

# アーカイブデータが切り拓く天文学

高田 唯史

〈国立天文台光学赤外線天文学研究系 〒181 東京都三鷹市大沢2-21-1〉

e-mail: takata@optik.mtk.nao.ac.jp

天文学において、観測データは、その時刻についての天体の情報を持つ唯一無二のものというもので、非常に貴重な存在といえます。このようなものは、いつどのような形で生かされるかという事は誰にもわかりません。従って、欲しいときにそのようなデータをすぐ引き出してこれるシステムを作っていくことは非常に重要なことです。世界的なアーカイブの現状からすると、現在の日本の現状の後進性は明らかです。しかし、今後アーカイブデータを含む天文データベースを用いた統計的天文学などが発展していくことは間違いがありません。そこで、我々のグループでは、可視光のデータアーカイブを手始めに、日本独自の全波長横断データアーカイブシステムの構築を目指し、様々な開発協力を始めています。

## 1. データアーカイブの意義

### 1.1 データアーカイブの役割

観測データは天文学のもっとも重要な基盤の一つです。夜ごと各観測所で蓄積されている観測データは、観測者がねらった天体の情報だけでなく、その他数多くの有用な情報を含んでいます。「新」天体と称される天体も、我々が認識を新たにただけであって、その天体は以前から存在し、過去の観測データの中に顔をのぞかせていることもしばしばです。もしデータアーカイブがあれば、天体の時間変化をさかのぼって調べることもできます。観測データはこの意味で、世界で唯一の貴重な財産なのです。アーカイバルデータは、過去の状態を知るばかりではなく、その豊富な情報量を生かして、さまざまな天体の統計的性質を調べることに役立ってくれます。全天には無数の天体が存在し、それぞれはまた広い波長域にわたって光を放射しています。一方、我々が一時に観測することができるのは、天空のほんの一部の領域のごく限られた波長域の情報にすぎません。蓄積されたアーカイバルデータを、時間、空間、波長と

いった多次元空間の上で組み合わせれば、元の観測者には思いもよらなかったような未知のフロンティアが開けてくるにちがいありません。新たなアイデアが思い浮かんだら、アーカイバルデータを用いて目的の天体のようすを知ることができます。そのデータで充分なら貴重な望遠鏡割当時間を専有しないですむし、新たな観測が必要であった場合も最適な観測方法が推測でき、観測の効率を大きく改善することができます。また、望遠鏡が大きくなればなるほど観測されるデータの質も上がり、従ってアーカイブデータを使う意味合いも大きくなるのが確実です。特に、現在世界的に建設が進んでいる8メートルクラスの大望遠鏡は、その建設費や運営費が非常に高価なため、データ一枚の取得にかけられる費用は莫大なものであり、従ってなるべく無駄のない運営が要求されますが、データアーカイブはその一翼を担うこともできます。観測所に対するメリットといえば、これ以外にも「データの整理庫」としてアーカイブを用いて、検出器の劣化の状況を調査するためのエンジンの一部にしたりすることなども考えられます。このように、アーカイブというものは、

天文学を遂行する上での様々な場面で用いることができ、一つ持っているとは非常に大きな恩恵に預れるものです。

## 1.2 有効利用の実例

では、実際にアーカイブされたデータなどを用いて、どのような研究がなされているかを2, 3紹介してみましょう。まずはパロマースカイサーベイ(通称パロマーチャート, Digitized Sky Surveyを含む)を用いた様々な研究があげられます。一見、おやっと感じられる方もおいででしょうが、パロマーチャートというのはまさに“人類の宝的”なアーカイブデータです。現在はデジタル化されてCD-ROMに収まったものまで登場していますが、これらを使って統計的な天文学を行なった例としては、銀河をプレートスキャン用の機械を用いて同定し、大規模な銀河サンプルを作り、我々の近傍の大規模構造を明らかにした研究<sup>2)</sup>などがあげられます。今後も、超新星の発見や、低表面輝度銀河の探査、低質量星の探査が行われる予定です<sup>3)</sup>。また、ちょっと私的なこととなりますが、銀河面に隠された系外銀河を探すなんて風変わりな仕事<sup>4)</sup>も、“真面目な日本人!?”が行なった例がいくつか存在しています。パロマーチャート以外のもので用いた仕事としては、様々な乾板データを用いて、写っている恒星状の天体の位置と明るさの変化を見ることによってクエーサーをさがしたり<sup>5)</sup>、1938年から1991年までに撮られた26枚の写真乾板からSculptor銀河の1177個の星の固有運動を同定し、メンバーシップと銀河全体の運動を調べた<sup>6)</sup>というような仕事も行なわれています。日本の代表的な銀河のサンプルであるKUG(KISO UV-excess Galaxies)<sup>7)</sup>なども、木曾の乾板がたくさん残されていたのでできた仕事です。やはり、統計的な仕事と時間変動を用いて行なう研究が多いようです。

## 2. データアーカイブシステムの世界的現状

世界的なアーカイブシステムの開発、運用の動向を見てみると、衛星からのデータ用としては、ハッブル望遠鏡のデータアーカイブが非常に洗練されたシステムとして有名です<sup>8)</sup>また、地上の望遠鏡としては、ハワイのCFHT(Canada France Hawaii Telescope)<sup>9)</sup>やチリのESO(European Southern Observatory)のNTT(New Technology Telescope)、カナリア諸島のラパルマの望遠鏡群のデータのアーカイブ<sup>10)</sup>などが有名です。特にCFHTやNTTのシステムは、WWWによるデータの検索、プレビューなどが簡単に行える、非常に進んだシステムとなっています。しかし、衛星のデータアーカイブに比べると、周囲の環境の変化の影響を受けやすい地上観測データについては、観測手順やデータの処理手順の確立が難しいので、遅れをとっている感は否めません。ここでは、衛星データの代表例として、HSTのアーカイブについて触れる他、地上データのアーカイブについて、システムの概要を説明することにします。

### 2.1 HSTのデータアーカイブ

非常に洗練されたアーカイブシステムの代表と言ってよいでしょう。抜群に解像度の良いイメージデータや、地上では得られない紫外線のスペクトルなどの魅力的な天体画像以外に、観測装置ごとに行なわれる特殊なキャリブレーション用のデータをも保持し、解析エンジンとあいまって、自動化された効率的なサイエンスデータの排出を実現しています。整約処理のやり直しなども比較的簡単に行なえる解析ソフト、及びユーザーに対するサポート体制に至るまでが充実しており、日本から使う上では、速度の遅さ以外はそれほど苦にならないシステムです。地上データのアーカイブシステムが目標とすべきシステムと思われる。実際の利用方法などは天文情報処理研究会発

行の『HST アーカイブデータの利用 (第1版)』を参照して頂きたいと思います。

## 2.2 CFHT のデータアーカイブ

米国ハワイ島はマウナケア山頂に存在する口径3.6mの望遠鏡のデータのアーカイブシステムです。現状では、地上観測データのアーカイブシステムとしては最も進んでいるシステムといって良いでしょう。具体的な特徴を以下にあげてみると、

- 1992年9月18日から運用開始。
- 山頂でとられたデータは、ヘッダー部は即時にftpでCADC (Canadian Astronomical Data Center) に転送されるが、データ部は一旦Waimeaに転送され、optical diskにコピーされた後、CADCでアーカイブされる。
- 観測者のデータ保有優先権は2年。
- データの量は半年で26GB程度。フォーマットはFITS。
- 観測装置の中の20パーセントはvisiting instrumentなので、それらの扱いが難しい。
- 現在のところ、基本的には生データしか扱っておらず、キャリブレーションは個々の観測者に任されているのが現状。数年後にはキャリブレート済みデータを扱う予定で、現在、IRAFなどをベースにしたパイプラインソフトの開発を行っている。
- データへのアクセスは現在のところWWW、独自のユーザーインターフェースのSTARCATのどちらか片方を用いて行えるが、今後2年間にすべてWWWに置き換わる。(図1)

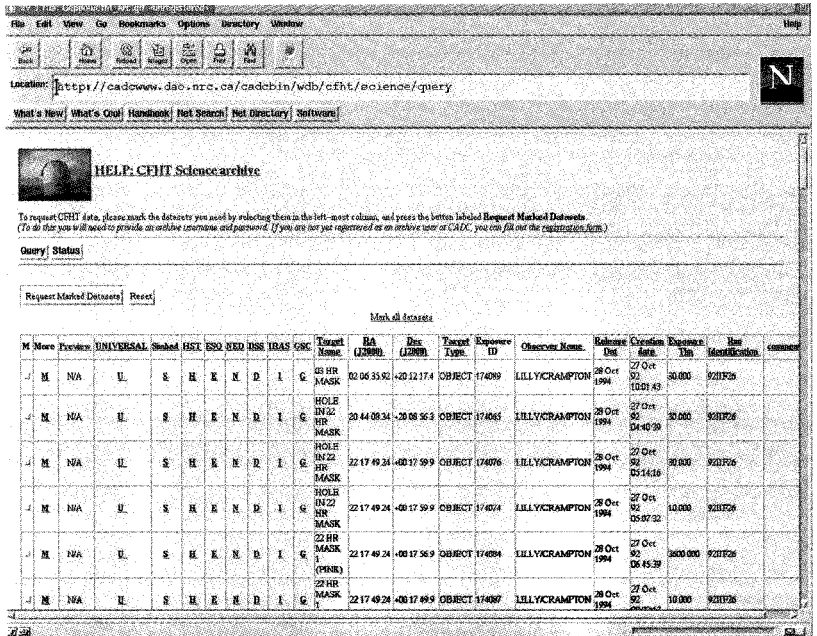


図1 CFHTのデータアーカイブシステムのユーザー画面

- URLはhttp://cadwww.dac.nrc.ca/cfht/cfht.html。実際のデータ取得にはアカウントを取得する必要がある。但し、申請後2、3日で取得できるのが通例。

というところでしょうか。

実際の使い勝手ですが、特にWWWでのアクセスの場合、検索フレームの領域についてDSSやIRASのイメージデータを参照できるようになっており、非常に便利です。また、HSTやESOのアーカイブにもリンクが張られていて、やはり検索領域について、HSTやESOでの観測例がないかを即座に調べることができます。これは非常に便利な機能で、今後我々の開発するシステムにも取り込みたいと思っています。データは各観測装置毎に分けてアーカイブされていて、SQL (Structured Query Language) などを知らなくても、検索は比較的単純に行うことができます。検索に使える項目は天体の座標や観測者名などで、装置の細かいパラメータについても指定ができるので、かなり特化した検索も可能です。イメー

ジのプレビューは、GIF と FITS が選択でき、GIF の場合は無設定で早見ができますが、FITS の場合は SAO image がクライアント側に必要となっています。検索の速度は非常に速く、画像の表示の遅さを我慢すれば十分使用に耐えるものです。

### 2.3 ESO (La Silla) のアーカイブ

南米はチリの ESO (European Southern Observatory) の施設の一つ、La Silla の観測所にあるアーカイブシステムです。このアーカイブは、基本的には NTT の可視分光器のデータのアーカイブを目的として構築され始めたのですが、現在ではそれ以外の装置についてもアーカイブを始めているようです (IRSPEC (近赤外分光器) など)。CADC と ST-ECF の協力を得て作っているのです。基本的には CFHT の作りと一緒です。特徴を以下にあげてみると、

- 観測者のデータ保有優先期間は1年間。
- NTT の建設にあわせて実装開始。1991 年から実際の運用が始まっている。
- 基本的には生データをアーカイブ。現在、パイプラインを開発中。将来的にはキャリブレーション済みデータも扱う予定。フォーマットは FITS。
- 1993 年の 4 月においてすでに 25000 フレームであった。現在の保有量は不明。
- URL は <http://arch-http.hq.eso.org> などがあげられるでしょう。

CFHT のアーカイブに比べると、検索について詳しい指定ができない反面、シーイングのデータなどの外部環境データが参照できるようになっているなど、少し違った特徴を持っています。データの取得にはアカウントが必要です。

### 2.4 La Palma のアーカイブ

アフリカ沖のカナリア諸島にある La Palma 天文台のアーカイブシステムです。ING (Isaac Newton Group) がシステムの構築、運用を行なっています。

- 1984 年からアーカイブを開始。地上データのアーカイブの先駆けといってもよいであろう。
- 生データをアーカイブしており、キャリブレーションについては観測者が行って、残す場合もある。
- データ取得は FITS で行い、3 ヶ月程度観測所で保存した後、観測者はコピーを持って帰れるというシステム運用になっている。オリジナルデータは RGO に送られて、そこでアーカイブされる。
- 観測者のデータ保有優先期間は1年間、但し希望によっては延長が許可される場合もある。
- データ量は 1993 年の 4 月時点では 150 GB で、1 年あたり 33 GB の割合で増加中。現在の状況は不明。
- アクセスは基本的にはケンブリッジの VAX マシンにログインして、そこで検索を行う。アドレスは [gxvq.ast.cam.ac.uk](http://gxvq.ast.cam.ac.uk)

このシステムは、WHT などの最新データを保持しており、サイエンス的には非常に魅力のあるシステムですが、検索、取得を行うためのユーザーインターフェースがキャラクターベースのものしか用意されておらず、アーカイブユーザーには、貧弱すぎて使いにくいものとなっています。VAX マシンをベースに運営していますが、速度的にも遅い気がします。検索についても、初心者のユーザーが覚えるのは面倒だと思ふような、特殊記号を用いた文法が採用されており、いまいち使いにくさを覚えます。

このように、地上データのアーカイブは、まだ画像処理を何も施していない生画像しか扱っていません。これは、処理済みデータにする段階で、周囲の環境の変化の影響などに対する考慮などが解析の手続きに入るの、解析の自動化が困難を極めるからです。データアーカイブが最大限に利用されるようにするためには、解析手順の標準化、解析の自動化など様々な難関をクリアせねばならず、非常に大変な仕事となっています。従って、

人、お金両方が大規模に投入できないと実現は不可能なわけですね。

### 3. 日本の現状とMOKA (Mitaka Okayama Kiso Archival System) の開発

一方、日本国内に目を移してみると、現在国内の観測所の多くでデジタルの観測データが得られています。一般に利用できるアーカイバルシステムが開発され運用されている例は存在していませんでした。木曾観測所のシュミット乾板検索システムが唯一のアーカイバルシステムといつてよいでしょう。アーカイバルシステムを早急に構築、確立し、データをアーカイブしないと、これら多くの貴重な観測データは再利用不可能になってしまうこととなります。そこで我々、天文情報処理研究会のアーカイブワーキンググループのメンバーは、国立天文台岡山天体物理観測所のスペクトロネビュラグラフ (SNG) および東京大学木曾観測所シングルチップ CCD カメラのデータのためのアーカイブシステム MOKA (Mitaka Okayama Kiso Archival System) を作成し<sup>3)</sup>、一般公開にほぼ耐えうるシステムにまで成熟させ、昨年6月より運用に入りました(図2)。また、国立天文台岡山天体物理観測所のマルチモード観測装置である OOPS (Okayama Optical Polarimetric and Spectroscopic System) のデータアーカイブシステムの MOKA への組み込みも行ない、様々な機能を実験的に実装しました<sup>10)</sup>。この他にも、検索を補助するための機能であるネームレゾ

ルバ(図3)や座標コンバータ(図4)も開発し、ユーザーにより使いやすい検索システムを提供しています。このシステムは、両観測所のユーザに役立つばかりではなく、現在ハワイのマウナケア山頂に建設中であるすばる望遠鏡のデータアーカイバルシステムのプロトタイプにもなると、我々は考えています。なお、現在の開発メンバーは市川伸一、伊藤孝士、高田唯史(以上、国立天文台三鷹)、吉田道利、西原英治(以上、国立天文台岡山)、吉田重臣(東大木曾観測所)、濱部勝(東大天文センター)、洞口俊博(国立科学博物館)の8人です。

我々は、今回の MOKA の開発を通じて、今後、データアーカイブシステムを構築していく上での様々な問題点に加えて、世界的観点にたつての日本の天文学の中でのデータアーカイブを用いた天文学、及びそれに対する天文研究者の認識の後進性を浮き彫りにしてきました。我々は、これらの認識を踏まえて、次のアーカイブシステムの開発を計画しています。それが MOKA 2 です。このシ

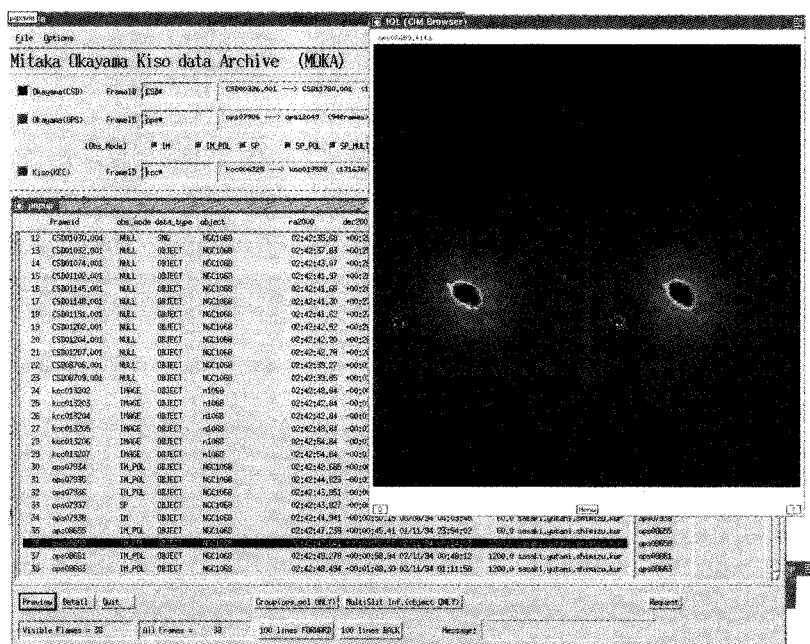


図2 MOKA のユーザー画面

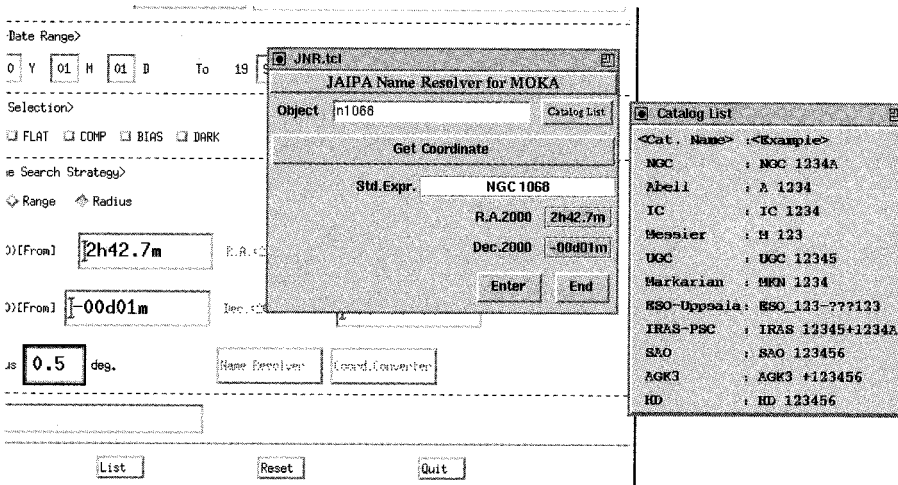


図3 ネームレゾルバの外観

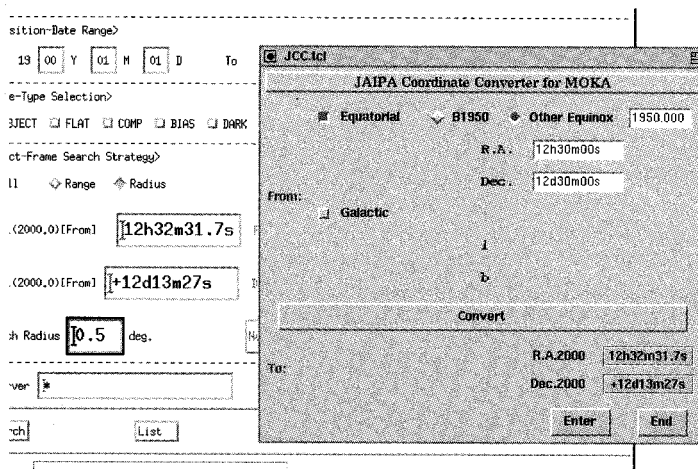


図4 座標コンバータの外観

システムは、国立天文台岡山天体物理観測所と東京大学木曾観測所のデータを扱うという意味ではMOKAと変わりませんが、実際にデータベース天文学を遂行する上で、非常に役に立つと思われる様々な新機能を備えたものに生まれ変わる予定です。この次世代のデータアーカイブシステムであるMOKA 2は、今後の日本におけるデータベース天文学の発展を占う上で重要な役割を占めることは明らかであり、どのような方針で開発して

いくかは重要な問題です。

#### 4. 全波長横断天文データベースの構築に向けて

今後の日本の観測天文学において計画されているプロジェクトは、すばるをはじめ、IRIS (Infrared Image Surveyor) や LMSA (Large Millimeter and Sub-millimeter Array) といった、深宇宙を対象にしたものが多く、従って宇宙論的な

研究が多くなっていくことは間違いのないことです。とすれば、当然サーベイのような統計的処理を必要とするアプローチが多くなるわけで、こういった研究を支えるものがまさにデータアーカイブ、データベースなのです。特に、波長を横断する形で、同じ天体群を様々な切り口でみれた場合には、確実に新しい発想もわいてきますし、また、次の観測へのステップともなることでしょう。このように様々な波長のデータをなるべく統一的なインターフェースで検索できたら、研究の効率が非常に良くなることは確実です。このために、我々のグループでは宇宙研のデータベース開発グループとも交流を持ちながら、少しずつ、全波長横断的天文データベースの開発のための下準備を始めています。勿論、たとえ同じようにデータが検索できても、各波長ごとに独特の解析手法がはびこっている現在、すぐに科学的アウトプットを出せるわけではありませんが、解析の自動化や汎用化をも含めて、少しずつ前進していこうと思っています。近い将来、ある波長の画像を見て、目的天体のところをマウスでクリックすると、観測データのある全波長について、その天体のスペクトルがグラフになってすぐさま現れてくれるようにしたいものです。そのための道のりは、日本では今まさに始まったばかりと言えます。

### 参 考 文 献

- 1) Crabtree, D. R., Durand, D., Gaudet, S., Hill, N., Morris, S. C. 1994, in *Astronomical Data Analysis Software and Systems III*, A. S. P. Conf. Ser. Vol. 61, 123
- 2) Maddox et al. 1989, *MNRAS* 242, 42
- 3) Horaguchi, T., Ichikawa, S., Yoshida, M., Yoshida, S., Hamabe, M. 1994, *Publ. Natl. Astron. Obs. Japan*, Vol. 4, 1
- 4) 「HST アーカイブデータの利用 (第1版)」, 1995年, 天文情報処理研究会 HST ワーキンググループ編
- 5) 「木曾・岡山観測データアーカイブシステム (MOKA) 開発報告書(総合版)」, 1996年, 天文情報処理研究会アーカイブワーキンググループ編
- 6) Long, K. S., Baum, S. A., Borne, K., Swade, D. 1994,

- in *Astronomical Data Analysis Software and Systems III*, A. S. P. Conf. Ser. Vol. 61, 151
- 7) Majewski, S. et al. 1991, in *The Space Distribution of Quasars*, A. S. P. Conf. Ser. Vol. 21, 55
- 8) Takase, B., Miyauchi-Isobe, N. 1993, *Publ. Natl. Astron. Obs. Japan* Vol. 3, 169
- 9) Takata, T., Yamada, T., Saito, M. 1996, *ApJ* 457, 693
- 10) Takata, T., Ichikawa, S., Horaguchi, T., Yoshida, S., Yoshida, M., Nishihara, E., Ito, T., Hamabe, M. 1995, *Publ. Natl. Astron. Obs. Japan* Vol. 4, 9
- 11) Zuiderwijk, E. J., Martin, R., Raimond, E., van Diepen, G. N. J. 1994, *PASP*, 106, 515
- 12) Schweitzer et al. 1995, *AJ* 110, 2747
- 13) Reid, N., Djorgovski, S. 1993, in *Sky Surveys*, A. S. P. Conf. Ser. Vol. 43, 125

### Archival data will open up a new field in astronomy

Tadafumi TAKATA

*National Astronomical Observatory of Japan*

Abstract: Since the observational data have the unique information of astronomical object at the observed time, they are very valuable in astronomy. Anyone does not know when they are needed in astronomical study, and constructing the data archival system, which makes it possible to get the data when the data are needed, is very important. In comparison with data archival system in other countries, it is obvious that Japan is still behind them. But statistical astronomy and so on, which based on astronomical database, including archived data, must have progress in near future. Then, we begin to construct astronomical data archival system which covers most of all electromagnetic wavelength. As a first step, we have constructed the data archival system of optical CCD data.