

すばる望遠鏡公開データアーカイブシステムの開発 4

出田誠, 榎基宏, 小澤友彦 *1, 吉野彰, 仲田史明, 奥村真一郎 *2, 山本直孝 *3, 古荘玲子 *4,
矢治健太郎, 山田善彦, 八木雅文, 洞口俊博 *5, 高田唯史, 市川伸一

(2005年4月28日受理)

Development of Public Science Archive System of Subaru Telescope 4

Makoto IDETA, Motohiro ENOKI, Tomohiko OZAWA *1, Akira YOSHINO, Fumiaki NAKATA,
Shin-ichiro OKUMURA *2, Naotaka YAMAMOTO *3, Reiko FURUSHO *4, Kentaro YAJI, Yoshihiko YAMADA,
Masafumi YAGI, Toshihiro HORAGUCHI *5, Tadafumi TAKATA, Shin-ichi ICHIKAWA

Abstract

We have improved the public science archive system, SMOKA (Subaru-Mitaka-Okayama-Kiso Archive system). In this paper, our focus is on Suprime-Cam data, which is most frequently used in SMOKA. We have developed and implemented two features to search Suprime-Cam data; one is a pin-point search, and another one is a function to use seeing and a limiting magnitude of each frame as search constraints. Accordingly, new interfaces to search Suprime-Cam data and to show corresponding search results were developed. In addition, new mosaicked image of each shot taken with Suprime-Cam is released. We have also developed interfaces to use all FITS keywords including comments as search constraints.

1. はじめに

天文学研究において観測は主要な柱であり、観測データはその天域・観測時間における唯一無二の貴重な記録である。そこで、それを散逸しないよう保存管理し、後の研究・教育目的の利用に供するのがデータアーカイブシステムの最大の存在意義である。天文データアーカイブは利用価値が高く、将来の天文学の発展に重要な役割を担うであろうことは既に過去の論文^{1), 2), 3)} (以下論文1, 論文2, 論文3)において述べてきたが、その重要性はアーカイブデータの蓄積・多様化とともに日に日に増している。そのような中我々は、国立天文台のすばる望遠鏡、岡山天体物理観測所188cm望遠鏡、ならびに東京大学木曾観測所105cmシュミット望遠鏡の観測データを保管・提

供するアーカイブシステム SMOKA (Subaru-Mitaka-Okayama-Kiso Archive system)^{*1(次ページ)}の開発・運用を行なってきた(論文1, 論文2, 論文3)。SMOKAでは、これら観測装置に共通の検索ユーザインタフェースを提供し、利用者はそれを通して観測データを検索・取得できる。

アーカイブデータを利用した天文学研究・教育をより効率的に進めるには、アーカイブシステムに観測データをまとめて保管しておくだけでは不十分である。さらに、利用者が目的に沿った観測データを容易に取得できるよう様々な検索条件での検索を可能としたり、データの品質を利用者が判断し使用の可否を決定できるようにしたり、取得した観測データを整約解析するのに必要な情報を整備・提供することも必要である。我々はこの

*1 みさと天文台 Misato Observatory

*2 宇宙航空研究開発機構 Japan Aerospace Exploration Agency

*3 産業技術総合研究所 National Institute of Advanced Industrial Science and Technology

*4 早稲田大学 Waseda University

*5 国立科学博物館 National Science Museum

ような目的のため、2001年6月のSMOKA運用開始以降も、論文1、論文2、論文3で述べられているような開発課題を解決してきた。その中でも特に重要な課題は論文3で実装した節約済みデータの提供であった。現在はすばる望遠鏡の主焦点カメラ Suprime-Cam⁴⁾の観測データの一部のみではあるが、位置較正・フラックス較正・節約処理を施した観測データを利用者が取得できるようになった。

SMOKAの公開から4年あまりが経ち、日本でもデータアーカイブを用いた天文学研究が浸透してきている。SMOKAは多くの利用者に活用され、それを利用した研究論文や学会発表も増加しており、SMOKAアーカイブデータを主として利用した学位論文⁵⁾も現れた。海外からの利用も増加の一途をたどり、世界的によく知られたシステムになりつつある。また、SMOKAのアーカイブデータを活用した教育活動も進んでおり、最近では、PAOFITSワーキンググループ(WG)²⁾によって高校生向け教材が作成され、実際の教育現場で活用されている。SMOKAはこのように様々な形で利用されるようになったが、天文学研究・教育をより推進していく上では、まだ開発・改良すべき余地が多くある。その大きな課題の一つとして、検索機能の強化が挙げられる。論文2において検索機能は一度は強化された。しかし、その後も膨大な観測データがアーカイブされ続けた結果、絞りこみ検索を行なっても観測データの候補が数多く検索にかかってしまい、その中から利用者が本当に必要なものを選び抜くのが困難になりつつある。また、利用者数の増加に伴い多様化するデータ検索への要求に応えられるよう、システムを強化する必要性も生じている。

特に、すばる望遠鏡のSuprime-Camデータは他観測装置に比べて利用頻度が高く、Suprime-Camデータに対する高度な検索機能を実現する必要がある。具体的には、各観測データのCCDチップ毎の視野に基づいた効率的な位置検索ができるピンポイント検索機能や、seeing値や限界等級といった、データ品質を検索条件にできる機能などである。またSuprime-Camに限らない検索機能強化として、これまでのSMOKAの検索では実現できなかった、FITSヘッダの全キーワードを用いた検索やコメント部まで含めた全文検索機能も求め

られている。また、従来はハワイ観測所の観測装置にしか対応していなかったカレンダー検索機能(旧称:プロポーザルカレンダー表示機能)の岡山天体物理観測所と木曾観測所の観測装置への対応も必要である。さらに検索以外の機能として、Suprime-Camのサムネイル画像は現在でも参照できるが、さらにバックグラウンドレベル補正の上、モザイク化した画像も観測データの中身を確認する手段として求められている。

既存のSuprime-Camの節約済みデータ提供機能は、ある一つのフレームに対し唯一の節約済みデータしか提供できない仕組みであった。しかし、全ての目的に対し万能の節約処理法がない以上、利用目的に応じた幾つかの節約済みデータを用意し、利用者にデータ取得の際選択させる仕組みが要求される。この複数パターンの節約済みデータの提供により、従来の単一パターンでの節約処理では実現できなかった高い精度の節約済みデータを生産することが可能になる。また、節約処理の自由度が増すことで、節約処理可能な観測データが一気に増えるという利点もある。

その他にも、SMOKAのより効率的な利用を促進すべく、世界の主なデータセンターへの情報提供やすばる望遠鏡の中間赤外観測装置であるCOMICS⁶⁾の簡易処理済みデータの公開、移動天体検索で用いる太陽系内天体の軌道要素テーブルの自動更新など、いくつかの既存課題にも早期に取り組む必要がある。

また、最近のSMOKA利用者の増加、および要求機能の多様化は、同時にシステム自身の複雑さや運用の煩雑さをも増加させた。このような状況下、限られた人数の運用者で対応するには運用の効率化が必要である。このためには、システム的设计、および運用に関する文書の作成や管理を今まで以上に強化したり、可能な限りの処理の自動化を実現することで、運用の安定性や継続性を確保しなければならない。我々はこの課題についても取り組む必要がある。

本論文では、以上の課題を解決すべくSMOKAに追加・補強された機能について報告する。2章では、本論文の主要な開発箇所である検索機能強化、特にデータ請求量の多いSuprime-Camに特化

*1 <http://smoka.nao.ac.jp/>

*2 <http://paofits.dc.nao.ac.jp/>

した高度検索機能について解説する。3章では検索以外の機能強化を、4章では煩雑化したSMOKAの運用面での効率化について述べる。SMOKAの利用状況と運用課題、また、教育利用の実態などについては5章、今後の技術的課題については6章で述べ、7章でまとめを行なう。

2. 検索機能の強化

本章では、SMOKAに新たに追加された検索機能について述べる。まず、特にデータ請求量の多いSuprime-Camデータのための高度な検索機能として、ピンポイント検索と品質指標で検索できる機能について2.1節で詳しく述べる。また、多くの情報を内包するFITSヘッダの全情報を検索可能にした全キーワード検索、および全文検索について2.2節で解説する。さらに、従来はすばる望遠鏡の観測データにしか対応していなかったカレンダー検索機能（旧称：プロポーザルカレンダー表示機能）を岡山天体物理観測所・木曾観測所のデータにも対応させたので、2.3節で述べる。

2.1 Suprime-Cam専用高度検索機能

すばる望遠鏡の共同利用観測装置の一つであるSuprime-Camは、8~10mクラスの光学赤外線望遠鏡の観測装置としては最も広い視野を誇る。この装置を使って行なわれているサーベイ観測は、アーカイブにおいても利用価値の高い観測データを供給している。このような点もあり、SMOKAでデータ請求されるフレーム数の内、Suprime-Camの占める割合は、他観測装置に比べ非常に高い（2004年度期では請求フレーム数の9割がSuprime-Camデータ）。しかし、Suprime-Camのデータだけでもそのフレーム総数は膨大で、利用者が本当に必要なものを選び抜くのは従来の検索機能だけでは困難になりつつある。そこで、Suprime-Camに重点を置いたより高度な検索機能として、ピンポイント検索機能とデータ品質指標による検索機能を開発した。各々について、以下で詳しく解説する。

(1) ピンポイント検索機能

Suprime-Camは10個のCCDチップをモザイク化して焦点面に並べた観測装置であるが、得られた観測データのFITSヘッダに記載されている座標値（FITSキーワードのRA, DECもしくはRA2000, DEC2000）は、全10チップに共通の赤経・赤緯の値（おおよそ全10チップの中心）でしかない。従来のSMOKAでは、赤経・赤緯を検索条件とした場合、FITSヘッダの値を基に検索

を行なうため、Suprime-Camデータについては必ず10フレームセットで検索にかかってしまう。指定した座標値を実際に含むフレームの少なくとも10倍のフレームが検索にかかり、無駄が多く、SMOKAサーバーへの負担にもなる。特に大規模サーベイ観測が行なわれた天域においては、膨大な数のフレームが検索にかかり、目的とする観測データに到達するのは困難であった。そこで我々は、この問題を解消し、より効率の良い検索が行なえるよう、ピンポイント検索機能を開発した。この機能は、データベース化した各CCDチップ毎の視野の位置情報を基に、ある座標値を含む観測フレームを選び出すものである。

すばる望遠鏡の典型的seeingサイズは0.5~1秒角であることから、本検索機能の位置精度として1秒角を開発目標とした。しかし、元々のFITSヘッダに記されている位置情報は精度が悪くなく、また光学系に起因する像面湾曲もあって、真値と数十秒角程度のずれが生じることがある。そのためSMOKAでは、論文3で述べたとおり、Suprime-Camの観測データについてはUSNO-B1.0カタログ⁷⁾とのマッチングにより位置較正を行なっている。その結果、精度良く較正されたフレームでは0.2~0.3秒角程度の位置精度を達成している。そこでピンポイント検索機能では、このような精度良く位置較正されたフレームのみ検索対象とした。また、2000年11月以前のSuprime-CamデータにはWCS関連のFITSキーワードが欠落しているため、検索の対象外とした。2000年12月以降の観測データに限ると、2005年6月末現在、位置較正はSuprime-Camデータのフレーム総数236,548の内166,774フレームについて行なわれている。

位置情報をデータベース化する方法としてはSDSS (Sloan Digital Sky Survey)⁸⁾などで用いられているHTM (Hierarchical Triangular Mesh)⁹⁾というアルゴリズムがある。このアルゴリズムは、天球をまず8個の象限毎の球面三角に分割し、その後、各辺の中点を結ぶ大円で再帰的に4つの球面三角に必要精度まで分割していき、その三角形で座標を表現する方法である。最も単純なHTMの実装としては、1秒角の位置精度が達せられる階層（18階層）で各ピクセルの赤道座標値をHTMを用いてインデックス化し、4つの子三角形がすべて視野内にあれば親三角形のインデックスに置き換えていくという方法である。しかし

この方法では、1フレーム当たりおよそ800万ピクセル分のHTMインデックスを作成する必要があるが、現状のSMOKAサーバーでは30分という運用上非現実的な時間を要する。これを解決する方法として、一つはより効率のよいHTMの実装を行なうことが挙げられる。前述のように最も深い階層のインデックスからボトムアップ的に作成するのではなく、浅い階層から三角形が部分的に視野に含まれている場合はより深い階層に分割し、三角形全体が視野内に入ったなら分割をやめることで、トップダウン的に作ることも考えられる。しかし、これはアルゴリズムとしては少々複雑になる。もう一つはHTMに頼らず、より実装が単純なアルゴリズムを採用する方法が考えられる。本論文では開発スピードを最優先し、後者の方法を選択した。実際に用いたアルゴリズムを以下で詳しく解説する。

まず、1秒角の精度を実現するため、全天域を天の北極から天の南極まで1/5000度=0.72秒角毎に赤緯一定の小円（赤道のみ大円）にスライスする。すると、各CCDチップの視野はこれらの小円のいずれかと必ず交わる。そこで、その交点の座標値で視野の位置情報のデータベース化を行なった。さて、特定の小円に対し交点がある場合、その交点の数は通常2個である。その交点の座標を図1のように (RA1, DEC), (RA2, DEC) とした時、
DEC, RA1, RA2
のようなデータを記録する。ここで、 $RA1 \leq RA2$ とする。視野が0hを跨ぐ場合、例えばRA1=0.1, RA2=359.9の場合、
DEC, 0.0, 0.1
DEC, 359.9, 360.0
と2行にまたがり記録する。また、光学系に起因

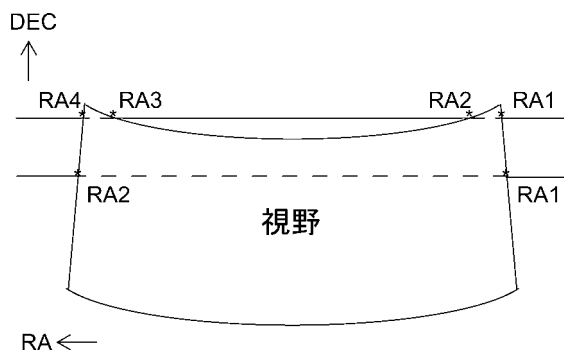


図1. 視野情報の例. 赤緯一定の小円とCCDチップ1枚の視野の交点が2点の場合RA1, RA2を、4点の場合RA1, RA2, RA3, RA4を視野情報とする。

する視野の湾曲のため、まれに交点が4個になる場合がある。この場合、残る2つの交点の座標を (RA3, DEC), (RA4, DEC) とし、
DEC, RA1, RA2
DEC, RA3, RA4
のように2行にまたがり記録する。なお、先と同様 $RA1 \leq RA2 \leq RA3 \leq RA4$ とする。さらに交点が多い場合についても、同様の方法で記録する。また、交点の数が奇数になることもありうるが、この場合は $RA1=RA2$ と対して対応する。このように視野の位置情報をデータベース化していく。実際にある座標値 (RA_IN, DEC_IN) で検索する場合、赤緯方向には1/5000度毎に離散化されて視野情報がデータベース化されているので、まず、DEC_INに対応するDEC_IN'を求め、
DEC_IN', RA1, RA2
のようなデータが視野データベースにあるかどうか判断する。次に、
 $RA1 \leq RA_IN \leq RA2$
が成り立っていれば、当該フレームの視野内に指定した座標値が含まれることになる。

本論文では、このように実装が容易なアルゴリズムを用いた。しかし、この方法では、Suprime-Camデータのようなほぼ四角形の領域とは異なる、例えば小円との交点が数多く生じるようなより複雑な形状を持つ領域を登録しようとする時、データベースに登録すべき情報が多くなってしまおうという問題がある。この点は今後の検討課題である。

さて、視野の外周部の情報だけしか保持しないとはいえ視野が広いため、各フレーム毎に多くのデータが存在し、総データ量は膨大となる。SMOKAで用いているデータベース管理ソフトウェアSybaseは巨大テーブルの検索を得意としないため、我々は独自仕様のバイナリファイルにデータベースを格納することとした。単一のファイルに位置較正済みフレームの視野データを書きこむと、現在位置較正されているおよそ17万フレーム分だけで3GB程度のファイルとなり、バイナリといえどファイル入出力に時間がかかる。そこで、天の北極から天の南極まで赤緯方向に一定の幅で1,800分割し、各々別個のファイルに保持するようにした。その結果、最も大きなファイルでも50MBと比較的小さいサイズに抑えることができた。このような方法はUSNO-B1.0カタログで用いられている。

各種データの流れを図2にまとめた。視野の位置情報のデータベースを作るため、まず、フレーム毎に位置較正済みデータの有無 (WCS_FLAG) をファイル管理テーブルFILEMNGより取得し、位置較正済みデータがある場合は、位置較正の情報を誤情報修正テーブルHeaderRecordCorr (図中 HRC) から取得する。また、視野を決めるのに必要な、EQUINOXやNAXIS1, NAXIS2等といったFITSキーワード値はHDI (Header Information) ファイルから取り出される。これらの情報を基に、視野の外周に相当する座標値が、赤緯方向に1/5000度毎に離散化された状態で、フレーム毎にアスキー形式の位置情報ファイル中に保持される。この時、既にファイルが生成されている場合には、そのファイルの更新日時と位置較正の行なわれた日時 (WCS_TIME) とが比較され、後者のほうが新しい場合には新たな位置較正情報を用いて位置情報ファイルが更新される。この処理時間はおおよそ毎秒2フレーム程度である。現在17万フレーム程の位置較正済みデータがあり、全てを処理するには30時間程度を要する。次に、位置情報ファイルからバイナリテーブルを作成するが、これも位置情報ファイルの更新日時とバイナリテーブルに保持されている日時とを比較し、前者が新しい場合にのみ更新を行なう。この処理にはおおよそ17万フレームで2時間半程要す

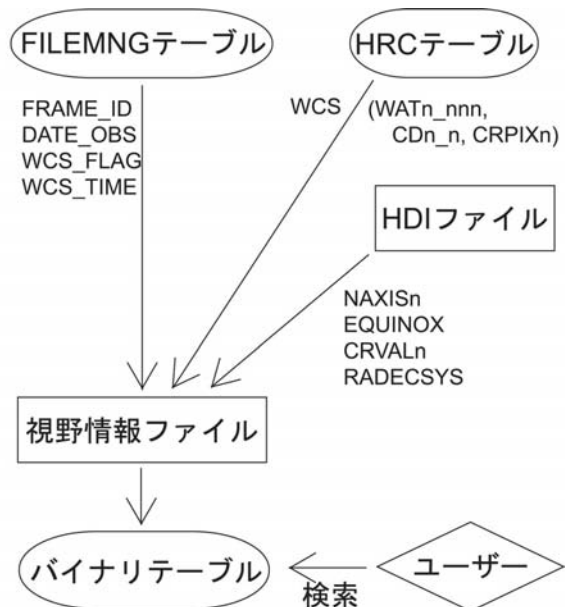


図2. 視野情報のデータベース化の処理の流れ。ファイル管理テーブル (FILEMNG)、誤情報修正テーブル (HRC) とヘッダ情報ファイル (HDI) から得た情報を基にして視野情報を求める。それをバイナリテーブルに変換して検索に用いている。

る。実際の運用としては、処理速度は現状のマシンで十分な速度が出ており問題ないと言えよう。

さて、実際の検索でのユーザインタフェースであるが、本検索機能や次で述べる品質指標による検索機能は現在Suprime-Camのデータにしか対応していない。そのため、従来の観測装置に共通のユーザインタフェースではなく、図3のようなSuprime-Cam 専用検索画面を作成した。ピンポイント検索機能は"Pin-Point Search"のラジオボタンを"on"にすることで使用され、"off"の場合は従来の検索機能が使われる。具体的には、J2000.0かB1950.0での赤経・赤緯の値を直接入力、もしくは天体名を入力し、ネームリゾルバを介して赤経・赤緯の値を取得することで、先のバイナリテーブルが参照され、検索が実行される。また、今までのアドバンスド検索と同様、露出時間や観測者などの情報と組み合わせて検索することも可能である。この場合、一旦位置情報だけで絞りこみ検索をかけ、さらに通常のSQLを用いて検索するという、二重の検索が行なわれる。これは前述のとおり、位置検索がSMOKAで通常用いられている Sybaseとは異なる独自のデータベースを使っているためである。また検索時間については、一つの座標値だけを指定して検索を行なった場合、天域によって異なるファイルにデータベース化されているため検索時間にもバラツキが生じるが、最も観測データの多い、従って検索時間のかかる天域でも0.3秒程度と、実用上問題ない速度が達成されている。従来の検索では、同様の検索条件で10秒程度要しており、検索時間は大きく短縮できた。これは、位置情報に限定したデータベースを作成した利点が出たといえる。さて、(RA, DEC) = (34.5, -5.2) というSXDS (Subaru/XMM-Newton Deep Survey)¹⁰⁾ 領域内のある座標値で検索を行なう場合、従来の検索方法では11,814フレームという膨大な数が検索にかかってしまう。このうち、ピンポイント検索の対象となる位置較正済みデータがあるフレーム数は 9,524である。これに対し、ピンポイント検索を用いると、744フレームとかなり絞りこんだ検索が可能となった。前述のように検索速度も向上しており、本機能の開発目的であった効率良い位置検索は達せられたといえる。

本開発で、観測データの視野の位置情報をデータベース化し、特定の座標を確実に含む観測データの検索が可能になった。しかし、特定の座標値

だけでなく、ある天域を含む観測データを検索することはまだできず、残された課題である。さらに、この検索の位置検索精度は位置較正の精度に大きく依存するので、それを如何に自動的に評価し、必要に応じどう修復するかは今後の重要な運用課題である。

(2) データ品質評価指標による検索

観測データの品質は、観測時の天候や観測システムの調子などに大きく左右される。しかし、従来のSMOKAには品質指標がなかったため、アーカイブ利用者が膨大な観測データの中から自分の研究目的に合った品質のものを探し出すには、一旦全てのデータについて解析処理を行なう必要があり、多大な労力が必要で無駄も多かった。それ故各データに品質指標をつけ、それを条件に検索を行なえる機能が求められていた。我々は、すばる望遠鏡のデータ品質評価システムを開発しており、これにより測光零点やPSF (Point Spread Function), バックグラウンド値などを求めている¹¹⁾。そこで、研究目的に合致した品質の観測データを検索できるようにするため、品質評価システムでSuprime-Camのデータの品質評価指標を予

め求め、これを条件として検索を行なう機能を開発した。その結果、例えばseeing値が1秒以下の観測データのみを取得するといった、有用な検索機能を実装できるようになった。

今回の開発では、品質評価指標として特に重要と考えられる、seeing値と限界等級を条件として検索できるようにした。品質評価システムで得られたフレーム毎のバックグラウンド値 (スカイノイズ値) とseeing値、及び測光零点を基に各フレームのS/N比を求め、限界等級を算出する。そして、この限界等級値とseeing値をデータベーステーブルに登録する。検索時に品質評価指標が条件として指定された場合、このテーブルを参照することにより検索を実現している。

なお、ユーザインタフェースは先のピンポイント検索のものと同様Suprime-Cam 専用検索画面で実現している (図3)。この画面中、"Data Quality"という項目の"Seeing", および"Limiting Magnitude"という部分に検索対象とするseeingの最大値、または限界等級を入力し、検索を行なうことで、データ品質指標による検索が可能となる。

図3. Suprime-Cam 専用検索画面。ピンポイント検索が使えるほか、データ品質 (シーイング, 限界等級) による制限づけ、較正済みデータの選択も可能になっている。

2.2 FITSヘッダ検索機能

観測データのFITSヘッダには数多くの情報が含まれているが、従来のSMOKAの検索ではこのうち一部の良く用いられるキーワードしか検索条件に使用していなかった。しかし、より多様な検索を実現するため、FITSヘッダの各キーワードを検索条件とすることができる、二つの検索機能を追加した。一つは全キーワード検索機能であり、FITSキーワードの「値」を検索に使うものである。もう一つは全文検索機能であり、FITSヘッダ部分の HISTORYやCOMMENT、各キーワード行のコメント部といったものまで「語」として検索するための機能である。各々について詳細を以下にまとめる。

(1) 全キーワード検索機能

本機能は、観測データのFITSヘッダ上の各キーワードの値、およびそこから導出された重要なパラメータ値での検索を可能とするものである。これまでのSMOKAの検索では、それらの中で大部分の利用者が求めるであろうものだけを対象としてきた。検索項目を最小限にすることで、検索プログラムの過剰な肥大化やユーザインタフェースの複雑化を防ぎ、かつ検索速度を確保するためである。しかしSMOKAの利用者が増え、その要求も多様化してきたために、これまでの検索項目だけでは不十分な場合が生じてきた。そこで本論文では、すばる望遠鏡の観測装置毎のユーザインタフェースを作ることで複雑化を最低限に抑えつつ、本機能を実現した。

本開発では、全キーワード検索機能を以下のように実現した。まず、予めキーワードの一覧を作成しておく。この一覧には、すばる望遠鏡の観測データのヘッダ情報の一部を登録しているデータベーステーブルXXX_MASTARS (XXXは観測装置：論文2, 7.1節参照) から得た各キーワード毎の値の型 (文字, 実数, 整数) の分析結果、すなわち、

- ・ 値が文字：値の種類が64通り以下の場合にはそのリスト、それより多い場合には頻度の多い方から3つのリスト、ただし、全フレームで値が異なるなど頻度の違いが顕著でない場合にはサンプル3つのリスト。例外としてOBJECTとOBSERVERの場合には、値の内容だけでなく記述形式の多様

性が極めて多く分析が複雑となるため、分析を行っていない。

- ・ 値が実数：値の種類が1通りならその値、2つ以上の場合には最小値と最大値。
- ・ 値が整数：値の種類が32通り以下の場合にはそのリスト、それより多い場合には最小値と最大値。

を含めている。値の型によって扱いが異なるのはユーザインタフェースに表示した場合の見やすさと、利用者の値入力の利便性を考えた結果である。ただし、「64通り」、「32通り」といった数については検討の余地が残る。この一覧作成作業は、新たな観測データが入力された際に行なっている。なお、FITSヘッダ上の全キーワードを対象にするとと言っても、検索する意義が低いもの (SIMPLEやEXTEND, OBSERVATやTELESCOPなど) や全文検索機能に委ねる方が効率的であるもの (HISTORYやCOMMENT) は除いている。一方、FITSヘッダ上にはないが、含めると便利だと思われるXXX_MASTARSテーブル上の導出パラメータ、例えばRASEC (RAの秒単位表記) やX_2000 (2000年分点での方向余弦のX成分) なども含めている。

利用者がWeb上で本機能を選択すると、このキーワード一覧をもとにして各観測装置対応のユーザインタフェースが表示される (図4)。前述の値分析結果に応じて、

- ・ 値が1通りの場合にはその値 (利用者の値入力はなし)
- ・ 値が64または32通り以下の場合には値選択プルダウンメニュー
- ・ 最小値・最大値の表示と検索条件としての最小値・最大値入力窓
- ・ 値のサンプル3つと検索条件入力窓

のいずれかが表示される。値分析を行っていないOBJECTとOBSERVERはあらかじめ決めておいた値の例を示している。FITSヘッダ辞書^{*3}掲載のキーワードには辞書へのリンクが設定され、キーワードの意味を調べられる。検索は入力された検索条件すべてのANDで行なわれ、検索条件が入力 (選択) されないキーワードに対しては条件付けがされない (つまり全ての値が選ばれる)。検索結果は、該当するFRAME_ID (フレーム番号) とそのDATE-OBS, 検索条件が設定されたキーワードの値の一覧として表示される。

*3 <http://smoka.nao.ac.jp/fitsdic.jsp>

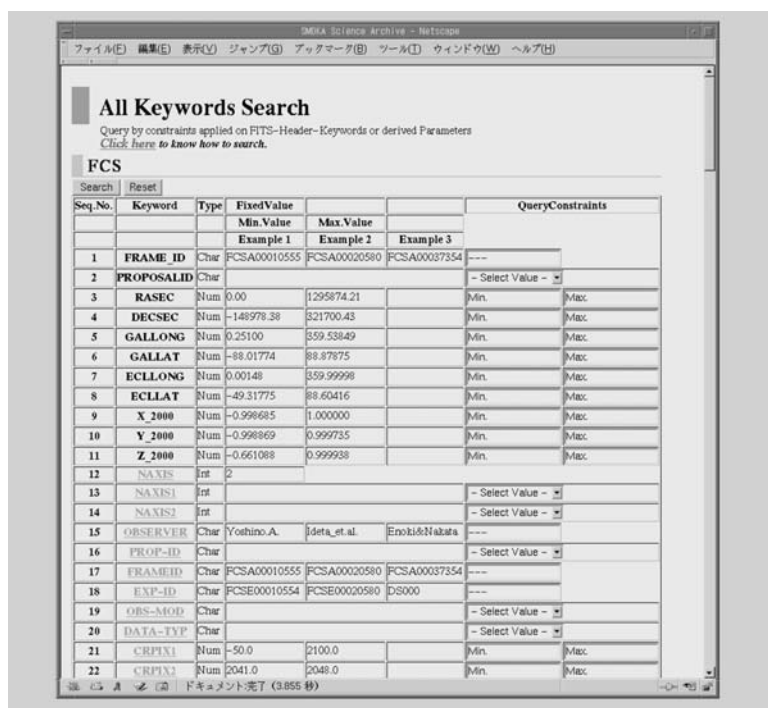


図 4. 全キーワード検索画面の例。FITS キーワードの大部分（MASTARS に未登録のもの、検索に意味がないものは除いている）と有益な導出パラメータによる検索が可能になっている。この例は FOCAS についての検索画面である。FITS ヘッダ辞書掲載のキーワードには、辞書へのリンクが設定され、キーワードの意味が調べられるようになっている。

今後は、本機能とこれまでの検索の関係や、後述する全文検索機能との使い分けなど、より使いやすい検索の実現を求め、運用を通じてさらに検討を深めていく予定である。

(2) 全文検索機能

前述の全キーワード検索ではデータベーステーブル XXX_MASTARS からキーワード一覧を取得している。この XXX_MASTARS というテーブルはすばる望遠鏡システムの一部として存在する観測者向けアーカイブシステム MASTARS¹²⁾ のテーブルのコピーである。従って、全キーワード検索では、MASTARS の枠組とは関係のない岡山天体物理観測所と木曾観測所の観測データを検索することができない。また、FITS キーワードとしては存在するが XXX_MASTARS には存在しないキーワードも検索条件にできない。例えば、HDS¹³⁾ の I2 Cell mode を使用したかを示す H_I2CELL キーワードは HDS_MASTARS には存在しない。更に、全キーワード検索は各キーワードの「値」を検索対象としていたが、HISTORY や COMMENT、各キーワードのコメント部分の検索のためには「語」として検索できる全文検索機能が必要となる。そのため、観測データの FITS ヘッダ部分の全文を検索する機能を構築し

た。実際に検索対象とするのは、FITS ファイルからヘッダ部分のみを切り出した HDI ファイルである。本機能では、全文検索システムとして namazu⁴⁾ を用いている。インデックス化は namazu で提供されているツール mknmz により行なっている。このインデックス化は、インデックスファイルの肥大化を緩和するため、観測装置毎に行なっている。

実際に検索する際は、検索キーワードを入力し、検索対象とする観測装置を選択する（図 5）。また利用者によどのような FITS キーワードがあるのか示すため、FITS ヘッダ辞書へのリンクも張られている。検索を行なうと FRAME_ID（フレーム番号）と、FITS ヘッダ中どの行が検索にかかったかが表示される（図 6）。さらに各フレームにはチェックボックスがついており、必要なフレームにチェックすると、従来の検索結果一覧表示画面に移動し、データ取得などができる。

運用としては、cron によって定期的に新たな、もしくは更新された HDI ファイルの有無を検出し、新しい、もしくは更新されたファイルがあった場合、そのファイルの情報を装置毎のインデッ

*4 <http://www.namazu.org/>

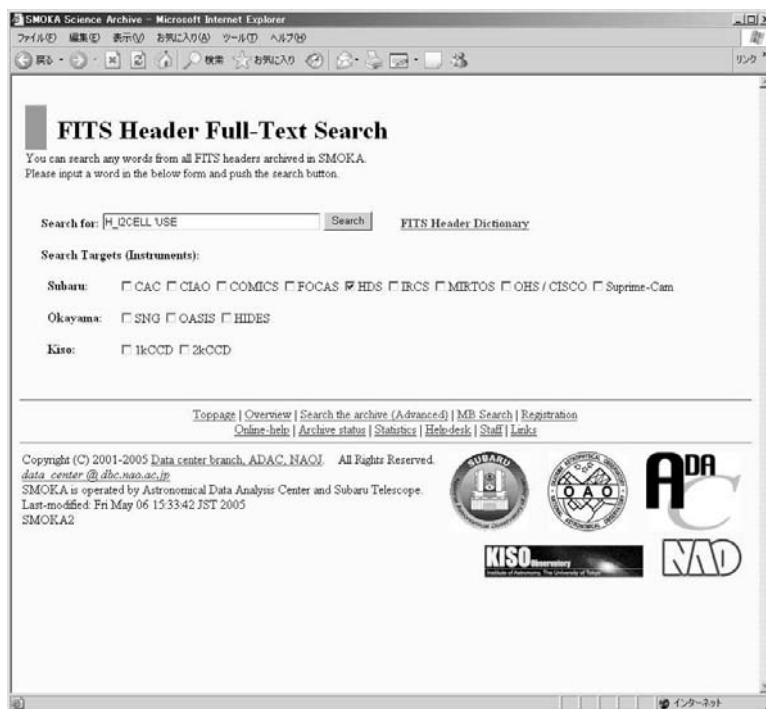


図 5. 全文検索画面の例。FITSヘッダの全文検索ができる。検索は、FITSキーワードとその値だけでなく、各行のコメント部、COMMENT行や HISTORY行などFITSヘッダ内の全てのテキストを対象にしている。

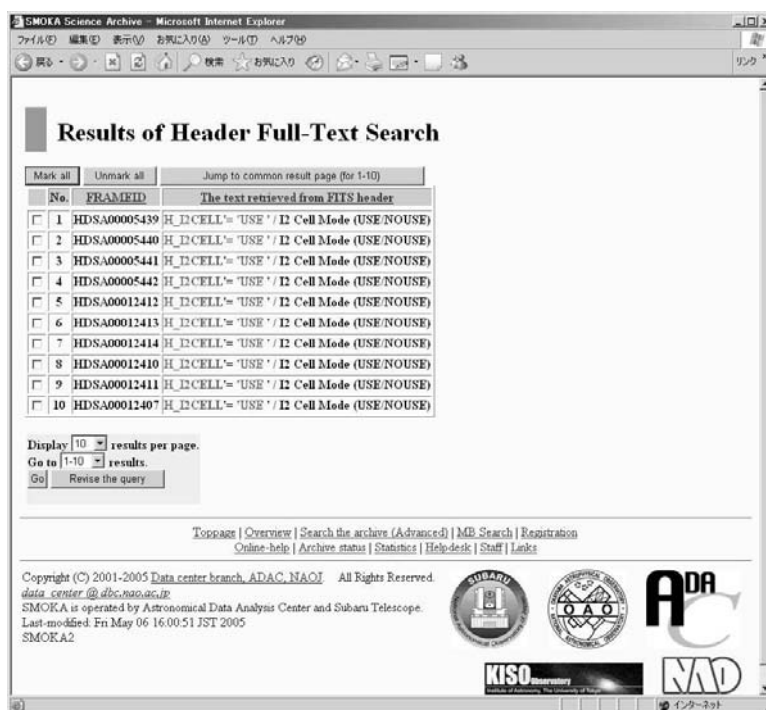


図 6. 全文検索結果一覧表示画面。FITS ヘッダ全文検索の結果一覧表示画面を示した。入力した語を含むフレームのフレーム番号と、その語を含む FITS ヘッダ行が表示される。

クスファイルに追加するようにしている。

この全文検索システムは、利用者の指定した検索キーワードと完全一致、もしくは部分一致する語句をFITSヘッダに含む観測データを検索することが目的である。すばる望遠鏡FITSヘッダ辞

書への届出がされないままに追加されたようなキーワードはXXX_MASTARSへ登録はされない。従って、このようなキーワードを検索するには、FITSヘッダに書かれているものすべてが検索対象となる、この全文検索機能で検索する以外方法はない。ただし、「語」としてしか検索できない

ので、キーワード値の範囲指定などはできない。そのような目的には、可能であれば前述の全キーワード検索などを用いることになる。全文検索と全キーワード検索をどのように連携させるか等については、その利用価値なども含めて今後の検討課題である。

2.3 カレンダー検索機能

地上での観測データの品質は天候に大きく左右されるため、アーカイブユーザは観測日の天候をデータ取得の判断基準とすることが多い。そこで我々は、観測日毎に気象情報を表示し、そこからデータを直接検索できる機能を開発した。以前我々は、すばる望遠鏡の気象情報などステータス情報を表示する機能として、プロポーザルカレンダー表示機能を開発した（論文2，4.2節参照）。この機能では、すばる望遠鏡の観測データに対して観測装置・観測年月を指定すると、カレンダー

形式の表に観測日・プロポーザル番号毎に分類された総フレーム数やフレーム番号の範囲（最大/最小番号）、気象データなどの情報が参照でき、そこからその観測日に取得された観測データを検索する機能を持っていた。

本論文では、すばる望遠鏡の観測データだけでなく、岡山天体物理観測所や木曾観測所の観測データにも対応すべく、カレンダー表示による検索機能の再開発を行なった。プロポーザルカレンダー表示機能は、アーカイブ利用者から高い評価を受けている。そこで、各観測所の観測データに対応させる際、ユーザインタフェースの大きな変更は行なわないようにした。以前のプロポーザルカレンダー表示機能では、観測装置の追加・変更・削除、または気象データなどの他の付加情報を追加する場合、HTMLファイルを生成するプログラムの変更が必要であり、運用上大きな負担であっ

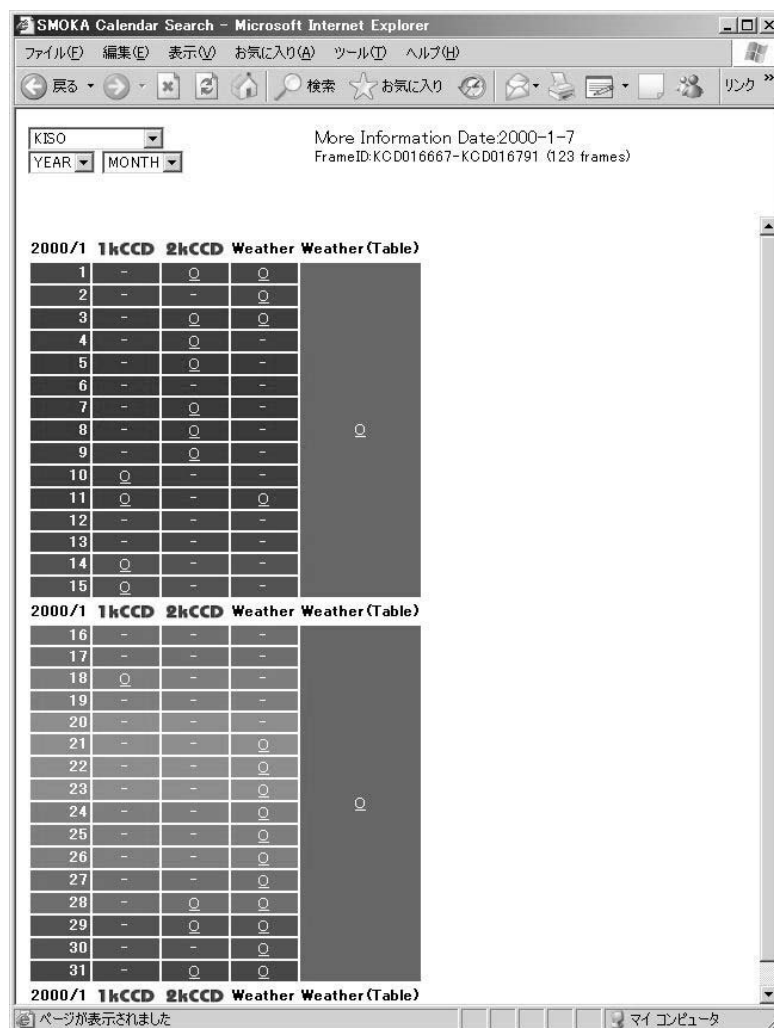


図7. カレンダー検索画面の例。左上のプルダウンメニューは、観測所、年、月の選択のためのもの。表の右端の欄には、気象データや天候表などへのリンクがある。画面構成を従来のプロポーザルカレンダー表示機能のものに合わせている。

た。これを解消するため、観測装置や気象データなどの情報はテキストファイルから読み込むことにした。その結果、観測装置の追加や変更の必要が生じた場合、テキストファイルを編集するだけで対応できるようになった。

本検索機能では、週に一回の頻度で日毎の観測や気象データの有無、観測フレーム数のカウントなどを行ない、カレンダー検索画面を構成するHTMLファイルを作成している。その例を図7に示す。この画面の左上にプルダウンメニューが3つあり、それぞれ観測所、年、月を選択するものとなっている。論文3において追加された岡山天体物理観測所や木曾観測所の気象データ、天候表や野帳などの付加情報は、表の最も右側の列からリンクされている。また、以前のシステムにおけるHTMLファイル作成には2時間程度の処理時間が必要であったが、SQL文を見直すことにより、この処理が数分程度に抑えられ、運用の効率化がはかられた。

3. 検索以外の機能強化

2章で述べた検索機能以外にも数点の機能強化をSMOKAに施した。まず、我々は Suprime-Camデータを各チップ間でバックグラウンドレベル補正した上でモザイク化した、新たなショット画像を作成・公開した。これについては3.1節で解説する。次に3.2節では、複数版の整約済みデータの公開を可能とするSuprime-Cam専用の検索結果一覧表示画面について示す。また、仏CDS^{*5}のVizieR^{*6}からSuprime-Camデータを検索することができるようにした。これについて3.3節で解説する。COMICSの生データに簡易チップピング処理を施したCOMQデータの公開について3.4節で、論文2で述べた移動天体検索機能で用いる太陽系内天体の軌道要素テーブルの自動更新について3.5節で述べる。

3.1 Suprime-Camデータのショット画像

Suprime-Camは10個のCCDチップのモザイクで構成されている。ここでは、同時に露出して得られた一群の観測データ（通常は全10チップ分）を1ショットと呼ぶことにする。Suprime-CamデータのFITSヘッダ上では、EXP-IDというキーワ

ードでショットが識別されている。

SMOKAでは、観測データを画像として閲覧・評価するためにQLI (QuickLook Image) を用意している (論文1)。さらに、チップ毎のサムネイル画像も用意している (論文3)。前者は観測データを統計情報を含めて詳細に閲覧するのに適し、後者は多数の観測データをチップ単位の表示ではあるが同時に閲覧するのに適している。しかし、フレーム数が膨大となってきたSuprime-Camデータにおいて、必要な観測データを効率良く見いだすためにはこれらだけでは不十分となり、より迅速かつ体系的に観測データを閲覧する手段を提供する必要性が生じてきた。そこで本論文では、1ショットをまとめたjpeg画像を作成して、閲覧に供することとした。

本機能の必要要件として、バイアス値の違いや感度 (AD変換係数の違いや光学系による周辺減光を含む) の違いによるチップ間のバックグラウンドレベル差を補正して見やすい画像にすること、Web上で容易に見ることができるjpeg, gif, pngなどの一般的な画像形式にすること、できるだけ人の手を介さずに画像を作成できることが求められる。また、フレーム番号や取得日時、天体名などの情報を画像と同時に表示することも必要である。

この要件はIRAF¹⁴⁾とds9¹⁵⁾を用いることで満たすことができた。まずIRAFを用い、1ショット分の観測データに対して、各フレームのバイアス値 (オーバースキャン部の平均値を採った) を求めて各フレームから差し引く。次に、ドームフラットデータを解析して得ておいた各フィルター毎のチップ間レベル補正計数を乗ずる。そして、IRAFのdisplayタスク中のzscaleアルゴリズムを用いて表示レベル (最小カウントと最大カウント) を定め、ds9上の各チップの相対位置に対応した場所に表示する。なお、視野端のチップでは周辺減光によりチップ内でのカウントの差が生じて、そのままでは見にくくなってしまいうため、こうして計算した表示レベルよりも低いカウント値 (約10%) まで表示するようにしている。1ショット分のフレームを並べて表示した上下のマージン部分には、チップ番号 (0~9) に加えて、FITS

*5 <http://cdsweb.u-strasbg.fr/>

*6 <http://vizier.nao.ac.jp/vizier/>

ヘッダから読み出した観測日、フレーム番号、データ種別 (OBJECT/SKYFLAT/STANDARDなど)、観測時刻、赤経・赤緯、フィルター名、露出時間、観測天体名の諸情報を書き込んでいる。次に、ds9のXPA (X Public Access)^{*7}機能を用いて、コマンドからds9の画面のハードコピーを取り、それをjpeg形式に変換して磁気ディスク上に蓄える。それをWebから参照できる所定の箇所に移動すればショット画像の完成である (図8)。

これらの処理はIRAFのclスクリプトとしてまとめられており、多数ショットを連続的に実行することができるようになってきている。つまり、ショット画像作成対象のリストを準備して起動すれば、あとは人手を介さずに必要分のショット画像の作成ができるようになってきている。

このように、ショット画像は上記の必要要件を満足し、充分実用になるものとなったが、運用的には2つの課題がある。1つは、ds9に表示された画面のハードコピーを取るため、ワークステーションの画面を起動し、かつ、しばらく占有する必要があるということである。もう一つの課題は、観測データ本体を読み出してds9に表示するため、

時間がかかることである。現在用いているワークステーションでは、1ショットあたりおよそ1分の時間を要している。1観測ラン (連続した一連の観測: 1ヶ月当り1週間程度) の観測データでは、ショット画像作成におよそ半日を要する。これらは現時点では大きな問題となっていないが、さらなる効率化を図る際には問題となろう。

3.2 Suprime-Cam一次処理済データの複数版公開

従来、SMOKAでは一切の解析処理を施していない生データのみを提供していた。従って、整約作業は利用者が全て行なわなければならなかった。装置固有の状況やデータ取得の状況を十分に把握できないSMOKA利用者にとっては、整約作業は容易ではなく、天文学的研究成果を得るのは困難である場合も多かった。そこで、利用者からの請求が最も多いすばる望遠鏡のSuprime-Camの観測データを対象に、整約処理を行なったデータの公開を2004年9月より開始した。ここでは、整約処理として、いわゆる一次処理 (バイアス差引、フラットフィールド処理)、および位置較正とフラックス較正の処理を行なっている (詳細は論文3参照)。

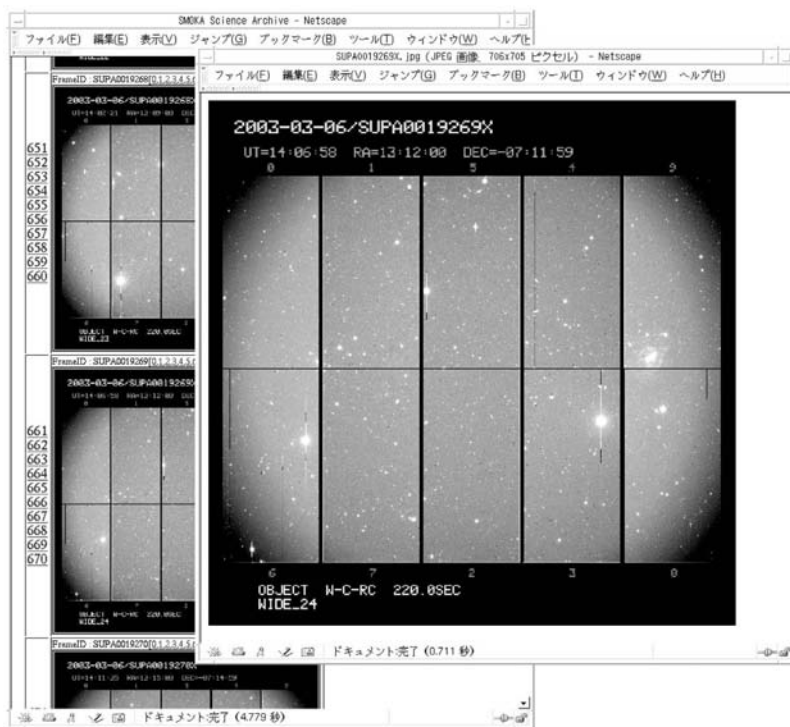


図8. ショット画像の例。背後の一部隠れた画面は、検索結果一覧表示画面の中のショット画像表示ボタンを押した際に表示されるショット画像の一覧である。その中の一つをクリックすると上に重なった画面のように拡大して表示される。Suprime-Camの10チップ分の画像が合体され、その上下に属性情報が表示されている。

*7 <http://hea-www.harvard.edu/RD/>

SMOKAで行なっている整約作業は、なるべく単純かつ標準的な処理をすることで可能な限り自動化を行ない、運用として現実的な作業量・作業難度に止めるように留意している。そのため、観測データの品質や研究目的によっては、従来の単一のパターンで行なっていた整約処理が最良であるとは限らない場合もあった。そこで、SMOKAでは複数の整約パターンで処理を行なったデータを用意し、利用者がそれらの中から自分の研究目的に合った整約済みデータを選択できる機能を追加した。具体的な整約パターンとしては、1 観測ランの中のフレームから作成されたSuperFLATを用いるパターン (S)、星や銀河にマスクをかけて作成したSuperFLATを用いるパターン (M)、複数観測ランのフレームから作成されたHyperFLATを用いるパターン (H)、の3つが現時点では採用されている。また同じ処理方法でも、SuperFLATやHyperFLATを作成するのに用いられるフレームの取捨選択を変更することによって、整約精度が向上する場合がある。この場合には、版番号 (バージョン) で区別するようにしている。

SMOKAでは現在、すばる望遠鏡の9つの観測

装置^{*8}、岡山天体物理観測所の3つの観測装置^{*9}、東大木曾観測所の2つの観測装置^{*10}のデータを提供している。他の多くの天文データアーカイブと異なり、SMOKAではこれら複数の装置の観測データを同時に検索し、結果一覧を同時に表示できるように統一的なユーザインターフェースを用意しており、SMOKAの大きな特徴の一つとなっている。しかし、整約済みデータ公開はSuprime-Camのみで行なうこと、更に複数のパターンの整約済みデータを用意することから、検索結果で表示されるべき情報がSuprime-Camとそれ以外の装置では大きく異なることになった。そのため、結果表示を分離した方が利用者にとって分かりやすいと考え、新たにSuprime-Cam専用の検索結果一覧表示画面を作成した (図9)。また、複数の観測装置で横断的に観測データを検索できるSMOKAの特徴を残すべく、従来の検索機能 (シンプル検索、アドバンスド検索、カレンダー検索、移動天体検索) においてSuprime-Camは生データのみ検索可能とし、検索結果一覧表示画面ではSuprime-Camの整約処理に関する情報の表示をやめ、論文3以前のものに戻した。Suprime-Cam専用の検索結果一覧表示画面 (図9) が従来のものと異なる点を図中丸で囲んで示した。具体

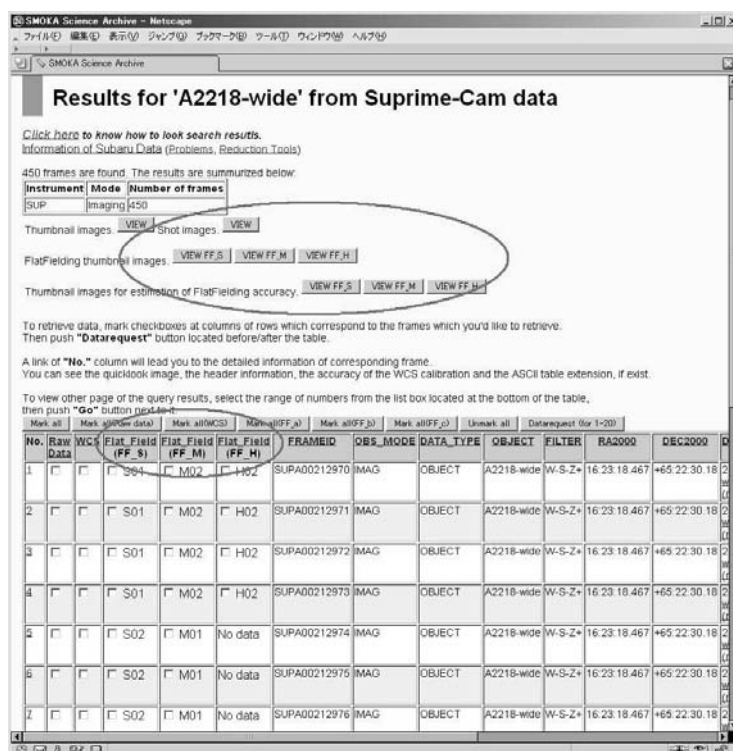


図9. Suprime-Cam 専用検索結果一覧表示画面。2.1 節で述べられている Suprime-Cam 専用検索画面から検索を行なった結果の表示画面である。整約・較正処理の情報が示されている。

*8 Suprime-Cam⁴⁾, FOCAS¹⁶⁾, HDS¹⁷⁾, OHS/CISCO¹⁸⁾, COMICS⁶⁾, CIAO²⁰⁾, IRCS²¹⁾, CAC, MIRTOS²²⁾

*9 SNG²³⁾, OASIS²⁴⁾, HIDES²⁵⁾

*10 1kCCD²⁶⁾, 2kCCD²⁷⁾

的には、整約済みデータの有無の表示がなされていること、一次処理の精度を表示する精度表示画面へのリンク、および3.1節で述べた Suprime-Camのショット図へのリンクがあることである。

SMOKAでは従来、フレーム番号のみを指定すれば観測データを唯一に特定できたため、データベース上でフレーム番号が主キーとして設定されてきた。しかし、複数版整約済みデータの公開により、同一フレーム番号のデータに対して複数パターンの整約済みデータが存在してしまう。そこで、整約済みデータを管理するため、ファイル管理テーブルFILEMNGとは別に、整約済みデータファイル名を主キーとする整約済みデータ管理テーブルFF_MNGを作成し、Suprime-Cam専用検索やデータ提供の際に用いるようにした。ここで主キーとなる整約済みデータファイル名の命名規則を新たに作った。従来は、例えばフレーム番号がSUPA00103701であるデータに整約処理を施したものは、SUPF00103701.fitsというファイル名であった。しかし、複数の整約パターンが存在したり、また同一の整約パターンでもバージョンの違いが生じたりするので、ファイル名にはパターンとバージョンを示す記号を含めることにした。上記の例においてはSUPF00103701_S02.fits等とし、“S”が整約パターン、“02”がそのバージョンを意味するようにした。また、このパターン+バージョン番号は整約処理済みデータのFITSヘッダにも明示している。

3.3 CDS VizieRへのリスト提供

仏CDSによって運用されているVizieRには世界中の天文カタログデータが蓄積されている。加えてVizieRは、世界の幾つかの望遠鏡について、得られた観測データの座標値や天体名、露出時間といったデータリストも保持しており、それを検索できるようにしている。SMOKAでもこのようなリストを作成し、多くの天文学者が利用するVizieRへデータリストの提供を行なうことにより、SMOKAの利用者を増加させ、科学的成果を増大させる効果が期待できる。しかし、Suprime-Camの生データの位置情報は2.1節でも述べたように、大きくずれているものが多く、位置の情報は10個のCCDチップすべてに共通である。このような位置情報を提供することは利用者には好ましくない。利用者には各CCDチップの精度良い位置情報を提供することが望まれる。そこで我々は、位置較正済みのSuprime-Camデータのデータ

リストをVizieRに提供することとした。VizieRに対し提供する情報を表1に示す。この中で、赤経・赤緯は各チップの中心ピクセルの赤道座標値(J2000.0)としている。

データ種別	キーワード
観測日	OBS_DATE
観測開始時刻	UT_START
フレーム番号	FRAME_ID
露出番号	EXP_ID
赤経・赤緯	RA, DEC
データ分類	DATA_TYPE
天体名	OBJECT
フィルター	FILTER
露出時間	EXP_TIME
エアマス	AIRMASS
チップ番号	CCD_CHIP
縦横 pixel 数	NAXIS1, NAXIS2
整約済みデータの有無	FF_FLAG
測光零点	ZP
Seeing	SEEING
星の楕円率	ELLIPTICITY

表1 データリスト一覧

3.4 COMICS の簡易処理済みデータ

SMOKAでは、これまでCOMICSの生データ「COMA」のみ公開していたが、これに加えて簡易処理済みデータの公開を始めた。これはCOMICSの観測直後に、ハワイ観測所の観測装置制御システムOBCP上のQDAS²⁸⁾によって生成される「COMQ」データを用いている。COMAデータは、検出器のx軸、y軸、時間軸(複数の露出フレーム)、検出器番号、の4次元データとなっている。それに対し、COMQはCOMAをもとに、時間軸に沿って同一チョッピングビームのフレーム(「ON」フレーム)については和をとり、異なるチョッピングビームのフレーム(「OFF」フレーム)については差をとって足し合わせたもの、つまり簡易チョッピング処理を行なって時間軸をまとめたものである²⁹⁾。従って、一つのCOMAデータに対して、一つのCOMQデータが生成される。COMQはCOMICSのQLIを作成するための元データとして既に運用側で利用されているが(論文3)、COMQが公開されたことにより、利用者はチョッピング差し引きの部分の整約の手間が省けるようになった。ただし、スカイフラッ

ト作成やデータチェックの用途にはチョッピング処理前の生データが必要なため、利用者へのCOMAの提供をすべてCOMQに置き換えることはできない。そこで、請求に対してCOMAとCOMQをセットで提供することとした。このようにしても、COMQデータは1ファイルあたり300~900KB程度（対応するCOMAの約十分の一）と比較的小容量であることから、運用上の負担にはならない。

3.5 軌道要素データベースの自動更新

移動天体検索では、太陽系内小天体の軌道要素を入力することでSMOKAにアーカイブされているSuprime-Camと2kCCDの観測データの中から目的の天体が含まれている可能性のあるフレームを検索することができる。その際、利用者が軌道要素の全てを入力するのは能率的でなく、また入力項目が多いので誤入力の可能性も高い。そこで、天体名を指定すると軌道要素を検索するネームリゾルブ機能を用意している。この機能を実現するために、観測データのテーブルとは別に天体名や確定番号、仮符号などを登録した軌道要素テーブル（ORBITS）を用意している。このテーブルは、MPC^{*11} Orbit Databaseより提供されている軌道要素ファイル（MPCORB.DAT）^{*12}を用いて作成している。しかし、移動天体検索で用いている位置推算コードは特定のエポック（元期）にしか対応していないため、観測日がエポックから大きくずれているような観測データに対して位置比較を行なおうとした際、位置推算の計算結果に大きな誤差が生じてしまう可能性がある。そこで、移動天体名を検索するレゾルブ機能では、ORBITSテーブルに登録されている複数のエポックから適切なものを利用者が選択する機能を提供している（詳細は論文2参照）。

移動天体の軌道精度は、その位置観測回数はもちろん、その観測期間によって大きく左右される。すなわち、観測期間が長ければ軌道精度を格段に向上させることができる。近年、移動天体サーベイ専用望遠鏡の稼働に伴い、年間の小惑星の発見個数が増加している。2004年に発見され、仮符号が付いている小惑星は3万弱存在する。これらの新天体をSMOKAにアーカイブされている過去の観測データの中から検出することができれば、観

測期間は一気に数年に広がることが期待され、より精度の高い軌道を決めることが可能となる。

移動天体をSMOKAで容易に検索するためには、ORBITSテーブルを常に最新に保ち、新天体の軌道要素が移動天体検索でレゾルブできるようにしておく必要がある。そこで、MPC Orbit Databaseが提供しているDaily Updateの情報を毎日ORBITSテーブルに登録する仕組みを実装した。Daily Updateでは、新発見された天体の軌道要素に加えて、既知天体の改良された軌道要素が提供されている。これらの混在した軌道要素をORBITSテーブルへ反映させるために、以下の処理を行なう。

- 1) Daily UpdateであるDAILY.DATファイルのダウンロード
- 2) 天体がORBITSテーブルに登録されているか否かの判定
- 3) すでに登録されている天体の場合は、
 - 3-1) 同一エポックでありながらORBITSテーブルと異なる軌道要素の場合は当該天体の軌道要素を更新する（改良軌道の登録）
 - 3-2) 異なるエポックの軌道要素が提供されている場合は、新規登録
- 4) 新天体であった場合は、ORBITSテーブルに当該天体を追加する。以上の処理によって最新の軌道要素を提供しつつ、過去のエポックの軌道要素も提供することができる。

上記処理3-2)が頻繁に繰り返されると、異なるエポックの軌道要素が数多く提供されることになる。エポックが異なる軌道要素を用いれば、位置推算に違いが生じる。しかし多くのメインベルト天体では、1年程度のエポックの違いによる位置推算誤差は典型的には数秒角程度であり、観測装置の視野に比べて充分小さい。そこで、異なるエポックの軌道要素が複数登録された場合、その頻度が年に一回程度になるようORBITSテーブルを定期的に整理することを検討している。

なお、論文2でも挙げた比較精度の向上と検索の高速化がなおも課題として残されている。

4. 運用の効率化

論文3で述べた整約済みデータの公開開始以来、SMOKAの運用項目に整約処理が加わり、

*11 <http://cfa-www.harvard.edu/cfa/ps/mpc.html>

*12 <ftp://cfa-ftp.harvard.edu/pub/MPCORB/>

SMOKAでのデータフローが非常に煩雑化した。一次処理がどのように行なわれたのかの履歴や、較正結果の精度を利用者が閲覧できるようにしたこともデータの流がより複雑化した要因の一つとなっている。現在数名のメンバーで作業を分担しSMOKAを運用している。しかし、ファイル数やデータベースの管理テーブル数が膨大であることや、しばしば起こる運用メンバーの交代のため、ミスが生じる可能性は以前より高くなっている。そこで、可能な限り人為的なミスをなくし、処理手順を把握しやすくするために、ファイル・データベーステーブルの依存関係図の作成や運用ドキュメントの整備を行ない、それらに沿った自動チェックプログラムを作成した。4.1 節から 4.3 節ではこのチェックプログラムについて述べる。4.4 節では、負荷が大きくなってきたSMOKAサーバーの複数サーバーへの負荷分散について解説する。

4.1 ファイル・テーブル依存関係図

まず、複雑化したSMOKAでのファイル生成やデータベース登録の流れを視覚的に把握するた

め、処理履歴や環境データを含めた全てのデータの流れを示した依存関係図を作成し、個々のファイルとデータベーステーブル、データ処理（プロセスと呼ぶ）の依存関係を明確化した。図10に Suprime-Camの例を示す。

この図には、MASTARSからFITSデータをコピーし、整約処理、HDI, QLI, サムネイルなどの作成を行ない、利用者にデータファイルを提供するまでの流れが示されている。図において、プロセス（矢印）とデータファイル（四角）、データベーステーブル（角の丸い四角）が区別されている。また、主要なプロセスは丸で囲んで他と区別している。さらにデータベースのうち、頻繁にアクセスが必要となるFILEMNGテーブルとHeaderRecordCorr テーブル、FF_MNGテーブルはそれぞれ星印と十字、三角で示されている。各プロセス（矢印）に付随するMNG01などの記号は後述の運用ドキュメント名を表している。

4.2 体系化された運用ドキュメントの整備

先に述べた依存関係図に従い、効率的かつ誤り

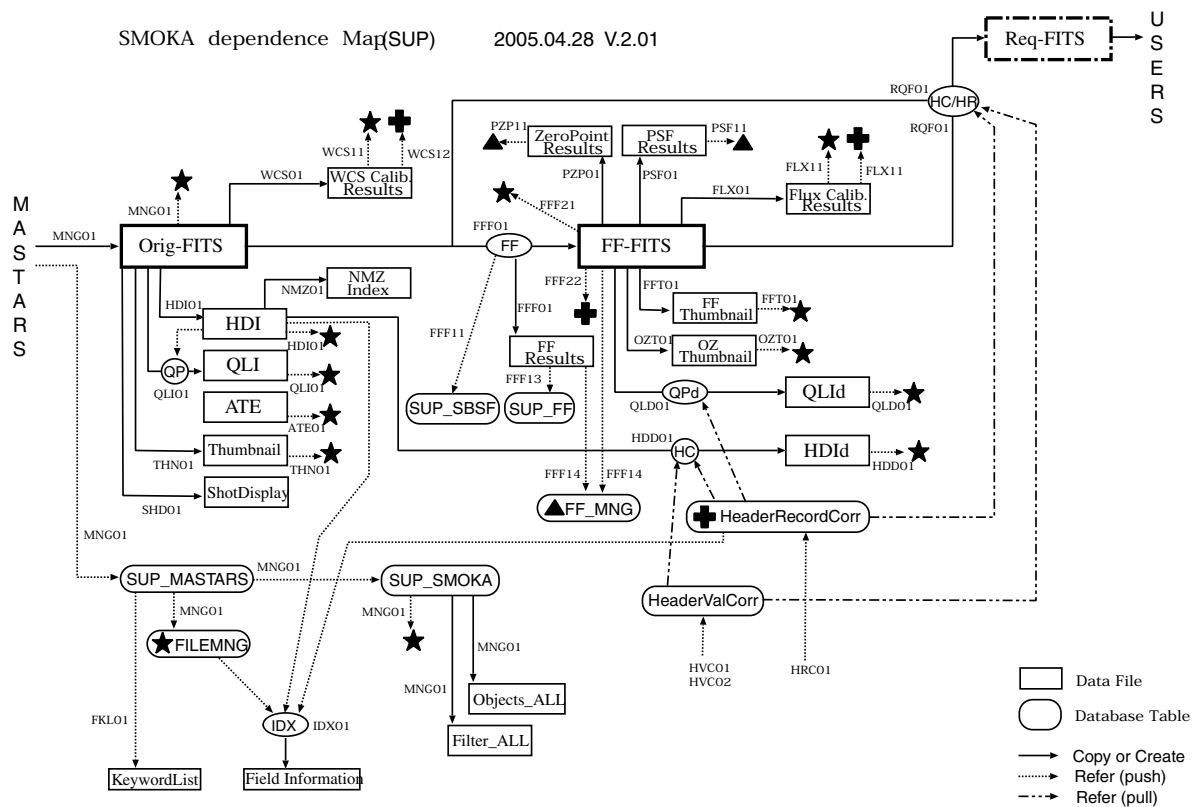


図 10. ファイル・データベーステーブル依存関係図。運用の効率化をはかるために作成された、ファイルやデータベーステーブル、それらを生成するプロセスの依存関係を表現した図。この図は運用・管理者用の Web に掲載されており、図をクリックするとその箇所に該当するドキュメントが表示されるようになっている。

無く運用するため、各プロセスを詳細に解説した運用ドキュメントを体系化して整備した。まず、ドキュメント命名規則と内容のフォーマットを定め、統一的なドキュメントを作成した。これにより、開発担当者と運用担当者が異なったり、プロセスで用いているソフトの改良を当初の開発担当者以外の者が行なう場合にも容易に対処できるようになった。また、提供されるデータの流れとは直接の関係はないデータのバックアップや利用者登録など、依存関係図に無い作業プロセスについても、同じ命名規則・フォーマットでドキュメントを作成した。前節で述べた依存関係図上の各プロセス名からそのプロセスの運用ドキュメントへリンクを張り、Webブラウザ上でプロセス名をクリックすると、簡単にその運用ドキュメントが閲覧できるようになっており、より効率的な運用に役立っている。

4.3 関連テーブル/ファイルのチェックプログラム

依存関係図、および運用ドキュメント整備により、SMOKAの運用を行なうにあたっての運用手順は明確になった。しかし、依然として人為的ミスにより運用方法に誤りが生じる可能性は否定できない。そのため、これら依存関係図、運用ドキュメントに基づき、各種テーブルやファイル群の互いの依存関係を自動的にチェックするプログラムを開発・実装した。このプログラムの具体的な機能は以下の通りである。

(1) FILEMNG テーブルと実ファイルとの整合性チェック

ファイル管理テーブルFILEMNGにはSMOKAで使われる1つの観測データに付随する様々なファイルについて、その存在の有無、存在する場合はファイルのフルパス、ファイルサイズ、タイムスタンプが記録されている。これらのファイル群には依存関係図に示してあるように互いに依存関係がある(例えばHDIはFITSよりも新しくなければならぬ)。そこで、このような整合性を確かめるプログラムを開発した。このプログラムでは、まず、FILEMNGテーブルに定義されたこれらのファイルがテーブルに記載されているパスに実在するか、確認する。実在するなら、ファイルサイズやタイムスタンプがテーブル記載のものと一致しているかチェックし、相違が認められた場合には自動的に修復する。最後に、各ファイル群間のタイムスタンプを比較し、所定の順序が順守されているか検査する。

(2) FILEMNGテーブルとSMOKAテーブルの整合

性チェック

SMOKAの主検索を担っているのは、MASTARSテーブルのミラーXXX_MASTARSの一部カラムをコピーしたXXX_SMOKAテーブルである。このXXX_SMOKAテーブルと、ファイル管理テーブルFILEMNGの二つがSMOKAの要を担っており、そのテーブル間での整合性がとれていないと、検索にはかかるのにファイルがないなどの問題が生じる。そこで、これらテーブル間の相互チェックが必要となる。まず、FILEMNGテーブルとXXX_SMOKAテーブル中のFRAME_IDを比較し、互いに過不足ないか検査を行なう。過不足がなければ引続きDATE_OBSとMJDが互いに一致しているかどうか、検査する。

(3) FILEMNG テーブルと HeaderRecordCorr テーブルの整合性チェック

SMOKAでは、Suprime-Camデータに対して一次処理、位置較正、フラックス較正を施している。それらの較正処理情報はHeaderRecordCorrテーブルに記録されており、利用者からのデータ請求時にFITSヘッダにその記録を付加して提供している。また、検索時にデータ情報表示画面で利用者に閲覧させているFITSのヘッダ部分を切り取ったHDIファイルも、HeaderRecordCorrテーブルを参照して較正処理情報を付加して書き換えて表示している(詳しくは論文3)。データの品質によってはそれらの較正処理が行なえないものもあるため、一次処理、位置較正、フラックス較正のどれを行なったか、書き換えたHDIファイルが存在するか、をFILEMNGテーブルに登録して管理している。しかしながら、較正処理が行なわれてHeaderRecordCorrテーブルに記録されているにも拘らずFILEMNGテーブルに記録されていないフレームが存在する、といった不整合が、処理の手順の誤り等により生じることがある。そこで、FILEMNGテーブルに記録されている較正処理を行なったかどうか、書き換えたHDIファイルが存在するか否かの情報と、HeaderRecordCorrテーブルに記録されている情報の整合性のチェックを行なっている。

(4) FILEMNG テーブルと SUP_FF テーブルの整合性チェック

整約済みSuprime-Camデータの有無の情報はFILEMNGテーブルにおいてFF_FLAGというカラムで管理されている。FF_FLAGの項が"0"の時には整約済みデータの存在を示し、Web上の検索結果一覧表示画面において整約済みデータ請求が可能となる。一方、整約済みデータがどのように

して作成されたかについての詳細情報は SUP_FFテーブルに登録されて管理されており、その情報はWeb上で整約処理詳細情報画面において閲覧できるようになっている。もしこの2つの情報に不整合があると、整約済みデータの請求か詳細情報閲覧のどちらかが不可能になる。そこで、SUP_FFテーブルに登録されているにも拘らず FILEMNGテーブルに登録されていない場合や、その逆の場合の不整合のチェックを行なっている。もし、ある整約済みデータについてそのような不整合がある場合には、そのフレーム番号と不整合情報が出力され、どのような修正が必要であるかが分かるようになっている。

このような機能を持ったチェックプログラムにより、依存関係図・ドキュメントに沿った整合性チェックが可能となった。実際の運用ではcronにより月に一度自動的に起動され、そのエラーレポートはエラーログファイルとして書き出される。そのエラーログを基に該当データ・データベースの修復、もしくは削除作業を行なう。現在、全ての検査を行なうのに4時間程度の時間を要しているが、現状では運用に支障を来す程のものではない。しかし当該チェックプログラムの性格上、処理時間は処理すべきデータベースの総量にほぼ比例する。そのため、観測データの蓄積に伴いデータベース総量が現状より著しく増加した際には、省力化が求められる。また、現状ではチェック自体は自動化されているが、検出された不整合の修復はほとんど自動化されておらず、人力に頼らねばならない。検出した不整合の自動修復機能を如何に実装するかが今後の最大の課題となる。

4.4 サーバマシンの負荷分散

SMOKA のサーバマシン (Sun Enterprise 4500, UltraSPARC-II 400MHz 4CPU, memory 4GB) は、これまで以下の役割を担ってきた。

- ・データ検索に対するWeb サーバー。
- ・データ請求に対するFTPサーバー。
- ・SMOKAに関する全ての情報を納めたデータベースサーバー。
- ・新規データ入力作業, HDI, QLI, ATE ファイル作成作業の大部分。
- ・観測データサムネイル画像の作成作業。
- ・カレンダー検索用情報ファイル作成作業。
- ・ヘッダ情報全文検索インデックス作成作業。
- ・ヘッダ全キーワード検索用リスト作成作業。

- ・Suprime-Cam データの位置較正作業。
- ・Suprime-Cam データピンポイント検索用インデックス作成作業。
- ・環境データ表示用静止画・動画作成作業。
- ・アーカイブ状況表示図 (時系列図, 天球分布図, フレーム数一覧) の作成作業。
- ・データベースと各種データファイルの整合性チェック作業。

つまり、SMOKAに関するサービス、作業の大部分を担ってきた。しかし、高度検索機能の充実や、一次処理済みデータの提供、環境データの提供など SMOKAの機能の増強が進み、SMOKAの利用者がしだいに増えてくるにしたがって、現サーバマシン1台に多くの役割を担わせ、大きな負荷を与えることは、利用者の使い勝手の面でも作業の効率の面でも無理が生じてきた。また、サーバマシンに直接接続され、HDI, QLI, ATEファイルや環境データなどを収めている磁気ディスク領域 (約2TB) も満杯に近くなってきたが、サーバマシンのハードウェア面での制約からこれ以上の磁気ディスク追加ができず、他のファイルサーバーを用意せねばならない状況になってきていた。

強力なサーバマシンを新たに購入して入れ替えることが、移行の手間の面でも、維持管理の面でも望ましいのであるが、強力なサーバマシンはかなり高価であり、予算的に不可能であった。そこで複数の安価なマシンに機能を分散することとし、図11のように再構成を行なった。

現サーバマシンには、引き続きデータベースサーバーとWebサーバーの役割を担わせ、その他の機能 (作業) を他のマシンに分散するという方針を採った。これはデータベースマネージメントシステム (Sybase) のライセンスの関係もあるが、移行を段階的に徐々に進めていくためというのが最大の理由である。したがって、データベースの更新をとまなう機能 (作業) は現サーバマシンに留め、データベースを使わない、または、データベースの参照だけの機能 (作業) を他のマシンに移行することにした。各種ファイルや画像の作成作業は、SMOKAで用いているデータベース (Sybase) と同等のデータベース (PostgreSQL) を置いた新作業サーバー (SunFire V250 :UltraSPARC-IIIi 1064MHz 2CPU : memory 2GB) に担わせている (Sybase でないのはライセンス上の都合)。観測データ本体は、

SMOKA サーバーの負荷分散

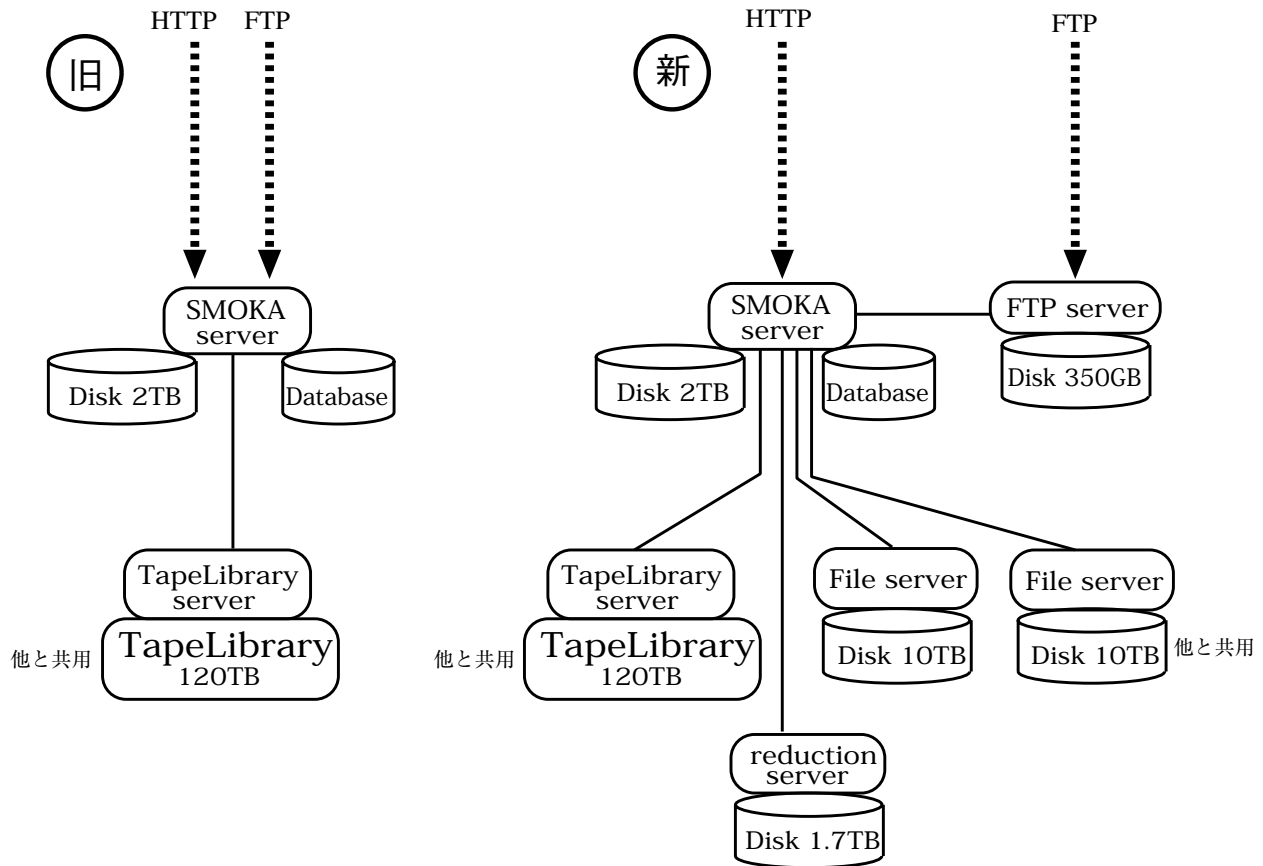


図 11. SMOKA サーバーの負荷分散. これまで全ての機能をほぼ 1 台のサーバーに担わせてきたが、複数台のサーバーに機能を分けて担わせ、負荷の分散をはかった。

当初より磁気テープライブラリ (PetaSite : 60TB x 2) に収められているが、利用者からの請求が多いSuprime-Camデータの2001年10月以降の分をファイルサーバーを設けてその磁気ディスク上に常駐させることとした。また、HDI, QLI, ATE ファイルや環境データなどの大半を別のファイルサーバーの磁気ディスクに移動し、NFS で参照するようにした。利用者からのデータ請求に応じるFTPサーバーも新たに用意し、利用者は新FTPサーバーにアクセスして観測データを取得するように変更した。これはセキュリティ強化の面でも有効であろう。

以上のようにSMOKAサーバーの機能を複数台のマシンに分散したが、どの機能をどのように分散するのが適当なのかについては、利用者のアクセス量・アクセス指向、運用作業の効率 (処理速度)、システムやプログラムの保守性をなどを考慮しつつ、今後運用を重ねつつ徐々に探っていくことになる。複数マシンへの機能分散は、設定や

管理に手間がかかるものの、小回りがきくという利点があり、今後の変更も比較的容易であり、最適な構成を目指して検討と実験を進めていく予定である。

5. 利用状況と運用面の課題

本章ではSMOKAの利用状況として、利用統計を5.1節で、SMOKAを利用した研究成果について5.2節で、そしてSMOKAの教育利用の現状について5.3節で述べる。5.4節ではSMOKAの運用課題として、ヘッダ情報不備の補填について論ずる。

5.1 SMOKA利用統計

2001年6月の公開開始以来、SMOKAの利用 (データ請求量, 登録利用者数, 検索数) は順調に伸びている。図12に1ヶ月毎の利用状況 (データ請求量) を示す。このデータ請求量の推移はWeb上でも公開されている¹³⁾。論文3でも述べたように、すばる望遠鏡の共同利用観測の観測データの公開が始まった2002年6月からデータ請求

量が急激に増加し、最近ではFITSファイルの月毎請求量は100GB以上、検索数は1500件前後となっている。以前は限られた利用者による請求が突出することも多かったが、最近は多くの利用者からのデータ請求が行なわれるようになり、SMOKAの利用が浸透してきたことを裏付けている。

SMOKAでは、検索やQLIの閲覧などを行なう限りは登録の必要はないが、観測データの取得は登録制にしており、そのための利用者登録は年度毎更新としている。この登録利用者数も年々増加しており、2001年度末には58名であったものが2004年度末では191名にまで増加した。

5.2 SMOKAによる研究成果

公開開始から4年近くが経過し、SMOKA利用者の増加とともにSMOKAで得られた研究成果も増えてきた。我々が把握している、SMOKAから取得した観測データを用いて執筆された査読論文は、2005年6月末現在13本に達し^{30) - 42)}、2003年以来一年あたり3～4本の論文が生産されるようになった。更に、2004年度にはSMOKAから得ら

れた観測データを主に使用した研究で学位を取得した利用者も生まれている⁵⁾。また、日本天文学会年会における発表も増えてきている。

今後も公開される観測データが増えるに従い、また現在は数が限られている整約済みデータが今後増えることで、更なる研究成果の増加が期待される。

5.3 SMOKA の教育利用

SMOKAのデータを請求する際に利用者登録が必要なことは5.1節でも述べたが、その際、氏名、所属等とともに、研究開発（太陽系、星、...）か教育普及（授業・講義、課外活動、...）といった使用目的の申告もあわせてお願いをしている（複数選択可）。申告された利用者の使用目的を見ると、2003年度は142名の登録利用者中53名（37%）、2004年度は191名中63名（33%）が教育普及を利用目的の1つにあげている。SMOKAデータが教育普及の面からも大きく期待されていることがこの数字からもわかる。

SMOKAのデータを活用した教育活動を展開し

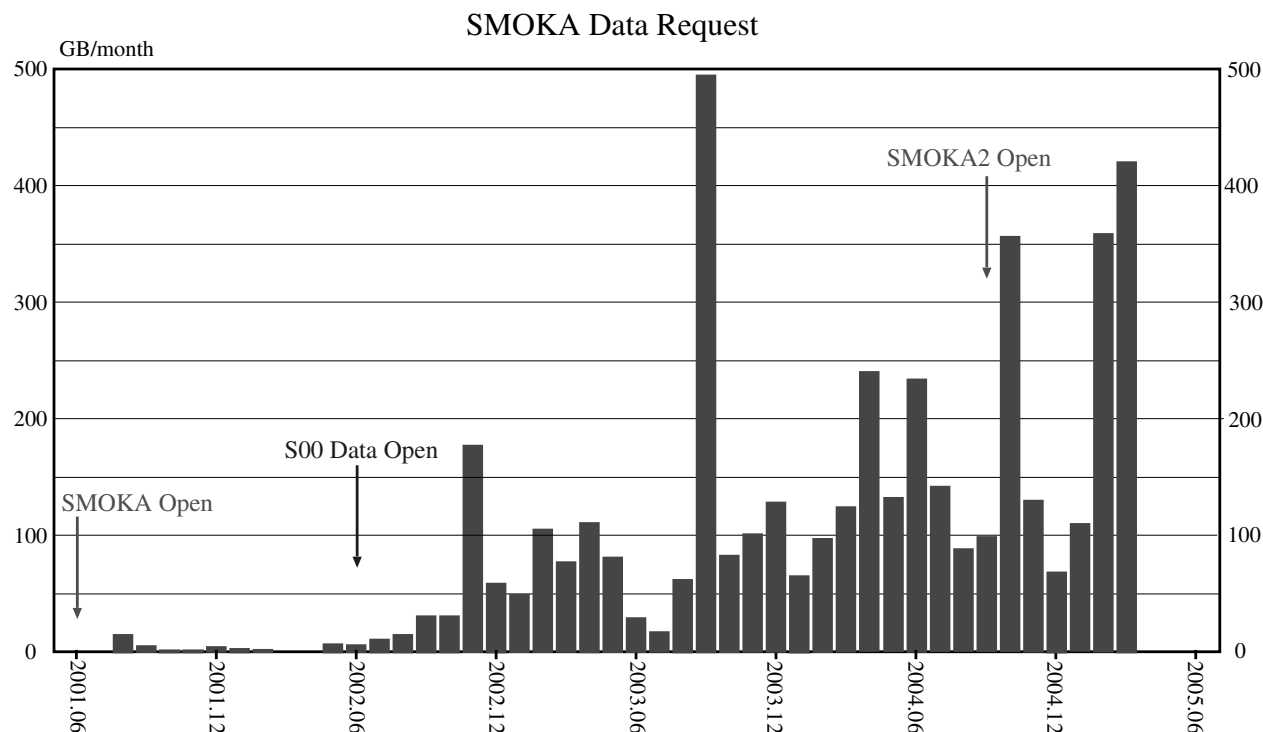


図 12. SMOKA 利用統計. SMOKA へのデータ請求量を月毎に示した図. 主な出来事を矢印で示した. S00 はすばる望遠鏡の最初の共同利用観測期間である. SMOKA2 は論文 3 で述べた整約済みデータの利用者への公開開始を示す.

*13 <http://smoka.nao.ac.jp/drstat.ja.jsp>

ている代表的なグループにPAOFITS WGがある。PAOFITS WGは、公開天文台ネットワーク (PAONET)⁴³⁾ の有志によって結成された、FITS 画像を教育に活用するためのワーキンググループである⁴⁴⁾。現在ではPAONET以外のメンバーも加わり、小・中・高校および大学の教員、国立天文台等の研究機関職員、科学館・博物館職員、公開天文台職員など、多彩なメンバーで活動を行なっている。

PAOFITS WGは、FITSデータを活用した6つの教材をこれまでに開発し、Webを通じて公開している^{*14}。最初に開発した「星団のHR図を作ろう」をテーマにした教材は、「球状星団編－星団の年齢を求めてみよう－」、「散開星団編－主系列星を見てみよう－」、「(球状+散開) 星団編－恒星の進化－」の3つの教材からなり、このうち散開星団のデータについては、木曾観測所で観測されたM38 (NGC1912) のデータをSMOKAから取得し、利用している。また、現在開発が進められている「距離はしご」をテーマにした教材では、「分光視差 (散開星団までの距離を求めよう)」、「ハッブルの法則」、「超新星の明るさと銀河の距離」の3教材の公開が始まっているが、そのうち「分光視差」教材では木曾観測所データが、「ハッブルの法則」教材では岡山天体物理観測所で観測された17銀河の分光データがSMOKAを通じて利用されている。

これらの教材は、上で述べたようにインターネットを通じて公開するだけでなく、利用講習会も開催され、その普及が図られている。講習会はこれまでに7回開催され、のべ97名が参加している。SMOKAデータの利用が多く、学校や博物館・科学館等の社会教育機関で始まりつつある。

5.4 ヘッダ情報不備の補填

観測中に望遠鏡・装置の不具合などにより観測時の望遠鏡や装置のシステムステータスを取り込めない事態が生じることがある。また、望遠鏡・装置制御用ソフトウェアのバグなどにより、設計上は取得されるべきステータス情報が正しく取得できない、もしくはFITSヘッダに書き込まれない場合がある。このような場合、観測データのFITSヘッダには正しいステータス情報が反映されない。このとき、観測時の望遠鏡や装置の状況

を観測データのヘッダから正しく抜き出すことができなくなるので、アーカイブシステムにおけるデータ利用に対して致命的な影響を及ぼすことになる。

我々は各観測装置のFITSヘッダ情報について詳細に調べ、特にFOCASについて、FITSヘッダ情報に不備が多いことを確認した。FOCASは撮像、ロングスリット分光、多天体分光や偏光観測を可能にする多モード装置であり、様々な点で再利用価値の高いデータを生産する可能性がある。このような装置のヘッダ情報の不備は、アーカイブを用いた天文学の発展の大きな障害となるだけでなく、誤った再利用で誤った結果を出してしまう可能性も生じるので、ヘッダ情報の不備を補填することは必要不可欠な作業である。

そこで我々はFOCASの全データについてFITSヘッダ情報の不備な部分をチェックし、必須キーワードが存在しない観測データや、明らかにキーワードの値が間違っているものを抜き出した。その結果、天文学を遂行する上で必須の情報である時刻、座標、および波長の情報に明らかに誤りであるものが多数あることが分かった。複数露出のデータが連続的に取得されているにもかかわらず、取得時刻が進んでいないもの、座標値が明らかに誤っているもの (赤経の値が "12:34:56.78"), 波長の情報が欠落しているものなどは、データベースに登録されているヘッダ情報から簡単に同定できた。一方、あり得ない値が入っている場合ではなく、装置としてはあり得る値であるが、その値が正しいかどうかの判定が難しい場合は、観測時に観測者もしくはサポートアストロノマーが書いた紙の観測ログにかかれていた情報と比較することが必要である。我々はFOCAS担当のサポートアストロノマーの協力を得て、紙の観測ログのほぼ全てを入手し、基本的な天文学遂行のために重要な、時刻の変化・波長情報などのチェックを開始している。チェックによって明らかになった誤情報は、観測装置の検索テーブルXXX_SMOKA上で修正され、誤情報修正テーブルHeaderRecordCorrに反映される。これによって、利用者は正しい情報に基づいてデータ検索・取得・解析を行なうことができるようになる。

この作業は今後もデータが生産され続ける以上

*14 <http://paofits.dc.nao.ac.jp/Materials/>

ずっと続くわけではあるが、問題となった情報の不備がなぜ起こったかを明確にし、再発を防止するためにサポートアストロノマーや装置開発者と協力していくことが極めて重要である。

6. 今後の課題

SMOKAではアーカイブデータを利用した天文学研究・教育を行なう上での有益性や利便性を向上するため、様々な機能を追加、補強してきた。しかしながら、利用者からの要望全てに答えられているわけではない。この章ではSMOKAで解決すべき主要な技術的課題について述べる。

6.1 カタログ化とそれを用いた天文学

論文3でも述べたように、SMOKAではSuprime-Camデータの一部について、整約済みデータの提供を行なっている。この整約済みデータから天体を抽出し、その座標値や等級、色指数などの物理量を抽出してカタログ化し公開することで、利用者がより多くの研究成果を挙げるのが期待できる。そのカタログの存在・利用価値を高め、より広く利用してもらう上で、我々はカタログ作成だけでなく、そのカタログを使った天文学的成果を挙げることも必要である。

カタログを用いた天文学の例として、大量のアーカイブデータを用いた統計的研究が挙げられる。そのような研究の際、天体検出などを一から作業しなければならないのは敷居が高く、天文学的成果に達する道のりは険しい。しかし予め、信頼に足る方法でカタログ化を行なっておけば、そのデータを利用するだけで比較的簡単に統計処理を行なうことができる。また、SMOKAでは現在も天体名での検索が可能であるが、それは観測者が自由に設定できるOBJECTキーワードを基に検索しているだけであり、視野内に別の天体が写っていてもその天体名だけでは検索できない。しかしカタログ化により、どのような天体が既にアーカイブデータの中にあるのか、座標の一致性から明確にすることができ、アーカイブ利用者にとって有益な情報となる。

6.2 検索機能のさらなる強化

強化すべき検索機能として論文3で述べた課題のうち、FITSヘッダの全キーワード検索機能、全文検索機能は開発・実装した(2.2節)。しかし、論文3で述べた重複領域検索機能については、実装できていない。この機能により、3色以上の観

測が行なわれている領域、積算積分時間が一定時間を越えている領域などを検索することができる。前者は天体の色指数を求めたい場合に有効であり、後者は位置や明るさの変化する天体を調べたり、深い撮像データを得たりする場合に有効である。この検索機能の実現のためには、多色で観測されている領域や積算積分時間が一定時間を越えている領域を抜き出し、それをデータベース化する必要がある。

カタログ連携検索機能もまだ実現されていない。この機能は、既存のカタログを基に、例えば固有運動が毎年5秒角以上の天体が写っているフレームを検索できるようにすることを目指す。また、先に述べたSMOKAのアーカイブデータを利用して作成したカタログとも連携を行ない、提供したカタログ天体を含む観測データを容易に取得することも目指す。さらに、2.1節で述べたピンポイント検索を用いることで、早見画像上、検索した座標値に対応する部分に矢印をつけることも可能であり、これもどこに対象天体が写っているか知る上で必要な機能である。

6.3 アーカイブデータに対する付加情報提供機能の拡張

(1) すばる望遠鏡オペレータ night log の活用

すばる望遠鏡では山頂オペレータが観測毎にnight logと呼ばれる観測記録をつけている。この観測記録には、天候の変化や天候不良に伴うドームのメインスリットの開閉といった作業情報が記載されている。この情報により、当日、観測に異常があったかどうかアーカイブデータの利用者にも分かる。そこで、観測データの取得や校正に役立つ環境データの一環として、SMOKAでnight logを公開する予定である。実際の公開は観測データと同様に観測日から1年半後である。すばる望遠鏡のnight logの公開開始は2004年11月の観測からであるので、SMOKAでの公開はその1年半後2006年5月から行なう計画である。

(2) すばる望遠鏡の新スカイモニターの活用

すばる望遠鏡のトップリングにはILCCDスカイモニターが設置されており、すばる望遠鏡の向いている天域の高感度な映像が取得されている。それにより、雲の到来や透明度の変化などの気象状態の変化を克明にとらえることができる。しかし、このスカイモニターの映像は山頂の観測室にあるモニターに映し出されているだけで、山麓や三鷹のリモート観測室の観測者はその映像を見る

ことができない。また、この映像は保存されていないため、観測者がその観測日の空の状態を後で確認することはできない。しかし、この I.I.CCD スカイモニターの映像を保存し公開することは、より良い観測データ較正に有益である。特に、その日の空の状態を直接知り得ないアーカイブシステムの利用者にとっては貴重な記録となり得る。そこで我々は、ハワイ観測所と連携し、現在特に整備されていない I.I.CCD スカイモニターの映像を記録し、アーカイブ化することを目指している。このデータと同じ晩の Suprime-Cam 観測の測光零点との比較から、利用者はより正確にその晩のデータの安定性・信頼性を確認できるであろう。

6.4 岡山天体物理観測所新観測装置データのアーカイブ化

岡山天体物理観測所では新たな観測装置が製作されている。188cm 望遠鏡では ISLE と呼ばれる J, H, K バンドの撮像、および分光のできる共同利用観測装置が、また 91cm 望遠鏡では 1 度の視野を持ち、Y, J, H, K バンドでの撮像を可能とする Wide Field カメラの開発が行なわれている。共に本格稼働は 2006 年以降であり、2 年間の占有期間があるため、実際に SMOKA で公開されるのは 2008 年以降である。

このような観測装置の観測データをデータアーカイブとして公開する上で、観測データの FITS ヘッダは観測時の状況の分からないアーカイブ利用者にとって非常に重要である。したがって、新観測装置データのアーカイブ化にあたっては、試験段階から観測装置グループと連携し、不備のない FITS ヘッダを作ることが望ましい。このことは論文 3 で述べたすばる望遠鏡の次期共同利用観測装置、MOIRCS や FMOS においても同様である。

7. まとめ

我々は、すばる望遠鏡、岡山天体物理観測所 188cm 望遠鏡、および東京大学木曾観測所 105cm シュミット望遠鏡によって取得された観測データを公開するアーカイブシステム SMOKA の運用を行なっている。また運用だけでなく、SMOKA の利用者がより効率的に観測データを取得し、天文学的成果を得られるよう、様々な開発・改良も行なってきた。論文 3 においては、利用者が観測データを解析する上で必要な情報をより効率的に得ることに重点がおかれたが、本論文では膨大にな

った観測データから目的の観測データを効率良く取得できるよう、検索機能の強化を重点的に行なった。まず、Suprime-Cam 専用検索機能として、各 CCD チップの視野情報を基にしたピンポイント検索機能、および品質評価指標による検索機能を開発した。また、より細かな検索を行なう手段として、FITS ヘッダの全キーワード検索、および全文検索機能を実現した。さらに、すばる望遠鏡の観測装置にしか対応していなかったカレンダー検索機能（旧称：プロポーザルカレンダー表示機能）を岡山天体物理観測所、木曾観測所の観測装置にも対応させた。検索機能以外では、Suprime-Cam のモザイク化した画像データを公開したり、今までは一つの生データに対し一種類の整約済みデータしか公開できなかったのを、異なる整約手法を用いた複数の整約済みデータを公開できるよう改良した。加えて、CDS の Vizier への Suprime-Cam データの情報の提供、COMICS の COMQ データの公開開始、太陽系内天体軌道要素データベースの自動更新化を行なった。また運用面としては、整約済みデータの公開開始以来、煩雑化した運用手順の詳細なドキュメント作成、それに基いた各種データベースやファイルの整合性チェックプログラムの開発を行なった。これにより運用の大幅な効率化、安定化および継続性が実現され、今後の SMOKA システムの運用の発展に見通しがつけられた。

SMOKA の利用者は年々増加しており、アーカイブデータを用いた研究成果も着実に増えて来た。また、天文学教育の目的での利用も進みつつある。しかし、6 章で述べたように SMOKA にはまだ解決すべき運用・開発課題が残っている。我々はこれらの課題の解決を目指し、今後も引き続き開発を進めていく次第である。

謝辞

本研究は、天文情報処理研究会のワーキンググループ活動の一環として行なわれた。開発作業にあたっては、天文学データ解析研究センター DB/DA プロジェクトの計算機資源を活用した。国立天文台の西村史朗氏、一橋大学社会学部の中嶋浩一氏、ハワイ観測所の能丸淳一、小杉城治、高遠徳尚、青木賢太郎の各氏、東京大学宇宙線研究所の安田直樹氏、宇宙航空研究開発機構の大山陽一、多賀正敏の各氏、早稲田大学教育学部の馬場肇氏の助力に感謝する。また、SMOKA の開発・運用を影に日向に支援して下さっている、ハ

ワイ観測所, 岡山天体物理観測所, 木曾観測所のスタッフの方々にも感謝する。有益な助言を下された匿名の査読者にも感謝したい。

本アーカイブ運用には, 仏CDSによって運用されている SIMBADデータベース, および米 JPLによって運用されているNEDデータベースを利用している。また, 移動天体検索に必要な軌道要素データベースを作成するにあたっては, 米MPC から提供されているMPCORBを利用している。有用な各種フリーソフトウェアを提供している開発者にも感謝したい。

本研究は, 国立天文台天文学データ解析計算センタープロジェクト経費の援助を得て行なわれた。

参考文献

- 1) 馬場肇, 安田直樹, 市川伸一, 八木雅文, 岩本信之, 高田唯史, 洞口俊博, 多賀正敏, 渡邊大, 奥村真一郎, 小澤友彦, 山本直孝, 濱部勝: すばる望遠鏡公開データアーカイブシステムの開発, 国立天文台報, **6**, 23–36 (2002).
- 2) 山本直孝, 野田祥代, 多賀正敏, 小澤友彦, 洞口俊博, 奥村真一郎, 古荘玲子, 馬場肇, 八木雅文, 安田直樹, 高田唯史, 市川伸一: すばる望遠鏡公開データアーカイブシステムの開発2, 国立天文台報, **6**, 79–100 (2003).
- 3) 榎基宏, 多賀正敏, 小澤友彦, 野田祥代, 奥村真一郎, 吉野彰, 古荘玲子, 馬場肇, 洞口俊博, 高田唯史, 市川伸一: すばる望遠鏡公開データアーカイブシステムの開発3, 国立天文台報, **7**, 57–84 (2004).
- 4) S. Miyazaki, Y. Komiyama, M. Sekiguchi, S. Okamura, M. Doi, H. Furusawa, M. Hamabe, K. Imi, M. Kimura, F. Nakata, N. Okada, M. Ouchi, K. Shimasaku, M. Yagi, and N. Yasuda: Subaru Prime Focus Camera – Suprime-Cam, *PASJ*, **54**, 833–953 (2002).
- 5) A. Yoshino: Observational Study of Bulge Evolution in Nearby Spiral Galaxies, Doctoral thesis, Tohoku University (2005).
- 6) H. Kataza, Y. Okamoto, S. Takubo, T. Onaka, S. Sako, K. Nakamura, T. Miyata, and T. Yamashita: COMICS: the cooled mid-infrared camera and spectrometer for the Subaru Telescope, *Optical and IR Telescope Instrumentation and Detectors*, ed. M. Iye and A. F. Moorwood, *Proc. SPIE*, **4008**, 1144–1152 (2000).
- 7) D. G. Monet, S. E. Levine, B. Canzian, H. D. Ables, A. R. Bird, C. C. Dahn, H. H. Guetter, H. C. Harris, A. A. Henden, S. K. Leggett et al.: The USNO-B Catalog, *AJ*, **125**, 984–993 (2003).
- 8) D. G. York, J. Adelman, J. E. Anderson, Jr., S. F. Anderson, J. Annis, N. A. Bahcall, J. A. Bakken, R. Barkhouser, S. Bastian, E. Berman et al.: The Sloan Digital Sky Survey: Technical Summary, *AJ*, **120**, 1579–1587 (2000).
- 9) P. Z. Kunszt, A. S. Szalay, I. Csabai, A. R. Thakar: The Indexing of the SDSS Science Archive, ADASS IX, eds. N. Manset, C. Veillet, D. Crabtree, *ASP Conference Proceedings Series*, **216**, 141–144 (2000).
- 10) T. Kodama, T. Yamada, M. Akiyama, K. Aoki, M. Doi, H. Furusawa, T. Fuse, M. Imanishi, C. Ishida, M. Iye et al.: Down-sizing in galaxy formation at $z \sim 1$ in the Subaru/XMM-Newton Deep Survey (SXDS), *MNRAS*, **350**, 1005 (2004).
- 11) 仲田史明, 出田誠, 八木雅文, 榎基宏, 吉野彰, 山田善彦, 高田唯史, 市川伸一: すばる望遠鏡観測データ品質評価システムの開発, 国立天文台報, 印刷中
- 12) T. Takata, M. Yagi, N. Yasuda, R. Ogasawara: Distributed data archive systems for Subaru Telescope, ed. P. J. Quinn, *Proc. SPIE*, **4844**, 242–249 (2002).
- 13) K. Noguchi, W. Aoki, S. Kawanomoto, H. Ando, S. Honda, H. Izumiura, E. Kambe, K. Okita, K. Sadakane, B. Sato, A. Tajitsu, T. Takada-Hidai, W. Tanaka, E. Watanabe, and M. Yoshida: High Dispersion Spectrograph (HDS) for the Subaru Telescope, *PASJ*, **54**, 855–864 (2002).
- 14) D. Tody: The IRAF Data Reduction and Analysis System, in *Instrumentation in Astronomy VI*, ed. D.L. Crawford, *Proc. SPIE*, **627**, 733–756 (1986).
- 15) W. A. Joye, E. Mandel: New Features of SAOImage DS9, ADASS XII, ed. H. E. Payne, R. I. Jedrzejewski, and R. N. Hook, *ASP*

- Conference Proceedings Series*, **295**, 489–492 (2003).
- 16) N. Kashikawa, K. Aoki, R. Asai, N. Ebizuka, M. Inata, M. Iye, K. S. Kawabata, G. Kosugi, Y. Ohyama, K. Okita, T. Ozawa, Y. Saito, T. Sasaki, K. Sekiguchi, Y. Shimizu, H. Taguchi, T. Takata, Y. Yadoumaru, and M. Yoshida: FOCAS: The Faint Object Camera and Spectrograph for the Subaru Telescope, *PASJ*, **54**, 819–832 (2002).
 - 17) K. Noguchi, W. Aoki, S. Kawanomoto, H. Ando, S. Honda, H. Izumiura, E. Kambe, K. Okita, K. Sadakane, B. Sato, A. Tajitsu, T. Takada-Hidai, W. Tanaka, E. Watanabe, and M. Yoshida: High Dispersion Spectrograph (HDS) for the Subaru Telescope, *PASJ*, **54**, 855–864 (2002).
 - 18) F. Iwamuro, K. Motohara, T. Maihara, R. Hata, and T. Harashima: OHS: OH-Airglow Suppressor for the Subaru Telescope, *PASJ*, **53**, 355–360 (2001).
 - 19) K. Motohara, F. Iwamuro, T. Maihara, S. Oya, H. Tsukamoto, M. Imanishi, H. Terada, M. Goto, J. Iwai, H. Tanabe, R. Hata, T. Taguchi, and T. Harashima: CISCO: Cooled Infrared Spectrograph and Camera for OHS on the Subaru Telescope, *PASJ*, **54**, 315–325 (2002).
 - 20) M. Tamura, H. Suto, Y. Itoh, N. Ebizuka, Y. Doi, K. Murakawa, S. S. Hayashi, Y. Oasa, H. Takami, and N. Kaifu: Coronagraph imager with adaptive optics (CIAO): description and first results, *Optical and IR Telescope Instrumentation and Detectors*, ed. M. Iye and A. F. Moorwood, *Proc. SPIE*, **4008**, 1153–1161 (2000).
 - 21) N. Kobayashi, A. T. Tokunaga, H. Terada, M. Goto, M. Weber, R. Potter, P. M. Onaka, G. K. Ching, T. T. Young, K. Fletcher, D. Neil, L. Robertson, D. Cook, M. Imanishi, and D. W. Warren: IRCS: Infrared Camera and Spectrograph for the Subaru Telescope, *Optical and IR Telescope Instrumentation and Detectors*, ed. M. Iye and A. F. Moorwood, *Proc. SPIE*, **2008**, 1056–1066 (2000).
 - 22) D. Tomono and T. Nishimura: Mid-infrared test observation system (MIRTOS) for SUB-ARU, *Infrared Astronomical Instrumentation*, ed. A. M. Fowler, *Proc. SPIE*, **3354**, 1150–1157 (1998).
 - 23) G. Kosugi, H. Ohtani, T. Sasaki, H. Koyano, Y. Shimizu, M. Yoshida, M. Sasaki, K. Aoki, and A. Baba: Spectro-nebula graph: A Tridimensional Spectroscopic System Based on a Local Area Network of Personal Computers, *PASP*, **107**, pp.474–482 (1995).
 - 24) S. Okumura, E. Nishihara, E. Watanabe, A. Mori, H. Kataza, and T. Yamashita: OASIS: A Multi-Purpose Near-Infrared Camera and Spectrograph, *PASJ*, **52**, 931–942 (2000).
 - 25) H. Izumiura: HIDES: a High Dispersion Echelle Spectrograph, *Observational Astrophysics in Asia and its Future*, 4th East Asian Meeting on Astronomy (4th EAMA) - held 3-10 Feb., 1999 in Kunming, China, ed. P. S. Chen., Yunnan Observatory, Chinese Academy of Sciences, **77** (1999).
 - 26) S. Yoshida, T. Aoki, T. Soyano, K. Tarusawa, M. Sekiguchi, M. Doi, N. Kashikawa, S. Okamura, K. Shimasaku, M. Yagi, and N. Yasuda: CCD Cameras for the Kiso 105 cm Schmidt Telescope, *Future Utilisation of Schmidt Telescope*, ed. J. Chapman, R. Cannon, S. Harrison and B. Hidayat, *Proc. of IAU Coll.148*, 33–37 (1994).
 - 27) N. Itoh, T. Soyano, K. Tarusawa, T. Aoki, S. Yoshida, T. Hasegawa, Y. Yadoumaru, Y. Nakada, and S. Miyazaki: A Very Wide-Field CCD Camera for Kiso Schmidt Telescope, *Publ. Natl. Astron. Obs. Japan*, **6**, 41–48 (2001).
 - 28) R. Ogasawara, G. Kosugi, T. Takata, J. Noumaru, T. Sasaki, Y. Chikada, Y. Mizumoto, M. Yagi, N. Yasuda, S. Ichikawa, M. Yoshida, K. Kawarai, Y. Ishihara, and A. Kawai: The Subaru Telescope Software Trinity System, *ADASS XII*, ed. H. E. Payne, R. I. Jedrzejewski, and R. N. Hook, *ASP Conference Proceedings Series*, **295**, 142–151 (2003).
 - 29) 片坐宏一: 地上からの中間赤外線観測の扉を開くCOMICS, *天文月報*, **97**, 210–215 (2004).
 - 30) S. C. Chapman, A. J. Barger, L.L. Cowie, D.Scott, C. Borys, P. Capak, E. B. Fomalont,

- G. F. Lewis, E. A. Richards, A. T. Steffen, G. Wilson, and M. Yun: The Properties of Microjansky Radio Sources in the Hubble Deep Field-North, SSA 13, and SSA 22 Fields, *ApJ* **585**, pp.57–66 (2003).
- 31) A. J. Barger, L. L. Cowie, P. Capak, D. M. Alexander, F. E. Bauer, E. Fernandez, W. N. Brandt, G. P. Garmire, and A. E. Hornschemeier: Optical and Infrared Properties of the 2 Ms Chandra Deep Field North X-Ray Sources, *AJ* **126**, 632–665 (2003).
- 32) T. Takata, N. Kashikawa, K. Nakanishi, K. Aoki, R. Asai, N. Ebizuka, M. Inaka, M. Iye, K. S. Kawabata, G. Kosugi, et al.: Dusty ERO Search behind Two Massive Clusters, *PASJ* **55**, 789–800 (2003).
- 33) P. Capak, L. L. Cowie, E. M. Hu, A. L. Barger, M. Dickinson, E. Fernandez, M. Giavalisco, Y. Komiyama, C. Kretchmer, C. McNally, S. Miyazaki, S. Okamura, and D. Stern: A Deep Wide-Field, Optical, and Near-Infrared Catalog of a Large Area around the Hubble Deep Field North, *AJ* **127**, 180–198 (2004).
- 34) E. M. Hu, L. L. Cowie, P. Capak, R. G. McMahon, T. Hayashino, and Y. Komiyama: The Luminosity Function of Ly alpha Emitters at Redshift $z \sim 5.7$, *AJ* **127**, 563–575 (2004).
- 35) A. Steffen, A. J. Barger, P. Capak, L. L. Cowie, R. F. Mushotzky, and Y. Yang: An Optical Catalog of the Chandra Large Area Synoptic X-ray Survey Sources, *AJ* **128**, 1483–1500 (2004).
- 36) I. Smail, S. C. Chapman, A. W. Blain, and R. J. Ivison: The Rest-Frame Optical Properties of SCUBA Galaxies, *ApJ* **616**, 71–85 (2004).
- 37) E. Egami, H. Dole, J.-S. Huang, P. Perez-Gonzalez, E. Le Floc'h, C. Papovich, P. Barmby, R.J. Ivison, S. Serjeant, A. Mortier, D.T. Frayer, D. Rigopoulou, G. Lagache, G.H. Rieke, S.P. Willner, A. Alonso-Herrero, L. Bai, C.W. Engelbracht, G.G. Fazio, K.D. Gordon, D.C. Hines, K.A. Misselt, S. Miyazaki, J.E. Morrison, M.J. Rieke, J.R. Rigby, G. Wilson: Spitzer Observations of the SCUBA/VLA Sources in the Lockman Hole: Star Formation History of Infrared-Luminous Galaxies, *ApJS*, **154**, 130–136 (2004).
- 38) C. Motch, K. Sekiguchi, F. Haberl, V.E. Zavlin, A. Schwobe, M.W. Pakull: The proper motion of the isolated neutron star RX J1605.3+3249, *A&A* **429**, 257–265 (2005).
- 39) T. Broadhurst, M. Takada, K. Umetsu, X. Kong, N. Arimoto, M. Chiba, and T. Futamase: The Surprisingly Steep Mass Profile of A1689, from a Lensing Analysis of Subaru Images, *ApJ* **619**, L143–L146 (2005).
- 40) S.C. Chapman, A.W. Blain, I. Smail, R.J. Ivison: A Redshift Survey of the Submillimeter Galaxy Population, *ApJ* **622**, 772–796 (2005)
- 41) R. Neuhaeuser, E.W. Guenther, G. Wuchterl, M. Mugrauer, A. Bedalov, and P.H. Hauschildt: Evidence for a co-moving sub-stellar companion of GQ Lup, *A&A* **435**, L13–L16 (2005).
- 42) A. Yoshino, and T. Ichikawa: Colors and Mass-to-Light Ratios of Bulges and Disks of Nearby Spiral Galaxies, *PASJ* submitted (2005).
- 43) J. Watanabe: PAONET: A Network for Sharing Newest Astronomical Results with Public Observatories and with the General Public in Japan, in *Astronomical Education with the Internet Proc. of the Misato International Symposium 1998 on Astronomical Education with the Internet*, Frontiers Science Series No. 27, eds. M. Okyudo, T. Ebisuzaki and M. Nakayama (Universal Academy Press), 113–116 (1998)
- 44) 古荘玲子, 原正, 洞口俊博, PAOFITS WG: リアルデータを教室に - 公開天文台ネットワーク PAOFITS WG の活動 - 天文月報, **97**, 149–155 (2004).