Initial Data Release from the INT Photometric Hα Survey of the Northern Galactic Plane (IPHAS)


¹ Institute of Astronomy, Madingley Road, Cambridge CB3 0HA, U.K.
² Isaac Newton Group of Telescopes, Apartado de correos 321, E30700 Santa Cruz de La Palma, Tenerife, Spain
³ Institut für Physik, Karl-Franzen Universität Graz, Universitätsplatz 5, 8010 Graz, Austria
⁴ Imperial College London, Blackett Laboratory, Exhibition Road, London, SW7 2AZ
⁵ Centre for Astrophysics Research, University of Hertfordshire, College Lane, Hatfield AL10 9AB
⁶ Royal Observatory Edinburgh, Blackford Hill, Edinburgh, EH9 3HJ, U.K.
⁷ Department of Physics, University of Warwick, Coventry, CV4 7AL, U.K.
⁸ Department of Physics, Faculty of Science, Naresuan University, Phitsanulok, 65000, Thailand
⁹ Department of Physics, University of Warwick, Coventry, CV4 7AL, U.K.
¹⁰ Department of Physics & Astronomy, University of Southampton, Southampton, SO17 1BJ, U.K.
¹¹ Instituto de Astrofísica de Canarias, Via Láctea s/n, E38200 La Laguna, Santa Cruz de Tenerife, Spain
¹² National Radio Astronomy Observatory, 530 Edgemont Road, Charlottesville, Virginia, 22903-2475, USA
¹³ Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, U.S.A.
¹⁴ School of Physics & Astronomy, University of Southampton, Southampton, SO17 1BJ, U.K.
¹⁵ School of Physics and Astronomy, University of Manchester, Stockport, PO Box 88, Manchester, M60 1QD, U.K.
¹⁶ Astrophysics Group, Department of Physics, Bristol University, Tyndall Avenue, Bristol, BS8 1TL, U.K.
¹⁷ School of Physics, University of Exeter, Stocker Road, Exeter, EX4 4QL, U.K.
¹⁸ Department of Physics, Macquarie University, NSW 2109, Australia
¹⁹ Anglo-Australian Observatory, PO Box 296, Epping NSW 1710, Australia
²⁰ Observatorio Astronómico de Valencia, Universidad de Valencia, 46071 Paterna-Valencia, Spain
²¹ Physics Department, University of Central Florida, Orlando, FL 32816, U.S.A.
²² Armagh Observatory, College Hill, Armagh BT61 9DG, Northern Ireland
²³ Departamento de Física Teórica y del Cosmos, Facultad de Ciencias, Av. Fuentenueva s/n, E18071 Granada, Spain

Received ..................; Accepted ..................

ABSTRACT

The INT/WFC Photometric Hα Survey of the Northern Galactic Plane (IPHAS) is an imaging survey being carried out in Hα, r' and i' filters, with the Wide Field Camera (WFC) on the 2.5-metre Isaac Newton Telescope (INT) to a depth of r' = 20 (10σ). The survey is aimed at revealing the large scale organisation of the Milky Way and can be applied to identifying a range of stellar populations within it. Mapping emission line objects enables a particular focus on objects in the young and old stages of stellar evolution ranging from early T-Tauri stars to late planetary nebulae. In this paper we present the IPHAS Initial Data Release, primarily a photometric catalogue of about 200 million unique objects, coupled with associated image data covering about 1,600 square degrees in three passbands. We note how access to the primary data products has been implemented through use of standard virtual observatory publishing interfaces. Simple traditional web access is provided to the main IPHAS photometric catalogue, in addition to a number of common catalogues (such as 2MASS) which are of immediate relevance. Access through the AstroGrid VO Desktop opens up the full range of analysis options, and allows full integration with the wider range of data and services available through the Virtual Observatory. The IDR represents the largest dataset published primarily through VO interfaces to date, and so stands as...
1 INTRODUCTION

The INT Photometric Hα Survey of the Northern Galactic Plane (IPHAS) is a 1800 deg² CCD survey of the northern Milky Way (|l| < 5°) using the r′, i′ and Hα passbands (table 1) down to r′ = 20 (Vega, 10σ for a point-like source in an aperture of 1.2 arcsec). The execution of IPHAS, the properties of its characteristic colour-colour plane, and the likely scientific scope of the survey were discussed by Drew et al. (2005)

This paper describes the Initial Data Release (IDR) of the IPHAS survey, containing observations for about 200 million objects and comprising 2.4 Tb of processed imaging data. The survey is being carried out using the Wide Field Camera (WFC) on the INT telescope. The WFC consists of four 2048x4096 pixel CCDs with a pixel scale of 0.33 arcsec and a field of view of 0.29 square degrees. The camera CCD detectors are configured in an L-shape and CCD number 3 is slightly vignetted in one corner.

The data are already being used to carry out a series of investigations: the interaction between the planetary nebula Sh2-188 and the inter-stellar medium (Wareing et al. 2006); the discovery of a new quadrupolar nebula (Mampaso et al. 2006); the study of the properties of a sample of cataclysmic variables (Witham et al. 2006); the interaction of the jet of Cygnus X-1 with the surrounding medium (Russell et al. 2007); brown dwarf searches (Martín et al. 2008, submitted); detection of nova progenitors (Steele et al. 2007, 2006); the study of the properties of symbiotic stars and first discoveries (Corradi et al. 2007), the study of young pre-main sequence stars in CygOB2 (Vink et al. 2008, submitted), the presentation of a sample of about 5000 Hα emitting stars (Witham et al. 2008) and the description of a method for selecting A stars and exploiting them to estimate distances and reddening (Drew et al. 2008). Several other investigations are ongoing and we hope the data release described here will lead to many data mining programs by the community.

In addition to the optical data provided by IPHAS, the Galactic Plane Survey (GPS; Lucas et al. 2007) of the UKIDSS (Dye et al. 2006) is also observing a similar area in J, H and K. This will be followed in the next few years by a survey of the southern Galactic Plane, VPHAS+, on the VLT Survey Telescope (VST) using OmegaCam (Kuijken et al. 2004). It serves as a digital successor to the UK Schmidt Hα Survey (Parker et al. 2005) and will reach about 2 magnitudes deeper. The UVEX project is targeting the same area in the U and g bands. The Herschel observatory (Pilbratt 2005) will also target the Galactic Plane. Together, these surveys will provide a better understanding of the stellar content and properties of our Galaxy.

The aim of this paper is to provide a guide to the Initial Data Release of the IPHAS survey, describe the data products and the data properties and usage. Section 2 summarises the observations, sky coverage and data quality. The data processing is briefly described in section 3. The data products contained in this release are described in section 4. Section 5 shows an example usage of the catalogue to extract sources detected in a large area of sky and cross match with the GPS survey. Finally in section 6 we describe the different user access interfaces to the data.

<table>
<thead>
<tr>
<th>Filter</th>
<th>Cen Wave (Å)</th>
<th>Width (Å)</th>
<th>AB (mag)</th>
</tr>
</thead>
<tbody>
<tr>
<td>r′</td>
<td>6254.1</td>
<td>1350.0</td>
<td>0.164</td>
</tr>
<tr>
<td>Hα</td>
<td>6568.3</td>
<td>96.0</td>
<td>0.352</td>
</tr>
<tr>
<td>i′</td>
<td>7771.9</td>
<td>1490.0</td>
<td>0.413</td>
</tr>
</tbody>
</table>

Table 1. Filter parameters. The second column lists the effective wavelength of the filter and the third its width. The last column lists the value to be added to the Vega magnitude to convert it to AB. Filter profiles are available from the ING web site (www.ing.iac.es) and discussed in Drew et al. 2005.

2 SKY COVERAGE AND DATA QUALITY

The IDR contains 1,600 deg² of imaging data in three bands covering the northern galactic plane at latitudes |b| < 5° and longitudes 30° ≤ l ≤ 220°. Figure 1 shows the distribution in the sky. The data have been obtained between August 2003 and December 2005 during a total of 212 nights (about 26 per cent of this observing time has been completely lost due to bad weather).

The area to be observed has been divided into “fields” each of them corresponding to the coverage of the WFC. Fields have been defined in a diagonal rather than rectangular geometry in order to take advantage of the L-shape of the detector configuration allowing for a 5 per cent overlap between adjacent pointings. Figure 2 provides an example of the observation configuration.

Gaps between detectors are typically 20″. Therefore the area not observed in each pointing due to the gaps is about 12 square arcmin (i.e. ~1 per cent). In order to cover the chip gaps and bad columns each “field” is observed in pairs which are referred to as “on” and “off” positions with an offset of 5′W and 5′S between the pointings. Most of the objects are thus observed twice. Field pairs are observed consecutively in the three bands before moving to the next field pair in order to minimize the effects of photometric variability. Fields with mean stellar seeing values larger than 2 arcsec, ellipticity larger than 0.2, r′-band sky counts larger than 2000 ADUs (which indicates observations too close to the moon or with clouds such that sensitivity is reduced to a significantly lower magnitude limit) or anomalous star ratios between the filters (which indicates variable clouds during observations) are scheduled for re-observation in the three filters. Each field pair is assigned an identification which we will refer to as “field number”. The average stellar counts is ~20,000 per square degree although this value is highly dependent on the position along the Galactic plane as seen in figure 3 which shows the number of sources detected in each field as a function of Galactic longitude.

The overall survey properties and quality of the data is summarised in table 2. The exposure time for each filter is 30s in r′ (except in the first year of observations, 2003, when it was 10s), 10s in i′ and 120s in Hα. All data obtained up to December 2005 have been included in the IDR regardless of seeing or photometric conditions. Figure 4 shows the seeing values for these data in the r′ band. Although the mean value is about 1.26 arcsec (and the median 1.18 arcsec) there are clearly observations obtained in poorer conditions. The maximum acceptable seeing value for the data to be accepted as final survey quality is 2 arcsec in the three bands. Data with larger seeing values have been included in this
Figure 1. IPHAS Initial Data Release sky coverage in galactic coordinates (degrees). The top figure displays the coverage of the full release and the bottom one the coverage of the PhotoObjBest best observations subset. Brightness indicates the total line-of-sight reddening as computed from the Schlegel et al. (1998) dust maps using 4 pixel interpolation.

Figure 2. Tiling configuration of 3 x 3 pointings covering an area of ~2 square degrees. Chips are outlined in black. Field numbers are also shown. An additional pattern with an offset of 5' x 5' is not included in order to make the figure clearer.

Figure 3. Source counts as a function of Galactic longitude (one data point per IPHAS field).

release but they do not appear by default to the user. For this purpose we are releasing a full catalogue, PhotoObj, with all the observations obtained and a PhotoObjBest subset which only contains the observations which pass our scientific criteria for this release (see section 4.2 for more details). As noted in table 2 the coverage of this latter catalogue is about 1,000 square degrees and contains about 90 million unique objects. The Final Data Release of the data
One of the criteria for inclusion in the PhotoObjBest catalogue is that objects have to be observed with seeing conditions better than 2.0 in the three bands (left of the dashed line).

Figure 4. Seeing distribution in the $r'$ band for the whole IDR.

will extend the defined scientific criteria to the whole 1,800 square degrees. Figure 5 shows the magnitude limit (5σ) for all the objects in the three survey bands. Note that $r'$ observations obtained in 2003 are shallower due to the smaller integration time used.

Figure 5. Magnitude limit distribution for the IDR in the three survey bands. The $r'$ band distribution has been split in two, the leftmost corresponding to the 2003 observations and the rightmost corresponding to post-2003 observations in which the exposure time was increased from 10 s to 30 s.

3 DATA PROCESSING

All the observations obtained with the INT telescope at La Palma are transmitted in pseudo-real time to the Cambridge Astronomy Survey Unit (CASU) for data archival via the internet. Although quasi real time data processing is possible it is preferable to perform the data processing on a run by run basis. The calibration is then performed using the whole run master flats and biases. For each observation the data are stored in a Multi-Extension Fits file (MEF) with a primary header describing the overall characteristics of the observation (pointing, filter, exposure time, ...) and four image extensions corresponding to each of the CCD detectors.

The data are processed by CASU as described in Irwin and Lewis (2001) following the same steps as for the processing of the Wide Field Survey (Irwin et al. 2005; McMahon et al. 2001) which used the same telescope and camera combination. We provide here a short description of the reduction steps; additional details are available from Drew et al. (2005).

The data are first debiassed (full 2D bias removal is necessary). Bad pixels and columns are then flagged and recorded in confidence maps, which are used during catalogue generation. Linearity tests using sequences of dome flats revealed that the CCDs have significant non-linearities at the level of 1 to 2 per cent depending on the detector and period of observation. Due to several changes to the WFC controllers the linearity values change from time to time so linearity measurements are updated periodically. A linearity correction using look-up-tables is then applied to all data. Flatfield images in each band are constructed by combining a series of sky flats obtained in bright sky conditions during the twilight for each run.

The standard catalogue generation software (Irwin 1985) makes direct use of the confidence maps previously generated (described in 4.1.1) for a variety of tasks. These include object detection and parametrization, producing quality control information, standard object descriptors and detected object overlay files. The possibly varying sky background is estimated automatically, prior to object detection, using a combination of robust iteratively clipped estimators. The image catalogues are then further processed to yield morphological classification for detected objects and used to generate astrometric and photometric calibration information.

Each detected object has an attached set of descriptors, forming the columns of the binary table and summarising derived position, shape and intensity information. During further processing stages ancillary information such as the sky properties, seeing and mean stellar image ellipticity, are derived from the catalogues and stored in the FITS headers attached to each catalogue extension. In addition to being the primary astronomical products from the pipeline processing, the catalogues and associated derived summary information form the basis for astrometric and photometric calibration and quality control monitoring.

Finally an astrometric solution starts with a rough World Coordinate System (WCS) based on the known telescope and camera geometry and is then progressively refined using the 2MASS catalogue. The WFC field distortion is modelled using a zenithal equidistant projection with polynomial radial distortion (ZPN; Greisen and Calabretta 2002). The resulting internal astrometric precision is better than 100 mas over the whole WFC array (based on intercomparison of overlap regions; see section 3.3). The object detection is performed in each band separately using a standard APM-style object detection and parametrization algorithm (Irwin 1985). The curve-of-growth, defined as the flux inside an aperture as a function of its radius, is calculated measuring the flux in a set of apertures of radius $r/2$, $r$, $\sqrt{2} r$, $2r$, $2\sqrt{2} r$ where $r = 3.5$ pixels (1.2 arcsec) and an automatic aperture correction (based on the average curve-of-growth...
nearest neighbour separation can then be calculated as

\[ r_{\text{NN}} = \frac{1}{2\pi p} \int_0^{\theta_{\text{NN}}} \theta P_{\text{NN}}(\theta) d\theta \]  

(1)

where \( p \) is the average density and \( \pi \theta_{\text{NN}}^2 \) is the number of neighbours in the interval \((0, \pi \theta_{\text{NN}}^2)\). The expected mean nearest neighbour separation can then be calculated as (Scott and Tout 1989)

\[ \langle \theta \rangle = \frac{1}{2\pi^2} \int_0^{\infty} \theta P_{\text{NN}}(\theta) d\theta = \frac{1}{2\sqrt{\rho}} \]  

(2)

For the space densities of the fields plotted in figure 6 we obtain \( \langle \theta \rangle = 2.5" \) for the crowded field and \( \langle \theta \rangle = 4.7" \) for the average density field. The maximum of the \( P_{\text{NN}} \) distribution is given by \( \theta_m = 1/\sqrt{2\pi \rho} \) and corresponds to \( \theta_m = 2.0" \) and \( \theta_m = 3.8" \) in the crowded and average density fields respectively. The large discrepancy between the random and observed distributions for the high density field at \( \theta > 2" \) is due to missing close pairs which produces an overestimation of the nearest neighbours at \( \theta \approx 2" \). We can calculate the fraction of missed objects due to crowding using the following expression (Irwin and Trimble 1984)

\[ \rho_{\text{corr}} = \frac{\log(1 - 4\pi p \theta_{\text{FWHM}}^2)}{4N \theta_{\text{FWHM}}^2} \]  

(3)

where \( \rho_{\text{corr}} \) is the corrected average density and \( \theta_{\text{FWHM}} \) the FWHM of the images (i.e. 1.2 arcsec). For the average density field the fraction of sources lost due to close pairs is 9 per cent while in the crowded field this percentage increases to 41 per cent. Note as well that non resolved objects pair, on average, with objects at a distance \( \sim \langle \theta \rangle \) occupying, and overestimating, the central bins of the nearest neighbour distribution (figure 6 top). We also plot in figure 6 the corrected nearest neighbour PDF as a dotted line using the corrected average density given by equation 3.

Note that the crowded field case represents a extreme case which has source density more than three times higher than the median source density of the survey.

The values obtained for the mean nearest neighbour separations are clearly well above our seeing values in both cases indicating that aperture photometry is appropriate for our case.

### 3.1 Morphological Classification

The morphological classification is based primarily on the aperture fluxes and the discrete curve-of-growth for each detected object defined by them. Ancilliary information such as object ellipticity derived from intensity-weighted second moments is also used but only as a secondary indicator. The curve-of-growth is a direct measure of the integral of the point spread function (PSF) out to various radii and is independent of magnitude if the data are properly linearised, and if saturated images are excluded. In using this property the classifier further assumes that the effective PSF for stellar objects is constant over each detector,1 although individual detectors are allowed to have different PSFs since all detectors, pointings and passbands are treated independently.

The average stellar locus on each detector in these parameter spaces is generally well-defined and is used as the basis for a null hypothesis stellarness test. Figure 7

---

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Telescope</td>
<td>2.5m Isaac Newton Telescope, La Palma</td>
</tr>
<tr>
<td>Instrument</td>
<td>Wide Field Camera</td>
</tr>
<tr>
<td>Field of view</td>
<td>0.29 sq. deg. (per pointing)</td>
</tr>
<tr>
<td>Detector scale</td>
<td>0.33 arcsec/pixel</td>
</tr>
<tr>
<td>Total area of survey</td>
<td>~1800 sq. deg</td>
</tr>
<tr>
<td>Total area in IDR</td>
<td>~1600 sq. deg</td>
</tr>
<tr>
<td>Area included in the PhotoObjBest catalogue</td>
<td>~200,000,000</td>
</tr>
<tr>
<td>Number of unique objects in the full catalogue</td>
<td>~90,000,000</td>
</tr>
<tr>
<td>Filters</td>
<td>( r', i', Ha )</td>
</tr>
<tr>
<td>Exposure time</td>
<td>30 s (( r' )), 10 s (( i' )), 120 s (( Ha ))</td>
</tr>
<tr>
<td>FWHM (median)</td>
<td>1.3&quot; (( r' )), 1.2&quot; (( i' )), 1.3&quot; (( Ha ))</td>
</tr>
<tr>
<td>Magnitude limit (Vega, 5( \sigma ))</td>
<td>21.8 (( r' )), 20.2 (( i' )), 20.6 (( Ha ))</td>
</tr>
<tr>
<td>Astrometry</td>
<td>&lt;0.1&quot; (based on 2MASS)</td>
</tr>
</tbody>
</table>

† 10 s in 2003

---

© 2008 RAS, MNRAS 000, 1-17

---

Table 2. Survey parameters and IDR quality summary.

---

1 In practice we find that the effects of the spatial variation of the PSF on the aperture fluxes at the detector level are generally negligible.
shows an example of this comparing the magnitude difference recorded in two apertures (1.2 and 2.4 arcsec diameter) as a function of the derived overall magnitude and illustrates how this relates to the final classification for this frame.

In practice, the aperture with radius $r = 1.2$ arcsec is used as a fixed reference and also defines the internal magnitude (flux) scale. The linearity of the system implies that the position of the stellar locus for any function of the aperture fluxes is independent of magnitude (at least until images saturate). Therefore marginalising the flux ratios over magnitude yields one-dimensional distributions that can be used to greatly simplify locating the stellar locus using an iteratively clipped median. With the location estimated, the median of the absolute deviation from the median (MAD) provides a robust measure of the scatter about this locus as a function of magnitude, at least until galaxies dominate in number. This process is repeated iteratively for each distribution, using 3-sigma clipping to remove non-stellar outliers, until satisfactory convergence is reached. After convergence the equivalent Gaussian sigma is estimated using $\sigma_{\text{gauss}} = 1.48\text{MAD}$ and by this means each of the image shape descriptors (in this case flux ratios or ellipticity) can be renormalised to follow a zero-median, unit variance Gaussian-like $N(0,1)$ distribution.

These measures are then directly combined to form the final classification statistic. The combination (essentially an unweighted sum of the normalised signed distributions) is designed to preserve information on the “sharpness” of the object profile and is finally renormalised, as a function of magnitude, to produce the equivalent of an overall $N(0,1)$ classification statistic.

In practice measures derived from real images do not exactly follow Gaussian distributions. However, by combining multiple normalised distributions (with well-defined 1st and 2nd moments), the Central Limit Theorem works in our favour such that the resulting overall statistic is Gaussian-like to a reasonable approximation and hence can be used with due care as the likelihood component of a Bayesian Classification scheme, making optional use of prior knowledge.

Objects lying within 2–3$\sigma$ of the stellar locus (i.e. of zero) are generally flagged as stellar images, those below – 3 to –5$\sigma$ (i.e. sharper) as noise-like, and those above 2–3$\sigma$ (i.e. more diffuse) as non-stellar. Although the discrete classification scheme is based on the $N(0,1)$ measure of stellar appearance it also has several overrides built in to attempt

© 2008 RAS, MNRAS 000, 1-17
to make it more reliable. For example, adjustments to the
boundaries at the faint-end (to cope with increased rms
noise in the statistic) and at the bright-end (to cope with
saturation effects) are also made, while the overall image
ellipticity provides a further check.

A by-product of the curve-of-growth analysis and the
classification is an estimate of the average PSF aperture
correction for each detector for those apertures (up to and
including 4r, which includes typically ~99 per cent, or more,
of the total stellar flux) used in deriving the classification
statistic. Accurate assessment of the aperture correction to
place the (stellar) fluxes on a total flux scale is a crucial com-
ponent of the overall calibration. We find that this method
of deriving aperture corrections contributes < ±1 per cent
to the overall photometry error budget and also provides a
useful first order seeing correction for non-stellar sources.
Further by-products of the morphological classification pro-
cess are improved estimates of the seeing and average PSF
ellipticity from making better use of well-defined stellar-only
sources. These parameters are required for quality control
monitoring of telescope performance and “atmospheric” seeing.

3.2 Photometric Calibration

Photometric calibration is done using series of Landolt stan-
dard stars (Landolt 1992) with photometry in the SDSS
system with additional stars from Stetson (2000). For each
night a zero point and error estimate using the observations
of all the standard fields in each filter is derived. For pho-
tometric nights the calibration over the whole mosaic has
an accuracy of 1-2 per cent. For the purpose of the pho-
tometric calibration, standard star observations have been
taken each night at an interval of 2 h and have been used
to calibrate the r' and i' frames. The Ha frames have been
calibrated using a fixed zero-point offset of 3.14 magnitudes
with respect to r', corresponding to the magnitude differ-
ence in r'−Ha for Vega. This value has been calculated
by convolving the HST Vega spectrum (Bohlin 2007) with
the atmospheric throughput, INT optical throughput (pri-
mary mirror and field corrector), CCD efficiency and the
WFC filter curves (available from the ING web pages at
www.ing.iac.es) so that r'−Ha = 0 for Vega.

All calibration is by default corrected during pipeline
processing for the mean atmospheric extinction at La Palma
(0.09 in r' and Ha and 0.05 in i' per unit airmass).

During non-photometric nights, in otherwise acceptable
observing conditions, we find that the derived zero point sys-
tematic errors can be up to 10 per cent or more. Although
the pipeline usually successfully flags such nights as non-
photometric (see figure 8) it still leaves open the problem
of what to do about tracking the varying extinction dur-
ing these nights. For this Initial Data Release we have not
included a global photometric solution for the whole sur-
vey. While it is possible to perform an offset to a common
reference system for a small subset of the observations
using overlaps, as demonstrated below in section 5, the ap-
lication of such methods to the whole survey area is more
complicated since one has to carefully select good reference
frames and allow for variation in parameters like extinction
or zero point over the whole survey period. Availability of
good photometric observations to tie the solution as much as
possible is crucial. Clearly this is a major task which we are
currently investigating and is best accomplished when the
survey is complete. Data from non-photometric nights have
been included in the release but flagged as such so that they
do not appear in our PhotoObjBest catalogue. Figure 9
shows the mean magnitude rms for the objects included in
this catalogue, zero point uncertainties are at the level of 2
per cent for r' and i' and 3 per cent for Ha. We also plot
in figure 10 the stellar counts in the three bands along two
different line of sights in the Galactic plane with different
reddening.

All magnitudes quoted are in the Vega system but we
give the conversion factors to the AB system in table 1.

3.3 Astrometric Calibration

Astrometric calibration is a multi-stage process and aims
to provide each image, and any derived catalogues, with a
World Coordinate System (WCS) to convert between pixel
Figure 9. Photometric measurement uncertainties as a function of magnitude in 0.5 mag bins for all the objects included in PhotoObjBest. The vertical error bars show the rms dispersion in the mean uncertainties in each brightness bin and the horizontal lines in each panel show the SNR=10 (dotted) and SNR=5 (dashed) values.

9 DATA PRODUCTS
4.1 Images

There are two types of imaging data products: (1) object images and (2) confidence maps. At this point we do not provide stacked images from the overlapping frames. We describe each of them separately.

As stated before, each image is a MEF file containing the four CCD frames. Each CCD is photometrically and astrometrically calibrated as described in the previous section. Photometry can be extracted from the images using the following expression in the instrumental system

\[ m = ZP - \kappa (X - 1) - 2.5 \log_{10}(f/t) + c \]  

where \( m \) is the calibrated magnitude, \( ZP \) is the zero point at unit airmass as extracted from the image headers, \( \kappa \) is the extinction coefficient, \( X \) is the airmass, \( f \) is the flux in counts and \( t \) is the exposure time. The \( c \) correction is based

and celestial coordinates. This happens in the pipeline in two generic stages.

An initial WCS based on knowledge of the instrument, e.g. orientation, field-scale, telescope pointing, is embedded in the FITS headers, with telescope-specific information in the primary header and detector-specific information in the secondary headers. This serves to locate each detector image to within a few to several arcsec, depending on the pointing accuracy of the telescope and model parameters. The essential information required is the RA and Dec of the pointing, a (stable) reference point on the detector grid for those coordinates (e.g. the optical axis of the instrument), the central pixel scale, the rotation of the camera, the relative orientation of each detector and the geometrical distortion of the telescope and camera optics, which defines the astrometric projection to use.

Given a rough WCS for the processed frames, a more accurate WCS can be defined using astrometric standards. We have based our calibration on the 2MASS point source catalog (Skrutskie et al. 2006) for several reasons: it is an all-sky NIR survey; it is calibrated on the International Celestial Reference System (ICRS); it provides at least 100 or more suitable standards per pointing; it is a relatively recent epoch (mid-1990s) minimising proper motion problems; the global systematics are better than 100 mas over the entire sky (Zacharias et al. 2003); and for 2MASS point sources with signal-to-noise > 10:1 the rms accuracy per source is \( \lesssim 100 \) mas.

By stacking the astrometric residuals from a series of independent pointings and CCD WCS solutions it is possible to assess the accuracy of the simple INT distortion model. This is illustrated in the figure 11 using the average residuals from a stack of a one week run. The end product of the full pipeline currently has an astrometric precision better than 100 mas over the whole array (i.e. across CCDs), as determined by analysis of independently calibrated adjacent overlapping pointings.

Figure 10. Star counts along two contrasted lines of sight with star densities of \( \sim 50,000 \) deg\(^{-2}\) and different reddening.
on the median dark sky recorded in science frames compared to the median for all the CCDs and as such is an ancillary correction to the gain correction derived from the flatfield (usually twilight flats). This correction is in the header keyword PERCORR and is typically smaller than 1 per cent and also corrects for the small quantum efficiency differences between the detectors.

Astrometric information is written in the FITS headers using a ZPN projection which allows for a description of the non uniform scale over the field of view without performing pixel resampling (necessary e.g. in the case of describing the distortions using a more usual tangential projection).

The aim of conventional flatfielding is to create a flat background by normalising out perceived variations from uniformly illuminated frames. If the sky area per pixel changes then this is reflected in a systematic error in the derived photometry. However the change of scale, i.e., the change of sky area per pixel, also creates photometric effects. The magnitude of the multiplicative correction to be applied to the measured flux can be modelled as (Irwin et al. 2008, in prep)

\[ d = \left( 1 + \frac{3P_2R^2}{P_1} \right) \left( 1 + \frac{P_2R^2}{P_1} \right) \]

where \( R \) is the distance from the optical axis and for the WFC the coefficients are \( P_1 = 1 \) and \( P_2 = 220 \) (corresponding to the PV2_1 and PV2_3 WCS keywords as defined in Greisen and Calabretta (2002)). Figure 12 shows the effect of this term in the photometry which is at a maximum level of only 2 per cent at the outer parts of the frames. The value displayed in the figure has to be subtracted from the magnitude calculated in equation 4 and is automatically included in the catalogue data products but has to be applied by the user when measuring fluxes from the images. Table 3 lists the main photometric keywords present in the image headers and needed to convert the fluxes to magnitudes as explained in this section.

### 4.1.1 Confidence maps

We define a confidence \( c_{ij} \) map, where \( j \) refers to pixel \( j \) of frame \( i \), that is an inverse variance weight map in units of relative variance denoting the "confidence" associated with the flux value in each pixel. The \( c_{ij} \) are normalised to a median level of 100 per cent (i.e. \( \langle c_{ij} \rangle = 1 \)). This has the advantage that the same map can also be used to encode for hot, bad or dead pixels, by assigning zero confidence.

The confidence map for each frame is derived from analysis of the flat fields and is unique for each filter/detector combination. It encodes individual pixel sensitivities and allows for vignettet or bad regions to be masked. These confidence maps are used in the object detection phase of the processing.

Use of the confidence maps for weighted co-addition of frames, or for object detection, simply requires an overall estimate of the average noise properties of the frame. This can be readily derived from the measured sky noise, in the
Poisson noise-limited case, or from a combination of this and the known system characteristics (e.g. gain, readout noise). All processed frames have an associated derived confidence map which is propagated through the processing chain. A detailed description of the process is given in Irwin et al. 2008, in prep.

4.2 Catalogues

The object catalogue is made available using a relational database system. We have followed the philosophy of reducing the number of tables presented to the user as much as possible. The database contains four tables which are summarised below.

- **FieldsAll** (table 4). Information about the field centres and number of objects detected. Each field pair has associated a unique field number (fieldno) and each of those an on/off position (onoff where a value of 0 has been assigned to the off position and 1 to the on position). These two are combined into the fieldID to provide a unique identifier for each pointing. Equatorial and galactic coordinates for each pointing as well as number of detections are also provided.

- **ChipsAll** (table 5). This table contains parameters about the observation (airmass, exposure time, World Coordinate System) as well as the Data Quality Control attributes for each chip, i.e., mean seeing and ellipticity, zero point, sky level, etc. Also for each pointing the Galactic coordinates are available as well as the Galactic extinction as measured from the Schlegel maps (Schlegel et al. 1998) using a 4 pixel linear interpolation for each position.

- **PhotoObj** (table 6). Main catalogue containing all detections with derived parameters and calibrated magnitudes. Each row has a set of keys (chip_r_id, chip_i_id, chip_ha_id, field_id) which allow the user to enquire about the photometric conditions (via the ChipsAll table) under which the measurements of a particular object have been done or to which field it belongs. Additional flags (e.g. seeing_flag) allow the selection of objects observed in particular (seeing) conditions. We have also included the Hierarchical Triangular Mesh index (htmid) as defined by the Sloan Digital Survey Science Archive ². The total number of detections in this catalogue is 411 million including duplicate observations. Due to the on/off observation strategy which leads to every source being observed twice this means about 200 million unique objects.

- **PhotoObjBest** (table 6). Subset of PhotoObj with additional constraints used to select only those observations obtained in good photometric conditions and thus will vary little between the IDR and the Final Data Release. The data obtained in the Aug 2003 and Dec 2004 observing runs have been discarded due to poor photometric conditions. This catalogue contains only observations for those nights for which the mean magnitude zero point rms in the r' band is better than 5 per cent, the mean r' zero point is greater than 24.2 and the individual frames have been obtained in seeing conditions better than 2 arcsec and ellipticity better than 0.2. Additionally only sources detected in all three bands have been included. The total number of detections in this catalogue is 181 million (approximately 90 million unique objects).

- **tGetNearbyObj**. This is a store procedure to perform a cone search without having to use the complete haversine formula explicitly to calculate angular distances. Given the coordinates (ra, dec) in degrees and a radius (r) in arcmin it returns a list of identifiers for the objects in such defined cone. The output can then be joined with one of the PhotoObj/Best tables to return the object parameters (see example in appendix A).

Tables 4, 5 and 6 list the column names and their description. Columns marked with an asterisk are indexed by the database software, i.e., searches in those columns are particularly fast. Users should use when possible at least one indexed column in order to speed up the queries.

Each entry in each table is assigned a unique ID number used for cross reference. Figure 13 shows an object ID in the PhotoObj/Best tables and how it is decoded. Each entry in the ChipsAll table is encoded in the same way but without the last five figures which represent the object number.

Individual detection star/galaxy classifications are combined into one single value following a similar approach to the WFCAM Science Archive (WSA; Hambly et al. 2007). A classification table is defined which assigns probability values for each classification code and then each classification is combined for a merged source using Bayesian classification rules assuming each value is independent:

\[
P(c_k) = \prod_j P(c_k)_j / \sum_k \prod_j P(c_k)_j
\]

where \(c_k\) is the classification flag and \(i\) denotes the \(i^{th}\) single detection passband measurement available. Decision thresholds for the resulting discrete classification flag are 90 per cent for definitive and 70 per cent for probable. An additional decision rule enforces class_flag=255 (saturated) when any individual classification flag indicates saturation.

5 CROSS MATCHING WITH UKIDSS GPS

In Drew et al. (2005) we describe the importance of the IPHAS data by themselves and the several papers published to date using just IPHAS data are clearly a demonstration of that. In this section however, we want to focus in the multwavelength aspect which the present IPHAS release will facilitate. We provide an example of cross matching

2 see http://skyserver.org/htm/
<table>
<thead>
<tr>
<th>Column Name</th>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>chipID*</td>
<td>long</td>
<td>chip ID</td>
</tr>
<tr>
<td>runno*</td>
<td>int</td>
<td>Run number</td>
</tr>
<tr>
<td>ra*</td>
<td>float</td>
<td>R.A. of the chip center (deg, J2000)</td>
</tr>
<tr>
<td>dec*</td>
<td>float</td>
<td>Dec of the chip center (deg, J2000)</td>
</tr>
<tr>
<td>glon</td>
<td>float</td>
<td>Galactic longitude (deg)</td>
</tr>
<tr>
<td>glat</td>
<td>float</td>
<td>Galactic latitude (deg)</td>
</tr>
<tr>
<td>ebv</td>
<td>real</td>
<td>E(B – V) from Schlegel dust map</td>
</tr>
<tr>
<td>naxis1</td>
<td>int</td>
<td>Chip size along x axis (pixels)</td>
</tr>
<tr>
<td>naxis2</td>
<td>int</td>
<td>Chip size along y axis (pixels)</td>
</tr>
<tr>
<td>mjd</td>
<td>float</td>
<td>MJD of observation</td>
</tr>
<tr>
<td>magzpt*</td>
<td>real</td>
<td>Photometric zero point (Vega)</td>
</tr>
<tr>
<td>magzrr</td>
<td>real</td>
<td>Zero point error (Vega)</td>
</tr>
<tr>
<td>extinct*</td>
<td>real</td>
<td>Atmospheric Extinction</td>
</tr>
<tr>
<td>airmass*</td>
<td>real</td>
<td>Airmass</td>
</tr>
<tr>
<td>exptime</td>
<td>real</td>
<td>Exposure time (seconds)</td>
</tr>
<tr>
<td>night*</td>
<td>int</td>
<td>Night of observation (YYYYMMDD)</td>
</tr>
<tr>
<td>band*</td>
<td>char</td>
<td>Filter name</td>
</tr>
<tr>
<td>maglim</td>
<td>real</td>
<td>Magnitude limit (5σ, Vega)</td>
</tr>
<tr>
<td>seeing*</td>
<td>real</td>
<td>Mean seeing (arcsec)</td>
</tr>
<tr>
<td>elliptic*</td>
<td>real</td>
<td>Mean stellar ellipticity</td>
</tr>
<tr>
<td>cdL1</td>
<td>float</td>
<td>CD Matrix</td>
</tr>
<tr>
<td>cdL1</td>
<td>float</td>
<td>CD Matrix</td>
</tr>
<tr>
<td>crval1</td>
<td>float</td>
<td>R.A. of reference pixel (deg)</td>
</tr>
<tr>
<td>crval2</td>
<td>float</td>
<td>Dec of reference pixel (deg)</td>
</tr>
<tr>
<td>crpix1</td>
<td>float</td>
<td>x coordinate of reference pixel</td>
</tr>
<tr>
<td>crpix2</td>
<td>float</td>
<td>y coordinate of reference pixel</td>
</tr>
<tr>
<td>percorr</td>
<td>real</td>
<td>Sky connection (Vega)</td>
</tr>
<tr>
<td>apcor</td>
<td>real</td>
<td>Aperture correction (Vega)</td>
</tr>
<tr>
<td>skylevel</td>
<td>real</td>
<td>Sky level (counts)</td>
</tr>
<tr>
<td>skynoise</td>
<td>real</td>
<td>Sky noise (counts)</td>
</tr>
<tr>
<td>nobj</td>
<td>long</td>
<td>number of detected objects</td>
</tr>
<tr>
<td>fieldJd*</td>
<td>long</td>
<td>link to FieldsAll table</td>
</tr>
</tbody>
</table>

Table 5. Column names, types and description of data quality control table ChipsAll. Columns marked with an asterisk are indexed.

<table>
<thead>
<tr>
<th>Field</th>
<th>Date Obs.</th>
<th>CCD</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>

Figure 13. The unique object ID encodes the field number, on/off flag (1=on, 0=off), date of observation (in the format YYYYMMDD), band (1=r'; 2=i'; 3=Ha) and CCD number.

We have selected a contiguous area of 2 square degrees centred at α = 20h48m, δ = 42°45' (l,b = 83.0, −0.5) in the Cygnus-X region and extracted all objects from the Pho­toObjBest catalogue. From the Schlegel et al. (1998) maps the extinction is E(B – V) = 3. The total number of objects in that area is 235,231. In order to assess the photometry the sources extracted from the IPHAS catalogue with those from the UKDSS GPS. Near infrared data are required to probe the many parts of the galaxy which suffer high extinction and an ESO-public near infrared survey of the northern plane, the UKIDSS Galactic Plane Survey, has recently begun.

We have grouped all the objects by field and used the field overlap algorithm as described in Glazebrook et al. (1994) in order to offset the photometry to a common reference system.

For this exercise we have just set one frame as the reference calibrated frame and allowed the others to vary and solved the matrix equation for each band. The resultant magnitude offsets between the different fields is smaller than...
objects brighter than $r' = 13$ and $K = 12$ to avoid saturation and finally we have removed all sources within 10 arcsec of any WFCAM chip edge (although problematic sources in the edges are anyway flagged by the pipeline classifier).

The main sequence of objects at zero reddening is well defined in the $(r' - H\alpha, r' - K)$ plane with a large population of objects up to the M stars. These data are for fields in the Cygnus-X part of the plane where there is a local essentially unreddened stellar population to a few hundred parsecs, and then a decline in stellar density and sharp rise in molecular gas and extinction to a distance of a kpc or so. Beyond this, the stellar density recovers as distant associations are traversed but of course it presents as a much more heavily reddened population. Furthermore stars later than spectral type mid-K are intrinsically faint and are only detectable at $E(B - V) < 0.8$ in the figure 16 where our limit is at $r' - K < 7$. This explains the lack of K and M dwarfs reddened objects in these diagrams.

Note that no artificial offsets have been introduced to the synthetic tracks to make them fit with the observed colours. This is further confirmation of the good photometric calibration of both the IPHAS and GPS catalogues.

6 USER ACCESS

User access to the catalogues and images has been provided by the AstroGrid project. There are two separate user interfaces to the IPHAS IDR. Users who only want to extract fluxes or images around a small position of sky (there is a limit on the radius of 1 degree) will find the web interface useful. Given a position or an object name the user will obtain a cone search of the PhotoObjBest catalogue in a variety of output formats (VOTable, Binary VOTable, Comma Separated Value or HTML). Input lists of objects are also supported. Together with the IPHAS catalogue we provide access from the same interface to other useful VO compliant archives like SDSS DR5 (Adelman-McCarthy et al. 2007), 2MASS Point Source Catalogue (Skrutskie et al. 2006), NED and the Carlsberg Meridian Astrometric Catalogue (Evans et al. 2002). Postage stamps generation for a particular input position is also available using the finding chart utility (see example output in figure 17).

Users who want more sophisticated queries on flags and algebraic combinations of parameters (e.g. colours) will find the AstroGrid Query Builder more suited to their needs. The Query Builder is an application inside the AstroGrid VO Desktop which allows to submit an arbitrary query to the database system. The Query Builder facilitates the task of building complex queries written in Astronomical Dataset Query Language (ADQL); an international agreed standard based on the Structured Query Language – SQL), save the queries and submit them asynchronously to the server (this facilitates sending long running queries and retrieve the results later).

---

3 STILTS is a set of command line tools to perform operations with tables like cross matching (http://www.star.bris.ac.uk/~mbt/stilts/)

4 Available from http://www.iphas.org/idr

5 http://www.ivoa.net/Documents/latest/ADQL.html
Figure 15. Reddening tracks in the $r - H\alpha$ versus $r - K$ plane for main sequence dwarfs (left) and giants (right).

Figure 16. Colour colour plots for point like sources detected in IPHAS and GPS with main sequence tracks overlayed (continuous line: main sequence dwarfs; dashed-line: giants) for reddening: $E(B - V) = 0$ (blue) and 2 (red))

Figure 17. Example postage stamps obtained using the finding chart utility. Each stamp is labeled with the run number, CCD number and filter. The object corresponds to the Principe de Asturias quadrupolar nebula presented in Mampaso et al. (2006).
Table 8. Synthetic tracks in the \((r' - J, J - K)\) plane, for main sequence dwarfs and giants, calculated for a range of reddening. A similar table with optical colours only can be found in table 2 of Drew et al. 2005.

<table>
<thead>
<tr>
<th>Spectral type</th>
<th>(E(B - V) = 0.0)</th>
<th>(E(B - V) = 1.0)</th>
<th>(E(B - V) = 2.0)</th>
<th>(E(B - V) = 3.0)</th>
<th>(E(B - V) = 4.0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>((r' - J))</td>
<td>((J - K))</td>
<td>((r' - J))</td>
<td>((J - K))</td>
<td>((r' - J))</td>
<td>((J - K))</td>
</tr>
<tr>
<td>O5 V</td>
<td>-0.955</td>
<td>0.331</td>
<td>1.415</td>
<td>0.246</td>
<td>3.742</td>
</tr>
<tr>
<td>O9 V</td>
<td>-0.826</td>
<td>0.252</td>
<td>1.543</td>
<td>0.323</td>
<td>3.869</td>
</tr>
<tr>
<td>B0 V</td>
<td>-0.720</td>
<td>0.139</td>
<td>1.649</td>
<td>0.437</td>
<td>3.976</td>
</tr>
<tr>
<td>B1 V</td>
<td>-0.674</td>
<td>0.107</td>
<td>1.696</td>
<td>0.467</td>
<td>4.011</td>
</tr>
<tr>
<td>B3 V</td>
<td>-0.706</td>
<td>0.283</td>
<td>1.658</td>
<td>0.294</td>
<td>3.980</td>
</tr>
<tr>
<td>B8 V</td>
<td>-0.249</td>
<td>0.000</td>
<td>2.030</td>
<td>0.531</td>
<td>4.423</td>
</tr>
<tr>
<td>B9 V</td>
<td>-0.145</td>
<td>0.003</td>
<td>2.108</td>
<td>0.573</td>
<td>4.530</td>
</tr>
<tr>
<td>A0 V</td>
<td>0.005</td>
<td>-0.005</td>
<td>2.366</td>
<td>0.569</td>
<td>4.684</td>
</tr>
<tr>
<td>A2 V</td>
<td>-0.139</td>
<td>-0.088</td>
<td>2.220</td>
<td>0.489</td>
<td>4.536</td>
</tr>
<tr>
<td>A3 V</td>
<td>-0.155</td>
<td>-0.131</td>
<td>2.202</td>
<td>0.446</td>
<td>4.517</td>
</tr>
<tr>
<td>A5 V</td>
<td>0.324</td>
<td>0.085</td>
<td>2.678</td>
<td>0.658</td>
<td>4.990</td>
</tr>
<tr>
<td>A7 V</td>
<td>0.390</td>
<td>0.132</td>
<td>2.739</td>
<td>0.705</td>
<td>5.047</td>
</tr>
<tr>
<td>F0 V</td>
<td>0.543</td>
<td>0.159</td>
<td>2.888</td>
<td>0.731</td>
<td>5.192</td>
</tr>
<tr>
<td>F2 V</td>
<td>0.475</td>
<td>0.057</td>
<td>2.819</td>
<td>0.632</td>
<td>5.122</td>
</tr>
<tr>
<td>F5 V</td>
<td>0.859</td>
<td>0.221</td>
<td>3.202</td>
<td>0.793</td>
<td>5.503</td>
</tr>
<tr>
<td>F6 V</td>
<td>0.909</td>
<td>0.218</td>
<td>3.249</td>
<td>0.788</td>
<td>5.547</td>
</tr>
<tr>
<td>G8 V</td>
<td>0.955</td>
<td>0.329</td>
<td>3.333</td>
<td>0.802</td>
<td>5.630</td>
</tr>
<tr>
<td>K0 V</td>
<td>1.210</td>
<td>0.310</td>
<td>3.543</td>
<td>0.882</td>
<td>5.836</td>
</tr>
<tr>
<td>K2 V</td>
<td>1.528</td>
<td>0.499</td>
<td>3.853</td>
<td>1.068</td>
<td>6.138</td>
</tr>
<tr>
<td>K3 V</td>
<td>1.899</td>
<td>0.556</td>
<td>4.220</td>
<td>1.215</td>
<td>6.501</td>
</tr>
<tr>
<td>K4 V</td>
<td>2.096</td>
<td>0.648</td>
<td>4.411</td>
<td>1.216</td>
<td>6.686</td>
</tr>
<tr>
<td>K5 V</td>
<td>2.245</td>
<td>0.674</td>
<td>4.560</td>
<td>1.241</td>
<td>6.835</td>
</tr>
<tr>
<td>M0 V</td>
<td>2.992</td>
<td>0.781</td>
<td>5.296</td>
<td>1.348</td>
<td>7.561</td>
</tr>
<tr>
<td>M2 V</td>
<td>3.418</td>
<td>0.807</td>
<td>5.961</td>
<td>1.375</td>
<td>7.971</td>
</tr>
<tr>
<td>M3 V</td>
<td>3.935</td>
<td>0.820</td>
<td>6.337</td>
<td>1.388</td>
<td>8.482</td>
</tr>
<tr>
<td>M4 V</td>
<td>4.560</td>
<td>0.832</td>
<td>6.830</td>
<td>1.399</td>
<td>9.062</td>
</tr>
<tr>
<td>O8 III</td>
<td>-0.813</td>
<td>-0.076</td>
<td>1.550</td>
<td>0.496</td>
<td>3.870</td>
</tr>
<tr>
<td>B1-2 III</td>
<td>-0.603</td>
<td>-0.198</td>
<td>1.766</td>
<td>0.376</td>
<td>4.092</td>
</tr>
<tr>
<td>B3 III</td>
<td>-0.427</td>
<td>0.018</td>
<td>1.935</td>
<td>0.591</td>
<td>4.253</td>
</tr>
<tr>
<td>B9 III</td>
<td>-0.206</td>
<td>-0.070</td>
<td>2.153</td>
<td>0.504</td>
<td>4.469</td>
</tr>
<tr>
<td>B5 III</td>
<td>-0.361</td>
<td>0.078</td>
<td>2.000</td>
<td>0.497</td>
<td>4.318</td>
</tr>
<tr>
<td>A0 III</td>
<td>-0.030</td>
<td>0.020</td>
<td>2.326</td>
<td>0.593</td>
<td>4.641</td>
</tr>
<tr>
<td>A3 III</td>
<td>-0.119</td>
<td>-0.139</td>
<td>2.238</td>
<td>0.437</td>
<td>4.553</td>
</tr>
<tr>
<td>A5 III</td>
<td>0.323</td>
<td>0.101</td>
<td>2.675</td>
<td>0.673</td>
<td>4.985</td>
</tr>
<tr>
<td>A7 III</td>
<td>0.285</td>
<td>-0.003</td>
<td>2.638</td>
<td>0.572</td>
<td>4.972</td>
</tr>
<tr>
<td>F0 III</td>
<td>0.565</td>
<td>0.199</td>
<td>2.911</td>
<td>0.770</td>
<td>5.216</td>
</tr>
<tr>
<td>F2 III</td>
<td>0.595</td>
<td>0.100</td>
<td>2.936</td>
<td>0.673</td>
<td>5.236</td>
</tr>
<tr>
<td>F5 III</td>
<td>0.876</td>
<td>0.291</td>
<td>3.214</td>
<td>0.861</td>
<td>5.510</td>
</tr>
<tr>
<td>G0 III</td>
<td>1.458</td>
<td>0.424</td>
<td>3.791</td>
<td>0.992</td>
<td>6.083</td>
</tr>
<tr>
<td>G3 III</td>
<td>1.681</td>
<td>0.477</td>
<td>4.098</td>
<td>1.046</td>
<td>6.294</td>
</tr>
<tr>
<td>G8 III</td>
<td>1.727</td>
<td>0.516</td>
<td>4.053</td>
<td>1.086</td>
<td>6.339</td>
</tr>
<tr>
<td>K0 III</td>
<td>1.844</td>
<td>0.541</td>
<td>4.167</td>
<td>1.110</td>
<td>6.449</td>
</tr>
<tr>
<td>K1 III</td>
<td>1.961</td>
<td>0.571</td>
<td>4.283</td>
<td>1.140</td>
<td>6.564</td>
</tr>
<tr>
<td>K2 III</td>
<td>2.188</td>
<td>0.671</td>
<td>4.506</td>
<td>1.239</td>
<td>6.783</td>
</tr>
<tr>
<td>K5 III</td>
<td>2.948</td>
<td>0.864</td>
<td>5.253</td>
<td>1.431</td>
<td>7.519</td>
</tr>
<tr>
<td>M0 III</td>
<td>3.190</td>
<td>0.968</td>
<td>5.493</td>
<td>1.534</td>
<td>7.756</td>
</tr>
<tr>
<td>M1 III</td>
<td>3.364</td>
<td>0.987</td>
<td>5.564</td>
<td>1.553</td>
<td>7.925</td>
</tr>
<tr>
<td>M2 III</td>
<td>3.586</td>
<td>1.068</td>
<td>5.881</td>
<td>1.633</td>
<td>8.135</td>
</tr>
<tr>
<td>M3 III</td>
<td>3.910</td>
<td>1.062</td>
<td>6.200</td>
<td>1.627</td>
<td>8.450</td>
</tr>
<tr>
<td>M4 III</td>
<td>4.464</td>
<td>1.086</td>
<td>6.744</td>
<td>1.652</td>
<td>8.983</td>
</tr>
<tr>
<td>M5 III</td>
<td>5.249</td>
<td>1.142</td>
<td>7.515</td>
<td>1.707</td>
<td>9.742</td>
</tr>
</tbody>
</table>
6.1 Image Access

The Simple Image Access Protocol (SIAP) defines a prototype standard for retrieving image data from a variety of astronomical image repositories through a uniform interface. The IPHAS images are available through a SIAP (ivo://uk.ac.cam.ast/IPHAS/images/SIAP) interface. The result of a query is, given a box centred at coordinates RA, Dec and a box size in degrees, a table of image CCDs which overlap the defined box. The table contains links to the processed image file itself. Alternatively the same list of images can be obtained with an appropriate SQL query (see below). At the moment only full image retrieval is supported. Although the graphical interface does not allow for a list of positions this is possible to accomplish from the command line interface.

6.2 Full Catalogue Access

The ingestion of all the catalogue data into a database provides the capability of using SQL in order to query the catalogue with any user defined constraint. The AstroGrid Data Set Access (DSA) provides a layer on top of any database system which makes it readily VO compliant and accessible from any VO compliant application. The examples in appendix A provide ADQL queries that can be used in the Query Builder to retrieve a variety of data.

6.3 Command line access

The AstroGrid Run Time (AR) provides a bridge between the user command line tools or programs and the VO. Catalogues can be accessed using this bridge from programming languages like Python, Perl, C or Java. There are several reasons to provide a command line interface to the VO. The more important one is the ability to perform batch queries for a list of objects and also to implement a call to a VO service inside a user written script that performs some analysis afterwards.

7 SUMMARY

We have presented the Initial Data Release of the IPHAS project. The total release contains photometry in $r^\prime$, $i^\prime$, and $H_0$ for about 200 million objects in the northern galactic plane comprising the largest photometric sample available imaged in narrowband $H_0$. We have also made available a subset of this catalogue, PhotoObjBest, which contains the best observations in terms of final survey quality as defined in the text and which comprises about 60 per cent of the final release. The legacy of the survey is increased by the extensive coverage appearing from other projects spanning the wavelength range range from X-rays to near-IR and sub-millimeter.

The data have been made available through a VO compliant interface provided by AstroGrid and accessible to all VO projects.

This initial release will be supplemented with data obtained from January 2006 onwards at a later date to be announced in the project web pages. At the time of writing, data-taking continues. There will also be a separate Final Release once the survey is entirely complete and all finally-accepted data have been corrected onto a fully-interlaced uniform photometric scale.


ACKNOWLEDGMENTS

This research has made use of data obtained using, or software provided by, the UK’s AstroGrid Virtual Observatory Project, which is funded by the Science & Technology Facilities Council and through the EU’s Framework 6 programme. This paper makes use of data from the UKIRT Infrared Telescope, operated on the island of La Palma by the ING in the Spanish Observatorio del Roque de los Muchachos of the Instituto de Astrofísica de Canarias. IPHAS has been a massive effort during 5 years of observations and the following are thanked for their contributions at the telescope: Andrew Cardwell, Jesus Corral-Santana, Pasqual Diago, James Furness, Antonio García, Wilbert van Ham, Marie Hrudkova, Avon Huxor, Mansura Jaigirdar, Rob King, Richard Parker, Mark Pringle, Aaron Robotham, Ernesto Rodríguez-Flores, Miguel Santander, Rachel Smith and Carolin Winkworth.

PJG, EvdB, GR and LMR are supported by NWO-VIDI grant 639.042.201 to PJG. DS acknowledges a STFC Advanced Fellowship as well as support through the NASA Guest Observer program. QP is supported by ANSTO Access to Major Research Facilities Programme (AMRFP).

This work is based in part on data obtained as part of the UKIRT Infrared Deep Sky Survey.

REFERENCES


© 2008 RAS, MNRAS 000, 1–17
APPENDIX A: EXAMPLE USAGE ACCESS

Example 1: Return coordinates, magnitudes and magnitude
errors of objects in a rectangular area.

This is the most basic query returning all objects in a
specific box with R.A. between 300 and 301 degrees and Dec
between 30 and 31 degrees.

SELECT P.ra, P.dec,
P.coremag_r, P.coremag_i, P.coremag_ha,
P.coremaggrr_r, P.coremaggerr_i, P.coremaggerr_ha
FROM PhotoObjBest as P
WHERE P.ra BETWEEN 300.0 AND 301.0
AND P.dec BETWEEN 30.0 and 31.0

Example 2: Return coordinates and magnitudes of objects in
a cone defined by its center at RA=300.0 and Dec=30.0 and
radius r=10 arcmin.

In this case instead of coding the whole trigonometry
in the query we use the auxiliary table tGet Nearby Obj
which returns the object IDs which satisfy a cone search
specified by the center coordinates (ra, dec) and radius (r)
in minutes of arc:

SELECT P.ra, P.dec,
P.coremag_r, P.coremag_i, P.coremag_ha,
P.coremaggrr_r, P.coremaggerr_i, P.coremaggerr_ha
FROM PhotoObjBest as P, tGetNearbyObj as G
WHERE G.ra=300.0 AND G.dec=30.0 AND G.r=10.0
AND P.objID = G.objID

Example 3: As previous example but only return objects clas-
sified as point-like in r and observed in conditions of seeing
best than 1.5 arcsec as measured in the r band.

In this query we are asking for information about the
quality of observations and this is stored in the ChipsAll
table. We need therefore to join the results with this table in
order to select only observation with good seeing conditions.

SELECT P.ra, P.dec,
P.coremag_r, P.coremag_i, P.coremag_ha,
P.coremaggrr_r, P.coremaggerr_i, P.coremaggerr_ha
FROM PhotoObjBest as P, tGetNearbyObj as G,
ChipsAll as CR
WHERE G.ra=300.0 AND G.dec=30.0 AND G.r=10.0
AND CR.seeing<1.5 AND (P.class_r=-1 OR P.class_r=-2)
AND P.objID = G.objID AND P.chips_r_id=CR.chipID

Example 4: Select all objects observed in a particular offset
field.

The following example returns all objects detected in a
particular observation from the (not cleaned) PhotoObj
catalogue.

SELECT * FROM PhotoObj as P
JOIN FieldsAll as F ON (F.fieldID = P.field_id)
WHERE F.field no=7012 AND F.onoff=0

© 2008 RAS, MNRAS 000, 1-17
Example 5: Return the number and position of fields observed in the best photometric conditions.

Section 4.2 explains the constraints used to define the catalogue of best observations. The following query returns the number, position and photometric characteristics of the fields which satisfy those conditions.

```
SELECT
   CC.night, F.fieldID, F.fieldno, F.onoff,
   F.ra, F.dec, F.glon, F.glat,
   T.magzpt_mean, T.magzrr_mean,
   CC.seeing, CC.elliptic
FROM
   ( SELECT
       C.night,
       sum(C.magzpt)/count(C.magzpt) as magzpt_mean,
       sum(C.magzrr)/count(C.magzrr) as magzrr_mean
   FROM FieldsAll F, ChipsAll C
   WHERE
       C.field_id = F.fieldID AND C.band='r'
   AND C.chipno=4
   GROUP BY night ) AS T,
   ChipsAll CC, FieldsAll F
WHERE
   T.night = CC.night AND CC.field_id = F.fieldID
   AND CC.chipno=4 AND CC.band='r' AND CC.seeing<2
   AND CC.elliptic<0.2 AND T.magzpt_mean>24.2
   AND T.magzrr_mean<0.05
ORDER BY night
```

The IDR web site (http://idr.iphas.org) provides further information as well as examples on accessing the catalogue from the command line.

This paper has been typeset from a \TeX/ \LaTeX\ file prepared by the author.