ヘルマナス)、アメリカ合衆国 (ボルダー) の14カ国にあり、宇宙天気情報サービスを行っている。図1に各予報センターから出されているフレアや地磁気活動などの予報の例と情報研究機構で提供しているその予測精度に関するWebページ (<http://swc.nict.go.jp/forecast/>) を示す。

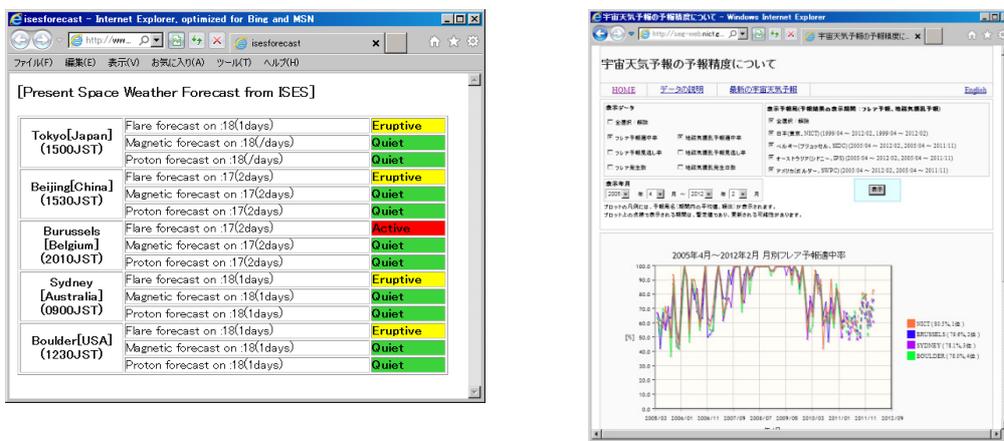


図1 各予報センターからの予報の例 (左図) とその予測精度に関するWebページ (右図)

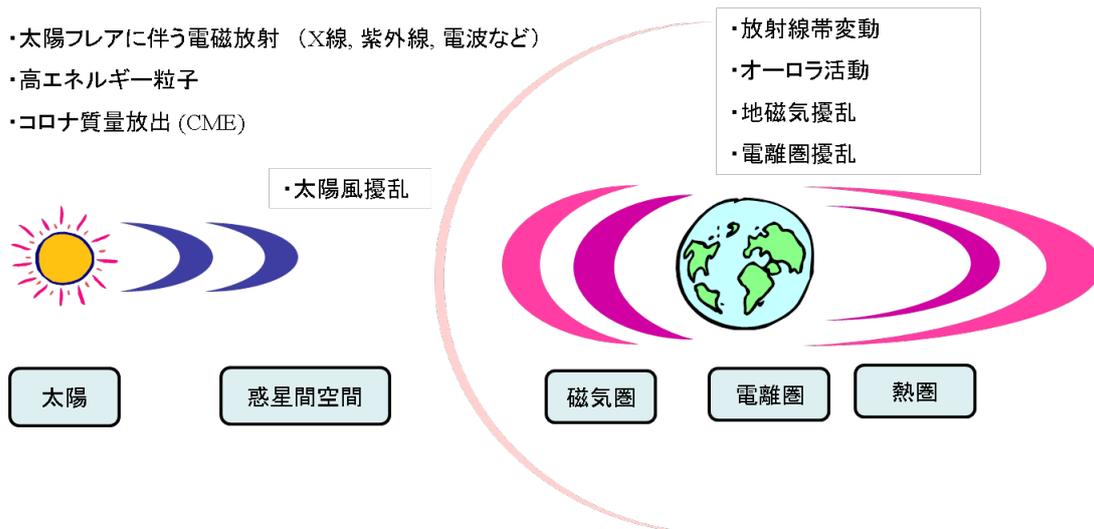


図2 宇宙天気のカバーする領域

宇宙天気は、図2に示すように、太陽、惑星間空間、磁気圏、電離圏、熱圏など太陽から地球までの広大な領域を扱う。さらに、宇宙天気に影響を与える太陽活動は、約11年の周期での変動に加えて、数百年、数千年という長期的な変動をしているので長期間のデータを扱う必要がある。そのため、宇宙天気の研究やサービスにおいては、観測データだけでなく、観測データと数値シミュレーションデータを融合した新たな情報処理環境の構築が必要である^{2,3)}。そこで、情報通信研究機構では、「宇宙天気クラウド」と呼ぶ、インフォマティクス技術を活用した新たな情報処理環境の構築を進めている。本報告では、宇宙天気クラウドを利用した宇宙天気情報サービスについて述べる。

2. 宇宙天気クラウドについて

現在、Grid Datafarm (Gfarm) アーキテクチャーによる大容量分散ディスクシステム^{4,5)}を中心として、スーパーコンピュータ、AVSやIDLなどをインストールした可視化サーバ群、ジョブサービスを行うR&D Chain Management (RCM) System, 様々な観測データをダウンロードしてプロットや解析を行う太陽地球系観測データ解析参照システム (STARS : Solar-Terrestrial data Analysis and Reference System) のサーバ、様々なデータのメタデータを自動収集するためのサーバ、動画配信のためのサーバなどが立ち上がっている (図3参照)。例えば、「宇宙天気クラウド」を利用することにより、数値シミュレーション結果の可視化や解析を一気通貫に行うことができる。「宇宙天気クラウド」の利用はインターネットからも可能であるが、新世代通信網テストベッド (JGN-X)⁶⁾の高速ネットワークによる接続も可能である。

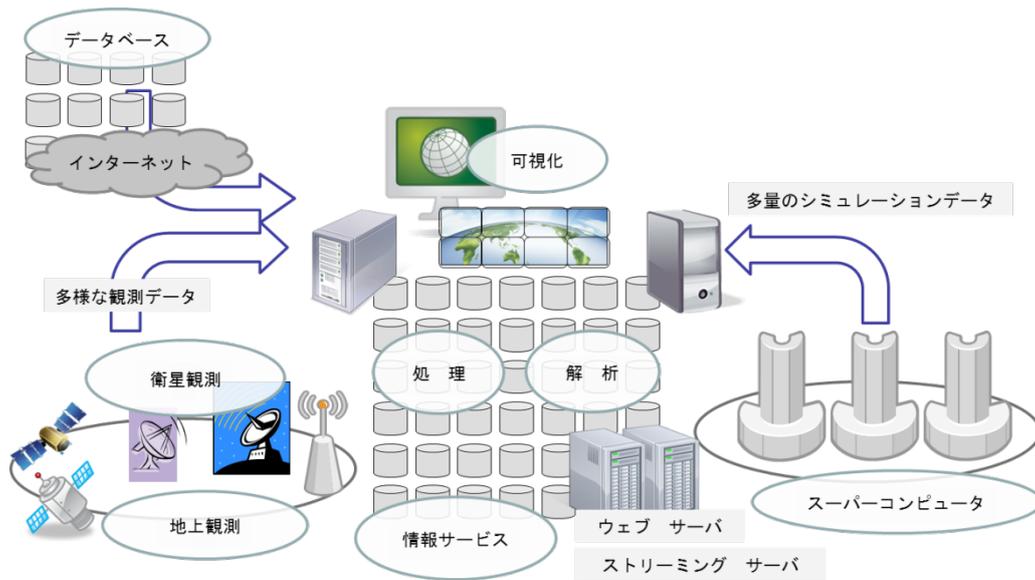


図3 宇宙天気クラウドの概念図

3. 宇宙天気情報サービスについて

「宇宙天気クラウド」を利用した宇宙天気情報サービスは、図4に示す e-SW の Web ページ (<http://e-sw.nict.go.jp>) から提供される。以下では、それらのサービスに関して具体的に紹介する。



図4 e-SW の Web ページ (<http://e-sw.nict.go.jp>)

3.1 メタ情報の収集

宇宙天気クラウドでは、インターネット上で公開・提供されている衛星観測や地上観測による様々なデータのメタデータの自動収集とデータベースへの登録に NiCTy (Network Infrastructure for data Collection Technology) と呼ばれるシステムを利用している⁷⁾。NiCTy は、図5に示すように http, ftp, sftp の通信プロトコルにより定期的にデータ公開・提供サイトにアクセスして情報を収集し、メタデータの生成・データベースへの登録を行う。また、データ提供サイトのサーバ上で NiCTy のアプリケーションを稼働させて、対象となるデータファイルを定期的にスキャンしてメタデータを自動的に生成し、生成されたメタデータを収集してデータベースへの登録を行う。NiCTy では、RSS1.0 (RDF Site summary) によりデータの見出し、データに関する説明の要約、更新時刻などの情報を xml 形式でメタデータとして記述している。

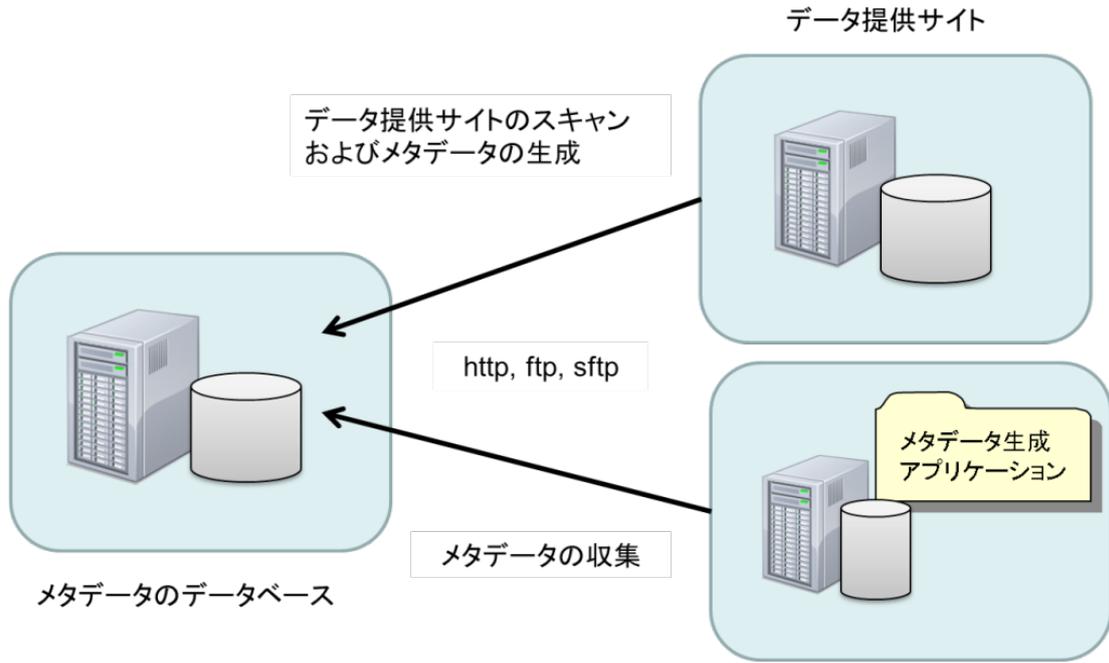


図 5 NiCTyによるメタデータの収集とデータベースへの登録

3.2 宇宙天気ボード

宇宙天気情報の利用者は、宇宙機の運用、衛星を使った測位、短波電波を使った通信、電力関係など様々であり、必要とされるデータも利用者によって異なる。例えば、静止軌道衛星の運用では、シングルイベントの原因となる太陽高エネルギー粒子や帯電の原因となる放射線帯粒子の情報が必要となる。一方、電力関係では、地磁気変動に伴う送電線への誘導電流の情報が必要である。このため、利用が自身に必要な情報を選択できるような情報のカスタマイズ機能を持つアプリケーションが便利である。そのようなニーズを考慮して、利用者が必要な宇宙天気データを選択してボード上にアレンジできるようにしたのが、図6に示す「宇宙天気ボード」である。

利用者は、右側のパネルを使って必要なデータを選択し、ボード上の好きな位置に好きな大ききでアレンジすることができる。データを選択する際にデータの更新時間を指定することができる。データに関しては、インターネットからの収集やNiCTyの機能と連動して収集することが可能である。利用者は、作成したデータのアレンジをサーバに保存することができる。この機能により、登録されたデータアレンジの利用者間での共有やあらかじめ作成されたテンプレートのデータアレンジの利用が可能となっている。

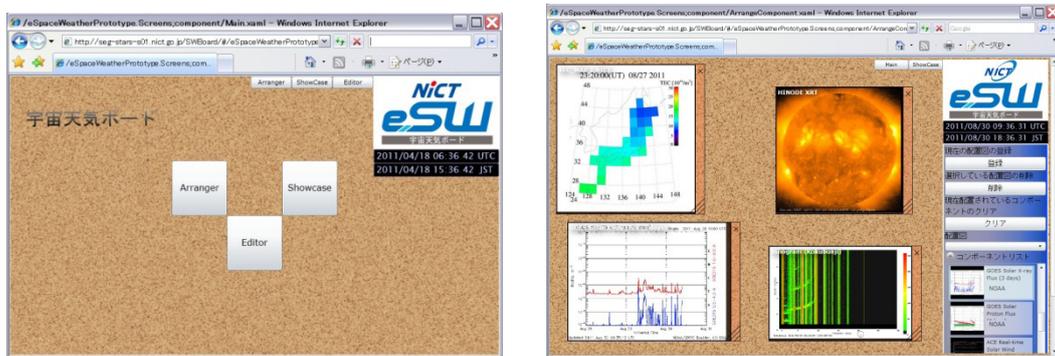


図 6 宇宙天気ボードのトップページ（左図）とアレンジされたデータを表示するページ（右図）

3.3 宇宙天気リアルタイムシミュレーション 3次元可視化

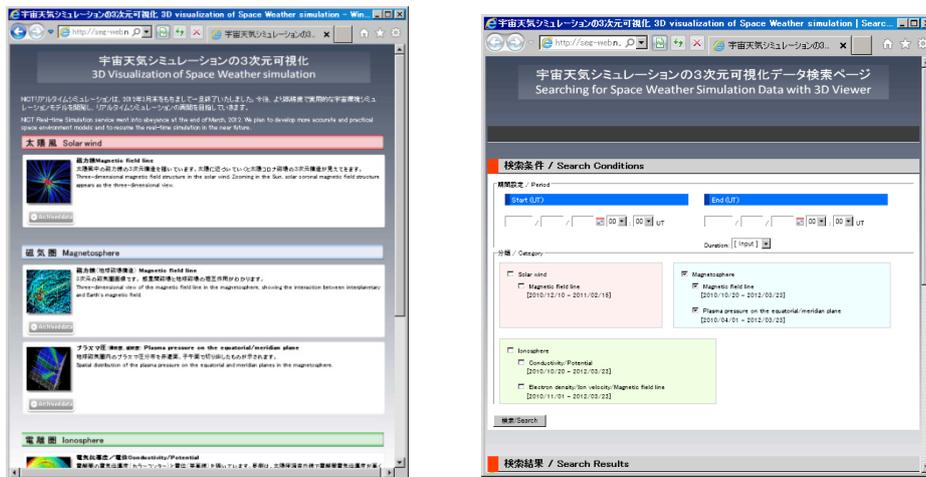


図7 宇宙天気リアルタイムシミュレーション 3次元可視化 Web ページのトップページ（左図）と検索ページ（右図）

情報通信研究機構では、宇宙天気の研究開発の一環として、太陽・太陽風、磁気圏、電離圏のリアルタイムシミュレーションを実施していた。それらの結果をクラウド内のストレージに保存し、「Virtual Aurora」と呼ばれる AVS を用いた 3次元可視化システムにより可視化して図7の Web ページから提供するというものである。図8に地球磁気圏シミュレーションの3次元可視化の例を示す。可視化されたシミュレーション結果は、リアルタイムのものだけでなく過去のものを検索して利用することもできる。宇宙天気のリアルタイムシミュレーションは、高精度モデルの開発のために2012年3月で終了したが、図7の Web ページから検索することにより過去の3次元可視化されたデータを参照することができる。

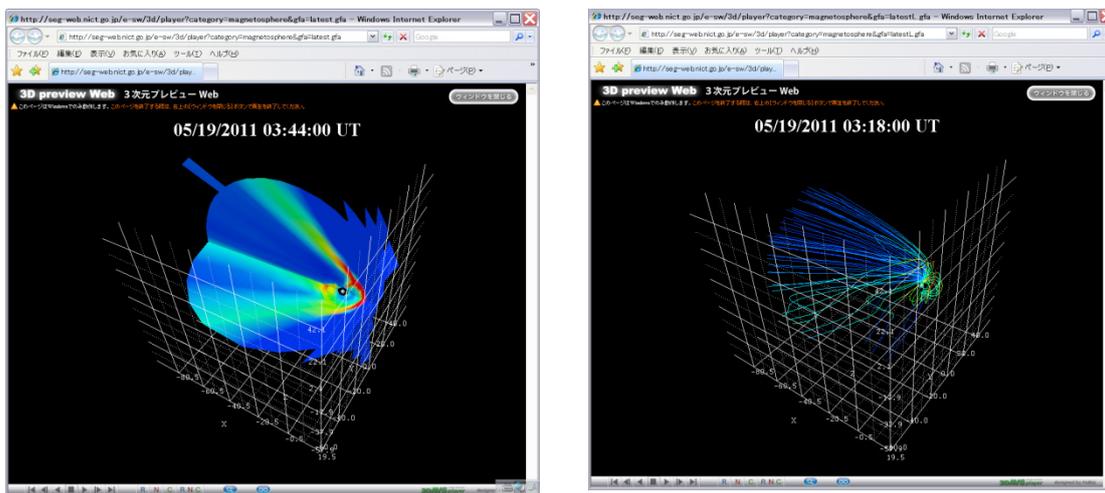


図8 地球磁気圏シミュレーションの3次元可視化の例

3.4 週刊宇宙天気ニュース



図9 週刊宇宙天気ニュース

一週間の宇宙天気の状況に関するサマリーと宇宙天気に関係する現象や用語などについての豆知識を動画で一般向けに配信するのが、図9に示す週刊宇宙天気ニュースである。週刊宇宙天気ニュースは、木曜日に収録を行い、毎週金曜日の午後に新しい動画番組がリリースされる。宇宙天気クラウド内の windows サーバによる HD 動画の配信や iPod 向けの動画の配信だけでなく、YouTube (<http://www.youtube.com/user/nictchannel>) からも動画の配信を行っている。宇宙天気豆知識に関して、現在、太陽黒点、太陽活動サイクル、電離圏、イオノゾンデ、デリンジャー現象、CME (コロナ質量放出)、磁気圏、SOHO 衛星、オーロラ、スポラディック E、太陽風、プロトン現象、放射線帯、F10.7、太陽フレア、GPS による TEC 観測、太陽電波バースト、プラズマバブル、地磁気指数、地磁気嵐、電離圏嵐の 21 のタイトルがリリースされている。

3.5 太陽地球系観測データ解析参照システム (STARS)

それぞれの研究機関や大学の研究室に分散している太陽地球系の観測データなどのメタデータを収集してデータベース化することにより、必要なデータを横断的に検索してダウンロードし、プロットの作成など解析することが可能になる。太陽地球系観測データ解析参照システム (STARS)⁸⁾ は、データベース化されたメタデータを使って、データプロットを作成するツールである (図10参照)。プロットされたデータの情報は、SPL(STARS Project List) と呼ばれるファイルに保存することができ、SPL ファイルを交換することにより、プロットの情報を共有することができる。また、STARS では、宇宙天気関連のイベントの登録や参照をすることができる。3.1 で述べた NiCTy が STARS でプロットに使うデータのメタデータの収集を行っている。

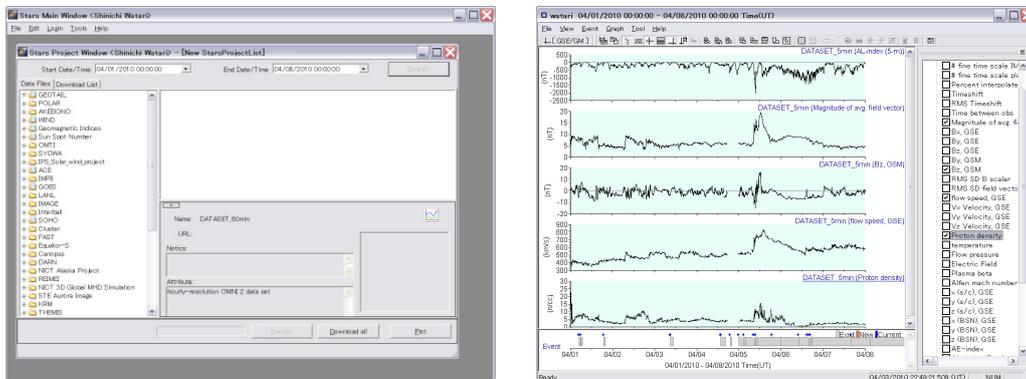


図10 STARS の期間・データ選択ウィンドウ (左図) とデータプロットの例 (右図)

3.6 ダウンロードサービス

図 11 に示すダウンロードサービスの Web ページでは、前述した STARS などのアプリケーションや情報通信研究機構がイオノゾンデと呼ばれる観測装置により国内で行っている電離圏の定常観測により得られる電離圏パラメータの月報や南極で行っている観測による電離圏パラメータの年報などをダウンロードすることができる。

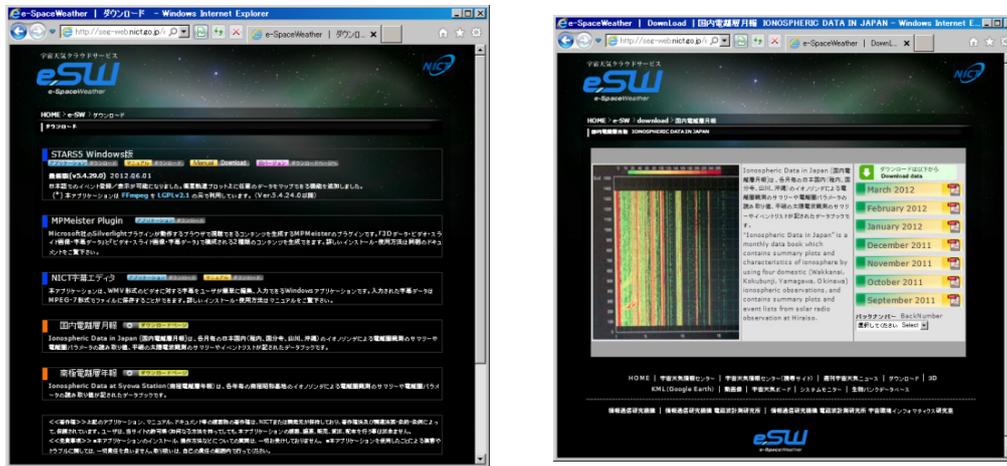


図 11 ダウンロードサービスの Web のトップページ (左図) と電離圏データ月報のダウンロードページ例 (右図)

4. まとめ

太陽から地球までの広大な領域のデータを扱う宇宙天気では、メタデータの活用により衛星観測や地上観測のデータを網羅的に収集する必要がある。一方、観測だけで全領域をカバーすることは不可能であり、数値シミュレーションとの連携が必須である。新たな衛星や地上観測網のデータ、数値シミュレーションのデータなどにより、宇宙天気での扱うデータ量は年々増え続けている。これらのデータを効率よく処理してデータ解析や情報サービスを行うためには、観測データと数値シミュレーションデータを統合的に処理できる「宇宙天気クラウド」のようなプラットフォームが必要である。「宇宙天気クラウド」を用いたデータインテンシブな解析^{9,10,11)}により、新たな知見の抽出や新たな宇宙天気情報サービスの創出を期待している。

太陽、太陽風、磁気圏、電離圏、熱圏などの領域を統合的に扱う宇宙天気においては、ネット上で公開されているデータベースを連携させる Linked Data¹²⁾ による領域の垣根を越えたコンテンツ間やサイト間のデータ統合の試みも重要だと思われる。また、大きな太陽フレアや地磁気嵐などの現象が発生したときに迅速に情報提供する必要がある宇宙天気では、インターネットによりほぼリアルタイムで収集されるデータストリームの中から、自動的に現象を検出して情報を配信するストリームマイニング技術^{13,14)} の応用も有効だと考えられる。

参考文献

- 1) Lanzoretti, L. J., Space Weather effects on technologies, in Space Weather (edited by P. Song, H. Singer, and G. Siscoe), AGU Geophys. Monogr. Ser., 125 (2001), pp.11-22.
- 2) Baker, D. N. and Barton, C. E., Informatics and the 2007-2008 Electronic Geophysical Year, EOS Transaction, AGU, Vol.89, No.48 (2008), pp.495-486
- 3) Rankin, R., Space science informatics: A Canadian approach, EOS Transaction, AGU, Vol.92, No.8 (2011), pp.61-62
- 4) 建部修見, 森田洋平, 平岡 聡, 関口智嗣, 曾田哲之, 広域大規模データ解析のための Grid Datafarm アーキテクチャー, 情報処理学会研究報, 2001-HPC-87, SWoPP2001 (2001), pp.177-182
- 5) 田浦健次郎, 頓 楠, 情報爆発時代のストレージ・データ集約的計算プラットフォーム, 電子情報通信学会誌, Vol.94,

No.8 (2011), pp.667-672

- 6) 新世代通信網テストベッド JGN-X
<http://www.jgn.nict.go.jp/>
- 7) 石倉諭, 村田健史, 久保卓也, 木村映善, 山本和憲, 篠原育, RSS1.0 を利用した科学衛星・地上観測データの自動収集, 電子情報通信学会論文誌 (B), Vol.91-B, No.4 (2002), pp.115-130
- 8) 村田健史, 岡田雅樹, 阿部文雄, 荒木徹, 松本紘, 太陽地球系物理観測の分散メタデータベースの設計, 情報処理学会論文誌 データベース, Vol.43, no.SIG12 (2002), pp.115-130
- 9) The fourth paradigm: Data-intensive scientific discovery (edited by H. Tony, T. Stewart, and K. Tolle), Microsoft Research, 2009
<http://research.microsoft.com/en-us/collaboration/fourthparadigm/>
- 10) 喜連川優, 情報爆発のこれまでとこれから, 電子情報通信学会誌, Vol.94, No.8 (2011), pp.662-666
- 11) 特集ビッグデータ革命ービッグデータは奇跡を起こすー, 日経コンピュータ 9月15日, No.791 (2011), pp.31-41
- 12) 森田武史, 山口高平, Linked Data を利用した情報統合, 人工知能学会誌, Vol.27, No.2, (2012), pp.189-199
- 13) 有村博紀, 喜田拓哉, データストリームのためのマイニング技術, 情報処理, Vol.46, No.1 (2005), pp.4-11
- 14) 櫻井保志, 時系列データのためのストリーミングマイニング技術, 情報処理, Vol.47, No.7 (2006), pp.755-761

月・惑星探査データ表示・解析システム「WISE-CAPS」

寺菌 淳也^{*1}, 中村 良介^{*2}, 児玉 信介^{*2}, 山本 直孝^{*2},
平田 成^{*1}, 小川 佳子^{*1}, 出村 裕英^{*1}

WISE-CAPS: Browsing and Analyzing System for Lunar and Planetary Exploration Data

Junya TERAZONO¹, Ryosuke NAKAMURA^{*2}, Shinsuke KODAMA^{*2}, Naotaka YAMAMOTO^{*2},
Naru HIRATA^{*1}, Yoshiko OGAWA^{*1} and Hirohide DEMURA^{*1}

Abstract

More and more data is obtained through the lunar and planetary explorations worldwide. Due to overwhelming amount of data and its complex nature, the methodology of the conventional desktop-based approach is becoming a tether. A network-oriented approach is an appropriate solution to solve this problem, and Web-GIS system is the best platform for data browsing, sharing and analyzing online. We are now creating a system called “WISE-CAPS” (Web-based Integrated Secure Environment for Collaborative Analysis of Planetary Science) in the University of Aizu to realize this idea. The system is composed of three servers (web server, database server and experiment server) with 12 Terabytes disk in the backend to ensure large data storage. The system uses FOSS (Free Open Source Software) entirely from the basic operating system to application software to utilize cutting-edge technology and to reduce the system cost. Currently, lunar data obtained by Clementine, American lunar explorer, is used for the base map data, and several image data obtained by Kaguya, Japanese lunar explorer, are added as separated layers. Our system also has user control function to confine data access to authorized users or groups. Utilizing this user control mechanism, users can limit (or permit) browsing of their data to designated users or groups with easy operation. Recently, the WISE-CAPS added several new capabilities. One of the new feature is data registration function from web browsers, enabling data registration to the system using only the web browser. Integrated display of LISM (Lunar Imager and Spectrometer, onboard optical instruments of Kaguya) data is another new feature. In this integrated display page, users can see three different kinds of data in one web browser page in the style of map. These new capabilities are purely unique, only in our WISE-CAPS.

概要

現在、月・惑星探査によって得られるデータは激増の一途をたどっている。これらのデータを研究に利用していく際、従来のようなデスクトップ中心の方法には限界があると考えられる。ネットワークを指向し、ユーザ(研究者)が扱いやすいシステムとして、ウェブGISを基盤とした月・惑星探査の閲覧・共有・解析システムが考えられる。我々は現在、そのようなシステムとして、会津大学内に WISE-CAPS (Web-based Integrated Secure Environment for Collaborative Analysis of Planetary Science) というシステムを構築している。本システムはウェブサーバとデータベースサーバ、そして新規実装試験用サーバの3台のサーバから構成され、バックエンドには12TBの巨大ディスクを装着し、大容量データに対応している。最新の技術を利用するためとシステム構築の費用を抑えることを目的として、システムは基盤からアプリケーションソフトウェアまですべてをオープンソースソフトウェアで構築している。ベースとなっているデータはアメリカの月探査機クレメンタインのデータであるが、その上に一部日本の月探査機「かぐや」によって得られたデータも搭載

*1 会津大学先端情報科学研究センター (CAIST/ARC-Space (Research Center for Advanced Information Science and Technology, Aizu Research Cluster for Space Science), The University of Aizu)

*2 産業技術総合研究所 (AIST, National Institute of Advanced Industrial Science and Technology)

している。また、本システムの特徴として、ユーザ管理機能を備えているということがある。これにより、データを一部のユーザだけに公開したり、グループを構成してその中だけで扱うこともできるようになっている。WISE-CAPS はさらに最近になり、いくつかの新しい機能が搭載された。例えば、データをウェブブラウザから登録する機能、「かぐや」の月撮像分光機器 (LISM) のデータ統合表示システムである。このように、WISE-CAPS では月・惑星探査データの解析に特化したユニークな機能を提供し、利用者の益を図ると共に、より強力な機能を装備し、最終的にネットワーク上の「仮想研究室」となることを目指している。

1 はじめに

1.1 月・惑星探査データに存在する問題点

月・惑星探査におけるデータ量は、特に 2000 年代になってから急速に増加しつつある。例えば、1994 年に打ち上げられた月探査機クレメンタインのデータ (探査機から取得された 1 次データ、公開分) は、CD-ROM にして 88 枚であり、データ量とすればたかだか 500GB 程度に過ぎない。一方で、2007 年に打ち上げられた日本の月探査機「かぐや」¹⁾ のデータは、生データに近い「レベル 0/1」データで約 15TB、処理後の「レベル 2」(後述) と呼ばれるデータの量にして 50TB に及ぶ²⁾。さらに、2009 年に打ち上げられ、現在 (2012 年 6 月) も月周回を行いながら観測を続けているアメリカのルナー・リコネサンス・オービター (LRO)³⁾ は、1 年間のデータ量が 133TB にも及ぶ。

他の火星探査なども含め、もはやテラバイトオーダーのデータは月・惑星探査において日常的となっている。今後、センサの高精度化、通信経路の効率化などの技術的な要素が進めば、この傾向はさらに加速度的に進行すると思われる。将来的にはペタバイト、エクサバイトオーダーとなることも想定しなければならない。

また、月・惑星科学データに特徴的なこととして、異なるパラメータによって処理されたプロダクトが多数存在することが挙げられる。

「かぐや」の月撮像カメラ (LISM) のデータの場合、到着したデータからさらに位置情報を補足し、修正を行ったデータとして、レベル 2a データ (L2a) が存在する。カメラに関するデータだけでもこれは 5 種類存在し、合計容量が 2.5TB となっている²⁾。L2a データを元にして、さらに画像をつなぎ合わせて地図上にした controlled mosaic データが作られていく。また、科学的に使用されるデータとして、地形カメラデータを元にした地形図 (DTM: Digital Terrain Map) なども作成される。これら、2 次的なデータは合計で 16TB に及ぶ⁴⁾。このように多種多様なデータセットの存在はさらにデータ量を増大させることへとつながる。

このような大規模データの処理に際し、科学分野、とりわけ月・惑星探査の分野では以下のような事情が存在

し、これが大規模データを利用した科学解析を困難なものにしている。

まず、1 人の研究者、ないしは数人の研究者グループが扱うには、1 人あたりのデータ量が多くなりすぎている。巨大なデータは、閲覧することはもちろん、解析プログラムへ投入することさえ時間と手間を伴う。また、急激なデータ量の増大は、研究者が従来 (主に自分たち自身で) 作成してきたツールでの解析に限界を生じてきている。

また、研究者自身が、どのようなデータが存在するかを把握することに困難を生じている。新たなプロダクトが次々に生み出されていく現状で、それがどのような内容であるのかを把握することは大変な労力を伴う。

データ解析の際、データを移動させることに多くの時間を必要とする。現在、ほとんどのデータ解析用ソフトウェア、プログラムはデスクトップで実行されるものであり、研究者はまずデータ配布用サーバへとアクセスし、データをダウンロードし、それを解析するという手順を踏んでいる。しかし、ネットワークの高速化にもかかわらず、データ量の増大により、データ取得 (ダウンロードなど) にかかる時間は増大しており、そのために貴重な研究のための時間が奪われるという問題が生じている。

研究者の解析データが個々のコンピュータに残ったままになってしまう。月・惑星探査データでは、探査機によって得られる観測データと、そのデータを研究者が解析することによって生じる 2 種類のデータが存在する。このうち、研究者が作り出す解析データは、上記のような問題点から、たいていの場合個々のコンピュータの記憶装置内にとどめられたままになる。しかし、何らかの理由で記憶装置が破損したり、誤って削除したりしてしまった場合、もしバックアップを適切に実施していない場合には、そのデータを回復させる手立てはなくなる。特に、適切にアーカイブされているケースがほとんどである探査データとは異なり、科学者による解析結果のデータは通常は 1 種類のみで、しかも個人のコンピュータの中に存在する場合はほとんどである。そうなると、せっかくの研究成果は永遠に失われてしまう。また、研究者が退職や異動などでそのデータを廃棄したり、時間

が経って記憶が失われてしまうことで、再解析や検証、さらには解析データを利用した応用解析などに必要となる重要なデータが失われることが発生し、データが存在しても事実上使用できないという問題が生じる恐れがある。

データを用いた共同研究などでの問題点。共同研究の際に、探査機が取得したデータや、実際に解析したデータを持ち寄り、議論するということはよく行われるが、データ量が增大しているために、持ち寄ることは困難になりつつある。また、データを持参することにより、データの破損、既存データとの重複(複製することによりデータが複製され、バージョンなどの管理が困難になる)などの問題点が危惧される。

1.2 これからの時代に求められる月・惑星データ解析システム

以上のような点を考慮すると、これからの時代に求められる月・惑星探査データ解析システムには、以下のようなポイントが必要になると考えられる。1.1節で述べた内容と対比しながら説明する。

- データの内容を一元的に把握できるシステム。特に、月・惑星探査データにおいてデータ量が最大となる画像データについて、それを素早く閲覧できるシステムが必要となる。
データ量が個人または小グループで扱うには大きすぎる場合には、それを適切にナビゲートする、あるいは個人でも把握できる形で提示することが必要となる。地図(画像)という形はナビゲートの面でも、またデータ提示の面でもっとも適切な形であると考えられる。さらには、ある科学データに関連した情報(広義のメタデータ)を容易に把握できるような仕組みもあると望ましい。
- 各ユーザ(科学者)が解析したデータを一元的に閲覧、集積すると共に、必要に応じて共有できるシステム。例えば、論文を執筆する際や、探査機器を開発する際には、グループを組んで作業を行うことが一般的だが、このグループ内では、必要に応じ論文や解析データが閲覧・共有できることが望ましい。一方で、グループに属さない他のユーザからはそれらのデータがみえないことが望ましい。同一の機器のデータを利用している科学者同士であるとはいえ、論文の成果の基盤となるようなデータはそれぞれの論文の執筆グループ、ないしは1つのテーマを解析しているチーム内でのみ閲覧できるようにすべきであるためである。このような柔軟かつ確実なユーザ管理を行えるような

システムが組み込まれていることが必要である。

- 機器を選ばないシステム。近年ではいわゆるパーソナルコンピュータに加え、可搬型ではあってもかなりの性能を発揮する機器、例えば、スマートフォンやタブレット端末などが市場に出回っている。これらの機器は現在のところは閲覧が主であるが、将来的には解析にも使われるようになっていくと考えられる。その場合、現在のようにデスクトップにおける解析を行うことは(少なくとも現在のようなアプリケーションの枠組みでは)到底不可能であって、このことから、サーバに閲覧や解析機能を持たせ、サーバ中心で処理させるシステムが必要となる。
- 研究者の解析データの消失、あるいは解析に必要な情報の消失を防ぐための機構が確立されていること。デスクトップコンピュータはたいていの場合バックアップが備わっていてもあまりそれを実行しないことが多い。そのことを考えると、サーバ上にデータを保管する方がより安全な解決策であるといえる。さらに、分散サーバなどの手法を用いれば、災害などによりサーバに損害が生じた場合でも、残りのサーバからデータを回復させることも可能である。また、データが個人のコンピュータ内に残される問題も、サーバ側にデータが存在することで解決できる。
- 複製・複製などによりバージョン管理が困難になる点については、バージョン管理をデータベースなどで厳密に行うとともに、正本を1つだけとし、それを各地から参照できる形とすることが解決策となる。この場合、正本はネットワーク上、すなわちサーバ上に置かれることになるため、サーバ上のデータをネットワーク経由で参照する形態をとらざるを得ない。また、複製やダウンロードができないようにするか、それらの(ローカル)データはキャッシュ扱いとして正しいデータとみなさないといった工夫も必要となる。

サーバ側にデータを置き、それをクライアント側(ローカル側)では閲覧するという形態をとるとした場合、閲覧するためのソフトウェアの問題が発生する。クライアントには様々なOSが存在するため、それらすべてで、できれば同様のルック&フィール、操作感で実行できるソフトウェアが望ましい。

この点については、どのシステムでも実行できるJavaなどのプラットフォームを用いるNASA WorldWind⁹⁾のようなアプリケーションも存在するが、スマートフォンやタブレット端末など、今後普及が予想される機器への対

応は、これらの小型機器が持つ性能を考えると難しい。また、デスクトップアプリケーションの場合にはその開発にも人的、資金的なリソースが必要となってくる。また、Javaなどをベースにした場合には、基本システムの性能を100%引き出せるわけではないので、パフォーマンスの面でも問題が生じる。

このため、基本システムを選ばない月・惑星探査データ解析システムの原型として、現時点ではウェブベースのGIS、すなわちWeb-GISが最適である。Web-GISは、ウェブサーバを通してデータの閲覧を行うシステムであり、ユーザはウェブブラウザを通してデータを閲覧することになる。従って、ウェブブラウザさえ用意できる(もちろんネットワーク接続が前提であるが)環境であれば、どこでもデータの閲覧が行える。そして、これに共有・解析環境が付加されれば、どこにいても、どのような端末からでも、データの閲覧、さらには解析などが行える。これが、我々の目指す理想的な月・惑星探査データの解析システム形態である。

ただし、現時点では、特に解析についてはWeb-GISはデスクトップベースのGISに劣るのが現状である。ウェブアプリケーションの場合、ブラウザが用意する機能を使用するか、JSP(JavaServer Pages)などを利用して機能を拡張するかどちらかの仕組みを選択して解析などの複雑な仕組みを実装する必要があるが、まだ通常の科学者が行う解析への要求を十分に満たせる状況にないことは確かである。それでも、将来的なウェブ技術の進歩(次々節で触れるようなHTML5などを含む)が進めば、将来的には複雑な解析なども含めてウェブ上に取り込めるのではないかと考えている。

1.3 他のシステム

このような月・惑星探査における大規模データ処理の問題については、日本のみならず世界的にも注目されており、多くのシステムが構築されている。また、その多くが、ウェブ(より正確に言えばHTTPを利用したデータ転送)をベースとしたシステムとなっている。ここでは、主にデータ閲覧を中心として開発されているウェブベースのシステムについて触れる。

古くから存在するシステムとしては、アメリカ地質調査所が公開しているMap-A-Planet⁶⁾がある。Map-A-Planetは、月に限らず、これまで探査が行われてきた多くの惑星、衛星についての地図が用意されており、それらの地図をブラウザ上で見ることができる。いわば月・惑星Web-GISシステムの基本ともいえる形態をとっている。また、解析データについても表示することが可能である。

一方、Web-GISというよりは探査立案・データ処理システムと一体になって開発されてきたシステムとして、LMMP(Lunar Modeling and Mapping Portal)^{7,8)}とJMARS(Java Mission Planning and Analysis for Remote Sensing)^{9,10)}である。名前からみてもわかるとおり、前者は月探査、後者は火星探査に特化した、Web-GISを含む解析・ミッション立案・解析ツールである。なお、JMARSはその姉妹サイトとして、JMOON¹¹⁾という月探査用のアプリケーションも存在する。

LMMPは、Web-GIS環境、スタンドアロンアプリケーションを利用した解析環境などからなる、その意味では月探査データ解析「ポータル」(入り口)を目指したサイトである。

Web-GIS環境は、LROのデータを基本とし、検索機能、マップのカスタマイズ機能などを有する。なお、後述するWMS機能も有しているとのことである。

解析については、Windows及びMacOS X上で利用できるIlliadsというソフトウェアをダウンロードすることによって実現され、本サイトと連動した解析が可能となっている。

LMMPは、LROによる探査をサポートするツールとして2007年にMSFCで開発がスタートし、当初はLMP(Lunar Mapping Project)として構築されていたが、その後解析機能などを付加することで、現在の名称となっている。現在では、NASA エームズ、アリゾナ州立大学、USGS、さらには軍なども加わった開発が進められている。

JMARSは、ミッション立案システムから進化してきたシステムで、アリゾナ州立大学で運営されている。その起源は、1997年のマーズ・パスファインダーミッションにまでさかのぼるとい歴史あるシステムである。

現在では、様々な形での火星データのマッピング(3Dデータ、科学データ、異なる探査機によるデータなど)が行えるようになっており、これはJavaベースで開発されたクライアントプログラムをダウンロードすることにより実現されている。

地球観測分野では、このようなGIS上での解析やデータ共有などを行うシステムとして、ERDAS APOLLO¹²⁾がある。本システムは製品として販売されているものである。データ配信をメインとしているが、後述のOGCプロトコルなどもサポートし、データ共有などを行うことが可能である。

以上のシステムをみると、最終的に表示・解析機能などを実現させるためにどうしてもデスクトップアプリケーションを導入せざるを得ないという点がある。この点は、1.2節でも述べた通り、特に今後モバイル機器が

ネットワークアクセスの主力となった際に大きな問題点となる可能性がある。今後は、機器を選ばないばかりか、アクセス手法さえも選ばないようなシステムが要望されていく可能性は高いといえよう。

また、商品としてシステムが販売される場合、データフォーマットや解析手法などにおいてベンダーが固定されてしまう「ベンダーロックイン」が発生してしまう可能性はどうしても否めない。また、月・惑星探査分野においては、世界中でデータを共有する必要があるだけでなく、探査データの保持に予算が潤沢に投入できる状況ではないことも考えなければならない。

1.4 システムの基本構想

1.2 節中程で述べた通り、これからの時代に向けた月・惑星探査データ解析システムを開発する際には Web-GIS を念頭に置いて実装を進めていくことが適切であると考えられる。

幸い、ウェブ技術はここ数年長足の進歩を遂げている。とりわけ、従来であれば Adobe Flash などのプラグインを必要としてきた高度な画像表示やウェブアプリケーション (RIA: Rich Internet Application) を、基本的なウェブ技術だけで実現する枠組みとして、HTML5¹³⁾ の普及が急速に進んでいる。

HTML5 は単にマークアップとしての HTML 言語だけでなく、周辺の CSS (Cascading Style Sheet) や JavaScript などを含めた包括的なウェブの枠組みを定義したものであり、特に動的なウェブサイト作りにとっては非常に大きな力を発揮すると考えられる。

ここでの「動的」なウェブサイトとは、ユーザからの入力に対して何らかの反応があることが前提となるシステムであり、CGI のようなものから、PHP、JavaScript などによるものまで幅広く含まれる。この分野においては、ユーザからの位置情報をもとにその地点の地図、科学データなどを表示する、あるいは希望する表示位置への表示を行うといったウェブサイト (ウェブページ) になる。

現時点で HTML5 はドラフトであり、すべての機能が実装できているわけではないが、将来的にウェブ技術がその方向に進むことを見越し、これらの、特に動的なウェブを実現させる技術を大いに活かしていくことが重要である。

最終的に、他のシステムでは実現できていない、すべての環境 (解析やデータ共有など) をもウェブにおいて完結させ、「ウェブブラウザさえあれば閲覧から解析まで研究に関わるすべての活動が行える」というようなシステムを目指すのがよいと考えている。

このような技術的な進歩をもとに、将来的な技術 (例えば上述の HTML5 など) を取り込みつつ、できうる限りオープンな規格でシステムを実装することがシステム構成の基本である。オープンな規格であれば他のシステムとの連携も容易に行うことができ、また規格は常に残るので (例えば会社内に秘匿されるということはないので)、将来誰かがこの技術を再度利用しようとしたとしても、問題なく行える。これは、何十年後かに再利用される可能性がある月・惑星探査データにおいて重要なポイントでもある。

上記の意味において、ウェブ技術の点からはプラグインなど、ブラックボックスになりがちな技術を除いた形での実装を行っていくことが望ましい。

以上、これからの時代に求められるシステムとしては、ウェブというオープンなプラットフォーム上に、オープンな技術により構築されるシステムであることが望ましいといえる。第 2 章で述べる WISE-CAPS システムは、まさにその思想を体現しているものといえる。

2 システム「WISE-CAPS」

第 1 章で述べた内容を実現させるため、我々は現在、WISE-CAPS (Web-based Integrated Secure Environment for Collaborative Analysis of Planetary Science)^{14,15)} というシステムを構築している。

本章では、システムの要件定義及び実装について述べる。

2.1 システムとして備える要件の定義

第 1 章で触れたシステムの構想に基づいたシステムとして、WISE-CAPS に関しては、以下のような要件を備えることが必要であると定義する。

- 基本的に、研究で行うすべての活動をネット上、とりわけウェブブラウザ上で実施できるようにする。
- 上記を実現するため、月・惑星探査データの閲覧はもちろんのこと、他のユーザとの共有、解析データのアップロード、これらのデータに基づいた議論、さらには解析、他のシステムとの連携が実現できるようにする。
- 利用するユーザは、月・惑星データを扱う研究者 (惑星科学者) を念頭に置く。
- これらのユーザは必ずしもコンピュータやソフトウェアの扱いに秀でていたわけではないことを考え、ユーザビリティを重視したインタフェースを設計する。

- プラグインなどに頼らず、ウェブブラウザが持つ機能でのみシステムを実現する。
- 他のシステムとの連携を考慮し、オープンな規格に準拠する。
- 新しい機能などが加わって行くであろうことを考慮し、システムの拡張性を重視する。また、システム全体(ハードウェア)を一度に整備することは予算規模から考えて困難であることから、将来的に拡張可能なシステムを構築する。

2.2 システムの概要

WISE-CAPS は、Web-GIS を実現するためのシステムであり、システムとしてはウェブサーバ部分とデータベースサーバ部分に分かれる。この両者は、負荷分散や故障の際の互いの代替などを目的として、ハードウェアとして用途を分けてある。

1.4 節で述べた通り、WISE-CAPS では、すべてを(ユーザからみた場合には)ウェブ上、さらにいえばウェブブラウザ上で完結させることを目指している。

このことは、すべての技術を広い意味でのウェブ技術で解決することを目指している。このため、システムはウェブブラウザを選ぶことがなく、標準技術に準拠した内容であれば、どの基本システム上のどのウェブブラウザでも同一のルック&フィール、同一の操作性が保証される。また、プラグインなどを用いないで実現させる方向性は安定性の向上にもつながる。

構成ソフトウェアはすべてオープンソースのソフトウェアを使用しており、優れた拡張性と安価なシステム構築を実現している。また、常に最新のシステムソフトウェアを利用することも利点として挙げられる。

以下、2.2 節ではハードウェアについて、2.3 節ではソフトウェアについて詳細を述べる。

2.3 構成ハードウェア

WISE-CAPS は、現在は 3 台のサーバと、1 台の外付け大型ディスクから構成されている。

3 台のサーバは、それぞれに役割が分担されている。3 台のうち 2 台はペアとなって活動するように設計されており、1 台がウェブサーバ、1 台がデータベースサーバとして動作している。残り 1 台は、新規実装試験用サーバとして、新しいシステムのテストや WISE-CAPS を応用した新しいシステムの導入などのために使用されている。この新規実装試験用サーバでは、ウェブサーバとデータベースサーバが 1 台のサーバ内で実行されている。

ウェブサーバとデータベースサーバは、共に、外付け

の大型ディスクに接続されている。この大型ディスクは、合計 48 台のハードディスクドライブを RAID 5 方式で統合して使用する形で、論理容量(実際にファイルシステムとして使用できる容量)として約 12TB の容量を確保している。但し、本サーバが採用している基本システムである CentOS(次節で詳細を述べる)で使用しているファイルシステム形式である ext3 のパーティション最大確保容量が約 6TB であるため、その限界に従う形で、ディスクを 2 分割し、6TB ずつをそれぞれのサーバにマウントして運用している。

なお、外付けディスクと 2 台のサーバ間は、それぞれを 4Gbps のファイバーチャネル(光ファイバー接続)で結ぶことで高速なディスクアクセスを可能としている。

外部とのネットワーク接続については、大容量のデータのやりとりが行われることを考慮し、JGN-X (Japan Gigabit Network eXtreme)¹⁶⁾との接続を行っている。但し、日常的なメンテナンスなどの便を考え、JGN-X に加え、会津大学内のネットワークにも接続している。

これらの装置は、会津大学情報センターの建屋内に設置されており、ハーフハイトのラック内に機器が納められている。

2.4 構成ソフトウェア

本システムの特徴として、基本システムからフレームワーク部分まで、すべてをオープンソースソフトウェアで実装している点が挙げられる。

ベースとなる基本システムは、Linux ディストリビューションの中でも幅広いシェアを占めている CentOS を使用している。現在、2 台のサーバ(ウェブサーバ、データベースサーバ)は CentOS 5、1 台の新規実装試験用サーバについては CentOS 6 を導入している。

この基本システムの上に、基本的なサーバソフトウェアが搭載されている。ウェブサーバは Apache httpd、データベースサーバについては、基本的には PostgreSQL を使用しているが、新規実装試験用サーバでは MySQL も併用している。

PostgreSQL は、後述する WFS プロトコルを使用するために、PostGIS を同時にインストールしている。なお、Apache httpd は後述するユーザ管理フレームワーク(GridSite)のためにバージョン 2.0 系列をインストールしている。PostgreSQL は現時点では 8.4 系列を使用している。

以上の基本システムの上に、Web-GIS を実現するためのアプリケーション、及びフレームワークを実装している。ここでは、MapServer¹⁷⁾と OpenLayers¹⁸⁾について述

べる。

MapServer は、Web-GIS におけるマッピング機能を提供の中核的アプリケーションである。ここでいうマッピングとは、画像ファイルを、別途提供される画像の座標情報や、投影方法に応じた形で表示することをいう。MapServer 自身は CGI として動作し、マップファイルと呼ばれる、レイヤー名や座標、マッピングすべきウィンドウの大きさなどを記載したファイルの内容を元に、レイヤーのマッピングを行う。マップファイルに記載された元画像の座標情報を元にして、指定された縮尺でのマッピングを行うため、画像の拡大縮小やタイリング処理などを行う。

また、MapServer の大きな特徴として、OGC (Open Geospatial Consortium)¹⁹⁾ が提唱する、各種のデータ交換用共通プロトコルに対応しているという点が挙げられる。OGC では、Web-GIS 間のデータ相互流通性を向上するために、システム間での共通プロトコルを提唱している。その中には、画像データの相互流通性を保証する WMS (Web Map Service)²⁰⁾、地図上に存在する各種の物体 (地物データ。地球の GIS では建物や道路などの、自然地形以外の情報を指す) の相互流通性を保証する WFS (Web Feature Service)²¹⁾ などが含まれている。現在、WISE-CAPS で実装している MapServer は、WMS、WFS を有効としている。従って、WMS をサポートする他のサーバとのデータ交換が行えるようになっている。

MapServer を使用するためには様々なライブラリが必要であるが、これらのライブラリもすべてオープンソースである。画像変換ライブラリの GDAL²²⁾ や幾何解析ライブラリの geos²³⁾、座標変換ライブラリの proj4²⁴⁾ などが MapServer を使用するためには必要になるが、これらはいずれもオープンソースである。システムを構築する際には、これらのライブラリをソースから順次コンパイル、インストールした後、最終的に MapServer をインストールするという形をとっている。

一方、OpenLayers は、オープンソースのレイヤ管理フレームワークである。OpenLayers は JavaScript を使用しており、サーバ側への負担がかからない。また、WMS や WFS などのプロトコルをサポートしているため、MapServer などをはじめとする、OGC のデータ標準をサポートする GIS サーバと連携して動作することが可能である。

また、OpenLayers では、レイヤをオブジェクトとして定義することができる。このため、レイヤを表示しようとする HTML ファイル内に JavaScript でオブジェクトとしてレイヤを定義し、名前や表示枠の大きさなどのパラ

メータを定義することにより、表示が簡単に行えるという利点がある。

さらに、表示されるレイヤーの拡大・縮小、移動などは、画面に用意されたスケールをクリックすることで行うことができ、さらにその際にはブラウザでページを再読込する必要がない。この機能は Ajax (Asynchronous JavaScript and XML) テクノロジーを用いて実現されているが、このコア部分は OpenLayers 内部に記述されているため、ユーザは Ajax に関するコードを意識することなく、この機能を利用することができる。

2.5 ユーザ管理フレームワーク

WISE-CAPS では、ユーザ管理システムを導入し、データ閲覧に際して制限を設けたり、グループ内でデータの閲覧を可能にするような枠組みを設けている²⁵⁾。

このユーザ管理システムの中核をなすものが、GridSite²⁶⁾ という、Apache httpd のモジュールである。

GridSite は、サーバ内のディレクトリに設置された GACL (Grid Access Control List) というファイルを読み取り、ユーザがそのディレクトリにアクセスしてよいかどうかを判断する。この際に、GridSite はユーザの証明書を読み取り、そのユーザが本人であるかどうかをチェックする。

証明書をウェブブラウザにインストールしていれば、職場であれ自宅であれ、あるいは PC であれタブレット端末であれ、どのようなコンピュータからでも認証を行うことができ、限定されたページにアクセスすることが可能である。

さらに、GridSite の特色として、ユーザでグループを組むことが可能となっている。このグループは VO (Virtual Organization) と呼ばれ、例えばある論文の共著者のグループ、開発中の機器の担当グループといったような形で、ユーザがグループに入ることが可能である。なお、1人のユーザは複数のグループに入ることが可能である。

このように、GridSite を利用することにより、柔軟なユーザ管理を実現することができるが、ここで問題となるのがデジタル証明書である。日本ではコンピュータ、あるいはウェブブラウザ内に自身のデジタル証明書を保持しているユーザはそれほど多くない。このため、証明書がないために利用できないという事態が発生することが考えられる。

このような状況に対応するため、ID とパスワードによる認証によって、証明書を発行する中間的なプロキシサーバ (OGC Proxy) が用意されている。WISE-CAPS のユーザ制限がかけられている部分にアクセスする際、

OGC Proxy でまず認証を行う。ID とパスワードでユーザ確認が取れた時点で、このプロキシサーバはリクエストを証明書つきで WISE-CAPS へリダイレクトする。これにより、WISE-CAPS ではユーザ管理システムが動作し、必要なコンテンツの閲覧が行えるというシステムである。

なお、これらの過程で、通信経路はセキュアプロトコル (HTTPS) で秘匿化されている。

以上の概念図を図 1 にまとめた。

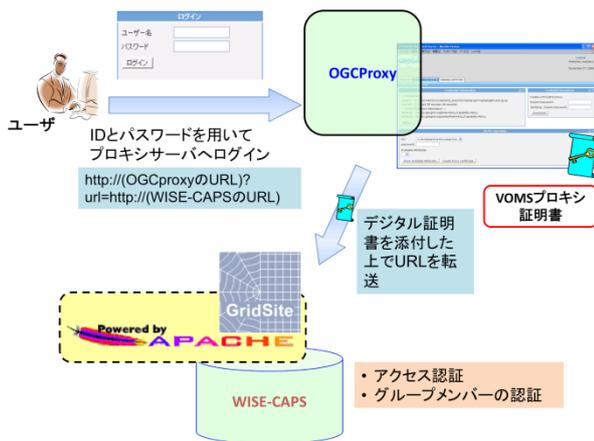


図 1 WISE-CAPS で採用しているユーザ管理フレームワークの概念図。ユーザからの ID/パスワード情報を OGC Proxy へ転送、デジタル証明書を付加して WISE-CAPS サーバへ転送するという構造になっている。

3 実際のシステムの内容

現在のシステムについては、参考文献^{14,15)}に詳細が記述されているので、ここではシステムを利用した代表的な例を記述する。

なお、本章で使用されている「マッピング」という言葉については、2.4 節の最初の部分で説明されているものと同一である。

3.1 システムの実際の例

図 2 には、WISE-CAPS の典型的な利用例として、月面の Diricilet クレーター及び Jackson クレーター周辺のマッピング例を示す。

本例では、アメリカのクレメンタイン探査機が取得した、UVVIS カメラの 750nm ベースマップ画像 (USGS にてモザイク化されたもの)²⁸⁾をベースマップとし、その上に、「かぐや」マルチバンドイメージャーが取得したこの地域の画像を重ね合わせている。それぞれが別レイヤーとなっている。

マッピングに際しては、OpenLayers の機能を用いて透明度を調整できるようになっており、複数レイヤーを重ね合わせた場合でも容易に下の地形との整合をとることができるようになっている。

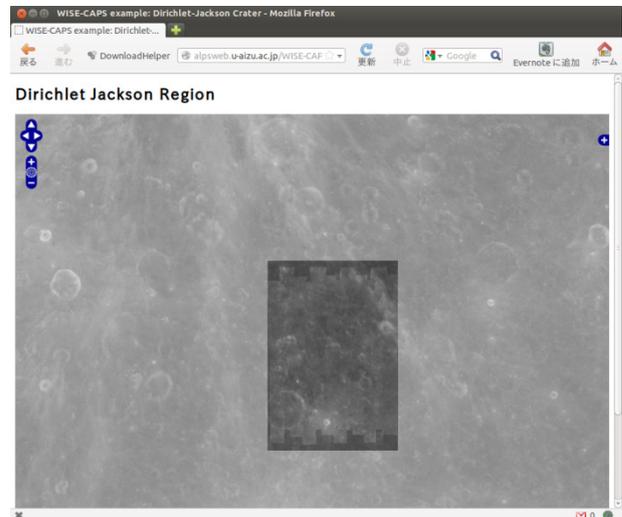


図 2 Dirichlet クレーター及び Jackson クレーター周辺のマッピング例。中央部にある詳細な画像が「かぐや」取得の画像。ベースマップはクレメンタイン UVVIS 画像。透明度は、画面下 (図ではウィンドウの下にあってみえない) の透明度ツールで調整できる。

図 3 では、より多くのレイヤーのマッピング、及びユーザ管理を行っているレイヤーマッピングの例として、Jackson クレーターのマッピング例を示す。

ベースマップは図 1 と同様であるが、その上に、「かぐや」の地形カメラによる画像を重ね合わせ、さらにその上に「かぐや」マルチバンドイメージャの画像 (本図では後述する比演算画像に重なってしまっていて見えない)²⁷⁾、そして同じくマルチバンドイメージャのデータを利用した比演算画像 (図 3 ではカラーで示されている縦に伸びる帯状の画像。疑似カラーで着色されている) を表示している。

なお、この比演算画像及びマルチバンドイメージャの画像は、ユーザ管理システムにより、特定のグループに属する人のみが表示可能となっている。また図 1 と同様に、それぞれの画像 (レイヤー) について透明度調整が可能である。レイヤー表示については、OpenLayers の表示レイヤー選択機能を用いており、この機能を使えば、透明度調整とは別に、直接表示したいレイヤーを指定することができる。

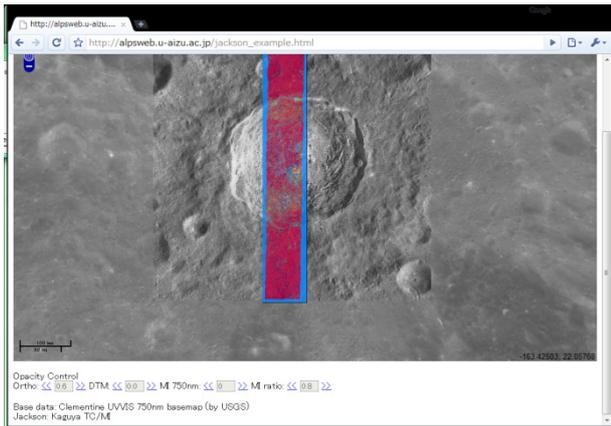


図3 Jackson クレーター周辺をプロットした例。ベースマップは図2と同様、クレメンタイン UVVIS 画像。写真中央部のより詳細な白黒画像は、Jackson クレーター周辺の TC 画像。中央部には、MI 画像より作成された比演算画像（上図カラー部分）が重なっている。本ページで合計 4 枚のレイヤーが重ね合わさっている。

4 新たな実装

本章では、WISE-CAPS の枠組みを利用して作成された、新たなシステムについて述べる。

4.1 LISM データ統合表示システム

LISM (Lunar Imager and SpectroMeter)²⁸⁾(月撮像分光機器)とは、「かぐや」に搭載されていた、カメラ及びスペクトロメータの総称である。

これらの装置とは、地形カメラ (TC: Terrain Camera), マルチバンドイメジャー (MI: Multiband Imager), そしてスペクトルプロファイラー (SP: Spectral Profiler) である。

TC 及び MI は画像 (ラスターデータ) を取得するタイプの機器である。一方、SP は月面の 1 点において、波長分解能が非常に高いスペクトルデータを取得することを目的としている。すなわち、得られるデータは測定位置及びスペクトルデータ (ベクトルデータ) となる。また、MI はその名の通り、マルチバンド、すなわち複数波長のスペクトル画像データを取得することを目的とした装置であり、単波長の画像を取得する TC とも異なる。

このように、LISM という 1 つの機器セットの中に、異なるタイプの機器が存在しており、しかも、特に MI と SP は、相互のスペクトルデータを結びつけて科学的な議論を行う「統合サイエンス」の要ともいえる機器である。にもかかわらず、従来この 3 種のデータを 1 つの画面上で統合して表示するというシステムは存在しなかった。

今回、WISE-CAPS に備わる基盤システムを利用し、

これらのデータを統合表示するシステムの作成を行った²⁹⁾。このシステムでは、TC、MI データをラスター (画像) レイヤーとして、そして SP データについてはその上に表示される測点をクリックすることでスペクトルグラフとして表示することが可能である。

図4にシステム概要を示す。TC、MI の画像データは、ウェブサーバ側に実際の画像データとして格納されている。これらのデータは、表示したい場所や縮尺に応じて MapServer で加工され、OpenLayers を利用して表示される。

一方、SP のデータは、観測地点の緯度・経度、及びその点の観測データが、データベースサーバ (PostgreSQL/PostGIS) に格納されている。このデータは、WFS プロトコルによってデータベースから引き出される。

引き出された測点データは MapServer 側の WFS 機能によって別レイヤーとして地図上 (TC、MI と同一の枠内) にプロットされる。そして、クリックするとその位置のスペクトルデータが読み出され、別枠にグラフとしてプロットされる。このグラフプロットは、JavaScript により記述されたウェブ用画像描画ライブラリである Dygraphs³⁰⁾ を使用している。

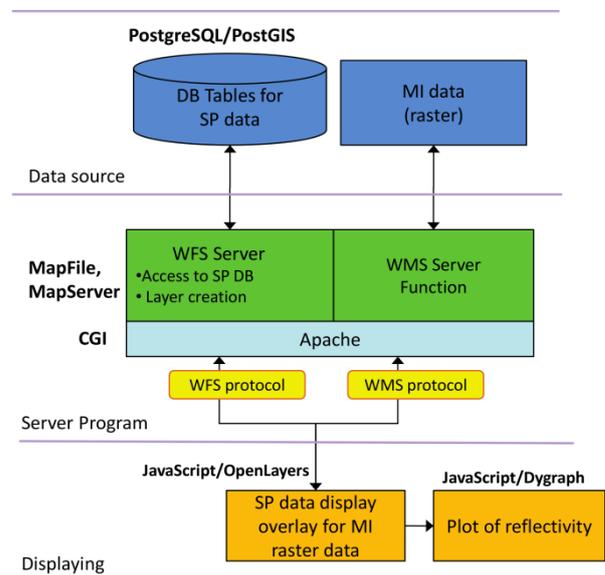


図4 LISM 統合表示システム構成図

以上の説明からもわかるように、本システムは Flash などを含めた追加のプラグインを一切使用せずに実現されており、事実上どんなウェブブラウザでも表示することが可能である。

実際のプロットの例は図5に示した。

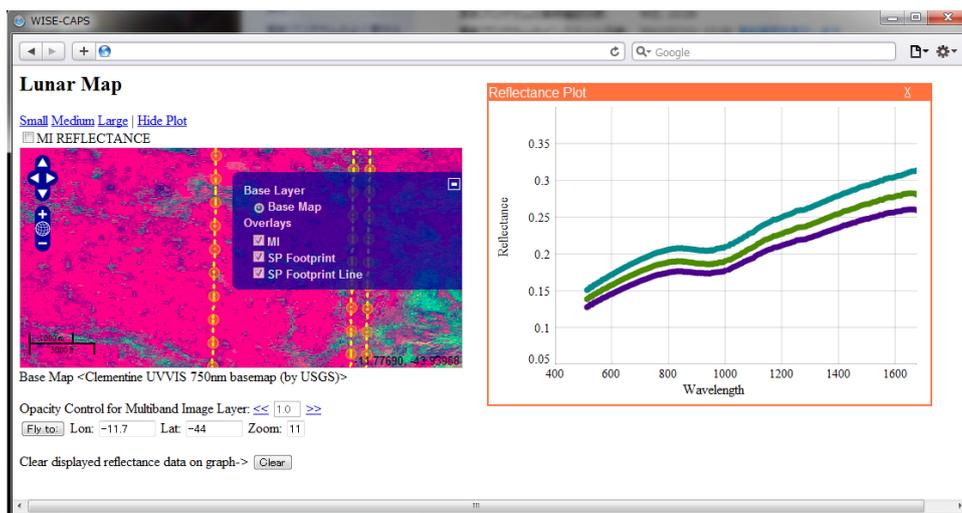


図 5 LISM 統合表示システムの例。左側には、ベースマップ（実際には隠れている）、及び MI による比演算画像（カラーの絵）が表示され、図の中心部には上下に SP の測線が走っている。測線には測点がプロットされ、それをクリックすることで右側のグラフ部分にスペクトルが表示される。この例では 3 点をクリックして、表示を重ねている。表示している部分は Tycho クレーター近辺。

4.2 データアップロードシステム

WISE-CAPS でユーザがデータをサーバにアップロードする際に問題となるのが、データの転送方法である。

現在、WISE-CAPS 上にデータを転送するためには、SFTP、または SCP などのセキュアファイル転送プロトコルを用いて、目的のサーバ上にデータを転送しなければならない。もちろん、サーバのどのディレクトリにファイルを設置すべきであるかといったことをユーザに対してあらかじめ伝えておかなければならない。

一方、システム管理者にとっては、このような状況はセキュリティ上あまり望ましくない。

まず、システムを本格的に使うことがないユーザのためにわざわざ UNIX アカウントを付与するということが自身がセキュリティの問題につながりかねない。例えば、UNIX アカウントをユーザに付与することにより、ユーザが弱いパスワードを使っていたりすれば、重大なセキュリティホールともなり得る。また、サーバ内部の構造をユーザに知らせることは、セキュリティホールにつながる可能性もある。

このような点を考えると、UNIX アカウントを使用せず、別のユーザ管理メカニズムを使った上で、SFTP や SCP などを使わないでファイルを転送できることが望ましい。

このようなシステムとして、ウェブブラウザ経由でファイルをアップロードするシステムを開発した³¹⁾。このシステムは単にファイルをアップロードするだけでな

く、ウェブブラウザ上でマップファイルや HTML ファイルなど、テキストファイルの編集も行うことができる。

MapServer を使って画像表示を行う場合、必要となるファイルの種類は以下の通りである。

- 図として表示したい画像ファイル (1 つとは限らない)
- 上記画像ファイルの四隅の位置 (緯度・経度)、及び画像の傾きを記した「ワールドファイル」
- レイヤー名や表示ウィンドウの大きさなど、マッピングに必要な情報を記した「マップファイル」
- 実際のマッピングの際に使用する HTML ファイル

なお、HTML ファイルは MapServer による画像表示とは直接関係しないが、表示させる画像や、OpenLayers によるレイヤー表示などの際に必要となる。

これらのファイルを数クリックの手順でウェブブラウザ上からアップロードすることが可能である。図 6 にそのインタフェース画面を示す。

図 6 では、各ファイルについて、そのファイルを指定してアップロードができるようになっている。なお、画像ファイルについては複数になる場合もあるため、ZIP 書庫としてアップロードすることもでき、その場合にはサーバ内で展開されて保存される。

アップロードしたファイルについて、特にマップファイルについては、正しくマッピングされない場合には再度編集することが必要である。従来であればローカルで

Mapファイルのアップロード
Mapファイルの拡張子は、txt形式でお願いします。

ファイルを選択 選択されていません

詳細:

アップロード

画像ファイルのアップロード
画像ファイルの拡張子は、jpg, png, gif, tif形式でお願いします。
Worldファイルもアップロードする場合はzip形式でお願いします。

ファイルを選択 選択されていません

詳細:

アップロード

HTMLファイルのアップロード
HTMLファイルの拡張子は、txt形式でお願いします。

ファイルを選択 選択されていません

詳細:

アップロード

図 6 ファイルアップロード用ページの画面

編集してサーバへアップロードし、正しいかどうかを確認した上で、正しくない場合には再度編集の上サーバへアップロードするという手順を繰り返すことになる。この手間を省くため、サーバ上のファイルを直接編集するシステムも導入している。この画面を図7に示す。

このシステムを利用することによって、マップファイルを編集し、表示させることができるようになってきている。また、同じシステムによって、HTMLファイルを編集することもできるので、HTMLファイルをアップロード後、修正が必要な場合でも、このシステムを使ってサーバ上のHTMLファイルを直接編集することができる。

後述するように、現在、サーバ上のレイヤー名の規約などが整っていないため、システムは試作段階にとどまっているが、規約などが固まり次第、早急にシステムを改良し、実用に供したいと考えている。

また、本システムはユーザ管理システムとも本来連動すべきものであるが、連動部分がまだ完成していないため、当面は切り離して運用することになる予定である。

アップロードによってユーザのデータがサーバを圧迫する可能性も考えられるが、現時点で想定されている利用法は、ベースデータのような大規模データではなく、ある領域を解析した科学データのような、小規模(大きくても数GBレベル)のデータを想定しているので、大規模バックエンド(数十TBレベル)を用意できれば、大きな問題とはならないと考えられる。なお、ユーザ増加などにより、データ領域が逼迫してきた場合には、

編集するMapファイルを選択してください

test.txt
test1.txt
samplemapfile.txt

ファイルを開く

ファイルを選択

概要: サンプルマップファイル

```
MAP # start MAP object
SIZE 600 300
#EXTENT -180 -90 180 90
LAYER # start LAYER object
NAME countries
TYPE POLYGON
STATUS DEFAULT
DATA countries_simple
CLASS # start CLASS object
STYLE # start STYLE object
OUTLINECOLOR 100 100 100
END # end STYLE object
END # end CLASS object
END # end LAYER object
END # end MAP object
```

ファイルを保存

概要: サンプルマップファイル

```
MAP # start MAP object
SIZE 600 300
#EXTENT -180 -90 180 90
LAYER # start LAYER object
NAME countries
TYPE POLYGON
STATUS DEFAULT
DATA countries_simple
LABELITER 'NAME'
CLASS # start CLASS object
STYLE # start STYLE object
OUTLINECOLOR 100 100 100
END # end STYLE object
LABEL
MINFEATURESIZE 40
END # end CLASS object
END # end LAYER object
END # end MAP object
```

ファイルを保存

Edit MapFile

図 7 編集するマップファイルを開く画面(上)と、実際に開いたマップファイルを編集する画面。ここではサンプルのファイルを編集している。画面左下には「ファイル保存」ボタンがあり、編集が終了したファイルはこのボタンを押すことでサーバ上に保存される。

Quotaのような機能を導入して一時的な制限を設けることも考えなければならない。

現時点では登録するデータについてマップ表示に必要なファイルのみに言及されているが、研究目的、あるいは永続的な保存などのためには、データに付随するメタデータの保存などの機能も将来的には実装したい。ユーザのフロントエンドをWikiベースにできれば、Wikiページをそういったメタデータ記述の場所とすることも可能

である。

また、データ量の増大だけでなく、誤ったデータ登録に対応するためにも、削除機能、及び削除・変更すべきデータなどを一覧できる機能を、将来に向けて準備しておくべきと考えられる。

5 将来的な展開

以上、WISE-CAPS システムの現状について述べてきたが、実際には本システムはまだまだ初歩的な機能の実装が完了しているという段階に過ぎない。

例えば、4.1 節で述べた統合表示システムは、現在のところ月面の一部領域でのみ稼働するだけでなく、対応ブラウザも Safari のみとなっており、他のブラウザで表示すると、プロットがずれたりするというような問題点を抱えている。一部領域での稼働となっている理由は、全データが膨大なためすべてをまだデータベースに投入し切れていないことと、データが膨大になった場合に現在のシステムで十分なパフォーマンスが出るかどうか不明なためという 2 点が大きな理由であり、この点については現在改良を進めている。将来的には月面全体で利用できるようにシステムとする予定である。

4.2 節で述べたデータアップロードシステムも、数クリックでレイヤーを登録できるという状態には未だ至っていない。これは、WISE-CAPS のレイヤー命名規則などが整っていないこともあり、どのような名前でもレイヤーを登録すべきであるかが確定していないことも大きな理由である。

レイヤーの名前を決定するのはマップファイルであるが、現在の仕組みではユーザが任意のレイヤー名をマップファイル内に指定することができるようになってしまっている。そのため、同じレイヤー名が別のマップファイル内に存在した場合、OpenLayers でレイヤーを指定する場合に問題が発生する可能性がある。このため、異なるユーザに対しても確実に一意で決定できるレイヤー名が指定できることが必要である。

WISE-CAPS が使用しているユーザ管理フレームワーク GridSite については、Apache httpd の 2.0 系列で安定して動くものの、2.2 系列(現在の主流)ではやや不安定であり、将来的に主流となる 2.4 系列では動作しないという欠点を抱えている。

また、ユーザ管理についても、現在は内部サーバで行うことができず、外部サーバにユーザの登録や削除などを委託しなければならないという状況で、自由な運用が行いにくい。

課題のいくつかについて述べてきたが、これらについて現在、鋭意解決策の用意を始めている。

ユーザ管理フレームワークについては、現在 OpenSSO³²⁾ についての検討を開始している。これは異なるシステムへのシングルサインオンを実現することができ、基本的なユーザ管理は LDAP (OpenLDAP) を利用している。LDAP が基盤となっていることからユーザ管理を内部で行うことも難しくなく、信頼性の高いシステムを構築することができると思われる。現に LMMP の内部で使用されているユーザ管理フレームワークが OpenSSO であり、類似システム使用の実績もあることから、今後ユーザ管理フレームワークを採用する際の選択肢として加えていきたい。

現在、WISE-CAPS 拡張として力を入れているのが、Web API の整備と、プログラミング環境の構築である。

Web API は、Google Earth や Google Map などでも使われているような枠組みで、サーバ側に用意されている JavaScript(など) を利用することで、そのサーバに用意されている機能を使用することができる機構である。例えば、表示されている地図の一部を取り込んだり、表示されているデータを数字(物理量)の形式で表示させたりすることも可能である。

Web API があれば、WMS や WFS のようなサーバ連携の仕組みとは別に、より多くのサーバで WISE-CAPS を使用してもらうことが可能になる。Google Maps でよく行われているように、地図の一部、あるいは該当領域を表示したり、必要なレイヤーだけを表示する、といったことが可能になる。

一方、サーバ上でのプログラミング環境は、WISE-CAPS の「A」、すなわち解析環境に欠かせないものである。具体的には、サーバ上にプログラムを書ける環境と、ある程度のライブラリ(初歩的な画像処理やデータ解析など)を用意した上で、サーバ上のデータを解析し、それを新しいデータとして蓄積できるような仕組みである。枠組みとしては Google App Engine³³⁾ に近いが、あくまでも月・惑星探査データに特化した枠組みでのプログラミング環境である。

これが整えば、データをローカルのコンピュータにダウンロードし、アプリケーションで解析する必要もなく、解析までもがすべてサーバ側で実施できる環境が整う。さらに、解析済みのデータをも共有できるようにすれば、データの閲覧、解析、そして議論をすべて WISE-CAPS 上で行うことができる。ここにコミュニケーション機能が実装できれば、WISE-CAPS はまさにネット上の「仮想研究室」として、地理的、時間的に離れた研究

者同士を結びつけて、月・惑星探査データの解析やその結果についての議論を行う場としての環境が整うことにある(もちろん、コミュニケーション機能は別途存在してもよい。例えば、Skypeなどを使えば現在でもリアルタイムで議論することは行える。ただ、実際のデータを前提とした、より研究用途に特化したコミュニケーションシステムは存在してもよい)。

ただ、この場合の問題点として、システムが巨大化することによって、サーバ等(CPU資源やディスク資源など)が大量に必要となり、結果的にコストの増大を招く恐れも否定できない。このような問題に対処するためにはいくつかの方法が考えられる。

- 仮想化サーバの積極的な活用。CPU資源などの消費の問題については、必要なときにのみ稼働する形での仮想化サーバを活用することによって、効率的な運用を行うことが可能である。
- 商用プラットフォームの利用。商用クラウドプラットフォームは現在非常に多くの会社が展開しており、価格も劇的に下がりつつある。上記の仮想化ソリューションとも関連するが、例えば一時的に多量のディスク領域が必要になった場合には商用クラウドプラットフォームからディスク領域を購入するという手段を執ることによって、管理コストと整備コスト両方を押さえることが考えられる。
- クライアント側リソースの利用。グリッド的な考え方であるが、クライアント側もシステムに組み込むことによって、クライアントのCPUやディスク資源を一時的に利用させてもらうことも考えられる。この場合のポイントは、ユーザ側にクライアント側を利用しているということ意識させてはならないということである。例えば、あらかじめクライアント側の資源を提供してもらい、それを分散型ファイルシステムやMapReduceのようなグリッド指向のシステムで結びつけることにより、十分に高速・大容量になったパーソナルコンピュータの資源を活用させてもらうことも可能である。ただしポイントは、あくまでもデータは中央(クラウド的なサーバ)に集中しているということであり、クライアント側(ローカル)には保存されていない、あるいは保存されていないようにみえるということがポイントである。

以上のような機能強化により、WISE-CAPSは、日本の月・惑星探査環境にマッチした月・惑星探査データシステムとしての進化を遂げ、より多くの人に使われるシステムとなっていくであろう。

また、機能拡張以外にもWISE-CAPSで行うべきことはまだまだ多数ある。例えば、データの登録数の増加、より新しいデータの登録なども行いたい。このようにデータの種類や数が増えることで、多くのユーザにとって魅力的なシステムともなるであろう。

WISE-CAPSを作り上げていく過程では、ユーザである惑星科学者からのフィードバックを得る必要があることはいままでもない。

我々は今後も、よりユーザが使いやすく、研究に貢献し、研究のための効率を上げることができるシステムを目指し、随時実装を重ねていく予定である。

謝辞

本システムのベース画像は、アメリカ地質調査所(USGS)が配布している³⁴⁾、クレメンタイン探査機の画像をベースとしたGIS用画像を使用している。また、「かぐや」画像は、JAXAが配布しているデータを使用している²⁹⁾。

参考文献

- 1) 「かぐや」プロジェクトウェブサイト, <http://www.kaguya.jaxa.jp>
- 2) Hirokazu Hoshino, Yukio Yamamoto, Shin-ichi Sobue, Katsuhide Yonekura, Mina Ogawa, Iwana, Kai Matsui, Hayato Okumura and Manabu Kato, Data Processing at KAGUYA Operation and Analysis Center, Space Science Reviews, vol. 154, No. 1-4, pp. 317-342, 2010.
- 3) Lunar Reconnaissance Orbiter, <http://www.nasa.gov/lro>
- 4) Naru Hirata, Jun'ichi Haruyama, Makiko Ohtake, Tsuneo Matsunaga, Hirohide Demura and LISM working group, Ground Data Processing of Lunar Imager/Spectrometer (LISM): A System Design and Core Algorithms, Proc. 23rd International Symposium on Space Technology and Science (ISTS), 2002
- 5) NASA WorldWind, <http://worldwind.arc.nasa.gov>
- 6) Map-A-Planet, <http://www.mapaplanet.org>
- 7) Cohen, B. A., Nall, M. E., French, R. A., Muery, K. G., Lavoie, A. R., The Lunar Mapping and Modeling Project (LMMP), 39th Lunar and Planetary Science Conference, #1391, 2008.
- 8) Lunar Mapping and Modeling Portal, <http://lmmpp.nasa.gov>
- 9) JMARS (Jaav Mission Planning and Analysis for Remote Sensing), <http://jmars.asu.edu>

- 10) Dickenshied, S., Christensen, P. R., Edwards, C. S., Prashad, L. C., Anwar, S., Engle, E., Noss, D., Jmars Development Team, Collaborative Planetary GIS with JMARS, American Geophysical Union, Fall Meeting abstract #IN32A-08, 2010.
- 11) JMOON (beta), <http://jmars.asu.edu/node/2055>
- 12) ERDAS APOLLO (日本インターグラフ株式会社), <http://www.erdas.jp/products/ERDASAPOLLO/>
- 13) HTML5, A vocabulary and associated APIs for HTML and XHTML, W3C Working Group, <http://www.w3.org/TR/2011/WD-html5-20110525/> (Working Draft as of Nov 25, 2011)
- 14) Junya Terazono, Ryosuke Nakamura, Shinsuke Kodama, Naotaka Yamamoto, Hirohide Demura, Naru Hirata, Yoshiko Ogawa, Jun'ichi Haruyama, Makiko Ohtake, Tsuneo Matsunaga. WISE-CAPS: Web-based Interactive Secure Environment for Collaborative Analysis of Planetary Science. Masatoshi Yoshikawa, Xiaofeng Meng, Takayuki Yumoto, Qing Ma, Lifeng Sun, Chiemi Watanabe (Eds), Database Systems for Advanced Applications: 15th International Conference, DASFAA 2010, International Workshops: GDM, BenchmarX, MCIS, SNSMW, DIEW, UDM, Tsukuba, Japan, April 2010, Revised Selected Papers, LNCS (Lecture Note in Computer Science) 6193, Springer-Verlag, pp.58-68, 2010
- 15) Junya Terazono, Ryosuke Nakamura, Shinsuke Kodama, Naotaka Yamamoto, Hirohide Demura, Naru Hirata, Yoshiko Ogawa, Tomoetsu Sugawara. WISE-CAPS: Data Archiving, Browsing and Analyzing Environment for Lunar and Planetary Data: Current Enhancement and Future Prospect. Proc. of 43rd Lunar and Planetary Science Conference, 2012
- 16) JGN-X (情報通信研究機構), <http://www.jgn.nict.go.jp/>
- 17) MapServer website, <http://www.mapserver.org>
- 18) OpenLayers website, <http://www.openlayers.org>
- 19) OGC, <http://www.opengeospatial.org>
- 20) Web Map Service, <http://www.opengeospatial.org/standards/wms>
- 21) Web Feature Service, <http://www.opengeospatial.org/standards/wfs>
- 22) GDAL, <http://www.gdal.org>
- 23) geos, <http://geos.osgeo.org>
- 24) proj4, <http://trac.osgeo.org/proj/>
- 25) 山本直孝, 田中良夫, 小島功, 関口智嗣, Tsukuba-GAMA: e-サイエンスのためのユーザ管理システムの設計と実装, 電子情報通信学会論文誌, 第 J93-D 巻, 10 号, 2292-2301 頁, 2010
- 26) GridSite, <http://www.gridsite.org>
- 27) かぐや (SELENE) データアーカイブ, JAXA 科学衛星運用・データ解析センター (C-SODA), <http://l2db.selene.darts.isas.jaxa.jp>
- 28) http://www.selene.jaxa.jp/ja/equipment/tc_j.htm
- 29) Tomoetsu Sugawara, Web-GIS Application for Analysis of Reflectance Spectra of Surface Materials on the Moon, Master Thesis, The University of Aizu, 2012.
- 30) Dygraphs JavaScript Visualization Library, <http://dygraphs.com>
- 31) Daisuke Tochigi, Fully Web-based Data Management System for Lunar and Planetary GIS, Graduate Thesis, The University of Aizu, 2012.
- 32) Oracle OpenSSO, <http://www.oracle.com/technetwork/jp/content/opensso-155204-ja.html>
- 33) Google App Engine, <https://developers.google.com/appengine/>
- 34) Lunar Clementine Basemap version 2, USGS, http://webgis.wr.usgs.gov/pigwad/down/moon_clementine_750nm_basemapV2.htm

モデル駆動型システムズエンジニアリングに基づくモデル管理およびデータ解析 — Nano-JASMINE データ解析への応用 —

初鳥 陽一*¹ 宮下 尚*¹ 清水 淳也*¹ 山田 良透*²

A Model Management and Data Analysis based on Model Driven Systems Engineering - Application for Nano-JASMINE data analysis -

Yoichi HATSUTORI*¹, Hisashi MIYASHITA*¹, Junya SHIMIZU*¹ and Yoshiyuki YAMADA*²

Abstract

Since data analysis for modern astronomy is getting more complicated than ever, effective information sharing among scientists and software engineers, and a distributed computing are important to compute large amount of data. In this study, a model construction and data analysis framework based on Model Driven Systems Engineering (MDSE) which makes it possible to effectively share information among project members are presented. Especially, we propose analytic framework based on a grey box modeling and least-square method used widely in data analysis. In particular, scientific theories and equations can be implemented as models of MDSE. In addition, proposed framework enables us to automatically generate effective machine codes from the models, which significantly reduces analysis efforts and development costs. Generated machine codes are computed by distributed processing technology based on MapReduce and GPGPU. As an example, proposed framework is applied to Japanese space astrometry project named Nano-JASMINE.

Keywords: Nano-JASMINE, Systems engineering, Data analysis framework, Distributed computing

概要

近年、天文学におけるデータ解析は複雑化しつつあり、科学者と解析担当者間の効率的な情報共有および大規模なデータの分散処理が重要な課題となっている。本研究では、プロジェクトメンバー間での効果的な情報共有を可能とする、モデル駆動型システムズエンジニアリング (Model Driven Systems Engineering, MDSE) に基づくモデル構築およびデータ解析のフレームワークの提案を行う。特に、データ解析で多く用いられているグレーボックスモデリングと最小二乗法に基づく解析フレームワークを提案する。具体的には、科学理論や方程式は MDSE のモデルとして実装し、そのモデルから解析コードを生成することで、開発のコストや労力を低減する。生成されたコードは、MapReduce および GPGPU と呼ばれる分散処理技術に基づいて処理を行う。一例として、提案したフレームワークを Nano-JASMINE と呼ばれる位置天文観測衛星のデータ解析へ適用する。

記号

| | | | |
|---------------|----------------------|---------|-----------------------|
| i | 星の ID | Au | 1 天文単位 |
| l | 観測の ID | $bG(t)$ | 時刻 t における衛星の位置 |
| $ui(t)$ | 時刻 t における星 i の位置 | $q(t)$ | 時刻 t における衛星のクォータニオン |
| α_i | 星 i の赤経 | an | B-Spline 係数 |
| δ_i | 星 i の赤緯 | $Bn(t)$ | B-Spline 基底関数 |
| ω_i | 星 i の年周視差 | η | CCD 座標における星の位置 (経度) |
| $\mu\alpha_i$ | 星 i の赤経方向の固有運動 | ζ | CCD 座標における星の位置 (緯度) |
| $\mu\delta_i$ | 星 i の赤緯方向の固有運動 | | |

*1 日本アイ・ビー・エム株式会社 東京基礎研究所 (IBM Research Tokyo, IBM Japan)

*2 京都大学大学院理学研究科 (Graduate School of Science, Kyoto University)

1. はじめに

近年の天文学におけるデータ解析は高精度化、複雑化が進んでいる。たとえば国立天文台、東京大学、京都大学が中心となって進めている JASMINE 計画¹⁾では、人工衛星を用いて銀河系内の星の観測を行い、得られた観測データを基に星の位置や年周視差をミリ秒角 (1 ミリ秒角は 1/3,600,000 度) やマイクロ秒角の精度で求めることが計画されている²⁾。JASMINE 計画では、主鏡口径 5cm の光学系を搭載する超小型衛星 Nano-JASMINE、主鏡口径 30cm の小型衛星 Small-JASMINE、主鏡口径 80cm の中型衛星 JASMINE という 3 機の衛星が検討されており、目標精度や観測範囲を改善しながら、最終的には銀河系中心のバルジと呼ばれる領域の位置天文観測を実施することを目標としている。表 1 に 3 機の衛星の目標精度、観測領域、打ち上げ年度について示す。このような高精度な位置天文観測を実現するためには、様々な解析手法を適用する必要や、人工衛星の姿勢パラメータや機器の劣化など観測結果に影響を及ぼす要素を洗い出し、解析の際に考慮に入れる必要がある。そのため、解析担当者がコードを実装する際には、複雑なモデルを理解し、大規模なデータを処理するための効果的なコードの実装を行うことが必要不可欠である。また、生成した解析コードを次のプロジェクトへ応用したいという科学者たちからの要求にこたえるためには、汎用的なコードの生成も必要となる。これらの要求を実現するためには、科学者と解析担当者との間における効果的な情報共有と大量のデータ処理を実現するための分散処理が重要な課題となる。そこで、本研究では、プロジェクトメンバー間での効果的な情報共有を可能とする MDSE の技術^{3), 4)} と、MapReduce^{5), 6)} および GPGPU^{7), 8)} の分散処理技術に基づく、モデルの実装およびデータ解析を行うフレームワークを提案し、すでにフライトモデルの開発が行われている Nano-JASMINE のデータ解析に対して適用する。

表 1 JASMINE 計画で計画されている 3 機の衛星の比較¹⁾

| | Nano-JASMINE | Small-JASMINE | JASMINE |
|------|--------------|---------------|-------------|
| 主鏡口径 | 5cm | 30cm | 80cm |
| 衛星質量 | 35kg | 400kg | 1500kg |
| 目標精度 | 3mas | 10 μ as | 10 μ as |
| 観測領域 | 全天 | バルジ数平方度 | バルジ 200 平方度 |
| 運用予定 | 2013-2015 | 2017-2019 | 2020 年代 |

(mas: milli arc-sec, μ as: micro arc-sec)

2. MDSE

MDSE とは OMG (Object Management Group²⁾) と呼ばれるモデリング言語の標準化団体によって仕様が策定された SysML^{3), 4)} (Systems Modeling Language) を用いて、自動車・航空機などの複雑なシステムを形式的に記述し設計するための新しい手法のひとつである。SysML はソフトウェアモデリング言語の一つである UML2 (Unified Modeling Language) の拡張付きサブセットであり、ハードウェア、ソフトウェア、データ、人、手順、設備を含むシステムの仕様化、分析、設計、確認、検証をサポートするグラフィカルなモデリング言語である²⁾。図 1 に SysML と UML との関係を示し、図 2 に SysML で用いられるダイアグラムを示す。MDSE の技術は従来自動車や航空機、組み込みシステムの開発に用いられてきたが、本論文では、衛星データ解析に対して MDSE によるモデリング手法を適用する。具体的には、ブロック定義図 (Block Definition Diagram, BDD) と内部ブロック図 (Internal Block Diagram, IBD) と呼ばれるダイアグラムを用いて科学者から提供された数式を SysML に基づいて表現することで、複雑化が進んでいる天文学におけるデータ解析に対して科学者と解析担当者との間における効果的な情報共有を実現する。衛星データ解析に対して MDSE を適用する潜在的なメリットとしては、科学者と解析担当者双方が共通理解できる形式で記述することによるコミュニケーションの向上、システムの妥当性確認と検証、開発効率の向上などが挙げられる。

1. <http://www.jasmine-galaxy.org/index-j.html>

2. <http://www.omg.org>

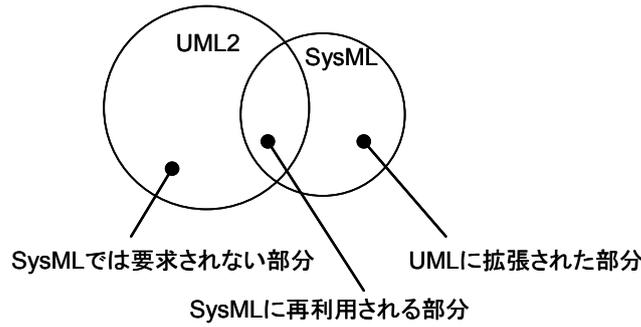


図 1 SysMLとUMLの関係³

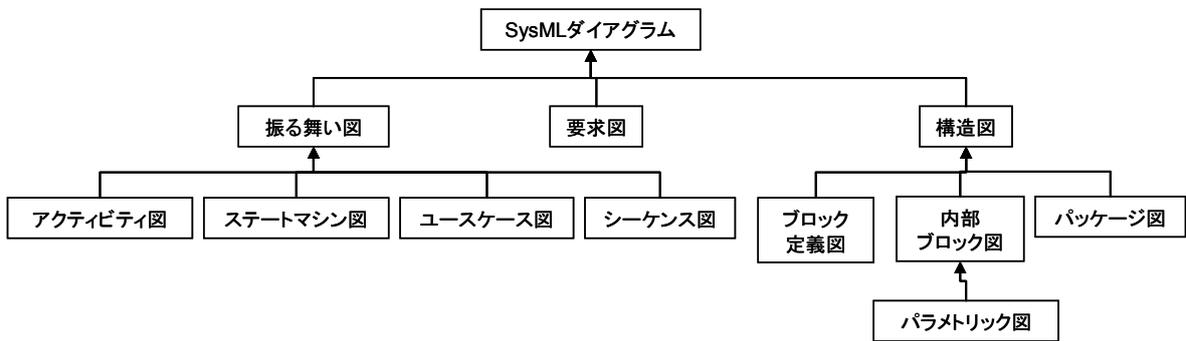


図 2 SysML ダイアグラムの分類³

SysMLによるモデル化では、システムの要素を記述する単位としてブロックが用いられ、値や制約、要求といった性質を定義することができる。BDDはブロックの特性と他のブロックとの構造的関係性を定義するためのダイアグラムであり、IBDはBDDで定義されたブロックの内部構造としてブロック間の接続方法を定義するためのダイアグラムである。BDDとIBDの関係性の例を図3に示す。図3左のBDDではAの構成要素としてB,C,Dを定義し、右のIBDではAの内部でB,C,Dがどのように接続されるかを定義する。本論文では、数式やパラメータといった最小の構成要素をブロックで表現し、構造と流れを定義するためにBDDとIBDを用いる。

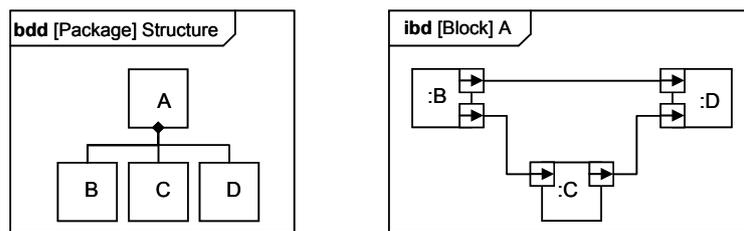


図 3 BDDとIBDの例 (左: BDD, 右: IBD)

3. JASMINE 計画におけるデータ解析

1節で紹介したJASMINE計画のデータ解析について述べる。JASMINE計画では、人工衛星を用いて銀河系内の星の位置・距離・固有運動を高精度に測定することを計画しており、地球近傍を周回する人工衛星から星を観測した場合、星の運動は星自身のもつ固有運動と地球の公転運動に起因するみかけの楕円運動により螺旋運動に見える。Nano-JASMINEでは2年間の観測を通して、同じ星を複数回観測することによって螺旋運動のパラメータを高精度に求めるこ

3. <http://www.omgSysml.org>

とを目的としている⁹⁾。図4にJASMINE計画における観測のイメージ図として、地球の公転運動に起因するみかけの楕円運動と、螺旋運動と観測の想像図を示す。

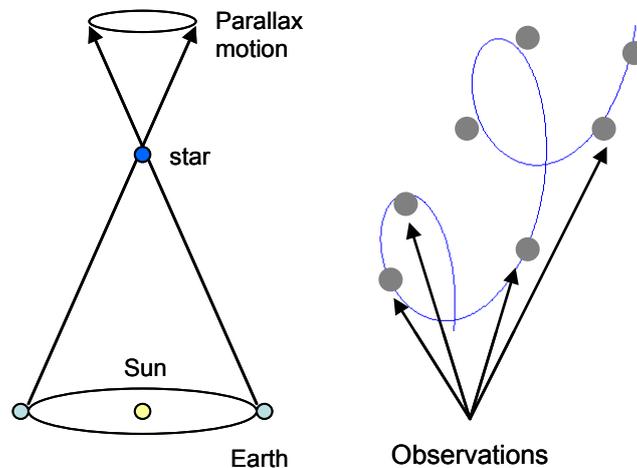


図4 地球の公転運動に起因するみかけの楕円運動（左）と、螺旋運動と観測の想像図（右）

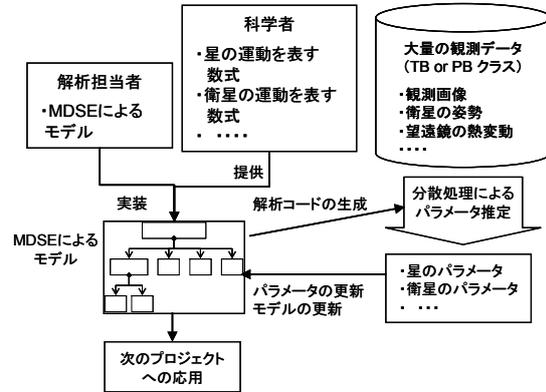
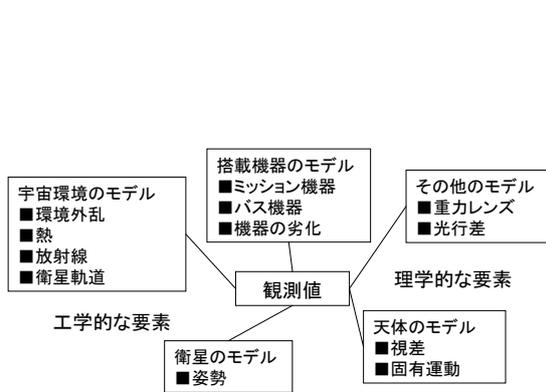
図4に示すように、JASMINE計画におけるデータ解析とは、星の動きは螺旋運動であることを仮定し、そのパラメータを求める問題である。推定すべき星のパラメータは、星の位置（2個）・距離（1個）・固有運動（2個）の5つのパラメータであり⁹⁾、Nano-JASMINEでは地球近傍の約50万個の星それぞれに対して、5個のパラメータを同定する。しかし、Nano-JASMINEデータ解析の特徴として、星のパラメータだけでなく姿勢のパラメータも同時に推定する必要があるという点が挙げられる。たとえば、Nano-JASMINEの目標とする要求精度は3masと非常に高い一方で、露光時間中の指向安定度要求は740masとなっている¹¹⁾。これは得られた画像から星の位置を3masの精度で求めたいにもかかわらず、衛星の姿勢運動によって画像が740mas変動している可能性があることを示している。最小二乗法を用いて星のパラメータと姿勢のパラメータを同時に解くことで要求される精度を達成することが考えられているが、そのためには望遠鏡の収差など観測機器のパラメータ、CCDの放射線劣化など機器の劣化のパラメータといったミッション機器にかかわるパラメータだけでなく、人工衛星の姿勢のパラメータなど観測結果に影響を及ぼすバス機器のパラメータも同時に推定しなければならない。このように、ミッション系、バス系のパラメータを同時に解くという複雑なデータ解析のための情報共有の一つの手段としてMDSEを適用した。ただし、前述のように星の動きは螺旋運動であることが既知で、そのパラメータを求める問題を仮定している。このように、背景にある数式は既知とし、そのパラメータなど一部が未知であるとするモデル化手法はグレーボックスモデルと呼ばれており¹²⁾、本研究ではこのグレーボックスモデルで記述された衛星データ解析を前提としている。

4. 本研究で用いる解析フレームワーク

4.1 全体像

位置天文データ解析では、星のパラメータの他に、観測に影響を及ぼすパラメータを同時に推定する必要がある。そのため、星の螺旋運動や重力レンズ効果による補正、光行差による補正など物理学の知見以外にも、衛星の姿勢運動を表現する数式、搭載機器、機器の劣化、環境外乱や熱、衛星軌道など周辺環境を表現する数式など工学的な知見も解析に考慮する必要がある。位置天文観測に影響する様々な要素の例を図5に示す。そのため、解析においては複数の学問領域にまたがる数式群が必要になることから、科学者たちと解析担当者間の効率的な情報共有が高精度な解析を行う上で重要な課題となる。そこで、MDSEに基づくモデルを構築し、解析コードを生成することで開発のコストや労力を低減する。また、生成されたコードを、分散処理を用いて大規模なデータを効率よく処理することを目指す。提案するフレームワークの全体像を図6に示す。図6において、科学者の提供する数式は解析担当者によってMDSEのモデルとして表現される。解析コードはMDSEによるモデルから生成し、MapReduceおよびGPGPUを用いて分散処理させることで大

規模観測データの効率的な処理を実施する。求められたパラメータを用いてモデルを更新し、また必要があればモデルを更新する。このモデルの一部もしくは全体を次のプロジェクトの解析の際に応用する。



4.2 分散処理によるパラメータ推定

本研究では大規模なデータから最小二乗法を用いてパラメータを同定する問題を前提しており、正規方程式と呼ばれる

$$y = Ax \tag{1}$$

の形式の連立方程式から x の推定値を計算する。ここで、 A は x の要素数を一辺とする正方疎行列で、Nano-JASMINE の問題では 5 節で述べるように A の一辺は 1000 万を超えるサイズの行列となるため、非常に膨大な数の要素を計算する必要がある。そこで、大量のデータを複数のプロセッサで計算させることが可能なフレームワークである MapReduce^{5),6)} を適用する。MapReduce は MapPhase と ReducePhase に分かれており、MapPhase では計算を独立に実行可能な複数のタスクに分割し、結果を Key / Value ペアと呼ばれる形式で出力する。ReducePhase では MapPhase の出力結果を一つにまとめて最終結果へ変換する。今回の問題では星の観測は一つ一つが独立しており、1 回の観測で行列 A の一部が埋まるため、観測ごとに処理を分割することができる。そこで、観測ごとの行列計算を MapPhase で実施し、行列のどの要素に対応するかを Key として出力し、計算された要素の値を Value として出力する。ReducePhase では、結果を全て統合することで行列 A を生成する。行列 A を計算した後、反復計算による x の推定値の計算には、行列演算に対して特に性能の高い GPGPU^{7),8)} を用いる。GPGPU は General Purpose computation on Graphic Processing Units の略で画像処理用のハードウェアとして用いられてきた GPU を数値計算などの用途へ応用する技術であり、CPU に比べて複数のコアを持つことから並列計算の効果が高いことが知られており⁷⁾、天文学への利用に対する研究も進められている⁸⁾。

5. Nano-JASMINE データ解析への適用

5.1 Nano-JASMINE データ解析の BDD

Nano-JASMINE のデータ解析に必要な要素を定義し、BDD を用いたモデル化を行う。科学者から提供される数式群は Lindegren らの論文⁹⁾ を参考にした。Lindegren らの論文には ESA で進めている Gaia プロジェクトにおけるデータ解析のための数式群が定義されており、観測手法が同じ Nano-JASMINE データ解析にも応用ができる。そこで、Lindegren らの論文の数式群の一部を利用して BDD と IBD の構築を行った。位置天文観測において考慮すべき要素は星のモデル、姿勢のモデル、較正のモデルに分類され、それぞれが数式と推定すべきパラメータを持つ。たとえば、星の螺旋運動は

$$\begin{aligned}
 \mathbf{u}_i(t) &= \langle \mathbf{r}_i + t \cdot (\mathbf{p}_i \mu_{\alpha_i} + \mathbf{q}_i \mu_{\delta_i}) - \varpi \mathbf{b}_G(t) / A_u \rangle \\
 \mathbf{p}_i &= [-\sin \alpha_i \quad \cos \alpha_i \quad 0]^T \\
 \mathbf{q}_i &= [-\sin \delta_i \cos \alpha_i \quad -\sin \delta_i \sin \alpha_i \quad \cos \delta_i]^T \\
 \mathbf{r}_i &= [\cos \delta_i \cos \alpha_i \quad \cos \delta_i \sin \alpha_i \quad \sin \delta_i]^T
 \end{aligned} \tag{2}$$

と記述される。ただし、 $\langle \rangle$ は規格化を表し、 $[\]^T$ は転置を表すものとする。衛星の姿勢はクォータニオンを 3 次の B-Spline 関数として、

$$\mathbf{q}(t) = [q_1 \ q_2 \ q_3 \ q_4]^T = \left\langle \sum_{n=l-3}^l \mathbf{a}_n B_n(t) \right\rangle \tag{3}$$

と記述される。よって、星の方向ベクトル $\mathbf{v}(t)$ は

$$\mathbf{v}(t) = [\cos \zeta \cos \eta \quad \cos \zeta \sin \eta \quad \sin \zeta]^T = f(\mathbf{q}) \mathbf{u}(t) \tag{4}$$

$$f(\mathbf{q}) = \begin{bmatrix} q_1^2 - q_2^2 - q_3^2 + q_4^2 & 2(q_1 q_2 + q_3 q_4) & 2(q_1 q_3 - q_2 q_4) \\ 2(q_1 q_2 - q_3 q_4) & -q_1^2 + q_2^2 - q_3^2 + q_4^2 & 2(q_2 q_3 + q_1 q_4) \\ 2(q_1 q_3 + q_2 q_4) & 2(q_2 q_3 - q_1 q_4) & -q_1^2 - q_2^2 + q_3^2 + q_4^2 \end{bmatrix} \tag{4}$$

から計算される。ここで式 (4) は非線形関数であるため、Lindgren らは非線形最小二乗法を適用することにより式 (1) の形式に変換している。具体的には、式 (4) の η に関する非線形方程式を星のパラメータ $\mathbf{s}_i = (\alpha_i, \delta_i, \omega_i, \mu\alpha_i, \mu\delta_i)$ とクォータニオン $\mathbf{q}(t)$ を用いて

$$\eta = h(\mathbf{s}_i, \mathbf{q}(t))$$

と表現し、この非線形の方程式を参照点 $(\mathbf{s}_i^*, \mathbf{q}(t)^*)$ 近傍で線形化する。観測値 η と参照点における値との差分 R_i に対してパラメータの偏微分を取ることで

$$R_i = \eta - h(\mathbf{s}_i^*, \mathbf{q}(t)^*) - \sum_l \frac{\partial R_l}{\partial p} R_l = \left[\sum_l \frac{\partial R_l}{\partial p} \frac{\partial R_l}{\partial p^T} \right] p \tag{5}$$

が得られる⁹⁾。ただし、 p はパラメータの参照点からの差分を示す。式 (5) の左辺は式 (1) の左辺 y の要素に対応し、式 (5) 右辺の第一項は式 (1) の右辺の A の要素に対応し、第二項の p は x の要素に対応する。ここで、パラメータの数は、50 万個の星それぞれが 5 個のパラメータを持つため星のパラメータが 250 万個、クォータニオンは式 (3) の通り 3 次の B-Spline で表現するため、2 年間のミッション期間を 30 秒間隔のスプライン関数で記述することにより 840 万個のパラメータで表現される。式 (4) より星のパラメータとクォータニオンのパラメータには関係性があるため独立に解くことができず、式 (1) のそれぞれの要素のサイズは、星と姿勢のパラメータ数を足して、式 (1) 左辺の y が 1090 万 \times 1、右辺の A が 1090 万 \times 1090 万、右辺の x が 1090 万 \times 1 となる。また、Nano-JASMINE は一つの星辺りの観測回数が約 100 回であるため、50 万個の星全ての観測回数は 5000 万回になり、5000 万個の観測値から 1090 万個のパラメータを推定する問題となる。

上記の計算に必要なコンポーネントを BDD で定義すると、図 7 が得られる。図 7 では、データ解析として定義されたブロックは観測値とモデルから構成され、モデルは星のモデル、姿勢のモデル、較正のモデルに分類される。また、それぞれのモデルは、与えられた数式と、数式を計算するために必要なパラメータに分類される。図 8 には、BDD における星のモデルの概念図を示す。星のモデルは、パラメータと数式から構成され、パラメータには同定すべき 5 つのパラメータが定義される。また、数式には星の螺旋運動として提供されたモデルを定義する。このように、BDD ではデータ解析に必要な要素を、階層構造を用いて定義する。

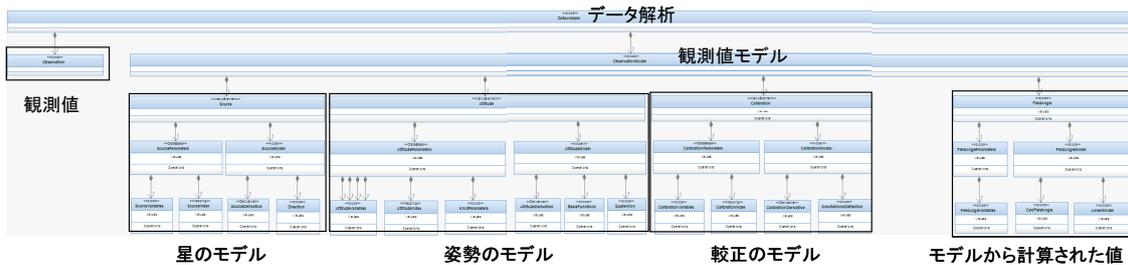


図 7 データ解析の BDD

5.3 連立方程式の生成と反復法による解法

5.2節で正規方程式の要素を計算するためのモデルを MDSE に基づいて構築した。Nano-JASMINE におけるデータ解析では、式 (1) に基づいて一辺が 1090 万の正規方程式を生成し、パラメータを同定する必要がある。まず、Map Phase では観測値ごとに式 (5) を計算する。Key には、観測が式 (1) のどの要素であるかを示す値を設定し、Reduce Phase で全ての観測を集計することで連立方程式を生成する。この連立方程式は 1090 万個のパラメータを持つため、直接法で解くことが困難であり、反復法を用いて解く必要がある。本研究では反復法として一般的な共役勾配法を用いて計算を行う。このとき、5.1 節、5.2 節で述べたように、解析のモデルは星、姿勢、較正の 3 つの要素からなるが、その疎行列の特徴は大きく異なる。たとえば、星の要素はブロック対角に配置され、姿勢の要素は対角を中心とする帯行列に配置される。そこで、図 11 に示すように式 (1) の正方疎行列 A を要素ごとに分割し、それぞれ最適な格納形式を選択することで計算効率の向上を行った。分割の様子と、選択した格納形式、分割したことによる効果を図 11 に示す。ただし効果の評価には、星は α 、 δ 、 ω の 3 つのパラメータを持つと仮定した簡易モデルを用いて、1 か月分の姿勢データに対して星の数を 3000、7000、15000、21000 と変化させたデータを用いた。全てを CSR で格納した場合を従来处理、分割した要素ごとに格納形式を選択したものを分割処理として比較を行った。その結果、分割した要素ごとに格納形式を選択した分割処理のほうが、全体を同一の形式で格納するよりも効率よく計算が可能であることが確認された。以降の計算では図 11 の格納形式を用いて計算を行うものとする。

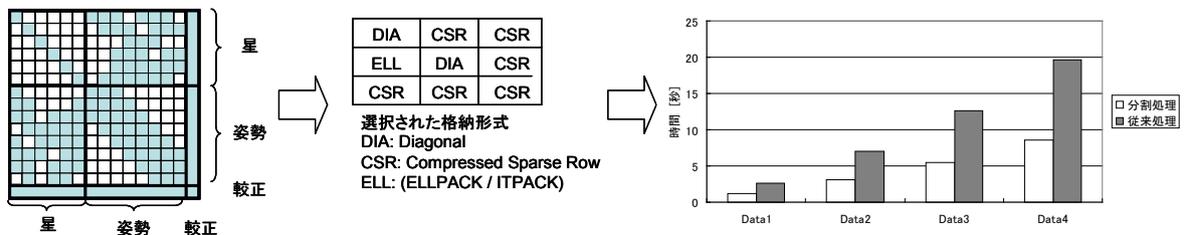


図 11 疎行列の分割と格納方法（左）と簡易モデルを用いた効果の検証（右）

5.4 位置天文データ解析の例

前節までの手法と数値解析による模擬データを用いて位置天文データ解析を行った。解析する星の数を 10312 個とし、それぞれが α 、 δ 、 ω の 3 つのパラメータを持つと仮定した。姿勢は、1 年間の姿勢データを 45 秒の knot で表現することで、666792 個のパラメータを持ち、較正の項は考慮しない。初期条件として、真の位置天文パラメータ α 、 δ 、 ω に対して 100mas, 100mas, 30mas (1σ) の標準偏差を持つ白色ノイズを付加し、姿勢には 50mas (1σ) の白色ノイズ、観測値には望遠鏡の収差などの影響を想定して 20mas (1σ) の白色ノイズを付加した。1 つの星は約 100 回観測される。ここで、5.1 節で述べたように、対象とする問題は非線形方程式を線形化し、式 (1) の正規方程式を用いてパラメータを求める非線形最小二乗問題である。したがって、式 (1) から得られた解を用いて参照点を更新し、更新された参照点を用いて再度式 (1) の正規方程式を生成して解を求めていく必要がある。そこで、線形化・式 (1) の生成・式 (1) を用いた求解・参照点の更新という一連の流れをイタレーションの 1 回とカウントし、イタレーションが進むにつれてパラメータの誤差がどのように変化するかを確認した。ただし、Lindgren らの論文⁹⁾に基づき、評価指標として RSE (Robust Scatter Estimate) と呼ばれる値を用いた。これは分散の上から 10% と 90% の値の差分に 0.390152 を掛けた値で、ガウス分布であれば標準偏差と等しくなる値として定義されている。そこで、1 回イタレーションを行うごとに全ての α 、 δ 、 ω の誤差、更新の量の RSE を求めた。結果の一例として α と δ の様子を図 12 に示す。図 12 の横軸にはイタレーションの回数、縦軸には RSE の値を示す。 α と δ は初期値として 100mas (1σ) の白色ノイズが付加されているが、10 回のイタレーションにより 40mas (1σ) まで減少することが確認できた。また、更新の量は 5 回のイタレーションで 1mas 以下の値となり、これ以降はイタレーションをしても結果は変化しないことがわかる。実際の観測では真の値は得ることができないため、Lindgren らの論文においてもこの更新の量が十分に小さくなることで収束の条件であると定義されていることから、収束したと判断した。

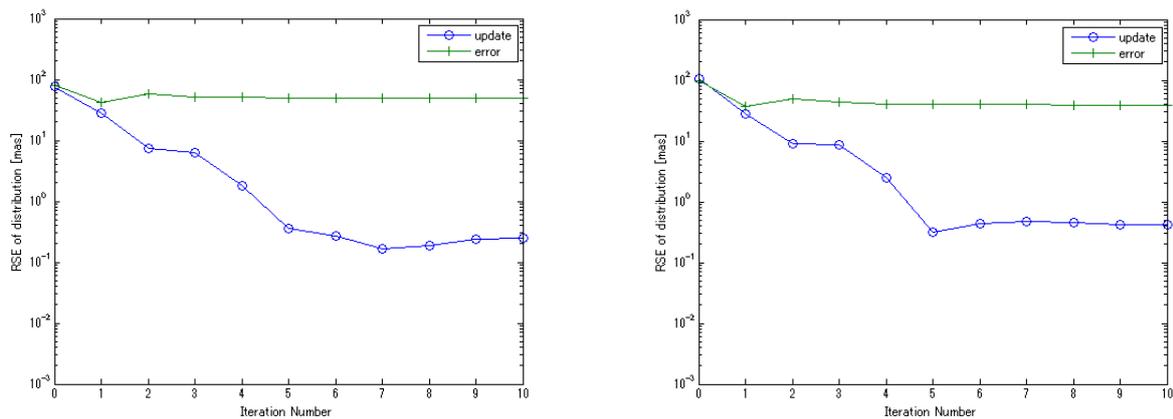


図 12 位置天文観測データ解析の例 (RSEと更新ベクトルのノルム, 左: α , 右: δ)

6. まとめ

大規模かつ複雑なデータ解析に対する科学者たちの要求を実現するため、解析担当者との効率的な情報共有および分散処理による効果的なデータ解析を可能にするモデル駆動型システムズエンジニアリングに基づくモデル管理およびデータ解析のフレームワークを実装した。一例として、位置天文観測衛星 Nano-JASMINE のデータ解析に適用し、解析モデルの生成、MapReduce による正規方程式の生成、GPGPU を用いた反復法によるパラメータ推定を実施した。現在は数値解析により生成した擬似観測データを用いて検証を行っているが、ノイズには白色性、正規性を仮定している。この仮定が成立しない場合、最小二乗法の解は真値に収束しないため¹¹⁾、今後は実機を用いた試験データからノイズの特性を分析することでモデル化誤差の有無を確認する予定である。ただし、地上で試験ができない姿勢制御の実データなどに関しては、打ち上げ後に得られるデータからモデル化誤差の有無を判断する必要がある。

参考文献

- 1) Gouda, N., Series of JASMINE missions, EAS Publications Series, Vol. 45, 2011, pp.393-396
- 2) Kobayashi, Y. et al., Nano-JASMINE: A 10-kilogram satellite for space astrometry, Proc. SPIE, 6265, 2006, pp.626544 1-10
- 3) Bock, C., Systems engineering in the product lifecycle, International Journal of Product Development, Vol. 2, 2005, pp.123-137
- 4) 西村他, システムズモデリング言語 SysML, 東京電機大学出版局, 2012, pp.2-14, 101-114
- 5) Jeffrey, D. et al., MapReduce: Simplified Data Processing on Large Clusters, Communications of the ACM, Vol. 51, No. 1, 2008, pp. 107-113
- 6) 白崎他, JVO 開発における大規模天文データ処理: 全天対応天文データ分散検索・解析機構の試験構築, 宇宙科学情報解析論文誌, 第一号, 2012, pp. 57-66
- 7) Narayanan, S. et al, A framework for efficient and scalable execution of domain-specific templates on GPUs, Proceedings of the 2009 IEEE International Symposium on Parallel & Distributed Processing, 2009, pp. 1-12
- 8) Christopher, J. et al, Astrophysical Supercomputing with GPUs: Critical Decisions for Early Adopters, publications of Astronomical Society of Australia, 2011, Vol. 28, pp.16-27
- 9) Lindegren, L. et al, The astrometric core solution for the Gaia mission Overview of models, algorithms and software implementation, Astronomy & Astrophysics, Vol. 538, id.A78, 2012, pp.1-48
- 10) 初鳥他, JASMINE 計画シリーズ: Nano-JASMINE から小型 JASMINE へ, 第 53 回宇宙科学技術連合講演会公演集, 2009, pp. 1101-1104
- 11) Inamori, T. et al, Compensation of time-variable magnetic moments for a precise attitude control in nano- and micro- satellite missions, Advances in Space Research, Vol. 48, Issue 3, 2011, pp. 432-440
- 12) 足立修一, システム同定の基礎, 東京電機大学出版局, 2009, pp. 4-5

xmlExcelHyper¹を用いた SIB2UI²の開発

¹ Excel 上で複雑な XML を扱う汎用ライブラリ

² 衛星情報ベース 2 作成ツール

松崎恵一^{*1} 山下美和子^{*2} 馬場肇^{*3} 西村佳代子^{*4} 宮野喜和^{*4}

Development of SIB2UI¹ with xmlExcelHyper²

¹ Spacecraft Information Base Version 2 Users' Interface

² General library on Excel which handles complicated XML data structure

Keiichi MATSUZAKI^{*1}, Miwako YAMASHITA^{*2}, Hajime BABA^{*3}, Kayoko NISHIMURA^{*4} and Yoshikazu MIYANO^{*4}

Abstract

Design of spacecraft and onboard components of ISAS satellites started after the BepiColombo/MMO project are stored in a database called Satellite Information Base Version 2 (SIB2). This database has a complicated data structure and its model itself is under development. In parallel activities, the spacecraft and onboard components must be developed with this model. In the developments of the spacecraft and onboard component, tool which support design with the model, SIB2UI, is needed and should also be developed. On the other hand, the model should be determined with feedback from design of spacecraft and onboard component. In the development of the model of SIB2, shortening of span of the feedback was the most important subject. To resolve the subject, we introduced a technique to implement design tool. This technique consists of a general library xmlExcelHyper which handles a complicated data structure on a spreadsheet program, Microsoft Excel. With this library, developer of design tool can easily implement Excel-based tool which visualizes, imports and exports tree-base data structure which has degree of freedom of selection and repeat.

Key Words: Development process, Modeling, Design Tool

概要

ISAS の衛星では、BepiColombo/MMO 衛星以降の搭載機器・衛星において、衛星情報ベース 2 (Satellite Information Base Version 2; SIB2) というデータベースに設計を記述する。このデータベースは、複雑なデータ構造を持ち、どのようにモデル化するか自身が開発対象である。他方で、このモデルを用いて搭載機器・衛星の開発も行う必要があり、モデルに従って設計を記述するツール (SIB2UI) が必要である。どのようにモデル化するかを決めるには、搭載機器・衛星の設計の実例を配慮する必要があり、このフィードバックを如何に早く収束させるかが課題であった。我々は、複雑なデータ構造を扱うモデルに従って設計するツールを効率よく開発する手法を考案し、適用することで、フィードバックの収束を加速した。その核となるのが、スプレッドシート (Microsoft Excel) 上で複雑なデータ構造を扱うための汎用ライブラリ xmlExcelHyper である。この、ライブラリを用いるとスプレッドシート上で選択構造や繰り返し構造からなる XML のツリー構造をそのデータ構造に応じ可視化・入出力するツールの開発を容易に実現できる。

*1 宇宙科学研究所 学際科学研究系

(Department of Interdisciplinary Space Science, Institute for Aerospace and Aeronautical Science (ISAS))

*2 宇宙科学研究所 Bepi Colombo プロジェクトチーム (Bepi Colombo Project Team, ISAS)

*3 宇宙科学研究所 宇宙科学プログラム・システムズエンジニアリング室
(Space Science Program Systems Engineering Office, ISAS)

*4 宇宙科学研究所 科学衛星運用・データ利用センター 衛星運用グループ
(Satellite Operations Group, Center for Science Satellite Operation and Data Archive, ISAS)

1 はじめに

ISAS (Institute for Space and Aeronautical Science) の衛星開発では、BepiColombo/MMO, SPRINT-A, ASTRO-H 以降、衛星情報ベース 2 (Spacecraft Information Base Version 2; SIB2) というモデルに基づいたデータベースに設計を記述する。このモデルはツリー状のデータ構造をもち、XML に格納される。他方、衛星設計の現場では、これをスプレッドシート (Microsoft Excel, 以下単に Excel) 上のユーザインタフェースから扱うことが求められている。そこで、我々 SIB2/GSTOS-1 プロジェクト¹⁾では、Excel 上で、SIB2 を扱うツール、SIB2 作成ツール (SIB2UI^{a)}) の開発を行っている。

モデルに基づく開発の実現においては、モデル自体を開発しながら個々の開発を実施することがある。この場合、モデル自身の開発と、モデルを扱うツールの開発、モデルを用いた個々の開発が並行で進むこととなり、リスクが高い。この状況を打開するには、モデルを扱うツールを効率よく開発することで、モデル自身の開発と個々の開発の議論を加速し、モデルの仕様を早く収束させることが重要となる。そこで、我々は、Excel 上で複雑なデータ構造を扱うための汎用ライブラリ (xmlExcelHyper^{b)}) を開発し、これを SIB2UI に適用することで、効率的な開発を実現した。

本論文では、xmlExcelHyper を適用したツールにおいて取扱い可能なデータ構造 (Excel での見え方、XML で可能な表現範囲) とその記述方法、xmlExcelHyper を用いたモデル・ツール開発の流れや効能について、SIB2UI の実例を用いて述べる。まず、2 章では、開発ツールの開発手法を実証の題材としたモデル SIB とこれを扱うユーザインタフェース SIB2UI について紹介する。3 章では、このようなモデルに基づく、設計ツールを開発する上での課題について示す。4 章では、我々が考案した汎用ライブラリ xmlExcelHyper に基づく設計ツールの構築について示す。5 章では、xmlExcelHyper が扱うデータ構造・XML 構造とこれに関わる機能について示し、6 章でまとめを示す。

2 Spacecraft Information Base (SIB)

SIB は、衛星の設計情報を記述するデータベースである。ISAS では、「のぞみ」衛星以降の「れいめい」を除く全衛星において、テレメトリとコマンドを SIB に登録し、SIB に従って動作する共通システムと共に利用することで、衛星システムの試験・運用を効率的に進めてきた。

SIB は搭載機器の設計を記述するものである。そこで、その入力、設計のタイミングでなされるべきである。しかし、従来は、搭載機器単体の開発に適用する共通システムも存在しなかったため、SIB への入力が搭載機器の衛星システムへの統合のタイミングとなってしまっていた。また、後追いで設計結果の入力作業を実施するため、衛星レベルの開発期間に、搭載機器単体の設計あるいは実装とデータベース入力間の整合性を確認する作業が必要となってしまっていた。

この状況を打破し、SIB を衛星設計段階からより積極的に適用することで衛星開発の効率化を図るため、山田、松崎らは、搭載機器の開発において、搭載機器単体の試験から適用可能な共通ソフトウェア GSTOS (Generic Spacecraft Test and Operations Software) と、新世代の SIB、SIB2 の検討を進めてきた。SIB2 においては、搭載機器・衛星を、山田²⁾³⁾が提唱する「衛星の機能モデル (Functional Model of Spacecraft; FMS⁴⁾)」と「衛星監視制御プロトコル (Spacecraft Monitoring and Control Protocol; SMCP⁵⁾)」に基づき設計する。ここで、FMS は、およそ、UML のクラスダイアグラム・ステートマシンダイアグラムのサブセットに衛星・搭載機器設計固有な定義情報を加えたものである。また、SMCP は、衛星の監視制御のためのプロトコルであり、FMS に従って設計された搭載機器・衛星において、パケットのユーザデータ部に詰めるメッセージを規定するものである。

SIB2 のモデルの定義 (Definition of SIB2; DSIB2⁶⁾) の抜粋を表 1 に示す。この表に登場する「機能オブジェクト」は、搭載機器や搭載機器内の機能の塊に対応する。また、「オペレーション」はコマンドにより駆動される動作、「アトリビュート」はテレメトリにより取得可能な値、「アラート」は、異常など機能オブジェクトから出力される間欠的なテレメトリを示し、これらが FMS の核となる。

SIB2UI の Excel ファイルは、1 つのファイルがサブシステムや搭載機器など衛星設計の一部に対応し、複数の Excel ファイルを取りまとめ衛星全体の設計を表現する。それぞれのファイルは表 2 に示すようにデータ構造及び衛星開発のステップを考慮して決められた 9 つのシートから構成される。これらのうち、最も基本的なのは functionalObject 及び attributeSequence シートである。functionalObject シートには、機能オブジェクト、オペレーション、アトリビュート、アラートを定義する。また、attributeSequence シートには、アトリビュートをテレメトリとして取得する際の並び順を定義する。

SIB2UI は、4 つの機能 Edit Support, Import, Export, Validate から構成される。SIB2UI のシートは functionalObject シー

a) 本稿執筆の時点での最新版は 3.1.2.0 版

b) 本稿執筆の時点での最新版は V11L01

トや attributeSequence シートなど、それぞれが複雑なデータ構造を担っている。そこで SIB2UI には、どのようなデータ構造かわかり易いよう可視化の機能 (Edit Support) を設け、罫線描画、セルの結合や背景色の設定を行っている。また、編集中の任意のタイミングでシート内・シート間のデータ構造の整合性が取れているかチェックする機能 (Validate) も持たせている。さらに、XML ファイルからの入力 (Import) と出力 (Export) の機能も持たせている。なお、SIB2UI では、実装を単純にするため、シート 1 枚を XML ファイル 1 つに対応させ、シート間の情報のマージは行っていない。

表 1 SIB2 のモデルの定義 (抜粋)

| |
|---|
| <p>機能オブジェクト定義は、機能オブジェクト毎に以下の事項を定義する。</p> <ol style="list-style-type: none"> 1) 名前 (1 個) 2) サブ機能オブジェクト (0~n 個) 3) アトリビュート (0~n 個) 4) オペレーション (0~n 個) 5) アラート (0~n 個) 6) 状態遷移関 (0~n 個) 7) 診断ルール (0~n 個) 8) この機能オブジェクトが有効となる親機能オブジェクトの状態 (1~n 個) <p>アトリビュート毎に以下の事項を定義する。</p> <ol style="list-style-type: none"> 1) 名前 (1 個) 2) 外部から値を設定可能か否か (1 個) 3) データ型 (1 個) 4) その値が有効となる状態 (1~n 個) 5) 初期値 (0~1 個) 6) 注意を喚起するためのリミット値 (上限 0~1 個, 下限 0~1 個) 7) 危険を知らせるためのリミット値 (上限 0~1 個, 下限 0~1 個) 8) 値が意味を持つ上下限值 (上限 0~1 個, 下限 0~1 個) |
|---|

表 2 SIB2UI を構成するシート

| シート名 | 概要 |
|-----------------------|---|
| 0.management | 衛星名を定義する。 |
| 1.enumeration | 数値と文字列の対応を定義する。 |
| 2.functionalObject | 機能オブジェクト (アトリビュート, オペレーション, アラート) を定義する。 |
| 3.attributeSequence | テレメトリとして受信するアトリビュートの並び順を定義する。 |
| 4.stateMachine | 状態を表す数値, 文字列を定義すると共に, 遷移の条件をオペレーションやイベントを参照することで定義する。 |
| 5.conditionAndEvent | アトリビュートの評価式の組み合わせ (論理式) により条件, イベントを定義する。 |
| 6.conversion | アナログ値の変換を定義する。 |
| 7.attributeLimit | アトリビュートのリミット値の危険値 (fatal limit), 警告値 (caution limit) を定義する。 |
| 8.attributeChangeRule | オペレーションに対するアトリビュートの変化則を定義する。 |

SIB2 作成ツール (SIB2UI) 入力リファレンス 第 1.5 版 2012 年 3 月 16 日から抜粋

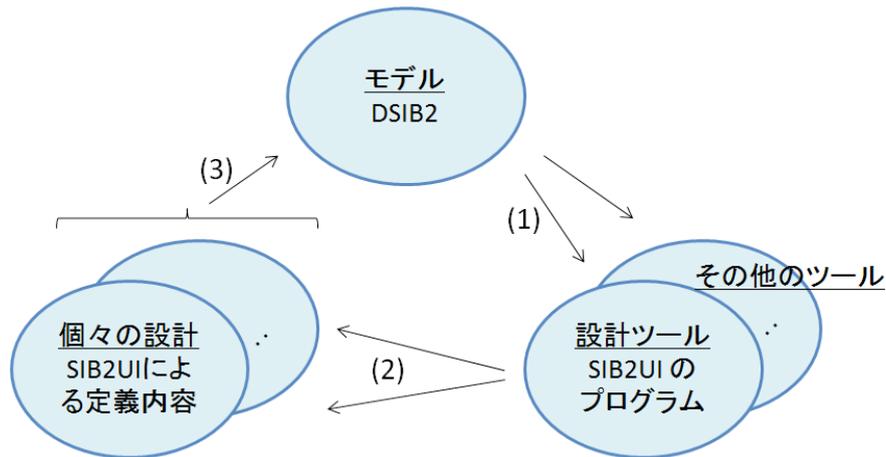
SIB2 及び SIB2UI は、ISAS の衛星に適用することを目的に開発を行っている。そこで、ISAS の衛星プロジェクト向けには web ページ^{c)}経由での配布を行っている。他方、SIB2・SIB2UI 自身は、任意な衛星に対応する汎用な技術である。そこで ISAS 以外の衛星プロジェクトに対しても、個別に条件を設定し配布を行っている。

SIB2 は表 1 に示されるように、ツリー状のデータ構造を柱に定義されている。ツリー状のデータ構造を扱う記法として、XML が普及している。そこで、SIB2 においても、データベースの源泉の表記として XML を採用した。RELAX NG⁷⁾ など、XML のスキーマ言語を用いると、XML 構造における多重度を定義できる。そこで、DSIB2 に定義される多重度も、スキーマ言語の記述に焼き直すことができる。

c) <http://www.c-soda.isas.jaxa.jp/software/project-tools.html>

3 設計ツール開発における課題

一般に、SIB2のようなモデルの策定においては、議論がデッドロックに陥り、仕様が収束しないというリスクが存在する(図1参照)。モデルが定まらなると、これをサポートする設計ツールが作れない。他方で、実際の設計で使うことができるツールが存在しないと、ユーザ(設計者)からは親身なレビューを受けられない。また、ツールのイメージを作り、ユーザとの議論を始めたのちも、ユーザとの議論はモデルをツールでどう表現するか(View)の話題に時間が費やされることになる。Viewはユースケースに対する依存度や個人の好みにも影響されるため、議論に多くの時間が費やされる。さらに、実例に基づかないと最適なモデルは作れない。実際、SIB2の開発も、SIB2UIのBepiColombo/MMO, SPRINT-A, ASTRO-H向け版の開発の初期段階(2010年4月頃)において、この種のデッドロックに陥っていた。



- (1) ツールの開発者は、モデルに基づき、設計ツールなど各種のツールを開発する。
- (2) 個々の機器の設計者は、設計ツールを用いて個々の機器の設計を行う。
- (3) モデルの制定者は、個々の機器の設計において有用なようモデル化を行う。

図1 モデルに基づく開発実現におけるスパイラル

モデルに基づき設計を行うツールを作成する手法は多様である。その一つとして、Excelは、各種の設計ツールを開発するためのインフラとして非常に良く利用される。これは、データの一覧性や、一括操作性、任意の付加情報の記述性をユーザに浸透したツールで実現でき、また、フルのGUIツールを開発するよりは低コストで開発できるためと考えられる。

実際、SIB2に基づく設計ツールにおいても、以下の4つの実装方法が試みられてきた。1) 山田による、汎用のUMLツールを利用する方法、2) 馬場による、Web用のフレームワーク(Ruby on Rails)を用い専用のツールを構築する方法、3) 山田による、スタンドアロンの専用プログラムを構築する方法、4) 松崎らによるExcel上で構築する方法。これらのうち、1)の方法は、ユーザ側において汎用で多機能なUMLツールの機能を学習する労力が高いため非採用となった。また、2)及び3)の方法は、専用のユーザインタフェースを作りこむコストが高くSIB2のモデル全体の実装には至らなかった。結局、SIB2に基づく設計ツールとして、実現に至ったのはSIB2UIに適用した4)のExcelを用いる方法のみであった。

Excelは、モデルに基づく設計ツールで利用されるものの、表現されるデータ構造のわかり易さには限界がある。スプレッドシートが素直に表現できるデータ構造は、表、つまり、1種類の構造体の繰り返し構造のみである。これを超えるデータ構造を扱う場合には、常に、どのように表現するのかというViewの問題に突き当たる。図2に、SIB2UIの開発段階で登場したユーザインタフェースの案を示す。これは、2種類の構造体が親子となった繰り返し構造を表すものであり、2案が併記されている。この種の表現の自由度は、ツールが出来上がった後も、ユーザに不利益を与える。つまり、どのようなデータ構造を表すのかマスを眺めても俄かにはデータ構造を把握しづらく、その理解に要する学習時間が個々の開発の設計のコストを押し上げることとなる。

| typeName | kind | description | kind property | | | | | | | |
|-----------------|---------------------|----------------------|---------------|-----------|---------|---------|---------|---------|---------|----|
| | | | k0 | k1 | k2 | k3 | k4 | k5 | k6 | k7 |
| /ERR_OK | enumeration | err ok 型 (システムが定義) | 0x00.OK | 0x01.ERR | | | | | | |
| /ENADIS | enumeration | ena dis 型 (システムが定義) | 0x00.DIS | 0x01.ENA | | | | | | |
| /HIGH_LOW | enumeration | high low 型 (システムが定義) | 0x00.LOW | 0x01.HIGH | | | | | | |
| /PCD_ERROR_CODE | enumeration | エラーコード | | | | | | | | |
| # | | error1 | 0x01.Error1 | | | | | | | |
| # | | error2 | 0x02.Error2 | | | | | | | |
| # | | error3 | 0x03.Error3 | | | | | | | |
| # | | error4 | 0x04.Error4 | | | | | | | |
| /TCIUXRX_COH | enumeration | 機種の項目へ同じ表示を割り当てる特殊な例 | 0x00.OK | 0x04.OK | 0x06.OK | 0x08.OK | 0x0A.OK | 0x0C.OK | 0x0E.OK | |
| # | | | 0TR.ERR | | | | | | | |
| /PCD | functionalObjectRef | Pxxxx Dxxxx Dxxxx | | | | | | | | |
| ERROR_CODE | enumeration | エラーコード | | | | | | | | |
| # | | error1 | 0x01.Error1 | | | | | | | |
| # | | error2 | 0x02.Error2 | | | | | | | |
| # | | error3 | 0x03.Error3 | | | | | | | |
| # | | error4 | 0x04.Error4 | | | | | | | |

図 2 複雑なデータ構造をスプレッドシート上に表した例

Excel を用いることで、Excel 自身が持つ GUI の機能を流用し、ある程度設計ツールの開発費を抑えることができる。しかし、その場合にも一定量の開発コスト・開発期間が必要である。表 3 に、SIB2UI の実際の開発規模感を示す。開発規模は、ほぼ表現しているデータ構造の大きさに比例 (1 カラムあたり ~0.5) しており、その低減が望まれる。

表 3 SIB2UI のこれまでのリリースと開発規模感
(単位は任意の相対値)

| |
|---|
| FY2005: 表 2 個, ~8 カラム, 開発規模 5 (*1) |
| FY2006: 表 2 個, ~13 カラム, 開発規模 +10 (*2) |
| FY2007: 表 4 個, ~30 カラム, 開発規模 +6 (*3) |
| FY2008: 表 4 個, ~30 カラム, 開発規模 +3 (*4) |
| FY2009-2010: 表 9 個, ~100 カラム, 開発規模 +26 (*5) |

- *1: SIB1.5 技術検討
- *2: SIB2 技術検討
- *3: BepiColombo/MMO EM ミッション総合試験向け
- *4: BepiColombo/MMO EM システム総合試験向け
- *5: BepiColombo/MMO, SPRINT-A, ASTRO-H 向け; 従来の開発手法の場合の見積もり
いずれも出力機能のみ実装, 入力機能を実装すれば規模は倍と見込まれる。

Excel から XML データを扱う手法は、幾つか存在する。図 3 の (a) は、単純な表形式の XML を扱うものである。この仕組みは Excel 2003 以降 Excel の標準機能として実装されている。一般にデータ構造は、キーを設定し、リレーションを張られた複数のデータ構造、最終的には単純な表にまで分割できるため、この方法のみでデータを扱うことは可能である。しかし、この方法では、情報が複数のテーブルに分かれることとなるため、設計ツールとしては使いづらい。図 3 の (b) は、逆に、Excel 上に XML の要素や属性名をそのまま見せることで任意の XML を扱うものである。この種のツールはいくつか存在するが、要素名や属性名は必ずしもわかり易いものではないし、全てをツリーで表現されたデータ構造はユーザにわかり易いものではないため、やはり、設計ツールにはなじまない。設計ツールの開発コストを下げるには、ある程度複雑なデータ構造をよりコンパクトに Excel 上に表現し、その操作をサポートする汎用ライブラリが求められる。そこで、我々は、2010 年 5 月に Excel ツールの開発方針の抜本的見直しを実施することとした。次章では、見直した開発方針と、その核となる汎用ライブラリ xmlExcelHyper を用いた設計ツールの構築について示す。

ユーザインタフェース

XML ファイル

| Column1 | Column2 | Column3 |
|---------|---------|---------|
| A | B | C |
| X | Y | Z |



```

<ROOT>
  <Row>
    <Column1>A</Column1>
    <Column2>B</Column2>
    <Column3>C</Column3>
  </Row>
  <Row>
    <Column1>X</Column1>
    <Column2>Y</Column2>
    <Column3>Z</Column3>
  </Row>
</ROOT>
    
```

(a)

| ROOT | A | | |
|------|---|-----|---|
| | | @A1 | 1 |
| | | B | |
| | | @B1 | 2 |
| | | C | 3 |
| | | B | |
| | | @B1 | 4 |
| | | C | 5 |



```

<ROOT>
  <A A1="1">
    <B B1="2">
      <C>3</C>
    </B>
  <B B1="4">
    <C>5</C>
  </B>
</A>
</ROOT>
    
```

(b)

図 3 XML データを Excel 上で扱う既存の手法

4 xmlExcelHyper による設計ツールの構築

ユーザインタフェースを無秩序に設計すると、ツールを作る側に労力を要するし、使う側のユーザビリティも悪いものとなる。そこで、我々が設定した開発方針のひとつめは、ユーザインタフェースにおけるデータの表現方法の自由度を減らすというものである。図 4 に、我々が選定したデータ構造の表現方法のイメージを示す。この表現方法では、ひとつとして、図 4 の (a) に示されるように、種別が固定的な階層関係は左右に並べることで示す (5.4 参照)。例えば、SIB2 においては「機能オブジェクトはオペレーション・アトリビュートを持つ」と言う関係は、SIB2UI において、この方法で表現した。また、図 4 の (b) に示されるように、種別が可変的な階層関係は入れ子構造で示す (5.3 参照)。例えば、SIB2 においては「オペレーションはブロックもしくはパラメータを持つ。ブロックはパラメータを持つ」と言う関係は、SIB2UI において、この方法で表現した。この方法を用いると、行ごとに表す対象が変わることとなり、表がまだらになる。いずれの場合においても、図 4 の (c) に示されるように、データ構造と同じ記法で表のラベルを設けることとした。

| functional object | name | kind | Information Definition | | | | attribute sequence property | |
|--|-------------------------------|-------------------|------------------------|-------------------|------------|----------------------------|--|---------------------|
| | | | octet position | bit position | bit length | data Type | | |
| functional object name lower APID = upper FOID and route MMO: N/A | attribute sequence name (def) | attributeSequence | | | | | Functional Definition (description of timing) | |
| | attribute name | attributeRef | | | | [automatically calculated] | [automatically copied] | |
| | # | reserve | | | | | [define here] | |
| | block name (def) | blockRef | | | | automatically calculated | | |
| | ... | ... | | | | | | |
| | attribute name | switch | | | | | | |
| | attribute value | osss | | | | | | |
| | ... | ... | | | | | | |
| | block name (def) | block | | | | automatically calculated | | |
| | ... | ... | | | | | | |
| (この行は使用しないで下さい) | | | | | | | | |
| /PCD | 0x01 | 0x01 | PCD_HK_pkt | attributeSequence | | | 衛星ネットワークのHK収集レートと同様 | |
| | | | PCD_HTR_Status | attributeRef | 7 | 0 | 16 | unsignedShort |
| | | | PCD_PWR_Status | attributeRef | 9 | 0 | 32 | unsignedInt |
| | | | PCD_PWS_PWR_Status | attributeRef | 13 | 0 | 16 | unsignedInt |
| | | | ... | ... | | | | |
| | | | /PCD_HK block | blockRef | 15 | 0 | 64 | |
| | | | ... | ... | | | | |
| | | | reserve | ... | 23 | 0 | 16 | |
| | | | # | ... | | | | |
| /PCD | DEI=0xFEFE Z=0xFF | 0x02 | HK_attrSeq | attributeSequence | | | | 衛星ネットワークのHK収集レートと同様 |
| | | | HK block | block | 7 | 0 | 64 | |
| | | | HV_Status | attributeRef | 7 | 0 | 16 | unsignedShort |
| | | | HV_ON_EnaDis | attributeRef | 9 | 0 | 16 | unsignedShort |
| | | | WDT_Status | attributeRef | 11 | 0 | 32 | unsignedInt |
| | | | # | ... | 15 | 0 | 16 | |
| | | | reserve | ... | | | | |
| | | | USER | attributeSequence | | | | 任意のタイミングで出力される |
| | | | EXPOSE_CNT | attributeRef | 7 | 0 | 32 | unsignedInt |
| /HKU | 0xFC | 0x04 | attrSeqCNT | attributeSequence | | | | テレメトリダウンリンクレートと同様 |
| | | | CNT | attributeRef | 7 | 0 | 16 | unsignedInt |

(a)

(b)

(SIB2UI functionalObject シートより)

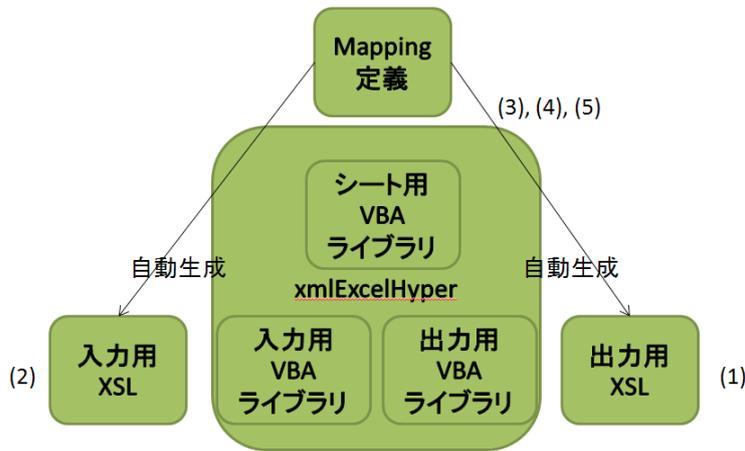
図 4 選定したデータ構造の表現方法のイメージ

我々が設定した開発方針のふたつめは、複雑な表と XML の対応の記述を把握しやすく、整合をとり管理しやすいものとする事である。この方法では、表の中身が XML の何に対応するかに対応 (マッピング定義) を、ユーザインタフェース上の表のラベルと上下にならべて表記する。なお、マッピング定義の記法については 5 章で、図 7、図 9、図 11 など例を用い示す。この記法を採用することで、項目の並び順の変更や、追加、削除に関し Excel ユーザインタフェースと XML ファイルの仕様策定を両者の対応の整合性をとりつつ実施でき、モデルとツールの開発が効率化される。SIB2・SIB2UI の開発では、この記法を投入することで、SIB2 のモデル決定におけるデッドロックを解消し、モデルが定まらないことによる開発コストの増大を抑制した。

マッピング定義は、XML の記法 (XML のスキーマ)、スプレッドシート上の記法 (スプレッドシートのスキーマ)、スプレッドシートと XML の対応関係の情報をすべて含んでいる。そこで、マッピング定義を参照すれば、原理的に、Excel から XML ファイルを読み書きすることが可能である。我々は、マッピング定義を参照し、Excel から XML ファイルを読み書きする汎用ライブラリ xmlExcelHyper を開発し、これを SIB2UI の開発に適用することとした。図 5 に xmlExcelHyper の機能・実装イメージを示す。xmlExcelHyper を用いるとサポートされるデータ構造の範囲内であれば、SIB2UI の 4 つの機能に対応する図中の (1)(2)(4)(5) の機能をツール毎のプログラミングなしで実現できる。

xmlExcelHyper は、Excel 上で動作するため、処理ロジックは基本的に VBA (Visual Basic for Application) で実現している。他方、スプレッドシートと XML 上の相互変換のロジックは、Excel に限らず一般的なスプレッドシートプログラムに適用可能なものである。そこで、将来、xmlExcelHyper を、他のスプレッドシートプログラムへも転用できるように、Excel からのファイルの入出力は表形式の XML ファイルとした。また、表形式の XML ファイルとマッピング定義で記される XML との相互変換のロジックについては、汎用な XML 変換技術である XSLT⁸⁾ のスタイルシート (以下単に XSL) を用いて実施することとした。xmlExcelHyper は、あらかじめ、マッピング定義に対応する XSL を自動生成し、ファイルの入出力時にこれを VBA から呼び出す。

スプレッドシート上のマッピング定義はツールの製作時にのみ必要なものであり、原理的には、Excel 上から消去してもツールとしては動作させるようにできる。しかし、まず開発した SIB2UI においては、ユーザに対し非表示とするようツール側で対処することとした。逆に、このようにしておくことで、ユーザであっても、xmlExcelHyper の仕組みを知っていれば、設計ツールに、任意の設計情報を追加できるものとなる。



- (1) マッピング定義から生成した出力用 XSL を使い、Excel から出力した表形式の XML をマッピング定義で規定する XML に変換する。(Export)
- (2) マッピング定義から生成した入力用 XSL を使い、マッピング定義で規定する XML を表形式の XML に変換し、Excel へ入力する。(Import)
- (3) マッピング定義を出力する。また、マッピング定義から XML 及びスプレッドシートのスキーマを抽出する。(スキーマ出力)
- (4) マッピング定義に基づき、シート上の入力値をチェックする。(Validate)
- (5) マッピング定義に基づき、階層を表す罫線描画を支援する。(Edit Support)

図 5 xmlExcelHyper の機能・実装イメージ

xmlExcelHyper の初版の開発は 2010 年度に終了し、SIB2UI も 2011 年 11 月のリリースにおいてこれを利用する実装に置き換えられた。開発した xmlExcelHyper は ISAS の衛星開発以外にも適用可能であり、web ページ^{d)} から一般に対し配布されている。SIB2UI への xmlExcelHyper の組み込みにおいては、特に問題は見いだされなかった。そこで、我々は適用した開発方針が妥当であったと考えている。ただし、構想段階で企画したものと出来上がった xmlExcelHyper の差異として、出力する XML スキーマは独自の記法のものとなったことがあげられる。これは、xmlExcelHyper の開発にかけた時間的な制約によるものであり、独自の記法は、原理的に RELAX NG に変換可能である。

xmlExcelHyper は SIB2UI の 4 つの機能 (Edit Support, Import, Export, Validate) のいずれにおいても、その実現を担う。ただし、xmlExcelHyper が担うのは複雑なデータ構造の中においても、SIB2UI に頻繁に登場したパターンの処理に限られる。これを超越する場合は、SIB2UI において外付けのプログラムを作成し、SIB2UI 全体の機能を実現している。SIB2UI・xmlExcelHyper の開発では、今回 xmlExcelHyper のサポート範囲に含めなかった範囲においても、今後、ライブラリ化の対象となりうる箇所も洗い出された。次章では、xmlExcelHyper が取り扱うデータ構造と XML 構造や、これらの間の対応を記述するマッピング定義の記法、SIB2UI を実現する上で洗い出された課題と複雑なデータを扱う機能のうち xmlExcelHyper で実現されたもの、されていないものについて述べることにする。

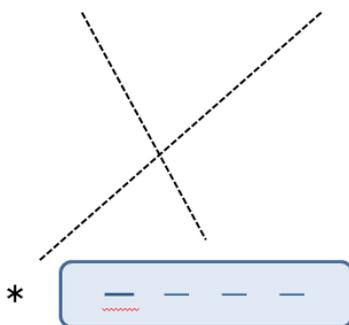
5 xmlExcelHyper が扱うデータ構造・XML 構造

xmlExcelHyper は、3 種類のデータ構造：単純テーブル (5.1)、選択テーブル (5.2)、階層テーブル (5.3) を基本とする。また、これらのテーブルに親子関係を規定し連結 (5.4) することができる。本章では、これらを順に示し、最後 (5.5) に xmlExcelHyper が対応しない表現方法について示す。

5.1 単純テーブル

xmlExcelHyper を用いて表現可能なデータ構造として最もシンプルなもの、単純テーブルと呼ばれるもののデータ構造とユーザインタフェース例をそれぞれ図 6 の (a) と (b) に示す。このデータ構造は複数のメンバを持つ構造体の 0~n 個の配列 (つまり表) からなるものである。SIB2UI は、この単純テーブルを 15 個有している。また、以降に示す選択テーブル、階層テーブルも含め、約 40 種の構造体を扱っている。

• 構造体の配列 (0~n個)



(a)

| name | attribute values for | | | | |
|-------|----------------------------|-------------|---------------|---------------|-------------|
| | attribute limit name [def] | (fatal min) | (caution min) | (caution max) | (fatal max) |
| BATL | | 10 | 15 | 25 | 30 |
| BATH | | 10 | 15 | 25 | 30 |
| RWA1 | | 0 | 10 | 30 | 40 |
| RWA2 | | 0 | 10 | 30 | 40 |
| RWA3 | | 0 | 10 | 30 | 40 |
| RWA4 | | 0 | 10 | 30 | 40 |
| STT | | -50 | 0 | 50 | 100 |
| CH08 | | -10 | 10 | 30 | 50 |
| STRPA | | -10 | 10 | 30 | 50 |
| STRPB | | -10 | 10 | 30 | 50 |
| SHNT | | -10 | 10 | 30 | 50 |
| PCU | | -10 | 10 | 30 | 50 |
| IRU | | -10 | 10 | 30 | 50 |
| SAP1 | | -50 | 0 | 50 | 100 |
| SAP2 | | -50 | 0 | 50 | 100 |
| GH16 | | -10 | 10 | 30 | 50 |

(SIB2UI attributeLimitt シートより)

(b)

図 6 単純テーブルのデータ構造, ユーザインタフェース例

図 7 に、単純テーブルにおける、マッピング定義と対応する表のラベルの例を示す。単純テーブルの定義は、構造体と XML との対応の記述を一つ含む。ここで、構造体と XML との対応 (以下、構造体定義) は ‘_ts’ と記述したセルの下のセルを左上、‘_br’ と記述したセルの上のセルを右下とする範囲に記述する。図 7 の例では、見やすさのためマッピング定義に罫線を描画してあり、構造体定義の範囲は太線の囲みで示される。ただし、罫線は xmlExcelHyper の処理には用いられない。

図 7 の構造体定義は 12 列からなり、右端の列はセルが空欄のみである。空欄のみの列は、データの表現には用いない。

d) <http://www.c-soda.isas.jaxa.jp/software/>

残りの 11 個の列が、それぞれ、構造体のメンバの定義となる。それぞれのメンバはテーブル定義の下に並べられた表のラベルで示される内容を保持する。テーブル定義の最下行 (5) には、構造体メンバの値に制約がある場合、その名称を $\${}$ で囲み記述する。それぞれの名称に対応する制約は RELAX NG のデータ型定義を記述したファイルを別途作成し、xmlExcelHyper に読み込ませる (図 15 参照)。なお、制約が無い場合は空欄とする。

| | | | | | | | | | | |
|--------|---------|--|--|--|--|--|--|--|--|--|
| (1)... | (3) (4) | | | | | | | | | |
| (2)... | (3) (4) | | | | | | | | | |
| (5)... | (3) (4) | | | | | | | | | |

| functional object | FO property | | description | ext1 | ext2 | ext3 | ext4 | ext5 |
|------------------------------|-----------------------------------|----------------------|-------------|---|------------------------|------|------|------|
| functional object name [def] | lower APID = upper FQID and route | lower FQID MMIO: N/A | (ack) | name of condition to be effective or always | (if functional object) | | | |

(SIB2UI functionalObject シートより)

図 7 単純テーブルのマッピング定義 (上段) と対応する表のラベル (下段) の例

それぞれの構造体メンバをどの XML 要素・属性に対応させるか、また要素・属性の親子関係は、(1)~(2) に示される最下行より上の行に記述する。この記法は図 4 の (a) と同様 (ただし、左右ではなく上下) に、要素・属性の親子関係を可視化したものである。それぞれのセルには、XML のノードを指し示す汎用な技術である xpath⁹⁾ の記法を用いる。つまり、要素の場合は要素名を記載する、属性の場合は属性名に @ を冠し記載する。また、テキストノードの場合 text() と記載する。図 7 の例は、最上位行 (1) に示す sib:functionalObject 要素が、中間行 (2) に示す name 属性、sib:lowerAPID 要素、IDValue 属性、sib:interaction 要素、sib:effectiveStates 要素、sib:description 要素、extension1 から extension5 属性を有すこと、また、sib:lowerAPID 要素及び sib:description 要素がいずれもテキストノードを有すること、さらに、sib:effectiveStates 要素が sib:conditionRef 要素を有し、sib:conditionRef 要素が name 属性を有することを表している。なお、空欄のセル、'*' のみが記述されたセルは、それぞれの列の定義において無視され、XML 上の構造としての意味はもたない。また、最下行 (5) は、XML 上の属性もしくは末端の要素の値に対応する。

要素もしくは属性の多重度は (3) に示すよう、正規表現の記号、すなわち、0 か 1 (optional) の場合 '?', 0 以上 (zeroOrMore) の場合 '*', 1 以上 (oneOrMore) の場合 '+' を xpath の記法の右に付加することで記述する。なお、属性については XML のデータモデルに合わせ '?' のみが指定可能である。一方、要素では zeroOrMore もしくは oneOrMore の指定も可能である。zeroOrMore もしくは oneOrMore を指定した場合、スプレッドシート上のセルの値を空白区切りで分解し、分解されたもの一つ一つを要素の値に対応させることを意図している。しかし、このような繰り返しの指定は SIB2UI において 2 個しか登場しないため、現状は不完全な実装^{e)} に留めている。この実装において xmlExcelHyper は一つのセルを一つの要素に対応させる。そこで、SIB2UI では、要素を分割・集約する処理を個別実装している (後述の個別対応項目 3 に含まれる)。

スプレッドシート上の並び順と XML の要素の並び順が一致するとは限らない。スプレッドシート上の並び順と異なる順序で XML の要素を並べる場合、(4) で示すよう、xpath の記法 (つまり要素名) の左に N. の形式で並び順を表記する。図 7 の例では、sib:functionalObject 要素の子として、順に sib:lowerAPID 要素、sib:effectiveStates 要素、sib:description 要素、sib:interaction 要素が並ぶこととなる。

データ型定義を用いると値がとりうる制約として、「N 個の値から選択する」という表記も可能である。この場合、Excel 上で、その選択肢をプルダウンメニューから選べるようにすることが望ましい。しかし、現状の xmlExcelHyper は対応しておらず、SIB2UI での個別実装となっている (個別対応項目 1)。

xmlExcelHyper では、XML の属性・要素が optional もしくは zeroOrMore の多重度が指定されていて、これに対応セルが空欄の場合、これらの属性・要素は XML に出力しない。さらに、この多重度を XML 上の親子関係において複数回指定した場合、スプレッドシート上のデータ構造と XML との関係が不定となることがある。例えば、表 4 のケースにおいて、XML の構造として親子関係にある fatalCondition 要素、minParam 要素のいずれも optional に指定するとする。このケースで、値 1 が存在しない場合 (ア) として扱うべきなのか (イ) として扱うべきなのか仕様の設定に判断が必要となる。

e) 現状の実装では '*' の指定は '?' の指定として扱われる。 '+' の指定は無指定として扱われる。

この判断が不要となるように xmlExcelHyper がサポートするマッピング定義では属性と末端の要素のみに多重度の指定が可能という制約を設けた。つまり、表 4 の例において (ア) のみを扱うこととした。

表 4 Optional の扱い

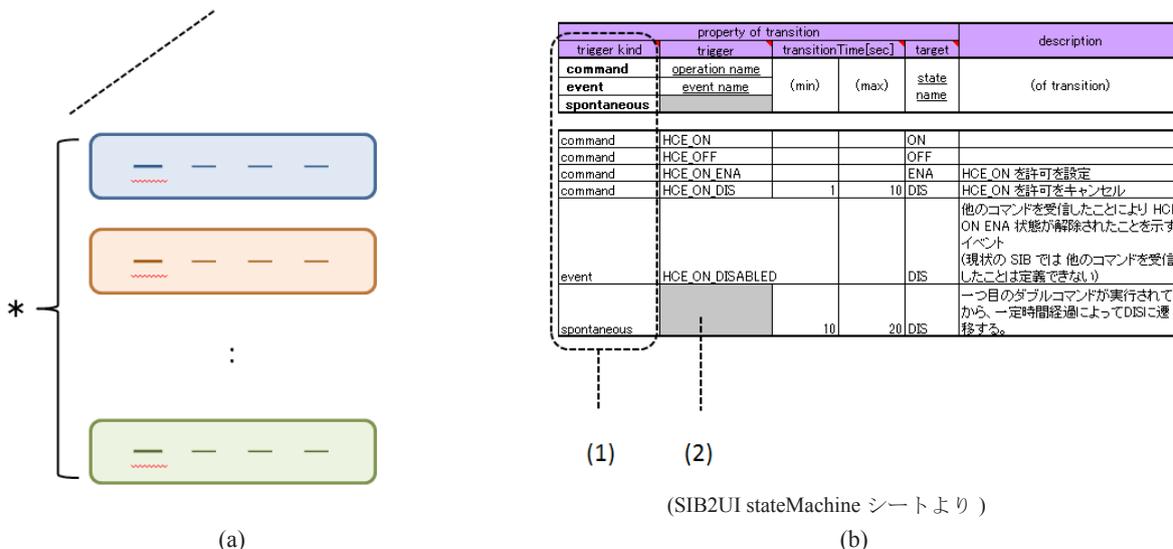
| 値がある場合 |
|--|
| <pre><attributeLimite> <fatalCondition> <minParam> 値 1</minParam> </fatalCondition> </attributeLimite></pre> |
| 値が無い場合 |
| (ア) |
| <pre><attributeLimite> <fatalCondition> </fatalCondition> </attributeLimite></pre> |
| (イ) |
| <pre><attributeLimite> </attributeLimite></pre> |

5.2 選択テーブル

xmlExcelHyper を用いて表現可能なデータ構造として 1 段複雑な、選択テーブルと呼ばれるもののデータ構造とユーザインタフェース例をそれぞれ図 8 の (a) と (b) に示す。このデータ構造では、複数の構造体を任意の順で並べ表を作ることができる。いずれの構造体を選んだかを示すため、表には構造体の種別を示す列 (kind 列) を設けることとする。この方法を用いると、行ごとに表す対象が変わることとなり、行ごとに入力可能な列が異なることとなる。

選択テーブルのユーザインタフェースにおいて、xmlExcelHyper はマッピング定義に従い、kind 列入力用のプルダウンメニューを生成する。また、この列の入力時に、それぞれの構造体に応じ入力不能なセルに入力が不能であることを示すなどのため色付けの処理を行う。SIB2UI はこの選択テーブルを 7 個 (遷移のトリガ、条件・イベント種別、条件式の演算子 (4 重)、アトリビュートの変化則) 有している。

• (構造体の選択)の配列 (0~n個)



- (1) kind 列 : それぞれの行が構造体の何れを選択したかを表すカラム
- (2) 構造体によりフィールドの数や意味が異なるため、行ごとにどの列が入力可能か異なる

図 8 選択テーブルのデータ構造, ユーザインタフェース例

図 9 に選択テーブルのマッピング定義と対応する表のラベルの例を示す。一つのテーブル定義は ‘_switch_tl’ と記述したセルの下のセルを左上, ‘_switch_br’ と記述したセルの上のセルを右下とする範囲である。選択テーブルの定義には 5.1 で述べた構造体定義を上下に複数積み重ねたものを含む。図 9 の例には、太線の囲みで示す 3 つの構造体定義が含まれている。これに加え、kind 列を ‘_case’ の記述により指定する。kind 列に含まれ、値の右端に ‘:’ と記したセルには、スプレッドシート上で構造体の種別を表す文字列を指定する。このセルは XML 上のデータ構造としては意味を持たない。構造体定義では最下段のセルに背景色を指定する。この背景色はユーザが kind 列を設定した際に、その行のセルの背景色として設定される。

| | | | | |
|--------------|---|-------------------------------|--------------------------|---------------------------------|
| switch_tl | | | | |
| _case | ts | | | |
| | sib:transition[descendant:sib:operationRef] | | | |
| | 1 sib:Trigger | 2 sib:transitionTimeCondition | | @target |
| | sib:operationRef | sib:minParam ? | sib:maxParam ? | * |
| | @name | text() | text() | * |
| command: | `\${structuralName data-def}` | `\${timeValue data-def}` | `\${timeValue data-def}` | `\${transitionTarget data-def}` |
| | ts | | | |
| | sib:transition[descendant:sib:eventRef] | | | |
| | 1 sib:Trigger | 2 sib:transitionTimeCondition | | @target |
| | sib:eventRef | sib:minParam ? | sib:maxParam ? | * |
| | @name | text() | text() | * |
| event: | `\${structuralName data-def}` | `\${timeValue data-def}` | `\${timeValue data-def}` | `\${transitionTarget data-def}` |
| | ts | | | |
| | sib:transition[not(descendant:sib:eventRef) and not(descendant:sib:operationRef)] | | | |
| | 1 sib:Trigger | 2 sib:transitionTimeCondition | | @target |
| | | 1 sib:minParam ? | 2 sib:maxParam ? | * |
| | | text() | text() | * |
| spontaneous: | | `\${timeValue data-def}` | `\${timeValue data-def}` | `\${transitionTarget data-def}` |
| | ts | | | |
| | _br | | | |
| | switch_br | | | |

| property of transition | | | | | |
|------------------------|----------------|---------------------|-------|------------|-----------------|
| trigger kind | trigger | transitionTime[sec] | | target | description |
| command | operation name | | | | |
| event | event name | (min) | (max) | state name | (of transition) |
| spontaneous | | | | | |

(SIB2UI stateMachine シートより)

図 9 選択テーブルのマッピング定義（上段）と対応する表のラベル（下段）の例

xmlExcelHyper のデータモデルでは、構造体の中身が構造体の登場箇所によらず、同じものである。他方で、xmlExcelHyper では、異なる構造体を同一の XML 要素に対応させることができることとした。表 5 に示す例では、異なる構造体を同一の XML 要素 <A/> に対応させている。この場合、<A/> 要素がいずれの構造体に対応するかは、子要素の違いを見なければ判断できない。そこで、XML の読み込みにおいて注意が必要となる。この種の判断を一般的に行うロジックは複雑なものとなるため、xmlExcelHyper では、冗長な情報ではあるものの、ツールの設計者において、それぞれ、A[B1] あるいは A[C1] などと、XML 読み込みの判定の際のヒントを xpath の述語の記法を用い、マッピング定義の要素名の指定の箇所へ記載することとした。この記述法は、図 9 の例においても、sib:transition 要素の指定で用いられている。

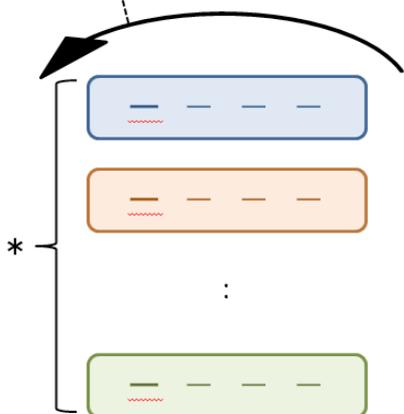
表 5 異なる構造体の同じ要素名へのマッピング

| |
|---|
| (a) |
| <A> <B1> 値 </B1> <B2> 値 </B2> |
| (b) |
| <A> <C1> 値 </C1> <C2> 値 </C2> |

5.3 階層テーブル

xmlExcelHyper を用いて表現可能なデータ構造としてさらに複雑な、階層テーブルと呼ばれるもののデータ構造とユーザインタフェース例をそれぞれ図 10 の (a) と (b) に示す。このデータ構造では、複数の構造体で任意の入れ子を組むと共に並べて表を作ることができる。いずれの構造体を選んだかを示すため、表に kind 列を設けることは選択テーブルと同様である。階層テーブルにおいては、親子関係を、階層的な段差により示し、同一種類の構造体の入れ子も可能となっている。なお、この段差を用い階層を示す列を階層化部と呼ぶ。xmlExcelHyper はこの階層関係が分かり易いよう可視化するため、罫線を描画する機能を持つ。SIB2UI は、この階層テーブルを 3 個 (機能オブジェクトの子、アトリビュートシーケンス、データ変換) 有している。

- (構造体の選択)の配列 (0~n個) の入れ子



(a)

| name | kind | type or parameter property | | operation property | | |
|----------------------|--------------------------|---|--------------------------|----------------------------|-------------------|--|
| operation_name {dir} | operation | | | attribute change rule name | (swarn, prohibit) | name of condition to be effective or always (初期値が既定/なしの場合) |
| # | constant | {byte,unsignedByte,short,unsignedShort,long,unsignedLong,float,double,hexBinary,variableLengthBinary} | (encode conversion name) | | | |
| parameter_name {dir} | operationAnalogParameter | | | | | |
| | enumParameter | enumeration_name | | | | |
| # | attributeRefParameter | attribute_name | | | | |
| block_name {dir} | block | | | | | |
| ... | | | | | | |
| ALL_AUTO | operation | | | | | always |
| DIR1_SETTEMP | operationAnalogParameter | short | /COS TPS RANGE1 | | | always |
| HITEMP | operationAnalogParameter | short | /COS TPS RANGE1 | | | |
| DIR3_SETTEMP | operationAnalogParameter | short | /COS TPS RANGE2 | | | always |
| HITEMP | operationAnalogParameter | short | /COS TPS RANGE2 | | | |
| SET_RANGE | operation | | | | | always |
| MOVE | attributeRefParameter | | MOVE | | | |

(1)

(2)

(SIB2UI functionalObject シートより)

(b)

- (1) 階層化部: 入れ子関係は表で階層的な段差により示す (同一種類の構造体の再帰的な入れ子も可能)
- (2) kind 列: それぞれの行が構造体の何れを選択したかを表すカラム

図 10 階層テーブルのデータ構造, ユーザインタフェース例

図 11 に階層テーブルのマッピング定義と対応する表のラベルの例を示す。階層テーブルの定義は、選択テーブルの定義に対し、階層化部の指定を加えたものである。階層テーブルの定義は、選択テーブルと同様に、複数の構造体定義と kind 列の指定を含んでいる。例では、7 個の構造体定義を含み、左から 3 列目に kind 列が指定されている。これに加え、階層テーブルでは、階層化部を ‘_level_s’ を記述した列から ‘_level_e’ を記述した列の範囲に記述する。例では、左端の 2 列が階層化部に指定されている。階層化部のマッピング定義のうち、構造体メンバと XML の対応は左端の列のみ記述し、他の列は空欄にしておく。

| | | | | | | |
|--|------------------------|------------------------------------|------------------------------------|------------------------------------|------------------------------------|--|
| switch.tl | | | | | | |
| level_s | level_e | case | | | | |
| ts | | | | | | |
| sib: function | | | | | | |
| @name | | | | | | |
| \$(userFunctionName_data-def) | function: | | | | | |
| ts | | | | | | |
| sib: attributeRefParameter | | | | | | |
| @name | | | | | | |
| \$(structuralName_data-def) | attributeRef: | | | | | |
| ts | | | | | | |
| sib: constParameter | | | | | | |
| text() | | | | | | |
| \$(constParameter_data-def) | constant: | | | | | |
| ts | | | | | | |
| sib: default[sib: polynomialParameter] | | | | | | |
| @condition | | sib: polynomialParameter | | | | |
| | | @c0? | @c1? | @c2? | @c3? | |
| | | @c4? | @c5? | | | |
| \$(sc_conditionName_data-def) | polynomial: | \$(polynomialCoefficient_data-def) | \$(polynomialCoefficient_data-def) | \$(polynomialCoefficient_data-def) | \$(polynomialCoefficient_data-def) | |
| ts | | | | | | |
| sib: case[sib: polynomialParameter] | | | | | | |
| sib: conditionRef | | | | | | |
| @name | | sib: polynomialParameter | | | | |
| | | @c0? | @c1? | @c2? | @c3? | |
| | | @c4? | @c5? | | | |
| \$(conditionName_data-def) | conditionalPolynomial: | \$(polynomialCoefficient_data-def) | \$(polynomialCoefficient_data-def) | \$(polynomialCoefficient_data-def) | \$(polynomialCoefficient_data-def) | |
| ts | | | | | | |
| sib: default[sib: expression] | | | | | | |
| @condition | | sib: expression | | | | |
| | | text() | | | | |
| \$(sc_conditionName_data-def) | expression: | \$(expression_data-def) | | | | |
| ts | | | | | | |
| sib: case[sib: expression] | | | | | | |
| sib: conditionRef | | | | | | |
| @name | | sib: expression | | | | |
| | | text() | | | | |
| \$(conditionName_data-def) | conditionalExpression: | \$(expression_data-def) | | | | |
| ts | | | | | | |
| restrictions | | | | | | |
| / := (function conditionalPolynomial conditionalExpression polynomial expression)* | | | | | | |
| function := (attributeRef)*constant)* | | | | | | |
| restriction_e | | | | | | |

| condition_or | kind | property of conversion | | | | |
|--------------------------------|-----------------------|------------------------|------|------|------|------|
| always [automatically defined] | polynomial | (r0) | (r1) | (r2) | (r3) | (r4) |
| always [automatically defined] | expression | expression | | | | (r5) |
| (name of condition, otherwise) | conditionalPolynomial | (r0) | (r1) | (r2) | (r3) | (r4) |
| (name of condition, otherwise) | conditionalExpression | expression | | | | (r5) |
| function name | variable parameter | | | | | |
| | constant parameter | attributeRef | | | | |
| | constant | | | | | |

(SIB2UI conversion シートより)

図 11 階層テーブルのマッピング定義（上段）と対応する表のラベル（下段）の例

XML 自身は、要素の無限な入れ子が可能である。他方で、階層テーブルにおいては、スプレッドシート上で視覚的に階層を表現するため、階層数に上限を設ける必要がある。xmlExcelHyper は、任意の上限値を許容する。SIB2UI においては、機能オブジェクトとアトリビュートシーケンスに対しては 9 段、データ変換に対しては 2 段と上限を設定した。ここで、機能オブジェクトとアトリビュートシーケンスにおいては、任意の入れ子があり得るので実用的に十分な段数を設定した。また、データ変換に対しては、原理的な最大値を設定した。なお、データ変換に関しては 5.4 に示すテーブルの連結を用いても表現可能であるが、表の幅がコンパクトになるよう階層テーブルを用いた表現を採用した。

各種のデータ構造においては、構造体の親子関係、並び順に、制約がありうる。そこで、マッピング定義には、表 6 の例に示すように、ある構造体に対し、子の構造体として許される並び順を正規表現にて示すことで、制約を記述することとした。この記述は、図 11 の例に示されるようにマッピング定義の ‘_restriction_s’ と書かれたセルの下のセル、 ‘_restriction_e’ と書かれたセルの上のセル、これらセルの間のセルに記述する。

表 6 制約の記述例

```

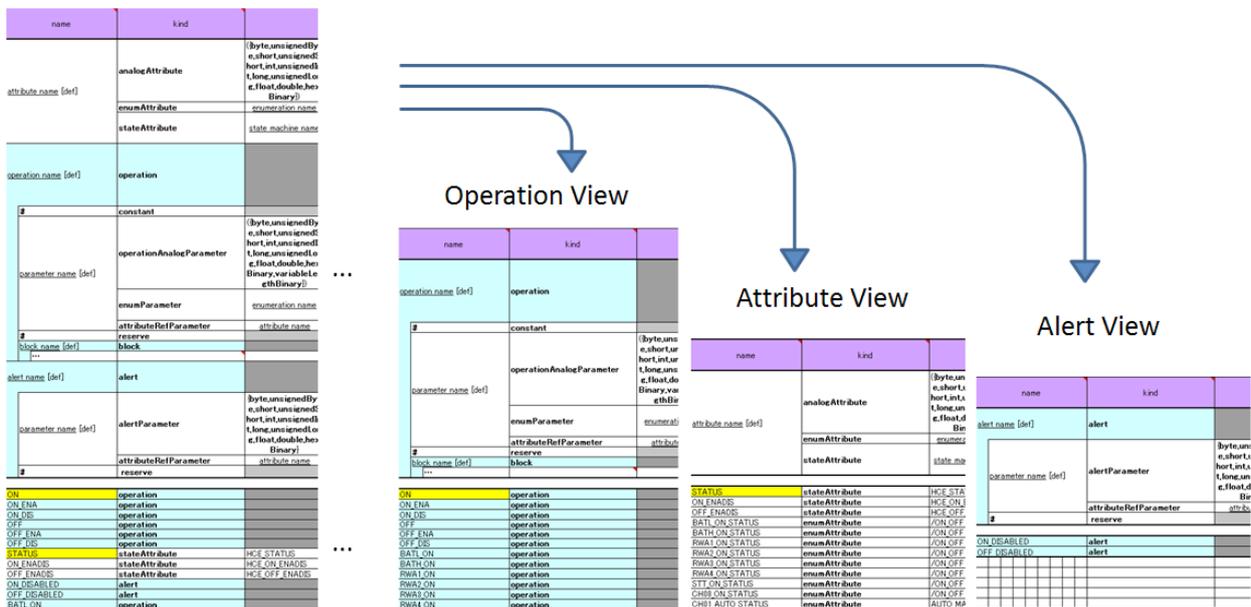
/ := ( attribute | enumAttribute | stateAttribute | operation | alert ) *
operation := constant ? ( block | operationAnalogParameter |
enumParameter | attributeRefParameter | reserve ) *
block := ( operationAnalogParameter | enumParameter |
attributeRefParameter | reserve ) *
alert := ( alertAnalogParameter | enumParameter |
attributeRefParameter | reserve ) *
    
```

この記法により、xmlExcelHyper が表現可能な XML 構造は、XML のスキーマの一つである、RELAX NG と比べ

- Mixed content (要素が、要素とテキストノードの双方持つ) は対象としない
- オptional と指定した要素・属性は値が空とならない
- 任意の要素名、属性名の登場を許容しない
- Interleave を許容しない
- 要素の入れ子の段数に上限がある

サブセットとなる。この種の表現が不能であっても、XML をデータの記述に用いる上で大きな支障とはならない。一方で、大きな階層化テーブルは必ずしも、把握しやすいものではないので、よりわかり易く表現する方法が必要である。

複雑な階層テーブルをシンプルに見せる手法のひとつとしてビューを設ける方法があげられる。SIB2UI においては、図 12 に示すように、選択された一部の構造体以外を行方向、列方向に非表示にすることで見える情報を厳選し、設計をより入力・把握しやすくする機能を設けている。この種の要求は、一般的な機能と考えられるが今のところ xmlExcelHyper には実装されておらず、SIB2UI での個別実装となっている (個別対応項目 2)。次節では、階層テーブルを用いずに構造体の親子関係を示す方法を示す。



(SIB2UI functionalObject シートより)

図 12 ビューによる表示情報の制限

5.4 テーブルの連結

親子関係の可視性を高めるために、xmlExcelHyper では、単純テーブル、選択テーブル、階層テーブルを左右に連結することで親子関係を表現することとした。図 13 の (a) に連結したテーブルのデータ構造の例、(b) に単純テーブル二つと選択テーブル一つを連結させたユーザインタフェース例を示す。xmlExcelHyper は、この親子関係がわかり易いように罫線を描画し、またセルを結合する機能を持つ。

| | | | | | | | | | |
|--|--------------------------------|-------------------|--------------------------|-----------------------|-----------------------|------------|------------|------------|---------------|
| <pre> namespace sib="http://oso.isas.jaxa.jp/sib/2.0" include include_dataPattern.rng start = element sib.defineEnumerations { functionalObjectRef-def* } </pre> | | | | | | | | | |
| ts | ts | ts | ts | ts | ts | ts | ts | ts | ts |
| sib:functionalObjectRef | sib:enumeration | | sib:enumerationLiteral + | | | | | | |
| @name | @name | sib:description ? | @value | @name | @stateLevel ? | @extensi | @extensi | @extensi | @extensionB ? |
| * | * | text() | * | * | * | * | * | * | * |
| * | * | | * | * | * | * | * | * | * |
| \$(structuralName_data-def) | \$(nonStructuralName_data-def) | | \$(enu_rawValue_data) | \$(nonStructuralName) | \$(stateLevel_data-c) | \$(string) | \$(string) | \$(string) | \$(string) |
| | br | | br | | | | | | br |
| | | | | | | | | | mark_br |

| functional object | name | description | kind property | | | ext1 | ext2 | ext3 | ext4 | ext5 |
|------------------------|------------------------|------------------|------------------------|--------------------------|---------------------|------|------|------|------|------|
| functional object name | enumeration_name [def] | (of enumeration) | { encode_value , OTR } | enumerationLiteral [def] | ((fatal, caution!)) | | | | | |

(SIB2UI enumeration シートより)

図 15 テーブルの連結のマッピング定義（上段）と対応する表のラベル（下段）の例

原理的に、階層テーブルは連結の任意の個所に登場して構わない。しかし、親子関係が簡単に把握できるよう SIB2UI では、階層テーブルは最も右にのみ登場することとし、xmlExcelHyper にも同様の制約を課すこととした。

5.5 xmlExcelHyper では対応できない表現方法

本節では、SIB2UI で用いている、あるいは、設計の際に議論されたデータの表現方法のうち、xmlExcelHyper ではカバーされていないものを示す。まず、5.5.1 でセル内のデータ構造の表現について述べる。次に、5.5.2 で参照関係を用いたデータ構造の分割について述べる。

5.5.1 セル内のデータ構造の表現

図 16 に、セル内のデータ構造の表現方法のうち、xmlExcelHyper の機能ではカバーされていないものの例を示す。(a) では、セル内で構造体や繰り返しを表現する。この例では、空白やカンマ区切りなどで繰り返しを表現するとともに、項目を区切り字（‘:’ など）で区切り、構造体のメンバを表現している。(b) では、セル内で、表現する構造体を選択するための文法を導入している。(c) では、複数のセルを列方向（さらに行方向）に並べ繰り返しを表現している。これらのうち、(a) (b) は SIB2UI でも利用されており、SIB2UI 側で外付けの XSL を作成し、処理が行われている（個別対応項目 3）。ただし、これらの表現方法も、セルの値が妥当かのみならば、データ型ライブラリの文字列型に正規表現の制約を適用することで、xmlExcelHyper のみで判定が可能となっている。

00:BLACK 01:RED 02:GREEN 04:BLUE 07:WHITE

(a)

| |
|-----------------------|
| command HCE_ON |
| command HCE_OFF |
| command HCE_ON_ENA |
| command HCE_ON_DIS |
| event HCE_ON_DISABLED |
| spontaneous |
| command HCE_OFF_ENA |
| command HCE_OFF_DIS |
| HCE_OFF_DISABLED |
| spontaneous |

(b)

| | | | | |
|-----------|----------|----------|----------|--|
| 00:BLACK | 01:RED | 02:GREEN | 04:BLUE | |
| 03:YELLOW | 05:MAGEN | 06:CYANT | 07:WHITE | |

図 16 対応していないデータ構造の表現（セル内）

5.5.2 参照関係を用いたデータ構造の分割

一般にデータ構造においては ID を用いた参照関係により二つのデータ構造に分割できる。親子のあるデータ構造においては、以下の参照関係がありうる。

A) 親 (1) に ID を定義し、子 (n) から参照する。

B) 子 (n) に ID を定義し、親 (1) から参照する。

親子のデータ構造を スプレッドシート上で、ID を用いた参照関係で表現するか、階層テーブルあるいはテーブルの連結で表現するかはシート内外の表現方法として多様性がある。また、XML 上で、ID を用いた参照関係で表現するか、要素の入れ子で表現するかにも、XML ファイル内外の表現方法として多様性がある。また、これら、スプレッドシート上と XML 上の表現方法は、各々、独立に決めることができる。xmlExcelHyper はこれらのデータ構造のいずれをサポートする機能ももたない。

SIB2UI は、複数のシートの各々を XML ファイルに対応させる。単純化のため、SIB2UI では基本的に、シート間の ID を用いた参照関係を XML ファイル間の参照関係に対応させている。xmlExcelHyper はこのデータ構造を扱わないため、SIB2UI では、参照整合性の確認を外付けのプログラムで実施している (個別対応項目 4)。また、シート間にデータ構造が分割されている場合、ユーザが設計を入力する際には、定義元・参照先に移動する機能、参照において記載可能な候補から入力する機能 (個別対応項目 5) が有用である。さらに、ユースケースに応じ、シート間の情報を集約したビューで見せる機能 (個別対応項目 6) も有用である。SIB2UI においては、これら個別対応項目 5 と 6 を、外付けの別プログラム SIB2Viewer にて実現している。SIB2UI においては management シートを除く 8 つのシートの間に、A) の参照関係が 7 個、B) の参照関係が 7 個、計 14 個と多くの関係が張られており、これらが個別対応項目 4 と 5 の個別プログラミングの対象となっている。そこで、これらの項目は、今後、xmlExcelHyper におけるライブラリ化の対象として優先度が高いと考えられる。

SIB2UI においても、例外的に、スプレッドシート上と XML 上で異なるデータ構造の表現を採用した個所がある。ステートマシンの XML ファイルにおいては、UML の XML 表現である XMI と構造が近くなるよう表現方法を設定したため、遷移と状態の間で ID を用いた参照関係を張ることとした。他方で、Excel 上での表現においては、設計のしやすさを考慮し、遷移の開始状態を元に、状態と親子関係を設定しテーブルの連結を用いて表現することとした。これらのデータ構造の変換は、外付けの XSL で行なっている (個別対応項目 7)。

6 まとめ

複雑なデータ構造を持つデータベース (SIB2) に対し、Excel 上で内容を記述し、XML に入出力するツール SIB2UI を開発した。スプレッドシートとの XML ファイルの対応は、スプレッドシート上にマッピング定義を記述することで、XML ファイルの仕様とスプレッドシートの仕様、両者の対応がいっぺんに視認でき、項目の追加、削除、並び替えにおいて整合性が担保されるようにした。また、マッピング定義に基づき、スプレッドシート上のデータの妥当性を確認、データの階層関係を可視化、XML ファイルとの入出力を行う汎用ライブラリ xmlExcelHyper を作成することで、アプリケーション側の開発の省力化を図った。

xmlExcelHyper が扱うデータ構造として、SIB2UI が扱うデータ構造を分析し、特に頻繁に登場するパターン、構造体を扱う 3 種類のテーブル (単純テーブル、選択テーブル、階層テーブル) を含めた。SIB2UI においては、この構造体が約 40 個登場する。また、単純テーブルが 15 個、選択テーブルが 7 個、階層テーブルが 3 個登場する。これらのうち最も表現範囲が広い階層テーブルを用いると、扱える XML は、RELAX NG スキーマよりは狭いものの、有限な階層という制約の下において、正規文法に沿った任意の XML を規定できる。ただし、階層テーブルは、そのままでは可視性に難がある。そこで、データ構造の入れ子関係の可視性を改善するためテーブルを連結できることとした。しかし、この方法を用いると、適用可能な XML は限定的なものとなる。そこで、現状の xmlExcelHyper が威力を発揮するのは、XML の仕様とスプレッドシートの仕様を同時に決められるケースと考えられる。

SIB2UI では、xmlExcelHyper には含まなかった幾つかの処理 (1: 選択項目のスプレッドシートと XML の表現の対応, 2: 階層テーブルをシンプルに見せるための View の作成, 3: セルでの繰り返し・構造体の表現, 4: 定義・参照関係のチェッ

ク, 5: 定義・参照関係の編集をサポートする機能, 6: 見やすいビューを表示する機能, 7: 参照を用いたデータ構造の表現方法の変換)を, 現状ではアプリケーション (SIB2UI) 側で個別対応としている。これらのうち, 項目4と5の対象であるデータ構造, すなわちシート間の参照関係については SIB2UI においては多数 (8つのシートの間に, 14の参照関係) 張られており, 今後の xmlExcelHyper サポート範囲の拡充として期待される。

謝辞

xmlExcelHyper の設計・製作は, 富士通株式会社, 富士通エフ・アイ・ピー・システムズ株式会社によって実施されました。山地尋之さん, 上妻巧哉さん, 川田耕一郎さん, 中野華奈さん, 境野淑介さん, 田中孝明, 広瀬智紀さんほかに感謝の意を表します。

参考文献

- 1) 馬場肇, 松崎恵一, 福田盛介, 山田隆弘, 飯塚祐介, 山地尋之, 大石克己: SIB2/GSTOS1 の開発, 第 11 回宇宙科学シンポジウム, 2011 年 1 月
- 2) Yamada, T.: Standardization of Spacecraft and Ground Systems Based on a Spacecraft Functional Model, SpaceOps 2008, Heidelberg, Germany, May 2008.
- 3) Yamada, T. and Matsuzaki, K.: Model-Based Development of Spacecraft Onboard Functions, International Council on Systems Engineering (INCOSE) International Symposium 2010, July 2010, Chicago USA.
- 4) Takahiro Yamada: Functional Model of Spacecraft (FMS) DRAFT, Issue 0.9, March 2011
- 5) Takahiro Yamada: Spacecraft Monitor and Control Protocol (SMCP) DRAFT, Issue 0.13, 10 August 2011
- 6) 山田隆弘, 松崎恵一: 衛星情報ベース 2 定義 (案) 第 0.9 版, 2009 年 11 月 11 日
- 7) RELAX NG Specification, Committee Specification 3 December 2001, Ed. James Clark and MURATA Makoto, The Organization for the Advancement of Structured Information Standards 2001 (<https://www.oasis-open.org/committees/relax-ng/spec-20011203.html>)
- 8) XSL Transformations (XSLT) Version 2.0, W3C Recommendation 23 January 2007 (<http://www.w3.org/TR/xslt20/>)
- 9) XML Path Language (XPath) 2.0 (Second Edition), W3C Recommendation 14 December 2010 (Link errors corrected 3 January 2011) (<http://www.w3.org/TR/xpath20/>)

科学衛星のための衛星時刻校正システムの整備

岡田 尚基^{*1} 山本 幸生^{*1}

Development of Spacecraft Time Calibration System for Science Spacecrafts

Naoki OKADA^{*1}, Yukio YAMAMOTO^{*1}

Abstract

Science spacecrafts have a spacecraft clock which indicates the time of command execution and telemetry data generation. In many cases, it is a counter which is incremented using onboard clock. Because it is a spacecraft specific time scale and its rate of increase is not constant, we have to calibrate it to a common time scale. We develop a time calibration framework for our future science spacecrafts in ISAS/JAXA. We will incorporate this common framework into our ground systems from spacecraft operations to telemetry data analyses.

Keywords: time calibration, spacecraft clock, system development

概要

科学衛星や探査機では、コマンドの実行時刻やテレメトリデータの生成時刻を扱うための時刻系として、衛星時刻が使われる。衛星時刻は、オンボードのクロックを使って増加するカウンタであることが多い。衛星時刻はその衛星・探査機に固有の時刻系であり、かつ進み方も一定ではない。そのため、これを共通の時刻系に校正して使うことが行われる。我々は、宇宙航空研究開発機構 宇宙科学研究所で科学衛星・探査機の時刻校正に適用する枠組みを整備している。この枠組みは、今後打ち上げられる科学衛星・探査機で、衛星運用からデータ解析に至るまでの各地上系システムに共通機能として組み込まれる予定である。

1 はじめに

科学衛星・探査機（以下、衛星）では、衛星時刻によってコマンドの実行時刻やテレメトリデータの生成時刻を管理している。宇宙航空研究開発機構 宇宙科学研究所では衛星時刻のことを Time Indicator(TI) と呼ぶ。TI は多くの衛星で、内部のクロックを使ってカウントアップする一つの固定長カウンタとなっている。TI は衛星内のデータ処理装置で作られて、各機器に配信される。TI の増加する速度（レート）は一定ではなく、温度等の要因により変化する。そこで、定期的に TI と標準的な時刻系を対応付けることで時刻を校正し、かつ絶対時刻で扱えるようにすることが必要である。

この時刻校正のために、衛星は時刻テレメトリというテレメトリデータを送るようになっている。時刻テレメトリには、あるフレーム（TI ラッチフレーム）が衛星内のデータ処理装置から送信機に出力された瞬間の TI 値（時刻テレメトリ TI）がラッチされ、そのフレームを特定するための情報と共に格納されている（図 1）。一方、地上の受信局で受信した

^{*1} 宇宙航空研究開発機構 宇宙科学研究所 科学衛星運用・データ利用センター 科学データ利用促進グループ
(Space Science Data Archive Promotion Group, Center for Science Satellite Operation and Data Archive, Institute of Space and Astronautical Science(ISAS), Japan Aerospace Exploration Agency(JAXA))

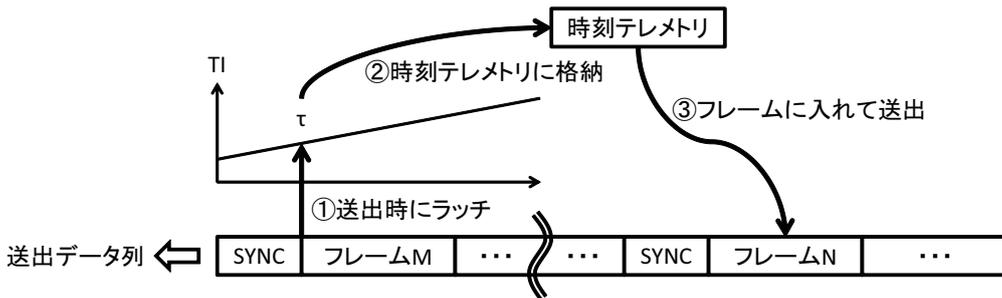


図1 時刻テレメトリの仕組み

テレメトリデータにはフレームごとにその受信時刻が協定世界時 (UTC: Coordinated Universal Time) で付けられるため、時刻テレメトリを使えば衛星送出時の TI と地上受信時の UTC の対応付けができる。後は、衛星送出時のラッチから地上受信時刻を付けるに至るまでの各種遅延を補正すれば、校正用の TI-UTC のペアを作ることができる。こうして作成した TI-UTC ペアを校正表にして用意しておけば、補間により任意の時点での TI と UTC の相互変換を行える。これが、時刻テレメトリを使用した衛星時刻校正の仕組みである。

TI のビット長や分解能は衛星により個別に定義されるが、この仕組みは共通である。従って、時刻について特別に高い精度要求がなければ、衛星ごとに時刻校正処理を用意する必要はない。また、データ解析をする上で、個別に異なる時刻校正処理が作られることにより、一つのデータに異なる時刻が付くことは望ましくない。宇宙科学研究所では従来、テレメトリデータの収集・配布を行う各システムで時刻校正を行ってきたが、システムごとに個別で実装されていたため、処理方式が異なっていた。また、その処理内容についても十分に明らかになっていない箇所があり、利用者がどの程度信頼して使用して良いものなのか判断できなかった。そこで現在、宇宙科学研究所で使う衛星時刻校正の共通フレームワークとして衛星時刻校正システムの整備を行っている。

2章ではまず、衛星時刻校正システムの全体像について述べる。3章で時刻校正処理について詳しく述べ、4章で時刻精度に関する考察結果についてまとめる。最後に、5章でまとめとこれからの整備方針について記す。

2 システム概要

衛星時刻校正システムの全体像を図2に示す。システムは、時刻校正表作成処理とパケット時刻校正処理からなる。

時刻校正表作成処理では、時刻テレメトリと TI ラッチフレーム、そして衛星-地上局間電波伝搬遅延補正のためにアンテナ予報値を入力し、衛星ごとの時刻校正表を出力する。この処理は衛星運用中に実行され、TI-UTC ペアを1レコードずつ追記していく。パケット時刻校正処理では、時刻校正表とうるう秒の挿入時刻が書かれたうるう秒ファイルを使い、各パケットの Time Code Field に格納された TI (パケット TI) を UTC に校正する。パケット時刻校正処理を行うためのプログラムはライブラリとして提供しているため、利用者は自身の環境にこのライブラリをインストールし、時刻

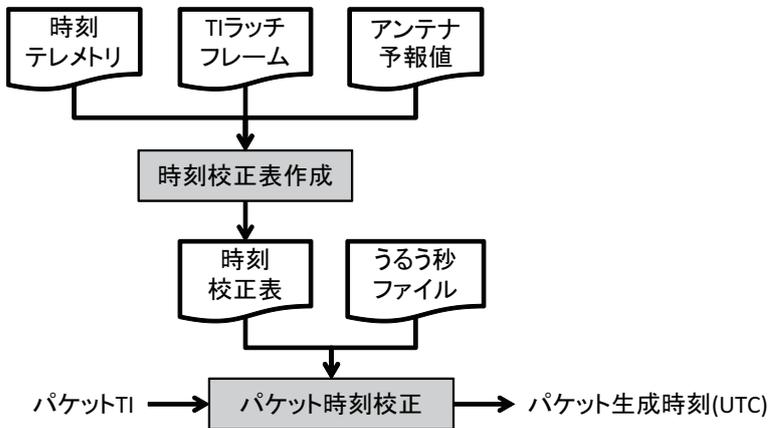


図2 システム概要

校正を行うことができる。衛星時刻を扱う上では、カウンタのロールオーバーやリセットに注意が必要である。衛星時刻は現在、約4.25年でロールオーバーする設計となっているものが多く、そのほとんどがこの期間を超えて運用している。また、コマンドやデータ処理装置の再立ち上げでリセットすることもある。

すなわち、衛星のライフスパンの中で、繰り返し同じ TI 値が現れることになり、このような場合 TI 値だけを見ても一意に UTC への変換が行えない。これを解決するため、本システムでは TI に上位桁を追加した Extended TI (ETI) を考える。カウンタのロールオーバーやリセットにより TI が不連続になったときに、追加した上位桁の部分 (ETI 拡張部) をインクリメントすることで (図 3)、ETI は衛星のライフスパンを通じて一意な値をもった衛星時刻として扱うことができる。

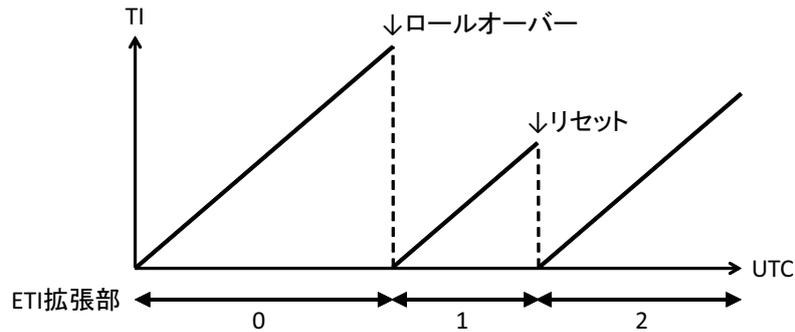


図 3 ETI 拡張部のインクリメント

衛星のライフスパンには、地上試験期間も含まれる。地上試験で出力されるデータは、不具合調査やデータの校正に役立つために保管しておくからである。地上試験データまで含めて一意の衛星時刻が割り当てられるよう、ETI 拡張部の最上位ビットはフライトデータと地上試験データを区別するフラグとして使用する。

なお、本システムを適用するためには、衛星時刻の管理について以下の条件を満たしている必要がある。

1. テレメトリ形式が Consultative Committee for Space Data Systems (CCSDS) の定める Space Packet[1] に準拠していること
2. 衛星時刻校正用に時刻テレメトリを用いていること
3. 各パケットの Time Code Field 及び時刻テレメトリの衛星時刻が、ある時点のエポックとする非分割のカウンタ (CCSDS Unsegmented Time Code (CUC) Level 2[2] 相当) であること

3 時刻校正の処理

3.1 時刻校正表の作成

3.1.1 時刻テレメトリの収集

時刻校正表を作るために、まずは時刻テレメトリを収集する。ただし、本システムでは衛星からの時刻テレメトリの送出間隔にかかわらず、基本的にはパスの先頭と末尾の二つの時刻テレメトリを収集する。また、深宇宙探査機ではパスが長時間に及ぶため、パスが1時間以上継続する場合には1時間ごとにも時刻テレメトリを収集する。これは、校正表のレコード数と時刻精度のバランスを考慮した標準値であり、時刻精度要求に従って衛星ごとに収集間隔は変更できる。こうして収集した時刻テレメトリから TI ラッチフレームを探すことにより衛星送出時 TI と地上受信時 UTC の対応が得られる。

3.1.2 計測タイミングの補正

宇宙科学研究所の衛星・地上システムでは、フレームの先頭で時刻を計測するか、フレームに付けられている同期コードの先頭で時刻を計測するかという点で、衛星送出時の TI ラッチタイミングと地上受信時の受信時刻計測タイミングが異なっている（図 4）。同期コードの長さ、及び通信のビットレートによってこのタイミングのずれは求められるので、これらの情報を使い計測タイミングのずれを補正する。

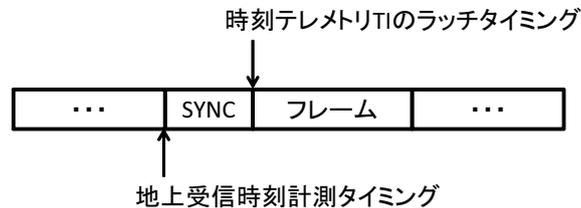


図 4 時刻計測タイミングの違い

3.1.3 電波伝搬遅延補正

衛星送出から地上受信までの遅延の主たる要因は衛星－地上局間の電波伝搬遅延である。この遅延量を求めるには、衛星－地上局間の距離（ダウンリンクレンジ）を知る必要がある。ダウンリンクレンジとしては実測値が最も正確と考えられるが、実際には必要な時刻の実測値がないことが多い。実測値から求めた軌道の確定値は、利用可能になるまでに時間がかかるため、衛星運用中にリアルタイムで使用されることをユースケースに含んでいる本システムでは間に合わない。そこで、本システムでは衛星運用時に確実に取得できるアンテナ予報値を用いる。アンテナ予報値には軌道予測を基に作られた、ある時刻での電波のダウンリンクレンジの予測値が、一定間隔で記載されている。TI ラッチフレームの地上受信時刻におけるダウンリンクレンジをこのアンテナ予報値から 2 次補間で求め、真空中の光速で除して電波伝搬遅延時間とする。これを受信時刻から差し引いて、TI-UTC ペアを作成する。アンテナ予報値の使用については、通常時は後述する時刻精度が得られるため問題ないと考えられるが、軌道が大きく変わる運用を行う場合には注意が必要である。

3.1.4 時刻校正表への記載事項

時刻校正表には以上で求めた TI-UTC ペアを記録として追記していく。このとき、各記録には直前の記録との間で計算した TI の増加速度（レート）、及び ETI 拡張部、それに時刻校正に関する補足情報も含める。補足情報には上述の補正に用いたビットレートやダウンリンクレンジの他、本システムでは補正に使用しないが、地上受信局名と通信バンド帯も含めている。これらの補足情報は、後で各記録の精度について評価するための判断材料として使えるほか、衛星プロジェクト側でさらに精度の高い時刻校正を行う際に有用と考えられる。

3.2 パケット時刻の校正

時刻校正表を用いて、各パケットのパケット TI を UTC に校正する。これは校正対象の TI の前後の TI-UTC ペアからの 1 次補間により計算される。時刻校正表の最終記録以降の TI が校正対象の場合は、末尾の 2 レコードからの 1 次補間となる。

この変換を行うプログラムはライブラリとして提供している。ライブラリには特定のエポックからの経過秒を介して、衛星時刻と UTC を相互変換する関数が用意されている（図 5）。TI にはロールオーバーやリセットによる値の重複があり得るため、衛星時刻としてはパケット TI だけだと情報不足である。そのためにライブラリではパケット TI に加え、

方法 1 パケット TI 付近と考えられる UTC（予想周辺時刻）を指定する方法

方法 2 ETI 拡張部を指定する方法

の 2 種類の方法で変換できるようになっている。各パケットの ETI は、テレメトリデータのアーカイブを行うシステム

である SIRIUS で作成する。SIRIUS では、テレメトリデータの受信時刻を参考にして方法 1 を使った時刻校正を行い、ETI 拡張部を作成する。これにより、SIRIUS からパケットを取得した場合にはパケットの付加情報としてこの ETI が得られるため、方法 2 を使った時刻校正ができる。なお、ライブラリはうるう秒ファイルを読み込むことでうるう秒の挿入に対応しており、うるう秒を 60 秒として出力する。

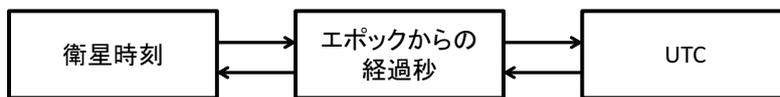


図 5 ライブラリ関数による時刻変換イメージ

4 時刻校正の精度

本システムを使用した場合の時刻校正の精度は、TI-UTC ペアの精度とそれを用いた 1 次補間の精度により決まる。

TI-UTC ペアを作成する際に考慮しなければならないのは、TI 及び UTC の時刻計測精度と、各時刻計測間の回線・処理遅延量である（図 6）。これらについては、衛星・地上局による違いはあるものの概ね表 1 に示したオーダーの誤差を有する。この表から、TI-UTC ペアの時刻精度は 1msec オーダーと言える。

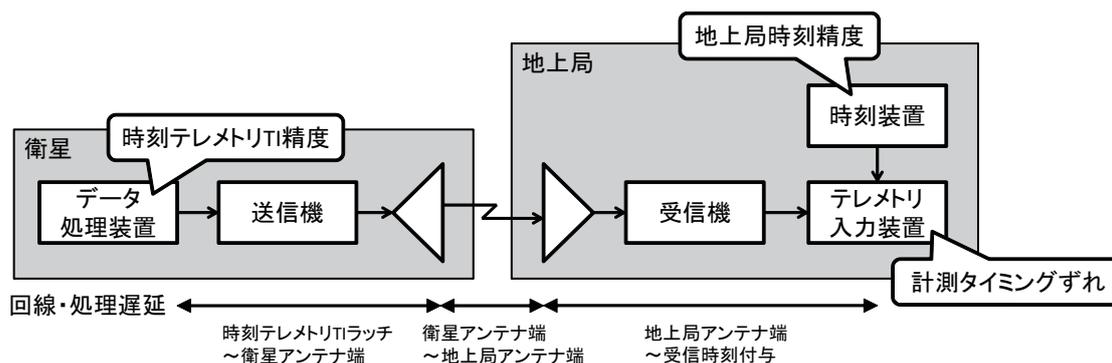


図 6 TI ラッチから受信時刻付与までの流れと時刻精度への影響箇所

表 1 TI-UTC ペア作成における時刻誤差オーダー

| 項目 | 誤差オーダー ^{*1} |
|--------------------------------|----------------------|
| 時刻テレメトリ TI 精度 | 1msec |
| 地上局時刻精度 | 1msec |
| 計測タイミングずれ | 0 (1sec) |
| 時刻テレメトリ TI ラッチ～衛星アンテナ端 回線・処理遅延 | 1msec |
| 衛星アンテナ端～地上局アンテナ端 電波伝搬遅延 | 1msec (1min) |
| 地上局アンテナ端～受信時刻付与 回線・処理遅延 | 0 ^{*2} |

^{*1} 括弧内は本システムによる補正前の誤差オーダー

^{*2} 上流のテレメトリ入力装置にて補正済み

1 次補間による誤差は、TI-UTC ペアの作成間隔と、その間のレートの変化量によって決まる。レートは主にデータ処理装置の温度に依存するため、衛星の軌道や運用方法により変化の仕方が異なる。いくつかの地球周回衛星で定常運用時のレートの最小値・最大値を調査したところ、どれも 1 時間当たり 10msec オーダーの違いを生む差があった。定常的

に運用される温度範囲内で、水晶発振器を用いて得られる時刻精度は、他の衛星でも同様の結果になることが予想される。

以上より、本システムの時刻校正精度は、TI-UTC ペアの作成間隔を1時間とすると概ね10msecのオーダーであると言える。多くの衛星でパケットTIの分解能が10msecオーダーであることから、この精度は利用者の要求を満たすものと考えられる。ただし、さらに高い時刻精度が必要な場合には、本システムを使う際に時刻テレメトリの収集間隔を短く設定するか、衛星プロジェクト側で高精度な時刻校正を行う仕組みを用意する必要がある。

5 おわりに

本稿では今後、宇宙航空研究開発機構宇宙科学研究所で打ち上げられる衛星に対して適用する、時刻校正のフレームワークについて紹介した。衛星の時刻管理方式について特定の条件を満たせば、その衛星で時刻校正システムを使用することができ、衛星運用からデータ解析に至る宇宙科学研究所の共通的な地上系システムを利用することができる。

一方で、ASTRO-Hや小型科学衛星では、従来とは異なる衛星時刻管理方式が採用されることが決まっている。これらの衛星では、時刻テレメトリを使った時刻校正が行える点で仕組みは従来方式と同じだが、衛星時刻についてGPS信号の利用を想定した設計となっているために、時刻テレメトリTIのフォーマットが異なるなどの差異があり、本システムをそのまま適用することができない。今後はこの方式を採用した衛星が複数打ち上げられることになるため、衛星共通の地上系システムを構築するために、本システムをASTRO-H/小型科学衛星の方式にも適用できるよう改修することが望まれる。

参考文献

- [1] CCSDS, Space Packet Protocol, Issue 1, September 2003.
- [2] CCSDS, Time Code Formats, Issue 3, January 2002.

衛星自動運用システムの開発

永松 弘行^{*1}, 齋藤 宏文^{*1}

Development of Automatic Satellite Operation System

Hiroyuki NAGAMATSU^{*1} and Hirobumi SAITO^{*1}

Abstract

We are in progress to develop a system for automatic operation of a satellite in order to reduce human load at satellite steady operation phase. The ground station for small satellite REIMEI is used as a test bench for verification of the proposed method. In our new automatic operation system, a scheduler software as a substitutive operator manages all the operations through a unified procedure, including sending command, receiving telemetry, and driving antenna in accordance with an operation time line which is prepared before the operation pass. The scheduler also performs diagnostics of satellite anomaly based upon the received telemetry data and status of the ground station. In case that some anomaly of the satellite is detected, the scheduler initiates an emergency schedule that was prepared depending on the emergency level. The automatic operation system is nearly completed for downlink operations of the data recorder that account for 80% of REIMEI steady operation. This approach is very effective to reduce psychological and physical load of operators.

概要

衛星の定常運用における運用者の負担軽減を目的とする自動運用システムを開発中である。「れいめい」衛星の運用システム(相模原局)を実験ケースとしてシステムを構築した実例を紹介する。提案手法の中心となるソフトウェアである「スケジューラ」が、オペレータに代わり、運用システムの機能ブロックを取り出したAPIを通じて運用システム全体を制御する。すでに「れいめい」の定常運用の約80%を占める観測データのダウンリンクの自動運用を実現、相模原局にて実験継続中である。

1. はじめに

衛星やロケットは高度に自動化されたシステムであるが、その運用においては、人間(オペレータ)が行うマニュアル操作が多くを占める。運用にはルーチンの作業がある一方で、システムの異常検知、それに対する適切な対応、復旧作業などの一連のタスクには、衛星および地上局についての高度な知識も求められる。衛星や地上局についての知識が豊富な人たちが必ずしも運用に従事しているわけではない。そのような場合は衛星および運用システムをブラックボックスとして扱うことになり、緊急時の対応には適切さを欠くこともありうる。地球周回衛星で夜中に運用パスがある場合など、オペレータの精神的・肉体的な負担は大きい。運用にかかるコストも無視できない問題である。一方で、今後ますます多くの衛星打ち上げプロジェクトが計画されており、それらに対応した運用体制を作り上げることも重要である。

衛星運用の現場にはさまざまな問題がある。すなわち、十分な運用要員が確保できるとは限らない、十分なスキルを

*1 宇宙航空研究開発機構 宇宙科学研究所
(Institute of Space and Astronautical Science, Japan Aerospace Exploration Agency)

持つ要員が確保できるとは限らない、運用時の精神的・肉体的負担が大きい、運用コストは削減傾向にある、などである。定常運用の負担軽減、効率化、コスト削減のための現実的な方策が求められると考える。運用システムを可能な限り自動化してオペレータの負担を軽減する方法がいくつか提案され、また、実施している例、機関もある [1][2][3]。

本稿では、衛星自動運用を実現するための一方法を提案する。本稿で紹介するシステムの最終目的は、「観測者(ユーザ)が運用計画を立て、それをサーバに登録すれば、以後の検証、運用、運用データのアップロード、運用ログ配信などの作業を自動的に行ってくれる」運用システムの実現である。その第一歩として、既存の設備をほぼそのまま使い、できる限りシンプルなシステムを用いて、コマンド送信、テレメトリ受信、衛星・地上局設備の監視・診断を自動的に行うシステムの実現を目指す。また、地球周回衛星における定常運用の自動化を主なターゲットとし、改修対象はソフトウェアのみとする。開発した自動運用システムは、ISAS/JAXAが開発した小型科学衛星「れいめい」の地上局(相模原局)をテストベッドとした運用試験に用いている。

本稿の構成は以下の通りである。第2章で、「れいめい」および相模原局の概要を述べる。第3章で、開発した衛星自動運用システムについて説明する。第4章で、自動運用の「れいめい」衛星・相模原局を用いた運用試験結果を紹介する。第5章で、自動運用システムの検証用に開発中の簡易衛星・地上局シミュレータについて紹介する。最後に第6章で、今後の課題を示し、本稿をまとめる。

2. 「れいめい」衛星と相模原局について

小型科学衛星「れいめい」(INDEX)はISAS/JAXAが開発した衛星である[4]。オーロラ観測と新規衛星技術の軌道上実証を目的とし、2005年8月24日、カザフスタンにあるバイコヌール基地からドニエプルロケットによって、光衛星間通信実験衛星「きらり」(OICETS)のピギーバッグ衛星として打ち上げられた。2012年6月30日現在、軌道上にて順調に稼働中である。図1に「れいめい」のフライトモデルを示す。

「れいめい」は専用地上局をISAS/JAXA相模原キャンパス内に保有しており、日常の運用はこの専用局(相模原局)を用いて行われる。相模原局は、簡易地上局による運用を開発当初からミッションの一部として捉え、ISAS/JAXA教職員、学生、およびベンチャーハード/ソフト開発者の協力で、インハウスの手法で開発を行った。また、運用計画コマンドの(半)自動生成、運用計画の事前検証の自動化など、運用を省力化するための様々な工夫が取り入れられている[5][6]。

相模原局の概略を図2に示す。運用計画コマンドを作成・検証する運用計画PC(Personal Computer)、軌道データサーバ、データ分配蓄積の役割を担うゲートウェイ、コマンド送信のためのコマンドPC、テレメトリ受信のためのテレメPC、テレメトリデータをコマンドPCやQL-PCに配信するサーバPC、オペレータが衛星の状態を監視するために用いるQL(Quick Look)-PC、アンテナ(3mφ、アップリンク/ダウンリンクともSバンド対応)のサーボ制御を行うアンテナ制御PCより構成される。また、衛星の動作を模擬する、搭載計算機のプロトタイプを使ったシミュレータ(試験運用システム)があり、これを用いたコマンド計画の事前検証も可能である。相模原局には、衛星運用に必要な設備一式が揃っており、これを用いた衛星運用システムを開発してゆけることが大きな特長である。ISASが運用する他の科学衛星の運用システムとは独立した環境でシステムを開発し、その成果は他の科学衛星への転用が期待できる。



図1 「れいめい」フライトモデル

3. 衛星自動運用システム

衛星自動運用の実現に必要な技術開発および実験検証[7]にあたり、既存設備を有効活用したシンプルなシステムによる実現を目指すという観点から、図2の相模原局をテストベッドとしている。相模原局では第2章で述べたように運用の省力化や自動化のための様々な工夫が取り入れられているが、それらをさらに一歩進める形で自動運用システムの実現を目指している。図3に自動運用システムの概略を示す。

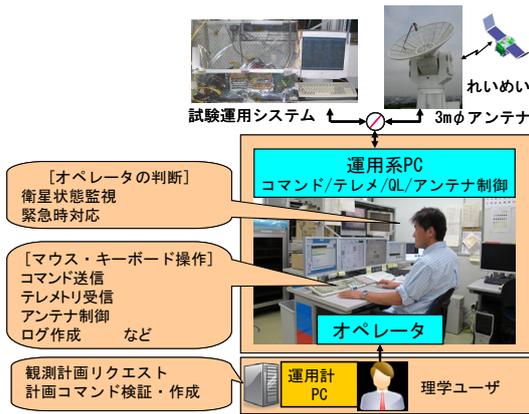


図2 相模原局の運用システム概略

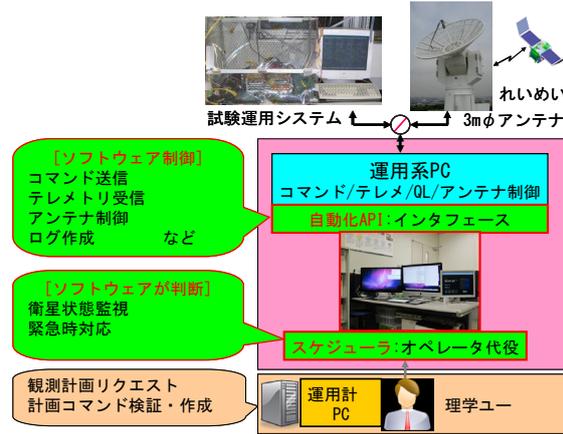


図3 自動運用システム概略

自動運用システムの開発方針は以下の通りである。1) 改修対象は運用システム「ソフトウェアのみ」とし、ハードウェアの改修、変更は行わない、2) 現用の運用システムソフトウェアに加える改修は最小限とし、現状の運用体制に支障をきたさないよう配慮する、3) 開発するシステムは現運用システム（特にコマンド・テレメトリ処理システム）とは独立性が高く、既存設備による定常運用を行いながら、自動運用システムの開発を可能とする。

自動運用システムの中心的役割を担うソフトウェア「スケジューラ」は、オペレータの代役として運用システム全体を管理・統括・制御する。スケジューラは、運用システム各PC（コマンドPC/テレメPC/アンテナ制御PC）と、運用システムの機能をブロック単位で取り出した自動化API(Application Program Interface)を介して通信（マウス・キーボード操作によるコマンド送信，データ受信をソフトウェアすなわちスケジューラが行う）する。図2と図3の違いを表1にまとめる。

表1 手動運用と自動運用の相違点

| 実施項目 | 手動運用 | 自動運用 |
|----------------------|--------------------------|-------------------------|
| コマンド送信 テレメトリ受信 など | キーボード操作， マウス操作による手動送信 | 自動化APIによる ソフトウェア制御 |
| 衛星状態監視 緊急時対応 など | 人間が判断し， 状況に応じて対応 | スケジューラが判断し， 状況に応じて対応 |

スケジューラは、あらかじめ準備された運用手順から生成されるスケジュールファイルにしたがって、自身の管理下にある運用システム各PCの動作を制御する。時刻情報については、入感/消感時刻、アップリンク(UPLINK)開始/終了時刻などをアンテナ予報値ファイルより抽出・整理し、ファイル(パス情報ファイル)にまとめる。可視時間帯の指定時刻に送信するコマンドがある場合には、運用計画から時刻情報を取得し、これもパス情報ファイルにまとめる。パス情報ファイルはスケジューラによる運用時間管理の基本情報を与える。スケジューラには運用システムのステータスやテレメトリ情報を読み込んで工学値変換する機能も実装し、状態監視による運用時間中の運用システムおよび搭載システムの監視・診断に用いる。この状態監視機能は、スケジューラの動作中はアクティブであり、例えば可視時間中にテレメトリデータのロックオフを検知して自動的に再ロック操作をすることで、データ欠損を防ぐ。また、可視時間中にチェック項目に異常を検知した場合、診断レベルに応じて、あらかじめ準備された「緊急スケジュール」を実行する。例えば衛星がセーフホールド状態で入感した場合、地上局のアンテナが停止した場合、受信機などが故障した場合などは、オペレータにメール通知し、オペレータに復旧運用を引き継ぐ。運用終了後は、運用データ(テレメトリデータ、コマンド送信履歴、地上局設備データなど)をデータサーバに登録するとともに、ユーザに運用ログとしてメール配信する。

図4にスケジューラによる自動運用シーケンスをまとめる。ただし、主だった機能の提示にとどめ、運用前後処理のいくつか(スケジュールファイル準備, アンテナ予報値取得, 運用データ登録, メール配信など)は省略した。

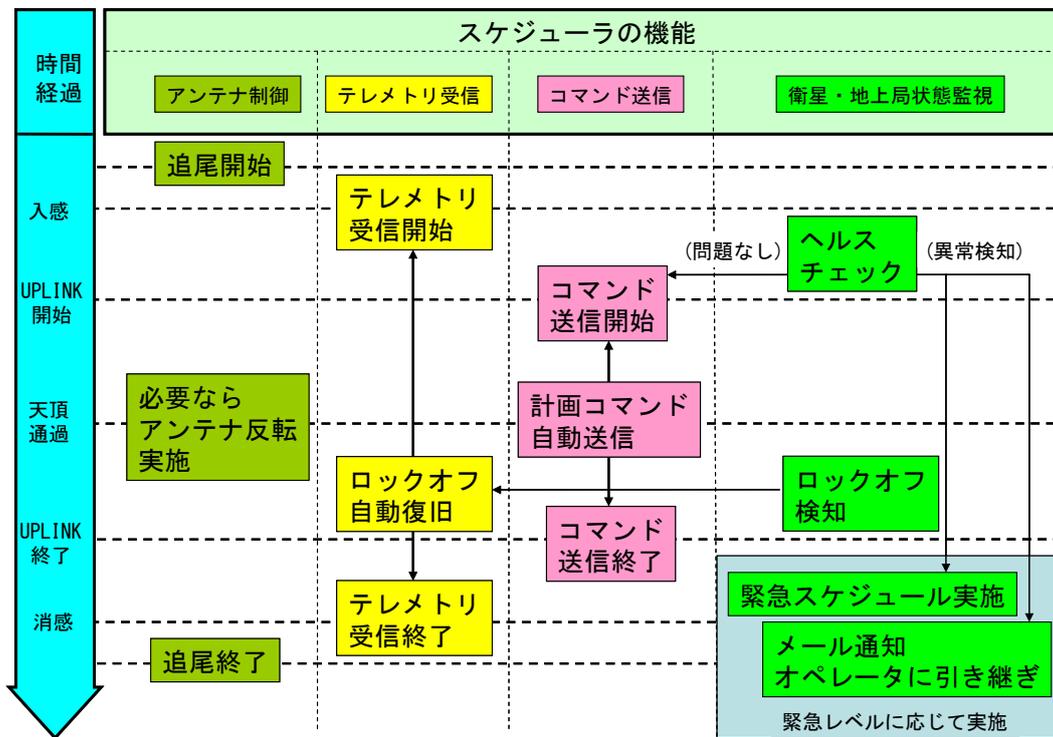


図4 自動運用システムによる自動運用シーケンス概略

4. 自動運用実験

自動運用システムの機能確認のため、「れいめい」実機を用いての自動運用実験を、2010年末より継続している(実験回数は2012年6月末現在、400回以上)。その結果を紹介する。「れいめい」の定常運用では、週単位では、月曜日・木曜日の昼可視時間帯で観測計画をアップロードし、火曜日・水曜日・金曜日の昼可視時間帯では、観測データのダウンリンクを行っている。前者の可視時間帯の一部、後者の可視時間帯を使って自動運用実験を継続している。

典型的な実験結果を、図5、図6に示す。図5、図6の縦軸は相模原局のアンテナ受信レベル([dBm])を、横軸は時刻([UT])を、それぞれ表す。図5はアンテナ最大仰角が約 40° 、図6はアンテナ最大仰角が約 89° である。図6の場合、仰角・方位角方式アンテナの駆動問題対策として、天頂通過時にアンテナ反転を行う。相模原局のアンテナはプログラム追尾方式であり、スケジューラはアンテナ予報値の登録、追尾モードへの切換を可視開始時間前に行う。スケジューラは可視開始時間になると衛星からの信号を自動的にロックオンし、データ受信を開始する。また、スケジューラは受信レベルが設定した閾値を上回ると判断したら、ビットレートを低速から高速に切り替え、テレメトリ受信モードも切り替え、観測データのダウンリンクを開始する。可視時間中、アンテナパターン、衛星の姿勢、アンテナ反転動作に起因するアンテナ受信レベルの変動がある場合でも(例えば図5の02:23:52付近、図6の15:41:23付近)、スケジューラが変動を予測し(受信レベルの絶対値[dBm]と絶対値の変動[dBm/sec]が、あらかじめ設定した閾値と比較して高いか低いかを判断する)、受信ビットレートや受信モードを切り替えることで、データ欠損を防いでいる。可視終了時刻になると、スケジューラは観測データの受信を中止し、運用終了の手続き(ビットレート、テレメトリ受信モードの変更)を行う。スケジューラは図4のシーケンスにしたがって人間のオペレータと類似の挙動を実現しており、この意味において実験結果は良好と言える。

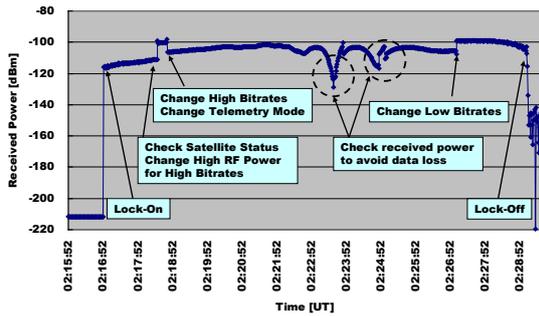


図 5 自動運用実験結果

(アンテナ最大仰角 :40° , アンテナ反転なし)

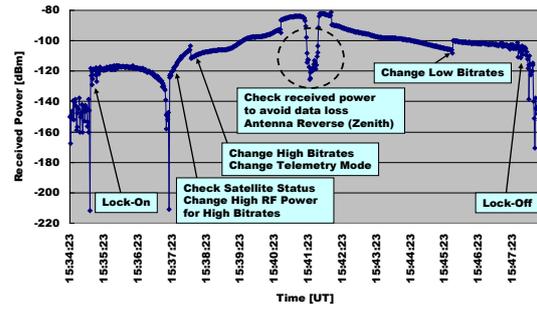


図 6 自動運用実験結果

(アンテナ最大仰角 :89° , アンテナ反転あり)

ここで、人間のオペレータの労力が自動運用システムによってどの程度軽減されるのか、簡単に考察してみる。人間のオペレータが運用する場合、図 5、図 6 中の四角で囲んだ操作を、すべてキーボード・マウス操作にて行う。コマンド送信は、コマンド選択、必要ならパラメータ入力、送信、結果確認、の 4 ステップで構成される。自動運用システムでは、この 4 ステップすべてをソフトウェアが制御しており、オペレータは「自動運用の様子を見ているだけ」である。また、オペレータは運用中に衛星・地上局の状態監視・緊急時対応を行うが、これは自動運用システムの衛星・地上局状態監視機能が代行し、例えば信号ロックオフ検知・修復は自動化される。軽度障害の場合は、あらかじめ準備してある(検証済みの)緊急スケジュールを自動送信する。地上局機器トラブル、セーフホールド時対応など、人間が介在せざるを得ない場合は、オペレータに通知・運用を引き継ぐ。2010 年末以来の 400 回以上の自動運用実験中、セーフホールド状態で衛星が入感し、人間が対応せざるを得なかったケースは一件だけであり(これ以外は、原則「見ているだけ」である)、自動運用による定常運用負担軽減の効果は大きいと考える。観測データのダウンリンクを例に、自動運用によってどの程度の労力が軽減されるか、キーボード・マウス操作数で手動運用と比較してみる。コマンドを一つ送信する場合、コマンド選択・送信の二回、パラメータ入力が必要な場合は、パラメーター一つにつきキーボード操作一回、とカウントした。結果を表 2 にまとめる。自動運用システムにより、オペレータが行うキーボード・マウス操作はすべてソフトウェア制御に置き換えられており、この点でも負荷軽減の効果は大きいと言える。

表 2 手動運用と自動運用のキーボード・マウス操作数比較

| コマンド入力 操作など | 手動運用 | | 自動運用 | |
|----------------|---------|---------|---------|-------|
| | キーボード操作 | マウス操作 | キーボード操作 | マウス操作 |
| 入感時 | 2 | 10 ~ 12 | 0 | 0 |
| UPLINK 中 | 2 ~ 4 | 6 ~ 30 | 0 | 0 |
| 消感時 | 2 | 10 ~ 12 | 0 | 0 |

今後も実績を重ねつつ、最終的には観測計画コマンドのアップリンクも含め、定常運用のほぼすべてのシーケンスを自動化し(セーフホールド状態からの復帰などを除く)、定常運用のさらなる負担軽減を目指したい。また、緊急時のリモート運用にも対応する予定である。

5. 簡易衛星・地上局シミュレータの開発について

現在の自動運用システムは、第 4 章で述べたように、実運用に適用するための一定の目処は立った。一方で、アンテナ受信レベルの変動に信号の欠損なく追従するためのアルゴリズム、衛星状態をモニタし、問題がある場合には適切なコマンド処理を施すためのアルゴリズム、などの自動化アルゴリズムの検証に限界がある、という課題が浮上してきた。その理由は、自動化アルゴリズムの検証に用いている相模原局・試験運用システムの出力信号が、必ずしも試験信号と

して期待するものにはなっておらず、試験パターンに限界があり、十分な検証ができていないと言いき難いからである。そこで、自動運用アルゴリズムの検証を容易にするため、衛星実機および地上局の挙動に近い(擬似)信号パターンを容易に生成できるシステム「簡易衛星・地上局シミュレータ」を開発中である(図7)[8]。擬似信号を発生できれば、自動運用アルゴリズムの検証がより確実になるし、自動運用システムでは現時点ですでに自動的に連続して複数の運用シナリオを走らせることができるので、試験もやりやすくなる。自動化アルゴリズム検証テストベッドとしての相模原局の有効活用も図れる。

この簡易シミュレータの一部機能は自動運用システムの改良に用いている。簡易シミュレータは、図7中にある「手のひらサイズ」の計算機に実装されている。これは安価だが十分な処理能力を持つ計算機で、一般に高価なシミュレータ導入の敷居を下げる効果(例えば開発メンバー一人に一台配布するなど)を期待している。また、簡易シミュレータはXML形式で記述される衛星・地上局データベースを書き換えることで他衛星・地上局への適用可能な汎用性を確保しており、他衛星・地上局への適用も検討中である。簡易シミュレータの詳細については、別の機会に発表させていただく予定である。

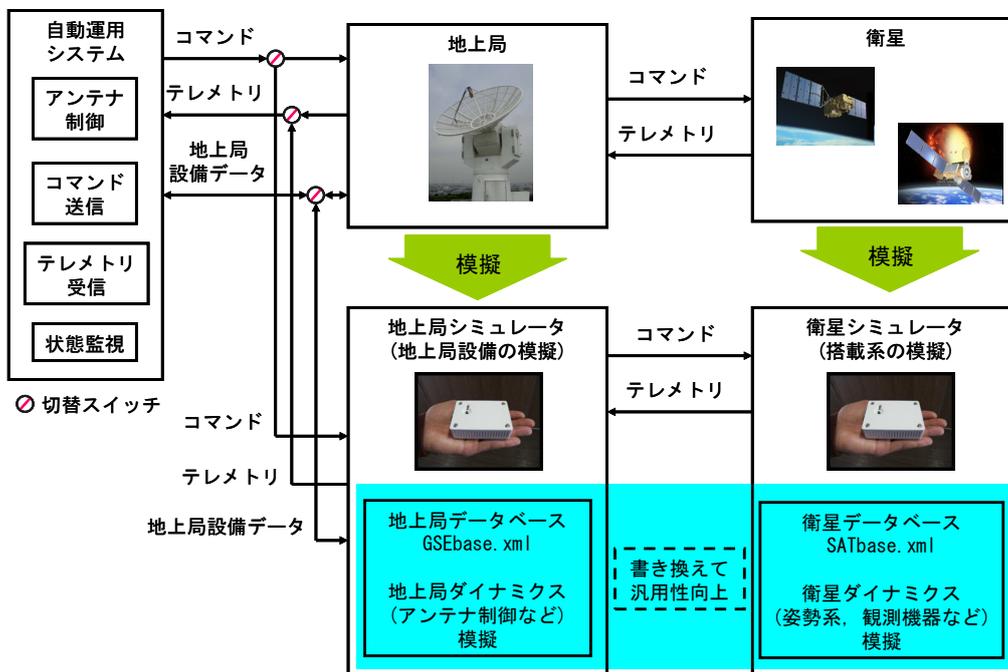


図7 簡易衛星・地上局シミュレータの概略

6. おわりに

本稿では、定常運用の負担軽減を目的として開発中の衛星自動運用システムの概略を紹介した。また、本システムを相模原局に実装し、「れいめい」衛星実機を用いての試験運用について述べた。今後は、観測計画コマンドの登録を含む定常運用の全自動化を目指す。緊急時に対応可能な、携帯端末を用いたリモート運用機能の検討も進める。他局・他衛星に転用可能なノウハウも蓄積したい。

謝辞

日頃からご指導ご鞭撻頂いております、「れいめい」開発メンバ、運用メンバの皆様、自動運用システムのソフトウェア開発でお世話になっております秋丸忠隆氏に、この場を借りてお礼申し上げます。

参考文献

- [1] Jean-Marc Soula: “Automation of CNES Ground Station Networks”, 8th International Workshop on RCGSO, Tsukuba, Japan, 2009.
- [2] D.Chung et.al.: “Concept Design of the Generic Mission Operation System”, 8th International Workshop on RCGSO, Tsukuba, Japan, 2009.
- [3] J.Jackson et.al.: “NASA’s Ground Network’s Right Sizing and Transition to Outsourcing” , 8th International Workshop on RCGSO, Tsukuba, Japan, 2009.
- [4] H. Saito et.al. : “An overview and initial in-orbit status of “INDEX”satellite”, 56th International Astronautical Congress, Fukuoka, Japan, Oct. 2005, IAC-05-B5.6.B.05.
- [5] S. Fukuda et.al.:“Flexible operation system for the microsatellite ‘REIMEI’ (INDEX)”, SpaceOps 2006, Rome, Italy, AIAA-2006-5895, June 2006.
- [6] 高原他 : れいめい衛星の統合化と効率的運用について, 第 50 回宇科連, pp.542-546, 2006.
- [7] H. Nagamatsu and H. Saito : “Development of Automatic Operation System-Using REIMEI Ground Station as a Test Bench”, 62nd International Astronautical Congress, Cape Town, South Africa, Oct. 2011, IAC-11-B4.3.4, 2011.
- [8] 永松 : 簡易衛星シミュレータの開発, SANE, SANE2011-163, 三重大学, 2012.

