

HSC-SSP データリリースへの道 (2)

データ公開とカタログデータベース

古澤久徳¹・高田唯史²

〈国立天文台 〒181-8588 東京都三鷹市大沢 2-21-1〉

e-mail: ¹furusawa.hisanori@nao.ac.jp, ²tadafumi.takata@nao.ac.jp

大規模サーベイプロジェクトにおいて、科学データアーカイブを通じた較正済みデータの公開は、観測データの価値を十分に引き出して科学的成果を促すうえで重要な要素です。特にカタログデータベースの構築は、研究者が自身の研究に必要な情報を適切な時間で取り出して科学研究を進めるために、欠かせない開発事項となっています。私たち HSC チームは、HSC 戦略枠観測のデータを処理し、科学や教育に利用できる形で一般公開しました。本稿ではカタログデータベースの開発意義を説明するとともに、第一回全世界向けデータリリースの内容についてご紹介したいと思います。

1. 科学データアーカイブとカタログデータベース

前稿 (HSC-SSP データリリースへの道 (1) データ解析) では、大規模サーベイ時代の観測プロジェクトでデータ解析が果たすべき役割を確認し、HSC 戦略枠観測プログラム (SSP) のデータ解析をご紹介しました。私たち HSC チームは、こうしたデータ解析を経て、2017年2月には最初の全世界向けデータリリース (Public Data Release 1: PDR1)¹⁾ を実施しました。現在までに延べ 800 名を超えるユーザー登録があり、このデータをもとにさまざまな科学成果が得られています。

しかし、HSC-SSP の膨大で複雑な処理済みデータ (データプロダクト) から、各研究者が自身のサイエンスに必要な条件に合致した情報だけを「迅速に」かつ「過不足なく」取り出すことは容易ではありません。HSC-SSP データプロダクトの科学的な利用価値を高めるうえで、科学研究に直ちに使える処理済みデータ等を保存・管理し

提供するためのカタログデータベースを中心としたシステム (科学データアーカイブと呼ぶ) は欠かせません²⁾。私たちは HSC プロジェクト立ち上げ当初からこの考えのもと、データ解析ソフトウェアと歩調を合わせて開発を進め、科学データアーカイブを構築・運用してきました。

本稿では、HSC-SSP のカタログデータベース開発の背景や開発を通じたチームの経験を振り返るとともに、PDR1 の内容をご紹介したいと思います。

2. カatalogデータベースの意義

データ解析ソフトウェアを実行すると、その出力として画像データファイルと画像から自動検出された各天体に対する測定情報 (具体的には位置、明るさ、形など) をまとめたカタログ情報を保持したファイルが得られます。5 年間の戦略枠観測を計画通り遂行した場合に得られる各画像やカタログに載ってくる天体の数を推定したものを表 1 に示しています。ここからわかるように、画

表1 HSC-SSP データリリースの大きさ。サーベイ最終時 (Final) の値は過去のデータリリースに基づいた推定値。階層はW=ワイド、D=ディープ、UD=ウルトラディープ、またDBはデータベースを指す。(国立天文台ハワイ観測所の山田善彦氏と林祐輔氏作成。)

リリース名	階層	CCD 画像数	CCD 総天体数	CCD 画像容量 (TB)	CCD DB 容量 (TB)	Coadd 画像数	Coadd 総天体数	Coadd 画像容量 (TB)	Coadd DB 容量 (GB)
PDR1	W	34万	7.7億	19	N/A	2万	5,266万	2.3	912
	D	7万	1.5億	4	N/A	0.8万	1,596万	0.8	337
	UD	4万	8,556万	2	N/A	0.2万	323万	0.2	71
Final	W	245万	300億	192	180	15万	5億	25	10,000
	D	35万	50億	27	30	1万	2,000万	1	400
	UD	20万	30億	16	18	0.2万	400万	0.2	100

像自体も容量が多い(たとえば各露出に対する CCD画像で約235TB, すべての露出を足し上げた Coadd画像で約26TB)ですが, ファイル数も途方もない数(それぞれ300万, 16万個)になっています。また, カタログされる天体数とそのパラメータ数の多さも今までには経験したことのない規模であり, 天体数は観測フィルター(バンド)ごとに約10億天体くらいで, 同じ天体が何回も観測されることを考えると, 延べ天体数は400億天体程度になることが容易に予想されていました。特に, 私たちの解析パイプラインは, 同じ観測領域に対する複数回の観測画像を足し上げた, 深い画像データに写る明暗すべての天体について, 各露出のデータに戻って測定するモードをもっています(この解析モードによる測定結果はPDR1には含まれていません)。これを私たちは Forced CCD測光と呼んでいます。各天体に対するパラメータの数が非常に多い(〜3,000個)ため, 縦にも横にも長いカタログデータができあがってきます。これをテキストファイルでユーザーに配布し, あとは各自でどうにかして, というわけにはいかない量でした。

そのため, ユーザー各自で必要な情報だけを可能な限り高速で取り出せるようなデータベースの導入が必須だったのです。欲しい情報の取り出し

方を決める場合に, 検索条件を定めるわけですが, それは取り出したい情報に依存, つまりは科学的目的に依存するものであり, ユーザー各自の自由な発想が生かされるべきところなのです。

逆に言えば, 検索の方法を限定するのはなかなか難しいことも意味しているので, その目的を統一的な文法で記述できるデータベース操作作用言語である「SQL」を利用できるリレーショナルデータベース^{*1}がこのサービスに適していることに疑いはありませんでした。それを後押ししたのが, 解析パイプラインの出力がリレーショナルデータベースを前提としたものとなっていたことです。リレーショナルデータベースでは, 画像ファイルや各天体の唯一無二性を保証する必要がありますが, 解析パイプラインにはそのための番号づけのアルゴリズムが導入されていました。したがって, まずは“パイプラインの言うがまま”にデータを保存管理するようにデータベース構造を設計すれば良いという予測がついた面はありました。

しかしそうした開発の糸口はあったものの, そもそも天体のカタログをデータベースに入れて配布するということが自体が, 日本においてはあまり例のないものでした。それまでにも, あかり衛星の FIS (Far-Infrared Surveyor) によるサーベイ

*1 関係モデルに基いたデータベースの総称で, 行と列からなる複数の表が特定の情報(カラム)で関連づけられる。

観測に対するデータベース³⁾がありましたが、私たちが開発しようとしているものは、想定されるレコード数(行数)だけでもその10-100倍に達するような規模であり、HSC-SSPのカタログデータベース開発がかなり挑戦的な計画であることは明らかでした。

3. カタログデータベースの実現に向けて

前述のように、HSC-SSPの科学データアーカイブが扱うプロダクトには画像や天体カタログを取めたファイル群と天体カタログの中身の情報があります。そこで、これらを管理するために、私たちは大きく分けて2種類のデータベースを用意する必要がありました。

一つは各画像やカタログのためのファイル(そのほとんどはFITSファイル)をユーザーが高速に探し出すために必要なパラメータを、ファイルごとに1レコード(1行)に記述した、メタデータ表と呼ばれる種類のものです。パラメータとしては、その画像の特徴を示すもの、具体的には、画像がどの天域を撮像したものかを示す天球上での位置情報や、取得時刻や露出時間、フィルター名など、その観測の目的に即して記録されたものや、画像処理の結果得られた情報、たとえば、測光ゼロ点やシーイングサイズ、画像上の各位置における点光源の画像上での広がり(Point Spread Function; PSF)を示すパラメータ群とそれらの誤差評価値などです。ユーザーはこれらの中から必要なパラメータを選び出し、目的とする画像を探し出すことができるようになります。

もう一つは、天体カタログの中身をデータベースに展開した、カタログデータ表で、画像ファイル上で検出された各天体について、その位置や形、明るさおよびそれらの測定誤差を記録したものです。さらにここには、解析処理において各天

体に付加されたフラッグと呼ばれるものも記録されています。このフラッグは、その天体の測定値を使用するうえで何らかの注意を必要とするかどうかを示しているものが多く、このフラッグをうまく使うことが、目的の天体リストを得るうえで必須であることを注意喚起しておきたいと思います。

もちろん、データベースだけではなく、ユーザーをうまく誘導するためのユーザーインターフェース(UI)の充実も必要でした。逆に言えばUIとの連動性も考えながら、データベースの設計を進める必要があります。特に、ユーザー認証によるアクセス制限の実装はたいへん細かいものとなりました。画像や天体情報へのアクセスを制御するためには、各ユーザーのアカウントとデータベースのテーブルやカラムへのアクセス制御との連動が必要であり、現在はそれを実装しています。

UIとの連動に関する実装の詳細について述べることは別に譲るとして、ここでは特に大きなデータベース表を扱わねばならないカタログデータベースについての工夫や苦労について述べようと思います*2。

筆者は天体数という意味では10億天体程度のカタログデータベースを作成・運用した経験は既にあつたものの、そのカラム数(パラメータの数)は多くて10個にも満たないものでした。HSC-SSPのカタログデータベースはこれらとは技術的にレベルの違うものが必要であろうということは容易に想像できたものの、それがどの程度たいへんで、どのような困難を伴うのかを完全には理解できていませんでした。しかも、以前開発したデータベースは使い方が限られていたため、特定の内容の検索を高速化するためのチューニングのみに注力すれば問題ありませんでしたが、HSC-SSPのカタログデータベースでは不特定多数の検索(クエリー)パターンがありうるので、

*2 画像データ管理には多くても数百万レコード程度の表で十分なことが多く、検索性能の点で大きな問題にはならないため。

単純な話ではないことも明らかでした。それらを1ステップずつ解決していくしかないというのが、経験値の少ない私たちには唯一の方法であったと言えます。

ユーザーに対してどのようなデータの取り出し方を提供するかは重要な点です。定型的な切り出し方だけでなく、ユーザーごとにフレキシブルに抜き出し方を指定できる必要があります。しかも、一般的にユーザーはそれほどSQLに詳しくない場合も多く、また複雑な記述はしばしば誤りの原因にもなります。文法的にそれほど難しいSQLでほとんどの検索が実現されたほうが良い、という要請もありました。

そのため、私たちはカタログデータベースの検索のために、Direct-SQLモードという、ユーザーが直接SQLを入力しフレキシブルに検索条件や出力を指定できるユーザーインターフェースを導入しました。また、データベーステーブルについては、縦（天体数方向）横（パラメータ数方向）に長いものを一つだけ用意するのは、検索効率の面から非常に効率が悪いことがあるので、測定パラメータのグループごとにいくつもの横方向に短い表に分割して、それらを仮想的に合体するようにして、ユーザーには一つの大きな表として見えるようにしました。このため、ユーザーに対しては、あたかも一つの大きなテーブルに対して検索をかけているように見せることができ、用意するSQL文の単純化を促すことができたのです。また、この方法のおかげで、必要なパラメータ以外が呼ばれないようにすることも可能となったため、比較的高速にカタログを検索することも可能となりました。

初期のHSC-SSP共同研究者向けデータリリースでは、実はデータベースのテーブル構造に対してそこまでの工夫をせずに取り組んでいたため、データ量が小さい間は良かったのですが、リリースを重ねデータ量が増える中で検索速度において行き詰まったこともありました。この問題を解決

するために導入した前述の実装には、峯尾聡吾さんや小池美知太郎さんをはじめとするHSCソフトウェア開発チームの大きな努力が払われたことを付け加えておきたいと思います。今後も、データが大きくなるにつれてさまざまな工夫や、大きな変更が必要になるやもしれません。

今後も観測が続く限り、さらにデータは増えます。しかも、使用するフィルターも当初計画より増えるかもしれません。将来的に検索速度を保てるのか、という問題もあります。加えて、Forced測光の各露出に対応するカタログデータベースの作成や提供は十分には実現できていません。この測定モードは天体の時系列測定データを扱いたいときには是非とも欲しいものですし、各露出の校正の精度を確認し、さらに高精度化するためのエンジンの役割をすることが期待されています。

また、サービスを維持・改善するためには、システムに過負荷をかける効率の悪いSQL発行を避ける工夫や、刻々と変化するデータ解析パイプラインに追従したデータベースの更新なども行わなければなりません。その対応は難解ではないが面倒なことが非常に多く、特にパイプラインの変更情報がなかなか確定しなかったりと、国際共同開発の難しさが顕在化する場合もあります。課題を挙げ始めればきりが無いものの、今後も開発・運用を地道に進めていかねばなりません。

4. 全世界ユーザー向けデータリリースPDR1

私たちは2017年2月に全世界向けHSC-SSPデータリリースであるPDR1を公開しました。同年6月と11月にはさらに測光赤方偏移（photo-z）や輝線銀河のカタログなどの情報を追加しました。これらのデータをもとに、本特集の成果を含む多数の研究が行われています。

PDR1で公開されているデータの詳細は論文¹⁾をご参照いただくとし、この場では簡単にその概要をご紹介します。

4.1 PDR1公開データの成り立ち

HSC-SSPのデータはいわゆるウェディングケーキ型と呼ばれるワイド、ディープ、ウルトラディープの3層構造の観測計画に基づいて取得されています⁴⁾。PDR1ではこれら3階層のデータセットごとに独立に解析を行いました(表2)。ワイド階層は、天の赤道帯の連続したフィールドを中心に約1,400平方度にわたる広領域を*i*バンドで26等程度の深さまで広帯域フィルター5バンド(*g, r, i, z, y*)で観測するプログラムです。ディープ階層は、可視光以外の波長によるサーベイが世界的に展開されているような科学的に重要な四つのフィールド(計26平方度)を広帯域と狭帯域フィルター(N387, N816, N921)を使って*i*~27等まで観測します。最後にウルトラディープ階層は多波長観測による深宇宙探査で重要な赤経2時(SXDS)と10時(COSMOS)付近の2視野をやはり広帯域+狭帯域(N816, N921,

N101)で、バンドによっては10-20時間の積分時間をかけ、さらに深い*i*~28等まで観測します。各階層の観測計画で想定されたサイエンスケースは多岐にわたり互いにオーバーラップしますが、ワイド階層は主に弱重力レンズ効果による宇宙論研究を、ディープやウルトラディープ階層は近傍から遠方に至る銀河・活動銀河核(AGN)の形成進化、宇宙の再電離史、さらに超新星・ γ 線バーストに代表される突発天体現象や太陽系内天体の探査を行うのに適したデータとなっています。

PDR1の公開データは、サーベイ開始から1年8か月後まで(2015年11月までの61.5夜)に観測されたデータを含めて構築しました。その内容は2016年1月にHSC-SSP共同研究者向けに公開したプロダクトがもとになっています。ワイド階層の目標サーベイ面積が1,400平方度であるのに対して、全5バンドで想定の高さに到達している領域の面積は約100平方度です。データ全体の星像品質は、合成画像のPSFサイズでおおむね0.8秒角以下くらいに収まっています。特にワイド階層の*i*バンドデータは弱重力レンズ解析に用いられるため、HSC-SSP共同研究チームの約束事でシーイングが約0.7秒角以下の天候条件のときだけ観測することになっています。そのため、ワイド階層*i*バンドの合成画像はPSFサイズ0.56秒角と精細な形状測定に適したデータとなっています。ディープ、ウルトラディープ階層の観測もまだ道半ばであり、達成度はバンドによりかなり差がありますが、広帯域データは目標とする積分時間のおよそ6分の1、狭帯域データは3-4分の1の達成状況です。387 nmと1 μ m帯を観測する狭帯域データはPDR1には含まれませんが、観測が進みつつあります。

データプロダクトに対する品質評価の結果、PDR1のプロダクトが統計的な科学研究を行うのに十分な天体測定精度をもつことが確認されました。測光較正の典型的な誤差は、同一天体の測定値の複数ショット間の内部比較から0.01-0.02等程度、

表2 PDR1公開データの概要。現在到達された積分時間や深さは*i*バンドの典型値で括弧内の値は目標値。N387とN101バンドのデータはPDR1には含まれません。

階層	天域名	バンド名	積分時間 (分)	深さ (等級)
W	XMM-LSS	<i>g, r, i, z, y</i>	20 (20)	26.4 (26.2)
	GAMA09H			
	WIDE12H			
	GAMA15H			
	HECTOMAP VVDS			
D	XMM-LSS	<i>g, r, i, z, y</i> , (N387), N816, N921	30 (126)	26.5 (27.1)
	E-COSMOS			
	ELAIS-N1			
	DEEP2-3			
UD	COSMOS	<i>g, r, i, z, y</i> , N816, N921, (N101)	130 (840)	27.0 (27.7)
	SXDS			

*) ワイド階層は観測途中のためいくつかの天域に分かれており、それぞれに便宜上の名前が付いています。

較正に使われた参照星を用いた Pan-STARRS1 (PS1) カタログ等級との外部比較では0.02-0.03程度です。位置較正についても同様に、内部比較で10ミリ秒角程度、PS1座標との外部比較で20-40ミリ秒角程度の誤差になっています。

PDR1の登録ユーザー数は公開時から順調に伸び、現在延べ800人を超えています。国立天文台の開発運用チームで種々の質問やトラブルシューティングにも対応していますが、幸い大きな問題もなく皆さんにデータを利用していただけています。まだ目標とするサーベイのサブセットの段階ですが、プロジェクト初期に掲げたレガシーに向けたデータ作りという壮大な目標の取っ掛かりくらいにはなったのではないかと考えています。観測装置とデータ解析とが密に連携して、データ取得現場からデータ公開までの仕組みの整備を目指してきたことが少し報われた思いです。

良かったことばかりを述べましたが、データ解析ソフトウェアを含めデータの取り扱い方にはまだ改善の余地が多く、今後のデータリリースで更新をし、測定・較正精度とデータの価値をさらに高めていくことになります。PDR1で既知の課題については以下で述べるデータ配信サイトや論文内で説明されており、その一部は2018年中に行われたHSC-SSP共同研究者内のデータリリースで修正されています。

PDR1に含まれる解析済みデータは、大きく分けて画像と天体カタログです。前稿でご説明したように、個々のCCDデータは較正された後、バンドごとにまとめられ、より深い画像データへと合成されます。較正されたCCD画像と合成画像それぞれについて天体の検出測定が行われ、天体カタログが作られます。

PDR1では、これら解析パイプラインの直接の出力であるFITSファイルと、測定された天体情報を格納したデータベースの2種類の形式でプロダクトが配布されています。画像ファイルは、CCDごとの画像、合成前のワープ済み画像（詳

細は前稿の第3章およびパイプライン論文⁵⁾を参照)、合成後の画像が利用できます(表1も参照)。天体カタログについては、CCDごとと合成画像の天体測定の両方についてファイルがダウンロードできますが、データ量と検索性能の制約から、現リリースのデータベースには合成画像の天体測定結果だけが登録されています。

4.2 PDR1 サービスの概要

PDR1のデータ配信は国立天文台が運営するHSC-SSPデータ公開サイト⁶⁾でユーザー登録をすることで、商用利用などを除いてどなたでもご利用になれます。このポータルサイトで利用できるサービスには以下のようなものが含まれています(図1)。

1. 画像検索 (Data Archive Service=DAS)
2. カタログ検索 (Catalog Archive Service=CAS)
3. ビューワ機能 hscMap

画像検索サービスDASは、興味のある天域や観測バンドで取得された画像のFITSデータの検索とダウンロードを行えるインターフェイスを提供します。また、PDR1の画像ファイルは天域を11分角程度のグリッドで分割した区画(patch)単位で提供されますが、それらから本当に必要な領域だけを切り出すためのQuarry(鉱石の切り出しなどの意味)と呼ばれる機能も提供しており、各バンドの天体の顔写真を取得したいときなどに便利です。ほかに、画像そのものではなく特定の場所のPSF情報を取り出すことも可能で、主に天体の形状や測光精度の評価などに使われています。

天体カタログについては、解析パイプラインが出力したFITS BINTABLE形式のファイルをそのままダウンロードすることもできますが、先に述べたように各天体の測定情報をデータベースに展開し、必要な天体情報を検索して取り出せるカタログ検索サービス(CAS)を提供しています。

CASの天体検索にはSQLを用います。最低限のSQL文法を覚えていただく必要はありますが、

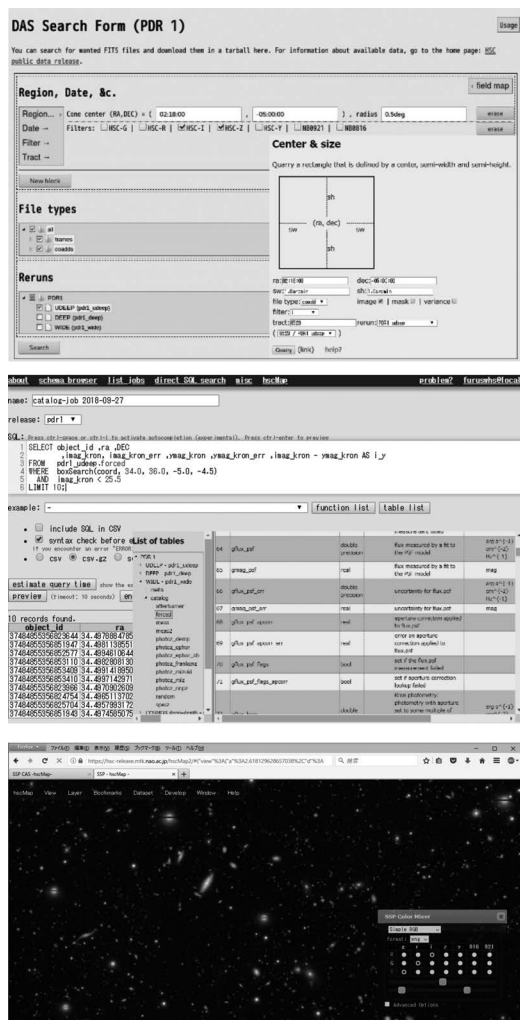


図1 (上) 画像検索DAS画面。右下は画像切り出し機能の入力画面。(中) 天体カタログ検索画面。右下はデータベース内のテーブル・カラム情報を調べるためのスキーマブラウザ。(下) 画像ビューワhscMap。

一般的な天文学研究のための天体検索を体験するには初歩的なSQLを知っていれば十分です。いくつか典型的な検索を行うためのSQLの文例も用意されています。少し慣れた方であればかなり込み入った条件を組み合わせた天体検索をすることで、直ちに科学研究に使える(サイエンスレディな)サンプル作りまで試みることもできます。一緒に提供されているスキーマブラウザ⁷⁾

を用いることで、カタログデータベース内にどのようなテーブルや関数が存在しどのような測定情報を利用できるのかを知ることができます。

CASには解析パイプラインによるHSCデータの測定情報だけではなく、HSC-SSPの共同研究チームによって求められた天体の測光赤方偏移やその較正に用いられた分光カタログなどの付加情報、さらにある天体がどの積分のどのCCDで観測されたのか、といったデータ解析の過程・履歴(ブックキーピング)の情報なども含まれています。これらの情報を興味のある天体の測定値と連結して表示したりすることもできます。PDR1のプラダクトを利用した研究活動の可能性を広げるため、付加的な情報を加える作業(インクリメンタルリリース)も半年から1年に一度のペースで継続して行っています。

最後に、較正済み画像とカタログデータベースを統合的に閲覧できる高機能なビューワ(hscMap⁸⁾)を提供しています。hscMapは一般的なウェブブラウザ上で動作するアプリケーションで、SSPサーベイ領域全体の擬似カラー画像の閲覧に加え、DASやCASと連動して、ビューワの画像上をマウスでドラッグして関心のある領域を選択し、そこに存在する天体リストや測定情報、あるいは合成前後の画像データを取得するといった込み入った作業も行うことができます。hscMapは本特集の小池さんの記事でも詳しく紹介されています。

5. まとめ

HSC-SSPのデータ公開の概要についてご紹介しました。PDR1は幸い現在まで安定した運用が続き、HSCチームは2019年に予定されているPDR2に向けた準備を進めています。今後も研究者ユーザー目線で考え、HSCのデータを最大限に利用するための科学データアーカイブを目指しますので、皆さんも機会があればぜひHSCのデータに触れてみてください。私たちは、データ

プロダクトやアーカイブ機能の向上はサーベイ終了後も継続して長期に取り組むべき活動と考えています。なお、信頼に足るデータアーカイブはソフトウェアを書いてデータベースを触るだけで実現するものではありません。きちんとしたデータを作るための観測現場から解析に至る継続的な取り組みがあって初めて成り立つのであり、また基盤である計算機システムの構築・維持・改良やユーザーサポートの努力も必須です。関係各所の皆さんの下支えが連動して本プロジェクトが成り立っていることを付け加えたいと思います。

謝 辞

HSCの科学データアーカイブの実現にご協力いただいた皆さんに御礼申し上げます。HSC科学データアーカイブを作り上げ運用に貢献されているHSCデータリリースチームの皆さん、特にデータベースの性能向上とアーカイブ機能充実に日々尽力されている峯尾聡吾さん、小池美知太郎さん、山田善彦さん、瀧田怜さん、田中賢幸さん、林祐輔さん、山野井瞳さん、に感謝します。山田さんと林さんには表1の値をご提供いただきました。安田直樹さん、プリンストン大学、SSP共同研究者の皆さんには議論を通してHSC科学データアーカイブの利便性向上にご協力いただきました。本稿を執筆する機会をくださり有用なコメントをいただいた小宮山裕氏に感謝いたします。本稿で述べた研究開発は科研費 (JP15H05887, JP15H05892, JP15H05893, 18072003, 22012007)

の補助を受けています。

参考文献

- 1) Aihara, H., et al., 2018, PASJ, 70, S8
- 2) 高田唯史, 2015, 天文月報, 108, 585
- 3) Yamauchi, C., et al., 2011, PASP, 123, 852
- 4) Aihara, H., et al., 2018, PASJ, 70, S4
- 5) Bosch, J., et al., 2018, PASJ, 70, S5
- 6) <https://hsc.mtk.nao.ac.jp> (2018.11.30)
- 7) <https://hsc-release.mtk.nao.ac.jp/schema/>
- 8) Koike, M., et al., 2014, SPIE, 9149, 914920

Road to the HSC-SSP Data Releases (2) Data Release and Catalog Database

Hisanori FURUSAWA and Tadafumi TAKATA

*Astronomy Data Center, National Astronomical
Observatory of Japan, 2-21-1 Osawa, Mitaka,
Tokyo 181-8588, Japan*

Abstract: Scientific data archives for providing processed and calibrated data are an important component in large survey projects, in order to exploit a full value of the data products and promote scientific activities. In particular, development of catalog databases is essential to enable researchers to derive necessary information for their sciences in a timely manner. HSC team has been processing HSC-SSP data and made the first public data release. The data products can be used for astronomical studies and education. We describe the importance of catalog databases and introduce the data products in the public data release.