Science-Grade Imaging Data for HAWK-I, VIMOS, and VIRCAM: The ESO–UK Pipeline Collaboration

Mark Neeser[1]
Jim Lewis[2]
Greg Madsen[2]
Aybuke Yoldas[2]
Mike Irwin[2]
Armin Gabasch[1]
Lodovico Coccato[1]
César Enrique García-Dabó[1]
Martino Romaniello[1]
Wolfram Freudling[1]
Pascal Ballester[1]

[1] ESO
[2] Cambridge Astronomical Survey Unit, Institute of Astronomy, University of Cambridge, United Kingdom

A new chapter for ESO science-grade data has begun with the implementation of three new pipelines developed for the HAWK-I, VIMOS and VIRCAM instruments. The HAWK-I and VIMOS image archives at ESO have been completely reprocessed using these new pipelines, and these data are now publicly available. This article introduces the work done to bring these pipelines to the level of science-grade, their use in reprocessing ESO archival data, and their dissemination into ESO science operations and to the ESO community.

In early 2012, a project was started with the Cambridge Astronomical Survey Unit (CASU), at the University of Cambridge, to make significant improvements to a number of ESO data processing pipelines. The goal was to develop state-of-the-art pipelines that can deliver science-grade data for the High Acuity Wide-field *K*-band Imager (HAWK-I), the Visible and Infrared Survey Telescope for Astronomy (VISTA) IR Camera (VIRCAM) and the VIsible Multi Object Spectrograph (VIMOS) imaging, and to use these pipelines to reprocess the entire HAWK-I and VIMOS (imaging) archives.

Expertise within CASU covers a broad range of ground-based and space-based projects ranging from data processing and image analysis techniques, through to data curation and access to UK facility data archives. Modern-era wide-field digital surveys produce vast amounts of data and CASU is at the forefront of automatically processing and archiving these legacy products. Data processed by CASU include all of the public surveys being carried out at the VISTA telescope (Emerson et al., 2006), some surveys from the Very Large Telescope (VLT) Survey Telescope (VST), the European Space Agency Gaia and Planck surveys, the Isaac Newton Telescope (INT) wide-field survey and the UK Infrared Telescope (UKIRT) infrared deep sky survey, to name but a few.

The ESO–UK project has now passed its final review. The deliverables have included three new pipelines for VIRCAM, VIMOS imaging, and HAWK-I, along with the associated Reflex[1] workflows that will allow an individual user to run these pipelines. Also included are the pipeline manuals and the tutorials for the Reflex workflows. Further, CASU has used these new pipelines to reprocess the HAWK-I and VIMOS imaging archives. They have delivered all of the pipeline products (science and calibration) from April 2008 to October 2015 for HAWK-I and from April 2003 to the end of 2015 for VIMOS. All these high-level science-grade images are now publicly available via the Phase 3 interface to the ESO Science Archive Facility (SAF[2]).

The project

CASU's first involvement with software at ESO was its delivery of an operational pipeline for the VISTA/VIRCAM imager in 2008. This pipeline is currently in operational use on Paranal and by Garching quality control (QC[3]), not to provide science-grade data but to quickly assess the quality of the observations and the health of the instrument. However, much of the functionality available to the VIRCAM pipeline currently used by CASU to process the data from the VISTA public surveys does not exist in the operational version.

The consortium agreed to upgrade the ESO operational pipeline to a level of sophistication equivalent to that used by CASU to process the data from the public surveys. In essence, CASU would improve on the ESO VIRCAM pipeline so that the data produced is science-grade, with de-striping across multiple detectors, advanced sky correction, on-field astrometric and photometric calibration, creation of source catalogues, and an optimised combination of multiple dithered images into tiles. According to the original agreement, this effort was also to be extended to the creation of completely new pipelines for HAWK-I and the VIMOS imager with pipeline routines to be written to ESO's Common Pipeline Library (CPL). Finally, once the new HAWK-I and VIMOS pipelines were completed, they were to be used to reprocess all of the data from these two instruments stored in the SAF to the end of 2015.

Using the VIRCAM pipeline as the starting point of the collaboration, ESO and CASU worked towards improving it, and adapting its modules to the peculiarities of the HAWK-I and VIMOS instruments. Among these improvements is the computation of an error budget through each stage of the data processing. Each pipeline, each Reflex workflow, and all of the reprocessed data submitted have been intensively tested at ESO and CASU.

The pipelines

All three pipelines offer significant improvements over those currently in ESO operation, and have been thoroughly tested to ensure that they deliver science-grade products. A major improvement in the new pipelines is that they combine dithered images to deliver tiles (HAWK-I and VIRCAM) or image stacks (VIMOS). The images are astrometrically and photometrically calibrated using catalogue standard stars in the field of the science exposure. This is a significant improvement as it allows images to be calibrated even when photometric conditions are less than ideal. Of course, the pipelines can also process dedicated standard star fields for instrument quality control and site monitoring, and for the cases in which the science fields do not contain catalogue sources.

To date, the calibration plan for HAWK-I was to observe a single standard star sequentially in each of its four detectors. With the new HAWK-I pipeline this very inefficient observing mode will be supplanted by observations of dedicated Two Micron All-Sky Survey (2MASS) touchstone fields. Not only is this a four-fold improvement in observing efficiency, but the multiplicity of stars observed (some fields have more than 1300 2MASS stars within the field-of-view of HAWK-I) will greatly improve the photometry statistics.

The relative astrometric accuracy achieved is about 0.3 arcseconds root mean square (RMS) for HAWK-I and 0.4 arcseconds RMS for VIMOS (see below for an analysis of the reprocessed archive data). This is at least a five-fold improvement over the current pipelines, which simply rely on the world coordinates written by the telescope and have no astrometric correction routines.

Further pipeline improvements include a de-striping routine to detect and remove the pick-up noise across the rows of the VIRCAM detectors, a master twilight flat routine that is significantly faster than what is currently available in the HAWK-I pipeline, a sophisticated sky background correction routine that can chose from a number of sky correction algorithms based on the type of science observation, and a fringe correction routine for VIMOS. The latter improvement is particularly relevant for data obtained with the *I*- and *z*-band filters using the old VIMOS detectors.

A comparison between several hundred HAWK-I images in the SAF and their counterparts processed with the new pipeline has revealed that improvements in the flat-fielding and sky background correction reduced the background standard deviation by, on average, 20 %. This improvement is visually evident in Figure 1.

Throughout the entire data reduction cascade, the new pipelines also compute an error budget that is saved as a variance map associated with the final science products. This can act as a useful diagnostic when comparing the actual image noise statistics to these computed values. The image stacks and tiles produced by the pipelines also include confidence maps.

Each pipeline product generated includes a comprehensive number of quality control parameters, useful to the operational monitoring of these instruments, and culminating in science product parameters, such as average stellar image full width at half maximum (FWHM) and ellipticity, limiting magnitude and sky background brightness, to name but a few. The image stacks and tiles culminate in source extraction catalogues that include 62 scientifically useful parameters, including a number of flux measures, a series of 13 aperture fluxes stepped through a progression of radii, object classification and statistics.

It is fortuitous that HAWK-I now has a new and vastly improved pipeline since the Adaptive Optics Facility (AOF) and the GRound layer Adaptive optics Assisted by Lasers (GRAAL) are expected to begin commissioning in early 2017 (see Arsenault et al., 2016). This four-laser guide star system and deformable secondary mirror, will provide HAWK-I with almost space-based observatory image quality. The new HAWK-I pipeline will be needed to assess the performance of the AOF and GRAAL during commissioning, and to provide science products worthy of the expected image quality.

The Reflex pipeline workflows

In order to make the new pipelines readily accessible to the entire ESO community, the CASU team has integrated them into the Reflex workflow environment (Freudling et al., 2013). The Reflex workflows allow the user to gain an immediate and intuitive understanding of the structure, flow and products produced by these complex pipelines. The workflows delivered by CASU are state-of-the-art in terms of their multi-level interactive actors and in the large number of ways by which a user can modify the underlying process and adapt it for use with the peculiarities of their data. For example, when running the HAWK-I Reflex workflow, the user can conveniently modify 46 individual pipeline parameters, and is presented with interactive product plots and diagnostics to optimise the science results.

These new Reflex workflows will be made publicly available following the April 2017 pipeline release[4]. They will include the Reflex workflow and its underlying pipeline, a workflow specific tutorial, and a demonstration data set that can be processed right out of the box. A screen capture of the HAWK-I Reflex workflow is shown in Figure 2.

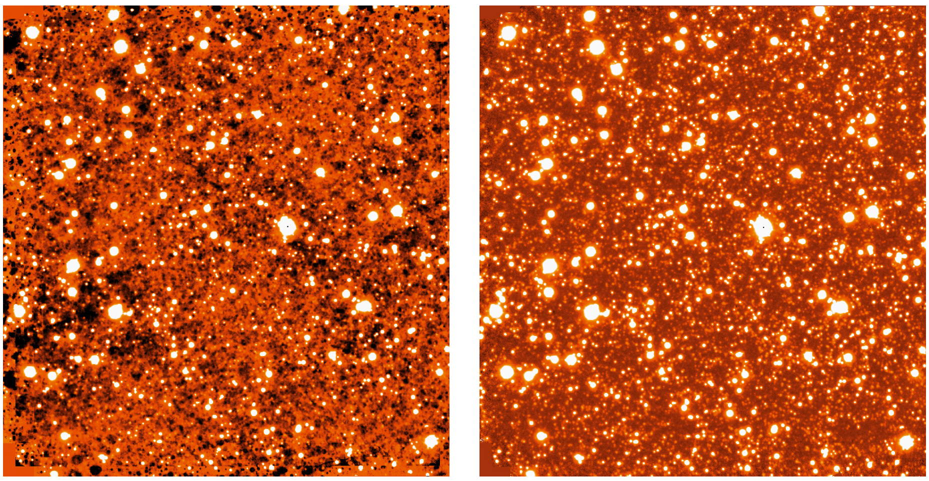


Figure 1. A comparison of a field processed with the current ESO HAWK-I pipeline (left panel) and with the new HAWK-I pipeline (right panel). The improvements in the twilight flat fields and in the sky background subtraction are evident in the new pipeline.

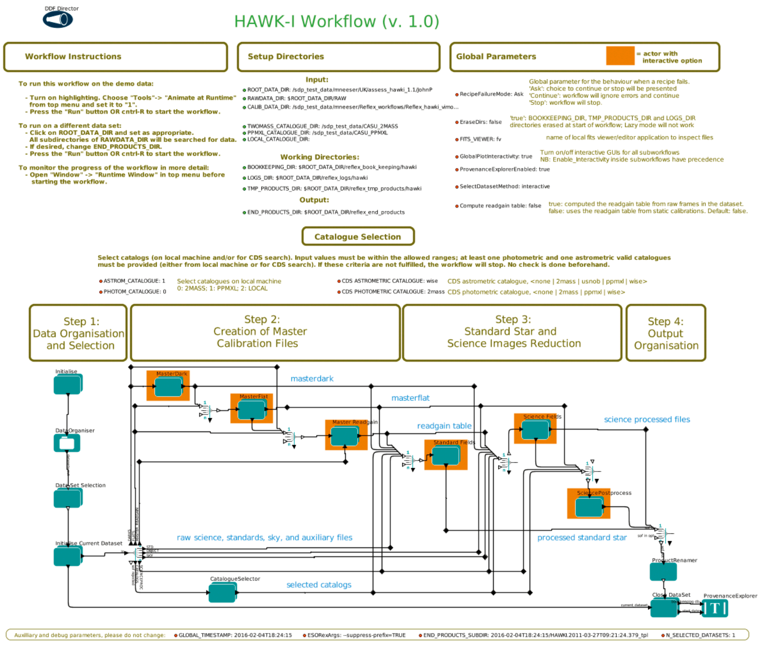


Figure 2. Screenshot of the Reflex workflow for the HAWK-I pipeline. Similar workflows also exist for the VIMOS imaging and VIRCAM pipelines.

HAWK-I and VIMOS imaging archive reprocessing

The CASU team has reprocessed the complete ESO image archives for HAWK-I (almost eight years of data) and VIMOS (12 years of data) using the new pipelines. Spanning the observing Periods P81 to P96, the more than 40 000 HAWK-I science frames consist of:
- stacked jittered images of the individual exposures at the detector level;
- variance arrays for the stacks;
- confidence maps for the stacks;
- source catalogues for the stacks;
- fully tiled images of all of the individual exposures;
- variance arrays for the tiles;
- confidence maps for the tiles;
- source catalogues for the tiles.

Spanning the Periods P71 to P96, the more than 37 000 VIMOS science frames consist of:
- stacked jittered images of the individual exposures at the detector level;
- variance arrays for the stacks;
- confidence maps for the stacks;
- source catalogues for the stacks.

Along with these science products, the CASU team has also delivered all of the master calibration frames (including the standard star fields and their source catalogues) used to process these data. The combined 5.1 Tb of reprocessed science data and 4.3 Tb of master calibrations are a significant contribution to the ESO science and calibration archives.

The challenge in processing these data lay in the tremendous inhomogeneity of the HAWK-I and VIMOS raw data. The data consisted of 290 and 456 unique observing programmes for HAWK-I and VIMOS, respectively, containing observations from many different projects and surveys, using various filters and observing methods. A summary of the statistics related to the science data from the archive reprocessing is given in Table 1.

To determine the overall astrometric quality of the images, a source-by-source match was made between all of the HAWK-I and VIMOS object catalogues and the 2MASS catalogue. The astrometric solutions are generally good to 300 milli-arcseconds, with no discernible systematic residuals; however, because some projects used defocused images, some images can give significantly worse solutions with residuals as large as 500 milli-arcseconds.

Similarly, a photometric comparison was made between the sources in the HAWK-I object catalogues and the 2MASS catalogue. This comparison was done using all of the 11 available HAWK-I filters (including the narrow-band filters). It was found that 82.4% of sources have magnitude differences of less than 0.1 mag. (92.3% have $|\Delta m| < 0.2$ magnitudes).

For both the astrometric and photometric quality, it is apparent that these results are slightly worse than the internal errors of the 2MASS catalogue (Skrutskie et al., 2006). The slightly larger mean internal RMS can be explained by the fact that the bright catalogue stars are invariably saturated in the HAWK-I and VIMOS images and therefore only the fainter, less reliable, catalogue stars are available for calibration. Also, a better photometric result can be achieved by limiting the matching to the core broad-band HAWK-I filters. These astrometric and photometric comparisons are summarised in Table 2 and in Figures 3 and 4.

	HAWK-I	VIMOS
Time span	April 2008 to October 2015	April 2003 to August 2015
Number of unique programmes	290	456
Number of raw science images	86 054	79 947
Number of delivered science products	40 552	37 168
Volume of science data products (Tb)	3.24	1.85
Non-contiguous sky coverage (square degrees)*	102.4	501.8
Number of detected sources†	19.1×10^6	150.6×10^6

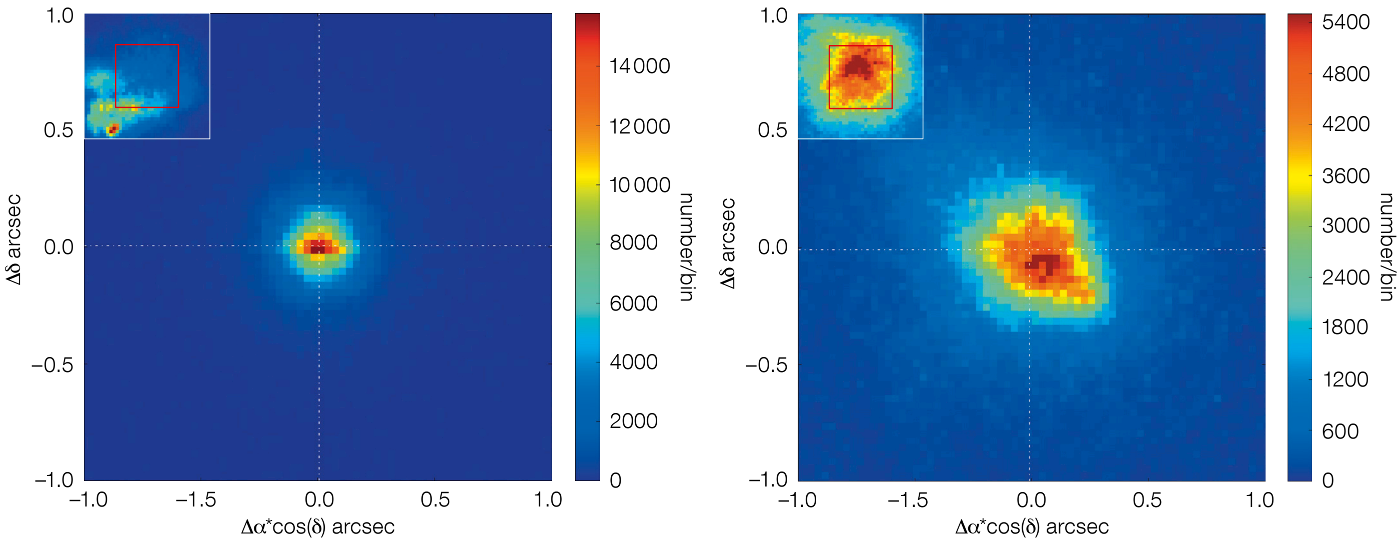
Table 1. Phase 3 statistics for HAWK-I and VIMOS science reprocessing.

* Computed from the sum of the area occupied by all HAWK-I tiles and VIMOS stacks.
† Sum of sources contained in all of the catalogues.

Table 2. Summary of the astrometric and photometric accuracies resulting from matching the HAWK-I tile source lists with the Wide-field Infrared Survey Explorer (WISE) and 2MASS catalogues, and the VIMOS stack source lists with the PPMXL catalogue (Roeser et al., 2010) and the American Association of Variable Star Observers (AAVSO) Photometric All-Sky Survey (APASS) catalogue.

	HAWK-I	VIMOS
Median $\Delta\alpha$ x cos(δ) (arcseconds)	0.006 +/− 0.3	0.000 +/− 0.4
Median $\Delta\delta$ (arcseconds)	0.008 +/− 0.3	−0.018 +/− 0.4
Median Δmag (mag)	0.003 +/− 0.2	0.02 +/− 0.3

Figure 3 (below). The astrometric quality of the HAWK-I reprocessed tiles (left panel) and VIMOS stacks (right panel). The HAWK-I source lists have been compared with the 2MASS catalogue and the VIMOS source lists with the PPMXL catalogue. The standard deviation of the $\Delta\alpha*\cos(\delta)$ and $\Delta\delta$ distribution is 0.3 arcseconds for HAWK-I and 0.4 arcseconds for VIMOS. The same plots for data from the old pipelines are shown as insets, where the scale of the insets is two times greater (4 arcseconds per side) to accommodate the standard deviation of the $\Delta\alpha*\cos(\delta)$ and $\Delta\delta$ distributions of > 1.7 arcseconds. The relative sizes of the new pipeline distributions are shown as red squares in the insets.



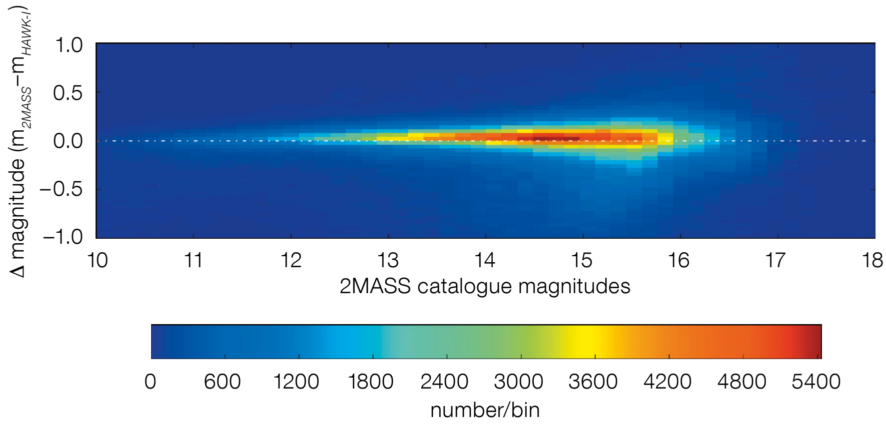


Figure 4 (right). A magnitude–Δmagnitude density plot showing the photometric quality of the HAWK-I source catalogues. The data are derived from sources in all 11 filters from reprocessed HAWK-I tiles matched with 2MASS catalogue stars. The standard deviation of the magnitude difference distribution is 0.2 magnitudes, with 82.4% of sources having $|\Delta m| < 0.1$ magnitudes.

Conclusions and access to products

The collaboration between ESO and CASU has been very positive and has provided ESO and its community with three state-of-the-art pipelines, three Reflex workflows, and a significant upgrade to the HAWK-I and VIMOS science archives. The success of this collaboration was due, in part, to intensive interactions on testing the pipelines, Reflex workflows and archive reprocessing at many intermediate steps and not only on the delivered products.

The design, testing, and data processing were always done from the perspective of an astronomer. For the Reflex workflows this meant that only the most relevant pipeline products and modifiable pipeline parameters are presented to the user, and that the default parameters are as close to their optimum as possible. For the archive reprocessing, this means that the data are as close to science grade as possible, irrespective of the scientific goals of the original programme; some example images are shown in the Appendix (Figures X.1 to X.4). As a concrete example, the archive reprocessing revealed about 2100 raw HAWK-I and VIMOS images with erroneous world coordinate headers. With considerable effort, the correct image pointings were deduced and the raw frame world coordinate headers were corrected. These corrected raw files have been re-ingested into the ESO SAF.

The new pipelines are now ready for operation and are being implemented by the quality control group in Garching and at the telescopes in Paranal. The intention is that ESO will continue to populate the Phase 3 science archive for HAWK-I and VIMOS where CASU has left off.

The reprocessed HAWK-I and VIMOS image archives are now publicly available via the ESO Phase 3 archive[2] or via the archive news release[5]. For individual pipeline users, the HAWK-I, VIMOS, and VIRCAM Reflex workflows and their associated tutorials can be downloaded from the Reflex webpage following the April 2017 release[1].

References

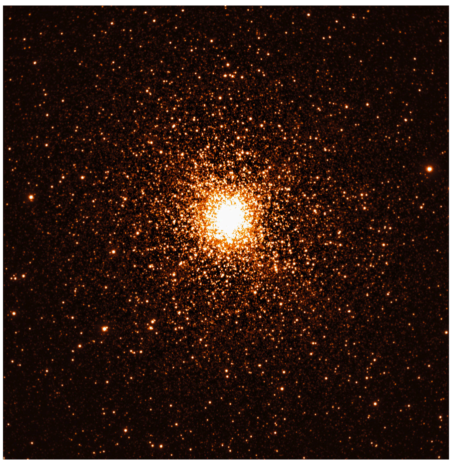
Arsenault, R. et al. 2016, The Messenger, 164, 2
Emerson, J., McPherson, A. & Sutherland, W. 2006, The Messenger, 126, 41
Freudling, W. et al. 2013, A&A, 559, A96
Roeser, S., Demleitner, M. & Schilbach, E. 2010, AJ, 139, 2440
Skrutskie, M. F. et al. 2006, AJ, 131, 1163

Links

1 ESO Reflex pipeline workflow environment: http://www.eso.org/sci/software/reflex
2 ESO Science Archive Facility: http://archive.eso.org/cms.html
3 VLT Quality Control: http://www.eso.org/observing/dfo/quality/
4 Instrument pipelines and manuals: http://www.eso.org/sci/software/pipelines/index.html
5 Archive news release: http://archive.eso.org/cms/eso-archive-news/Release-of-pipeline-processed-HAWK-I-images-as-part-of-the-UK-in-kind-reprocessing-project.html

Appendix

A number of HAWK-I and VIMOS example images from the archive reprocessing are shown to exemplify the quality of the delivered pipelines. It is important to note that these images are created using unaltered archive science data products. Other than a scaling of intensities and cropping, no other manipulation of the images was performed, attesting to the quality of these science pipeline products.



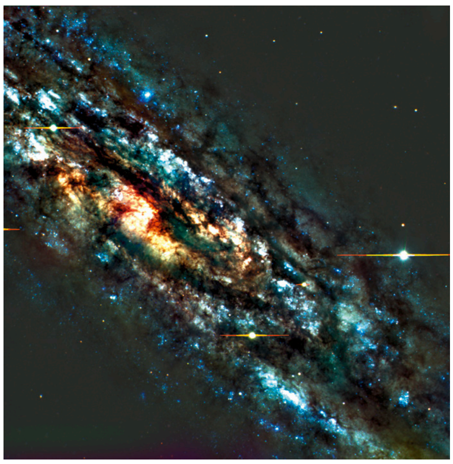


Figure X.1 (upper left). Composite HAWK-I *Y*-, *J*- and *H*-band colour tile of the spiral galaxy NGC 1232.

Figure X.2 (upper right). Composite HAWK-I *J*-, *H*- and *Ks*-band colour tile of the globular cluster NGC 1261.

Figure X.3 (bottom left). Single *R*-band VIMOS image of the globular cluster NGC 2808.

Figure X.4 (bottom right). Composite VIMOS *U*-, *B*- and *V*-band colour image of the southwest area of the spiral galaxy NGC 253.