THE TERAPIX T0002 RELEASE DOCUMENT

June 24, 2005

Y. Mellier[1,2], E. Bertin[1,2], H.-J. McCracken[1,2], M. Schultheis[3], M. Dantel-Fort[1,2], L. Domisse[1], F. Magnard[1], C. Marmo[1], G. Missonnier[1], J.-C. Malapert[1], A Rojbi[1]

[1]Institut d'Astrophysique de Paris, UMR 7095, Université Pierre & Marie Curie, 98 bis Boulevard Arago 75014 Paris, France
[2]Observatoire de Paris, LERMA UMR 8112, 77 avenue Denfert Rochereau, 75014 Paris, France
[3]Observatoire de Besançon UMR 6091, 41 bis, avenue de l'Observatoire BP 1615 25010 Besançon Cedex, France

1   Introduction

This document describes the data products and the tools developped at TERAPIX to produce the T0002 data release and the recent data candidate released by July 3rd, 2005. It aims at helping users to understand the way data have been processed as well as at better handling and analysing release data for their own scientific studies. The amount of images, catalogues, meta-data or even information web pages produced for each release is important and we are aware users may rapidly get lost. We therefore recommend users to read this document prior to start working on CFHTLS data release. We expect to improve this document in a proper fashion, so any comments/suggestions from users to `terapix at iap fr` are very welcome.

2   Overview

TERAPIX has four principal activities:

1. Development and release of software tools for processing panoramic images from mosaic detectors and of utilities to handle large images;

2. Development of a pipeline for processing automatically Megacam images and the CFHTLS data, starting from Elixir products that are sent to CADC;

3. Running and managing this pipeline, and releasing the CFHTLS products to CADC;

4. Finally, providing help and assistance to Megacam users. Another important task is to process, on request, PI data obtained with Megacam (25% of TERAPIX activities)

The TERAPIX team comprises 3 astronomers, Yannick Mellier (PI), Emmanuel Bertin (Project Scientist) and Henry Joy McCracken; 4 permanent position engineers, Mireille Dantel-Fort, Laurent Domisse, Frédéric Magnard (Project Manager) and Gilles Missonnier (leaving by June 2005), 2

non-permanent position engineers, Jean-Christophe Malapert (contract ending by January 2006) and Chiara Marmo (Starting by September 2005 for 3 years for WIRCAM) and one PhD student, Anthony Baillard. Several astronomers or engineers left TERAPIX over the past years and contributed to the work described below: Pierre Didelon, Bertrand Morin, Gérard Tissier, Mario Radovich.

The TERAPIX software tools are made publicly available once they have been tested and validated. Many (for example, Astrometrix / Photometrix, SExtractor, SWarp, QualityFITS, PSFex, STIFF, etc...) are already in use worldwide and have been included in several other public or private pipelines. In conjunction with the pipeline development and handling, software development and maintenance represents the largest fraction of TERAPIX activity.

Starting from the Elixir products, TERAPIX CFHTLS products can be separated in two principal components:

- a full set of individual 36 MEF weight and flag map images, created for each Megacam image received at TERAPIX, which are rapidly sent to CADC,

- processed images which correspond to stacked Megacam images which are sent to CADC as part of CFHTLS releases.

The TERAPIX processing hardware consists of a cluster of machines linked by gigabit Ethernet on a private sub-network separate from the IAP public computing facilities. In total there is 55 TB of fully secure RAID5 disk arrays distributed over 16 separate nodes running the LINUX GENTOO operating system. These CPUs provide a peak performance of about 150 GFlops. The TERAPIX computing facilities are totally managed by the TERAPIX team. As part of TERAPIX's service tasks, we make every effort to follow current trends in computing technology and assess which new product are best suited to our processing needs. This information is made publicly available on the TERAPIX web site.

3   Overview of TERAPIX pipeline

The TERAPIX processing pipeline which controls data and data base managements, file and pipeline configurations, process synchronization, data production and the user interface is called SPICA (Software Pipeline for Image and Catalog Analysis). The SPICA starting point is data transfer from input FTP disk space to TERAPIX network disk space, and the SPICA end point is data transfer to output FTP disk space. SPICA processes images in two steps, corresponding to two different "contexts", a data-ingestion context (step one) and a data processing context (step two). They represent two different pipelines that can operate asynchronously.

3.1   Processing "step 1": data ingestion context (Data transfer and QualityFITS-in)

The first context is the data ingestion of Elixir products from the CADC. At this level all images are processed using the same "on the fly mode" where affiliation to a particular survey program is ignored. The TERAPIX processing block tools involved in this step are "SNOOPPIX", "DataTransfer" (input), "QualityFITS-in" and the TERAPIX database. The original 36-MEF file structure is conserved.
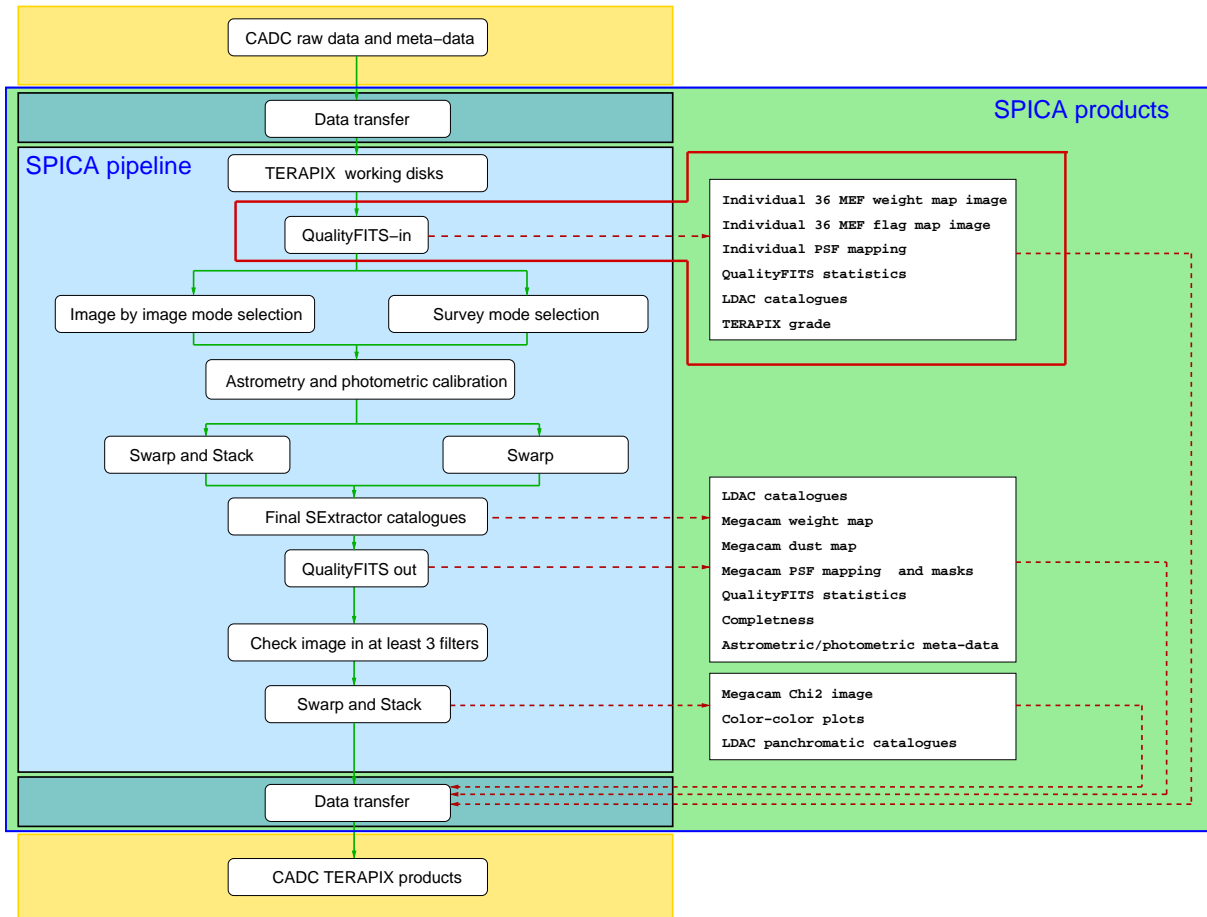
Figure 1: Flow chart of the TERAPIX pipeline and data transfer with CACD. Data received from CADC are first processed to produce quality assessments for each image and create the weight and flap map images for each (step1 of TERAPIX processing). These images, as well as primary catalogues and meta-data are immediately sent back to CADC. Hence, CFHTLS registered users can process any CFHTLS images on its own. Step1 is QualityFITS-in and data transfer to CADC of files inside the red line box. It looks like a single box though it indeed consists of several processing steps and several months of processing. Step2 defines the rest of the flow chart. It ends with the production of a "release" such as T0001.

3

Each incoming file is first tested for integrity and content and then registered in the database. For each science exposure, image quality is automatically evaluated on a per CCD level as well as globally over the Megacam field. Image evaluation and reliability validation for further scientific usage are measured through various statistical tools describing background homogeneity, shape and homogeneity of PSF, statistics of astronomical sources, and defects such as saturated pixels, cosmic ray impacts, glitches, trails, bright spikes and aureoles1.

At the same time, a weight-map and a flag-map image incorporating masks for the most obvious defects, the masks image and flat fields produced by Elixir are created and assigned to each exposure. These maps keeps the 36 MEF files format and are delivered to CADC, so they can be used by any CFHTLS users (or PI users for PI data processed at TERAPIX) that may wish to process CFHTLS data independently.

A shallow source catalog is then extracted using the weight-map to reject image artifacts. A PSF model is constructed for each CCD and used for tracking image sharpness and elongation over the focal plane.

The data ingestion mode products are also used to track several astronomical and technical image properties and to produce statistics that are continuously updated: image quality history, observing time distribution for each program and for each filter. These data are immediately made publicly available to Canadian and French institutes and will be available worldwide after the PI or CFHTLS proprietary period has elapsed. The QualityFITS-in information of each image and all relevant astronomical statistics are available to CFHTLS registered users at the URL `http://clix.iap.fr/steeringroup/`, using the userid and password sent to the CFHTLS mailing list.

The QualityFITS-in mode was also used during the tests of the wide field corrector by CFHT: test images were received at TERAPIX and immediately processed by QualityFITS in order to provide image quality information and PSF anisotropy mapping. Outputs were then returned to CFHT.

Over the past three years, TERAPIX delivered twice to CADC the "step one" products for immediate usage by CFHTLS registered users. For the first delivery, about 5000 weight maps, 5000 flag map images and 5000 catalogues were delivered. The second delivery was in January 2005. About 8000+8000 images were sent to CADC, as well as catalogues and meta-data (2/3 of them were part of the first 5000 but were re-processed to improve data quality).

3.2 Processing "step 2": Data processing context (SPICA and data transfer)

The data processing context uses either one of two processing modes: "image by image" or "TERAPIX release". TERAPIX processing block tools involved in this step are "SPICA", "QualityFITS-out", "DataTransfer" (output) and the TERAPIX database.

In contrast with the data ingestion context mode, only user-selected sub-samples of images may go through one of these two modes, depending on scientific goals. Though the two modes produce and handle MEF images, the original 36-MEF file structure of each image is no longer conserved, so individual CCDs of each Megacam image do not exist anymore: the quantum is now a full Megacam image.

Both modes use a global astrometric calibration process that can handle all images located near the target field.

- The "image by image" mode works on each single Megacam image that successfully passed the

data ingestion process regardless of its RunID. In this mode, each image is individually composed and warped into a pre-defined astronomical projection and astrometric system.

- The "TERAPIX release" mode manages image stacking (co-addition), warping and composition (image merging and projection over sky coverage much larger than one Megacam field) steps and can process together several sub-areas of the sky, each covering very large field of views observed in several filters over several years. In this mode, the final stack is also created according to a pre-defined astrometric projection and co-ordinate system.

In all modes, a (non-linear) astrometric solution is computed using an external or an internal reference astrometric catalog. In the "TERAPIX release" mode, the astrometric calibrations also use object cross-identifications from overlapping regions between Megacam images. A similar method is applied for photometric calibration. Exposures taken during photometric nights are used as anchor points for the zero-point re-scaling. Calibrated images can then be re-sampled, flux rescaled and co-added around each survey region2.

Finally, panchromatic source information are used for quality assessments in the final TERAPIX catalogue, provided enough data are available in several filters. Reference panchromatic images are the usual chi-squared images.

The "step two" products correspond to the TERAPIX releases of stacked images for the survey. In contrast with the "step one" production, TERAPIX is late in producing these data to users. T0001 was only released by November 2004 and T0002 has been delayed three times, due to several technical problems (see section describing the T0002 release).

3.3 Frequency and processing strategy for CFHTLS releases

Most survey programs will extend over several MegaPrime runs, and the data products extracted from stacked images will be derived from observing runs spread over several years. This strategy should allow the continual improvement of data quality and the stability of calibrations with time.

However, waiting for several years before releasing the whole survey would be scientifically counter-productive. Therefore, it was decided to run the second part of the pipeline (the data processing context) every six months on all the datasets released so far by CFHT. Successive TERAPIX data releases to CADC at six-month intervals will then be progressively better (deeper, better seeing and more homogenous). This incremental release process of CFHTLS products is optimum for most observing strategies until all data sets can be used together to produce the final versions of the CFHTLS data products. Based on this scheme, new versions of the data products will supersede the previous ones. Their accuracy and homogeneity will benefit from both the increasing number of pointings and calibration files and also from algorithmic improvements and bug fixes in our processing software. However, since Canadian and French astronomers may have already started scientific programs with past releases, all TERAPIX data releases will be archived at CADC for a long period.

4 CHTLS Data and meta-data products produced by TERAPIX

The bulk of data produced by TERAPIX and sent to CADC consists of images, catalogs and auxiliary data generated by the pipeline modules. The files format transferred will be multi-extension FITS

(MEF) for images, binary FITS tables or LDAC format for catalogs, PNG and PS images[1] and plots, and ASCII plain text or XML documents.

4.1   Step 1 QualityFITS-in products

QualityFITS produces a set of basic quality assessments on all input (QualityFITS-in) images and catalogs. Weight map FITS images are produced in QualityFITS-in using the WeightWatcher software tool. Its header contains some of the detrended image header keywords and new keywords. Flag map FITS images are produced in QualityFITS-in, also using WeightWatcher. The new header contains some all detrended image header keywords and new keywords. Header is the same as weight map images. To date, about 9000 weight maps and 9000 flag map images have been created. A part of them have been sent to CADC. The rest are PI data and are delivered to PIs only.

QualityFITS meta-data consist of PNG and PS plots and a summary table to evaluate whether images are valid data and do not significantly deviate from expectations. It helps to reject unexpected bad input images that cannot be used for scientific analysis (e.g. strongly defocused or multiply imaged objects) and to grade all images accordingly for further sample selection. Since it is blindly applied to all images, QualityFITS provides homogeneous and unbiased quality assessments. Note that QualityFITS does not provide astrometric calibration;

The plots produced by QualityFITS describe:

- The background histogram (for QualityFITS-in only), computed from the background estimated on each individual CCD,

- The FWHM histogram (for QualityFITS-in only), computed from the FWHM estimated on each individual CCD,

- Galaxy and star number counts over the whole Megacam field,

- A PSF model, with one sample point per CCD for QulaityFITS-in, and one sample point each 10 arc-minutes for QualityFITS-out.

- A series of half-light-radius magnitude diagrams showing sources selected for histograms and ellipticity map plotting.

The summary table provides more quantitative information on data shown in plots and basic statistics on images, for each CCD as well as for each Megacam field as a whole.

Data products consist of:

- Weight-maps, flag-maps FITS images, FITS tables and FITS (.ldac) catalogues, generated for each input science MEF image. All FITS images produced by TERAPIX are 32-bit floating point representations, except the flag map image which is in 8-bit integer;

- For each exposure, a full machine readable processing recipe in XML is provided as well as PNG thumbnail of the

---

[1]PNG is preferred to JPEG for quality control purposes, to prevent compression artifacts from DCT quantization that affect JPEG image thumbnails (lossless JPEG is not well supported and less efficient).

- Binned image content (image),

- Background level (image),

- Weight-maps (image),

- Point Spread Function model (image),

- Star counts with corresponding half-light radius diagram (plot),

- Galaxy counts corresponding Rh-mag. diagram (plot),

- Seeing histogram (plot),

- Background histogram (plot),

- PSF anisotropy mapping corresponding Rh-mag. diagram (plot),

Catalogs are produced by QualityFITS for each Megacam image. They are created by running SExtractor at the end of the data ingestion context "on the fly" mode (and also at the end of the "image by image" and "TERAPIX release" modes discussed below).

4.2 Step 2 data products

4.2.1 Processing modes and data preparation in Step 2

For each release, images will be spread into the three survey components, namely the Deep (D), the Wide (W) and the Very Wide (VW).
Because the Very Wide survey primary goals are detection and properties of TNOs, TERAPIX does not stack Very Wide survey images. Each image is calibrated and a single Megacam image will be produced in an astronomical projection system, according to an astrometric reference system. Since TERAPIX software tools and calibration files are all available to users, anyone has the possibility to stack these images, or to create another single Megacam image in another projection or astrometric reference system. TERAPIX is ready to provide stacks, on request.
In contrast, for the Wide and Deep surveys, TERAPIX produces stacked images. The main reason for this is the fact that stacked images are among the basic scientific data products necessary for the primarily science goals of these surveys, and also because TERAPIX must produce the colour-aligned chisquared image for these two surveys. Although, for CPU and disk space capacity reasons, we do not produce unstacked data of each Deep and Wide image, as for the Very Wide, it is worth noticing that the astrometric calibration data attached to each image are also sent to CADC, so anyone can use these calibration files to carry out an independent image stacking, if needed. This versatile approach allows users carry out their own image selection, that may be different from those taken for Legacy Data.
The following Sections describe each step of the processing mode.
subsubsectionSelection process of released CFHTLS data Once QualityFITS data products have been created, information regarding each image and relevant for running the next steps of the TERAPIX pipeline are stored in a local data base: location on TERAPIX disks of scientific images and their weight and flag maps, FITS header information, night by night CFHT queue observing metadata.
Prior to starting the processing mode, all images are visually inspected using the QualityFITS-in

quality assessment data products by TERAPIX astronomers. Each image is then graded A, B, C or D according to the following criteria.

- Grade A: image does not show any problem and QualityFITS evaluation are close to rough expectations

- Grade B: image is acceptable, though some minor problems have been detected. These problems could be technical, due to the TERAPIX processing during QualityFITS (e.g. wrong sky background estimate), or astronomical (galaxy counts show deviation from expectations)

- Grade C: image is not acceptable and cannot be injected in the full processing steps. For example, image with multiply-imaged stars due to telescope guiding problem, or image with extremely bright fuzzy halos.

- Grade D: image with partially wrong FITS astrometric keywords that positions one or several CCDs at a wrong position with respect to neighbour CCD. Grade-D image are presently removed from the input image lists, but when this problem is corrected they will be re-injected into the pipeline.

The Grade-A and -B data represent 95% of the sample; about 1-2% are graded C and will be definitely lost.

Beyond this step, only images with Grade A or B will be kept in CFHTLS releases. The image selection process can be refined in order to take into account the specifications of primary science goals (seeing, airmass, extinction, etc..). In the case of the Very Wide, all images will be processed, provided it passed the QualityFITS selection. For the Deep and the Wide, new specific selection criteria are added.

The selection criteria used for the T0001 release are:

- Observation period: From June 1st, 2003 to July 22, 2004

- Exposure time > 60 sec.

- Seeing <1.0" for g,r,i,z , and <1.3 for u* data

- Airmass <1.4

- TERAPIX grade A or B

The stack center position definition is an important point: for the four deep fields, stacks will be centered at the RA and DEC Deep center positions given in the CFHTLS documents and images will be included only if their RunID FITS keyword correspond to Deep survey data. For the wide, each stack will be centered according to the Wide target position provided earlier by the CFHTLS Wide coordinators, and only if the image center lies within 10 arc-minutes from a Wide target center position. This prevents the pipeline to select several times the same image for different Wide stacks, while still permitting to take into account the large dithering pattern.

Morevover, each stacked image will be only one squared degree in size to avoid the creation of giant images that would cover a whole wide field and to keep the CFHTLS Wide product simple

and logical with respect to the early survey design. This will avoid mixing Wide targets with different exposure time into a single giant image. Therefore, the TERAPIX Wide image products consist of square images, with a well-defined exposure time, and filling the Wide sky according to the http://terapix.iap.fr/cplt/Oldsite/Descart/summarycfhtlswide.html building block definition.

All selected images will be used in the TERAPIX release mode or will be processed according to the image by image processing mode. In addition, since TERAPIX astrometric calibration are computed globally all images can be used to perform the global astrometric solution. However, for practical technical reasons, TERAPIX only uses Megacam images within 2.5 deg radius with respect to a Very Wide image or a Deep or Wide center stack position. The accuracy of the astrometric calibration at a given position may therefore depend on the amount of data available around that position, but should increase with time as more images arrive from CADC. So, for any single image that belongs to a release, two image lists are generated: the "astrometric list", that corresponds to all files that will be used to perform the astrometric calibration of this image, and the "stack list" that correspond to all files that will be stacked together. The astrometric list may contain 10 or 100 times more images than the "stack list".

4.2.2   Construction of command line sequences

Once all images have been selected and the "astrometric" and "stack" lists done, SPICA automatically generates script-command files, one for each stack, for both the "image-by-image" and the "release" modes. The command file content includes the processing command line sequences, starting from loading and uncompressing all data and meta-data needed, continuing with the astrometric and photometric calibration, then handling the image warping and stacking (when necessary) and ending with the data compression of stacked image, copy to output disks of all data and meta-data sent to CADC. Since all command files are independent, one can process the release according to pre-defined priorities, for example starting with the Deep2-z band or the full set of Very Wide images.

The "image-by-image" and "TERAPIX release" modes can be run independently and simultaneously and on different CPUs as well. Once a stack is terminated, the TERAPIX release control team is informed automatically and SPICA launches the next stack.

The TERAPIX release control team checks each released image produced in both modes and grades it. Images with grade A or B are validated and can be delivered to CADC. Otherwise, the control team writes a very short problem report, the released image and its attached meta-data are examined in details and the SPICA process is ran again, once the problem is fixed.

SPICA operation interface displaying the T0002 release running on three nodes of the TERAPIX cluster. The green highlight bar shows the on going process of the processing pipeline. The tables at the top provide a continuously updated global status report of the release.

4.2.3   T0002 CFHTLS release : output products and filenames

QualityFITS is also the quality assessment tool for all images produced by the "image-by-image" and the "TERAPIX release" modes. The "QualityFITS-out" products are the same as QualityFITS in, except that the final weight map attached to each stacked image is already produced during the stacking step and that QualityFITS-out includes meta-data on completeness limits.

Prior to making data and meta-data available to CADC, final products are visually inspected using the quality assessment meta-data produced by the astrometric/photometric software SCAMP, by SPICA and by QualityFITS-out.

Scientific images produced at TERAPIX are all combined using the "SWarp" software (documentation on SWarp avaliable at the URL `http://terapix.iap.fr/rubrique.php?id_rubrique=49`) with a Lanczos3 interpolation kernel. A new FITS header associated with a processed image is created. This new header contains some existing detrended image header keywords (not modified) and some new keywords. Existing and new mandatory keywords are defined below. Default projection is tangent (TAN) and default astrometric coordinate system is Equatorial J2000 (FK5).

All source catalogues for CFHTLS releases are produced by SExtractor. The default format for RA-DEC data will be decimal degrees. The default equinox will be J2000.0 and default astrometric reference system is FK5. The default aperture diameter will be 3 and 4 arc-seconds, and the default mesh for background estimates will be 256 pixels. All catalogues have a default magnitude zero point set to 30 and the magnitude system is instrumental AB.

All catalogues produced by TERAPIX for CFHTLS releases will contain the 56 parameters listed below: Some new parameters may be added or removed in future catalogues. If so, TERAPIX will inform CADC by email. Users will be informed using on line explanatory tables about any new features in catalogues.

In general, processing software options in use correspond to those in the default software configuration files. If it is not the case, non-default options superseding those default options are be stored in the "COMMAND" tag of the processing meta-data XML file.

All software tools, configuration files, image lists, calibration files, catalogues and TERAPIX or CFHT meta-data that are used to release CFHTLS data will be sent to CADC. So, anyone will have the possibility to check the processing or to process the data in a different way. In particular, the .head file containing the astrometric calibration and the rescaled photometric calibration will be delivered to CADC.

The astrometric and photometric calibration software will be presented in the next section.

In the "image by image mode", the following products will be sent to CADC:

- Updated headers for the detrended images

- PNG/PS plots attached to the astrometric calibration step,

- PNG/PS plots attached to the photometric calibration step.

- QualityFITS outputs, as described below in the TERAPIX release mode,

- SExtractor source catalogue (.ldac file extension)

"Image by image" data products are available for download by CADC immediately after they are generated at TERAPIX.

In the "TERAPIX Release" mode, data products primarily consist of stacked images. In this mode output data will be generated every six months for archiving at CADC. TERAPIX release data products sent to CADC will be:

- Updated headers for each image contained in the stack;

SExtrator parameters in T0002 catalogues		
NUMBER		
X_IMAGE	Y_IMAGE	
X_WORLD	Y_WORLD	
XWIN_IMAGE	YWIN_IMAGE	
ERRAWIN_IMAGE	ERRBWIN_IMAGE	ERRTHETAWIN_IMAGE
ALPHA_J2000	DELTA_J2000	
MAG_AUTO	MAGERR_AUTO	
MAG_ISO	MAGERR_ISO	
MAG_APER(2)	MAGERR_APER(2)	
MAG_PETRO	MAGERR_PETRO	
FLUX_AUTO	FLUXERR_AUTO	
FLUX_ISO	FLUXERR_ISO	
FLUX_APER(2)	FLUXERR_APER(2)	
FLUX_PETRO	FLUXERR_PETRO	
THETA_IMAGE	ERRTHETA_IMAGE	
THETA_J2000	ERRTHETA_J2000	
THETA_WORLD	ERRTHETA_WORLD	
A_IMAGE	ERRA_IMAGE	
A_WORLD	ERRA_WORLD	
B_IMAGE	ERRB_IMAGE	
B_WORLD	ERRB_WORLD	
ISOAREA_IMAGE	ISOAREAF_IMAGE	
THRESHOLD	MU_THRESHOLD	
BACKGROUND		
FLUX_MAX		
MU_MAX		
KRON_RADIUS		
FLUX_RADIUS		
FLAGS		
CLASS_STAR		
X2_IMAGE	Y2_IMAGE	XY_IMAGE

- Weight-maps, background-maps and binary FITS table generated for each co-added image;

- PNG/PS plots attached to the astrometric calibration step;

- PNG/PS plots attached to the photometric calibration step;

- Co-added images in each filter and their associated weight-maps (both in FITS). Basic "image quantum" will be 1 deg × 1 deg. A 7 deg × 7 deg field will therefore produce 49 image quanta. Quanta will be non-overlapping areas on the sky. Each quantum will occupy a fixed area on the sky that will be covered by more MegaPrime pointings as the survey progresses;

- The chi2 images (see Szalay et al. 1999) combining all filters. This image will be provided even for single filter observations; * Source catalogs based on the chi2 images for detection and the co-added images for measurement;

- The "dust" FITS image is created using the Schlegel E(B-V) FITS maps interpolated and re-sampled according to CFHTLS image properties;

- A full machine-readable processing commands in XML format;

- PNG thumbnails of the

  - Binned image content (image)
  - Background level (image),
  - Weight-maps (image),
  - Point Spread Function model (image),
  - Star counts (plot),
  - Galaxy counts (plot),
  - Co-added and χ^2 images (image),
  - PNG plots related to the quality control of the "final" catalog (including star and galaxy counts, seeing histogram, magnitude completeness limit, two-point star angular correlation function, and, when available, color-magnitude and color-color diagrams[2]). The QualityFITS-out will include soon a PSF histogram (not implemented in QualityFITS V.10/V.11 used for T0001 and T0002).

The table given below summarizes the products delivered to CADC . They are discussed in more detail in the interface document. The CFHTLS products will reference to the survey type as follows:

- "W" for Wide,

- "preW" for Pre-survey attached to Wide,

- "VW" for Very Wide, and

- "D" for Deep,

	"On the fly mode" (Step1)	"Image by Image mode" (Step2 Very Wide)	"TERAPIX release mode" (Step 2 Deep & Wide)
Detrented image	546891p.fits	546891p.fits	546891p.fits 546890p.fits
Keyword used in header			Object name Filter
Processed data	546891p_head.txt[3]	546891p_V0001.fits 546891p_V0001_head.txt 546891p_V0001_ahead.fits for photometric images	CFHTLS_W_r_135934+533330_T0001.fits CFHTLS_W_r_135934+533330_T0001_546891p_head.txt and CFHTLS_W_r_135934+533330_T0001_546890p_head.txt ; CFHTLS_W_r_135934+533330_T0001_546891p_ahead.txt for photometric images
Weight-map and flag-map	546891p_weight.fits 546891p_flag.fits	546891p_V0001_weight.fits	CFHTLS_W_r_135934+533330_T0001_weight.fits
Catalogue	546891p_diag.fits 546891p_ldac.fits	546891p_V0001_diag.fits 546891p_V0001_ldac.fits	CFHTLS_W_r_135934+533330_T0001_diag.fits CFHTLS_W_r_135934+533330_T0001_ldac.fits
Auxiliary	546891p_bkg_histo.png 546891p_bkg_histo.ps 546891p_bkg_m.png 546891p_ell.png 546891p_fwhm_histo.png 546891p_fwhm_histo.ps 546891p_gal_histo.png 546891p_gal_histo.ps 546891p_m_med.png 546891p_m.png 546891p_psf_m.png 546891p_star_histo.png 546891p_star_histo.ps 546891p_wmm_med.png 546891p_wmm.png 546891p.xml 546891p_rhmag.png 546891p_rhmag.ps 546891p_rhmag_gal.png 546891p_rhmag_gal.ps 546891p_rhmag_star.png 546891p_rhmag_star.ps	546891p_V0001_astr_interror1d_%d.png6 546891p_V0001_astr_interror1d_%d.ps[4] 546891p_V0001_astr_interror2d_%d.png 546891p_V0001_astr_interror2d_%d.ps 546891p_V0001_astr_referror1d_%d.png 546891p_V0001_astr_referror1d_%d.ps 546891p_V0001_astr_referror2d_%d.png 546891p_V0001_astr_referror2d_%d.ps 546891p_V0001_bkg_m.png 546891p_V0001_cmd.txt 546891p_V0001_comptness.png 546891p_V0001_comptness.ps 546891p_V0001_distortion_%i.7png 546891p_V0001_distortion_%i.ps 546891p_V0001_ell.png 546891p_V0001_ell.ps 546891p_V0001_fgroups_%d.png 546891p_V0001_fgroups_%d.ps 546891p_V0001_gal_histo.ps 546891p_V0001_gal_histo.png 546891p_V0001_m_med.png 546891p_V0001_m.png 546891p_V0001_phot_error_%d.png 546891p_V0001_phot_error_%d.ps 546891p_V0001_phot_zpcorr_%d.png 546891p_V0001_phot_zpcorr_%d.ps 546891p_V0001_phot_zpcorr3d_%d.png 546891p_V0001_phot_zpcorr3d_1%dps 546891p_V0001_protopsf.fits 546891p_V0001_psf_m.png 546891p_V0001_referror1d_1.png 546891p_V0001_referror1d_1.ps 546891p_V0001_referror2d_1.png 546891p_V0001_referror2d_1.ps 546891p_V0001_rhmag_gal.png 546891p_V0001_rhmag_gal.ps 546891p_V0001_rhmag.png 546891p_V0001_rhmag.ps 546891p_V0001_rhmag_star.png 546891p_V0001_rhmag_star.ps 546891p_V0001_star_histo.png 546891p_V0001_star_histo.ps 546891p_V0001_scamp_accept.xml 546891p_V0001_scamp.xml 546891p_V0001_spica.xml 546891p_V0001_wmm_med.png 546891p_V0001_wmm.png 546891p_V0001_reg.txt 546891p_V0001_dust.fits 546891p_gri_V0001_diagcc.png8 546891p_gri_V0001_diagcc.ps 546891p_gri_V0001_chi2.png[5] 546891p_gri_V0001_chi2.ps 546891p_gri_V0001_chi2.fits	CFHTLS_W_r_135934+533330_T0001_astr_interror1d_%d.png CFHTLS_W_r_135934+533330_T0001_astr_interror2d_%d.png CFHTLS_W_r_135934+533330_T0001_astr_interror1d_%d.ps CFHTLS_W_r_135934+533330_T0001_astr_interror2d_%d.ps CFHTLS_W_r_135934+533330_T0001_astr_referror1d_%d.png CFHTLS_W_r_135934+533330_T0001_astr_referror2d_%d.png CFHTLS_W_r_135934+533330_T0001_astr_referror1d_%d.ps CFHTLS_W_r_135934+533330_T0001_astr_referror2d_%d.ps CFHTLS_W_r_135934+533330_T0001_bkg_m.png CFHTLS_W_r_135934+533330_T0001_cmd.txt CFHTLS_W_r_135934+533330_T0001_comptness.png CFHTLS_W_r_135934+533330_T0001_comptness.ps CFHTLS_W_r_135934+533330_T0001_distortion_%i.png6 CFHTLS_W_r_135934+533330_T0001_distortion_%i.ps CFHTLS_W_r_135934+533330_T0001_ell.png CFHTLS_W_r_135934+533330_T0001_ell.ps CFHTLS_W_r_135934+533330_T0001_fgroups_%d.png CFHTLS_W_r_135934+533330_T0001_fgroups_%d.ps CFHTLS_W_r_135934+533330_T0001_gal_histo.png CFHTLS_W_r_135934+533330_T0001_gal_histo.ps CFHTLS_W_r_135934+533330_T0001_m_med.png CFHTLS_W_r_135934+533330_T0001_m.png CFHTLS_W_r_135934+533330_T0001_phot_error_%d.png CFHTLS_W_r_135934+533330_T0001_phot_error_%d.ps CFHTLS_W_r_135934+533330_T0001_phot_zpcorr_%d.png CFHTLS_W_r_135934+533330_T0001_phot_zpcorr_%d.ps CFHTLS_W_r_135934+533330_T0001_phot_zpcorr3d_%d.png CFHTLS_W_r_135934+533330_T0001_phot_zpcorr3d_%d.ps CFHTLS_W_r_135934+533330_T0001_protopsf.fits CFHTLS_W_r_135934+533330_T0001_psf_m.png CFHTLS_W_r_135934+533330_T0001_referror1d_1.png CFHTLS_W_r_135934+533330_T0001_referror2d_1.png CFHTLS_W_r_135934+533330_T0001_referror1d_1.ps CFHTLS_W_r_135934+533330_T0001_referror2d_1.ps CFHTLS_W_r_135934+533330_T0001_rhmag_gal.png CFHTLS_W_r_135934+533330_T0001_rhmag_gal.ps CFHTLS_W_r_135934+533330_T0001_rhmag.png CFHTLS_W_r_135934+533330_T0001_rhmag.ps CFHTLS_W_r_135934+533330_T0001_rhmag_star.png CFHTLS_W_r_135934+533330_T0001_rhmag_star.ps CFHTLS_W_r_135934+533330_T0001_star_histo.png CFHTLS_W_r_135934+533330_T0001_star_histgo.ps CFHTLS_W_r_135934+533330_T0001_scamp_accept.xml CFHTLS_W_r_135934+533330_T0001_scamp.xml CFHTLS_W_r_135934+533330_T0001_spica.xml CFHTLS_W_r_135934+533330_T0001_wmm_med.png CFHTLS_W_r_135934+533330_T0001_wmm.png CFHTLS_W_r_135934+533330_T0001_reg.txt CFHTLS_W_r_135934+533330_T0001_dust.fits CFHTLS_W_gri_135934+533330_T0001_diagcc.png CFHTLS_W_gri_135934+533330_T0001_diagcc.ps CFHTLS_W_gri_135934+533330_T0001_chi2.png CFHTLS_W_gri_135934+533330_T0001_chi2.ps CFHTLS_W_gri_135934+533330_T0001_chi2.fits CFHTLS_W_gri_135934+533330_T0001_chi2_weight.fits CFHTLS_W_gri_135934+533330_T0001_chi2_ldac.fits

Table 1: Summary table of data and meta-data content and filename of TERAPIX releases. All data listed above are produced by TERAPIX, starting from CFHT/Elixired input data, and are delivered to CADC. CADC organises the data distribution for all CFHTLS registered users.

Initial CFHT image name, for the "on the fly" (step1) and "image by image mode" The meta-data files related to the stacked and composed images are generated by pipeline modules. Filenames will be :

- %s_bkg_m.png for PNG sky/detector background images,

- %s_psf_m.png for PNG medium sized PSF map images,

- %s_fwhm_histo.png for PNG FWHM histogram (PS format also available),

- %s_star_histo.png for star number count (PS format also available),

- %s_gal_histo.png for galaxy number count (PS format also available),

- %s_wmm_med.png for PNG medium sized weight map images,

- %s_wmm.png for PNG weight map images,

- %s_ell.png for PNG stellar ellipticity map images (PS format also available),

- %s_rhmag.png for Rh-mag. selection diagram of ellipticity map (PS format also available),

- %s_rhmag_gal.png for galaxy-filtered Rh-mag. diagrams used for galaxy count (PS format also available),

- %s_rhmag_star.png for stellar-filtered Rh-mag. diagrams used for star count (PS format also available),

- %s_cmd for all command files that launch stack processes,

- %s_m.png for PNG stacked images ,

- %s_m_med.png for PNG medium sized stacked images,

- %s_protopsf.fits for PSF fitting image maps,

- %s.xml for command XML file,

- %s_astr_interror1d_%d.png for internal error: 1 dimension astrometric plots (PS format also available),

- %s_astr_interror2d_%d.png for internal error: 2 dimension astrometric plots (PS format also available),

- %s_astr_referror1d_%d.png for external astrometric error with respect to an astrometric reference catalogs (like USNOs or GSCs) ; (PS format also available),

- %s_distortion_%d.png for astrometric distortion map PNG images (PS format also available),

---

[2]two-point angular correlation function, color-magnitude and color-color diagrams will be implemented later.

- %s_fgroups_%d.png for image overlays projected in the sky (PS format also available),

- %s_phot_error_

- %s_zpcorr_%d.png for zero point re-scaling plots (PS format also available),

- %s_scamp_accept.xml for astrometric and photometric summary tables,

- %s_scamp.xml for XML format astromeric/photometry processing history logs,

- %s_comptness.png for magnitude completeness plots (PS format also available),

- %s.reg for ASCII polygon mask files,

where %s is the filename of a TERAPIX release stacked/composed image and %d is the number of photometric group

For each release, TERAPIX may produce new types of data and new files that are not listed below or discussed in previous sections. For example, it is foreseen to produce PSF-homogenized survey images in the future.

The %s.reg masks are ASCII files readable with DS9 or SAOimage. Each mask contains points that draw polygons in WCS coordinate system. They are only delivered for the Deep and Wide stacks, one per field. Most are produced using the i-band image, when available, with the new TERAPIX "Automask" software tool, the oldest hand-made masks will be re-processed later.

4.3   Assigment of dates for the Public Data Releases

For the "on-the-fly" and "image-by-image" mode, the release date of FITS data products (images and tables) will be identical to the release date written in the FITS header of the input detrended images. For the "TERAPIX release" mode, the release dates for the stacked images and their associated source catalogs correspond to the release date of the most recently-released input detrend images.
All TERAPIX data products will provide release dates at the end of their processing.

4.4   Data transfer mechanism

All data product transfers between TERAPIX and CADC are carried out over a network connection using the BBFTP protocol. We do not use tape media. We selected this method of data transfer based on the following criteria:

- Calibrated images should be transferred to CADC as soon as possible to allow science exploitation to begin immediately;

- The weight and flag map images produced by QualityFITS for each individual incoming image are put in the TERAPIX outgoing data transfer area immediately after they are generated. CADC must be able to access these files for transfer using their tools or using the SNOOPPIX tools developed by TERAPIX;

- Other output files may need more time for processing before they are sent to CADC. The expected time between the arrival of an image at TERAPIX and the delivery to CADC of its complete set of "on-the-fly" and "image by image" data products will be at most 3 weeks;

- RICE data compression will be used whenever possible;

- Data should be transferred automatically.

- Transfer cost should be minimal.

Network transfers are the only method which fufills all these criteria. Over the period 2002 we measure an actual sustained data transfer rate between CADC and TERAPIX of 2.5MB/second, which is fast enough for our purposes.

TERAPIX downloads meta-data archived at CADC via proxy access. SNOOPPIX continuously monitors a CADC "TERAPIX-permitted" archive area and downloads available data whenever necessary. We have allocated a directory space (monitored by CADC software) for outgoing data transfers. The storage capacity at the TERAPIX FTP staging area is 3.6 TB. Once data has been copied to the CADC, TERAPIX will delete data from this staging area to allow transfers to continue.

We also supply to CADC a series of c-shell command files which indicate the sequence of commands used to generate each stack. This file is supplemented by a series of XML files which lists the input images and what processing was applied to each file. Configuration files used in each of TERAPIX modules, and the software version numbers, is also provided. Taken together, this will allow any remote facility to reproduce exactly the series of processing steps used in any TERAPIX data release.

5 Astrometric/photometric calibration of T0002 and next releases

SCAMP (Bertin et al 2005) is the new astrometric and photometric calibration software developed at TERAPIX which will be used for all releases beyond T0001. It replaces Astrometrix and Photometrix which were too slow and difficult to install and maintain (and not very portable). In comparison, SCAMP is about 100 times faster than Astrometrix and is fully parallelised. SCAMP was developed by E. Bertin and uses plain-text configuration files similar to those of SExtractor and SWarp and can be installed on any POSIX-compliant platform. SCAMP takes as input SExtractor catalogues and produces as output WCS image headers that can be read by SWarp to re-scale, re-sample and co-add the images. SCAMP has been extensively tested by the TERAPIX team over the period July 2004 to January 2005 and is expected to be released publicly in Q3 2005.

5.1 Astrometric calibration

With the current generation of TERAPIX hardware (the amount of physical memory is the main limitation) astrometric and photometric solutions can be computed simultaneously for around 2,000 exposures of 30,000 detections each, at once. Once called with a series of catalogues, the first task for SCAMP is to examine the image headers and group exposures by sky area, "astrometric context", and "photometric context". An "astrometric context" defines a set of exposures that have been observed with a common, mechanically stable configuration. For the CFHTLS, these are exposures which share the same filter and QRunID. There is one astrometric solution per astrometric context. Hence, a tight group of overlapping exposures in a given sky area will provide better constraints to the astrometric solution that may benefit more isolated exposures in a different sky area, if they share the same astrometric context. Likewise, a "photometric context" characterises a set of exposures with the same

spectral response, that is, with the same filter.

Before proceeding further, the CCDs in each available astrometric context are re-positioned to a common, "median" focal plane, to correct the strong CCD-to-CCD jittering introduced by the Elixir astrometric solutions.

A subset of the reference astrometric catalogue (USNO-B1, Monet et al. 2003) surrounding each identified sky area is automatically downloaded from CDS. Cross-correlation is used to match each exposure with respect to the reference catalogue: pixel scales, frame position angles and shifts in alpha-delta are determined within around 1" at this stage. Pattern matching is followed by a cross-identification with a search radius of 2" between all overlapping detections.

In the context of the CFHTLS, SCAMP computes a polynomial model of Megacam image distortions by minimizing a weighted, quadratic sum of differences in position between both overlapping detections and matched sources from a reference astrometric object catalogue. Astrometric catalogues currently available at the depth of the individual CFHTLS exposures ($17 < i < 22$) have a typical accuracy of 0.3" on small scales, which is clearly insufficient. Combining the constraints given by the relative positions of overlapping detections and those from the reference catalogue provides both a very accurate adjustment of exposures on small scales, and a correct positioning on the sky on larger scales.

The solution is particularly vulnerable to matching errors and other non-Gaussian sources of noise and so it is very important to achieve maximum robustness at this stage by firstly minimizing the number of degrees of freedom, secondly by maximizing the number of constraints in the overlaps and lastly by clipping outliers.

The first condition is fulfilled by assuming that the astrometric solution is stable within a given QRunID. This proves to be a good assumption for Megacam, once linear distortion terms which apply to the whole focal plane (mostly differential refraction effects) have been taken out. In practice a 3rd order polynomial per CCD, that is, 720 parameters for the whole Megacam field, is used.

In order to maximize the number of detections in the second step, galaxies and all other faint non-saturated, non-corrupted objects are kept and properly weighted in the final solution. Galaxies are the dominant population at faint fluxes and moreover provide a stable reference frame for future proper motion measurements. Therefore, the centroiding routine implemented in SExtractor for astrometric calibrations has been specifically re-designed to produce robust results with non-stellar profiles.

The last condition is met by iterating the solution twice and clipping four-sigma outliers. Less than 1%-2% of detections are typically rejected from the final solution.

Residual position errors give us an approximate idea of quality of the final astrometric solution, although the true quality depends a lot on the exposure dithering strategy and is not easily estimated (from the point of view of astrometric calibration, the dithering strategy of CFHTLS fields is far from optimum). Roughly, large scale (greater than the typical dithering step) errors are dominated by those of the reference catalogue which is typically a few tenths of an arc-second. At the smallest scales, the relative astrometric accuracy is typically 10 to 30 *mas* RMS, that is, photon-noise-limited for most sources. Note that for release T0002, all positions are expressed in the J2000 system, and ignore proper motion effects (which we expect to add in T0003).

5.2 Photometric calibration

Many CFHTLS exposures are taken in non-photometric conditions. Thus different magnitude zero-point corrections must be computed and applied prior to image co-addition (the magnitude zero-point of the final stacked image is set to 30.0). These zero-point corrections are computed by SCAMP in a way similar to the astrometric distortion coefficients by minimizing the weighted, quadratic sum of MAG_AUTO magnitude differences from overlapping detections. MAG_AUTO was chosen as it is fairly robust and insensitive to the seeing or the sizes of sources.

Contrary to the astrometric calibration, wide-field reference catalogues are not available. Some exposures are flagged by CFHT as "photometric" however, which means that they were observed under photometric conditions and that their zero-points are reliably known. They can therefore be used as "anchors" to fix the global zero-point of an image stack and correct for potential magnitude drifts in wide contiguous survey areas like the CFHTLS-VW (but note that some T0002 frames do not overlap any photometric field and cannot be photometrically calibrated yet).

Adding these photometric constraints in the solution works well, although small zero-point discrepancies of the order approx 1%-2% can be seen amongst overlapping "photometric" fields. Therefore we chose to relax the condition that the zero-points of "photometric" fields are known and fixed by allowing them to vary by a few percents around the values derived by Elixir. The resulting residuals on "bright" object magnitudes (S/N>100) are typically 0.02-0.05 magniutes RMS, but the relative zero-point uncertainties are much smaller. These zero-point uncertainties are dominated by the effects of CCD-to-CCD variations, seeing changes and residual gradients within a Megacam field, the sum of which are at or below 0.01 mag. This demonstrates the excellent quality of the illumination correction carried out at CFHT.

5.3 SCAMP output headers

SCAMP creates output headers that contain astrometric and photometric calibration data, and a series of quality assessments output files. The .head or .ahead files created by SCAMP contains astrometric and photometric calibration data.

The .head or .ahead files are ASCII files that contain a multi-FITS header with keywords providing updated WCS astrometric data and photometric rescaling coefficients attached to each CCD. These headers are produced and delivered to CADC for all input images processed using the image-by-image mode (i.e. one header per image) and for all stacked output images of the TERAPIX release mode. (i.e. one header per stack). Input files are object catalogues produced by SExtractor, so LDAC catalogs produced by QualityFITS can be used.

The PV_{i_j} astrometric keywords that are derived by SCAMP are the parameters of the WCS projection system described by Greisen & Calabretta in 1999. The equinox is currently J2000.0 and the astrometric reference system is FK5.

5.4 SCAMP quality assessment meta-data

Quality assessment data comprise a series of figures and plots showing the astrometric and photometric internal and external errors, and the relative photometric re-scaling of stacked images. Plots showing the astrometric and photometric are 2-dimension and 1-dimension error distributions along the two

18

axes are also generated.

Default formats are PNG and PS files. Filenames and file descriptions are given below (see also Section 2.4 and the summary table given in Section 2.6):

- %s_astr_interror1d_%d.: PNG and PS file format, contain internal error: 1 dimension astrometric plot

- %s_astr_interror2d_%d. : PNG and PS file format, contain internal error: 2 dimension astrometric plot

- %s_astr_referror1d_%d.: PNG and PS file format, contain external astrometric error with respect to an astrometric reference catalog

- %s_distortion_%d. : PNG and PS file format, contain the distortion map

- %s_fgroups_%d. : PNG and PS file format, are image overlays projected in the sky

- %s_phot_error_%d. : PNG and PS file format, contain the photometric error plot

- %s_zpcorr_%d. : PNG and PS file format, contain the zero point error plot of each image

The TERAPIX pipeline (SPICA) provides a series of XML files that summarizes the SCAMP processing history: filename table listing the input files used to produce the global astrometric solution and the photometric re-scaling, summary statistic tables describing filter and sky position distributions of input files, and a concise summary table on astrometric and photometric statistics of output data. The summary tables corresponding to the astrometric and photometric calibrations shown above for CFHTLS_W_i_143007+533430_*, are given below.

FIG Summary tables provided by SCAMP to SPICA operation interface and CFHTLS. These tables provides internal and external statistics on the astrometric and photometric calibrations and provide hyper links to more detailed pages.

6 TERAPIX release description

6.1 Short reminder on the T0001 release

The first TERAPIX provided a pre-release in March 2004 (information on T0001 avaliable at the URL `http://terapix.iap.fr/rubrique.php?id_rubrique=170`). TERAPIX delivered to CADC images, catalogues and few quality assessment meta-data for the Deep D3 and D4 fields and of the Wide W3 field. These data were made public to all CFHTLS registered users in order to familiarize them with the survey data. The format of this release did not follow any specifications agreed between CADC and TERAPIX.

The first T0001 CFHTLS release (`http://terapix.iap.fr/rubrique.php?id_rubrique=150`) was made publicly available to CFHTLS users on Nov.23, 2004, but the "step one" data products were send to CADC much earlier. These comprised 5000 weight maps and 5000 flag map images and also includes relevant catalogues and meta-data.

The T0001 release comprised only CFHTLS Deep fields, D1, D2, D3 and D4 in all filters. As described

before, each weight and flag map image had already been delivered in an earlier step to CADC and QualityFITS-in information regarding each image is already available on-line to any CFHTLS registered users. Information regarding all releases and recent news about TERAPIX processing, new software tools can be found here: `http://terapix.iap.fr/rubrique.php?id_rubrique=169`
Individual images were selected according to the selection criteria listed in Table 1. The seeing criterion for the u* band data was relaxed in order to increase the numbers of images. An unexplained astrometric residual was found in D4-z data, so it was decided to remove D4-z from this release.

Science verification studies on stellar populations, weak lensing, photometric redshift analyses and deep galaxy photometry demonstrate that the CFHTLS data are indeed very good and the CFHTLS-TERAPIX products are reliable and can be used for science.

6.2 Release T0002

6.2.1 Overview

The T0002 release contain data from the three components of the CFHTLS survey. The amount of images being processed is around 16000 (about 8000 science images and 8000 weight maps attached to RunIDs [03/04][A/B]L[01-06]). Since the Very Wide data are not stacked, the data reduction factor will be smaller than T0001 (a factor 100) and the amount of data sent to CADC will be more than one order of magnitude higher. The selection criteria are the same as T0001 except that the observing period is extended to Nov. 22, 2004.
All images will have 19354×19354 0.186 arc-second pixels (1 deg^2). As in T0001, the astrometric reference catalogue is USNO-B1.
The "step 1" processing corresponding to the data ingestion context that produces all weight and flag maps has been completed and all individual image maps have been sent to CADC. They have been available to CFHTLS registered users since Jan. 15, 2005 and consist of about 8000 weight map and 8000 flag map images and relevant catalogues. QualityFITS-in information for each image is also available on line.
The "step 2" processing corresponding to the "image-by-image" (Very Wide) or "Release" (Deep and Wide) processing modes has been unfortunately delayed by several months. It was originally scheduled to be completed by December 4th, but was rescheduled due to several diverse problems: technical difficulties related to the integration of the latest hardware components in the TERAPIX cluster, changes in the stacking strategy of wide data in order the stacks to follow the target lists supplied by the wide coordinators, delays in the database, the pipeline and the processing software components and difficulties in achieving the expected percentage accuracy in the photometric rescaling. Because we moved from Astrometrix to SCAMP, we also had to check that data processed with both software tools provide the same results. The most critical issue was the analysis of systematic residuals for cosmic shear. This control demands to analyse at least 10 deg^2 (because the signal to noise per Megacam field is low) each time new releases of software tools are produced, to measure B-modes and to check they are statistically zero with not systematic discrepancies between SCAMP and Astrometrix (we know from weak lensing analyses of VIRMOS-Descart and T0001 data that Astrometrix does not produce B-modes). So far, we have not found significant differences between both software tools. Similar checks are done to check consistency between photometry and object positions in T0001 and T0002 stacks.

The selection criteria for the T0002 release are the following:

- Observation period: From June 1st, 2003 to November 22, 2004

- Exposure time > 60 sec.

- Seeing Deep <1.1" for g,r,i,z , and <1.3" for u* data

- Seeing Wide <1.1" for g,r,i,z , and <1.4" for u* data

- Seeing VeryWide <1.2" for g,r,i,z , and <1.5" for u* data

- Airmass <1.4

- Skyprobe <2.0

- TERAPIX grade A or B

The Deep, Wide and Wide Pre-survey will consist of 208 stacked images of 1 deg^2 :

- 20 A/B graded Deep stacks (4 fields, 5 filters),

- 123 A/B graded Wide stacks, spread in

  - 20 W1,
  - 8 W2 and
  - 13 W3 fields.

- 10 W1 fields are produced in ugriz, none for W3 and W2 (gri at best),

- 65 Grade A/B Wide Pre-survey (r-band , only one 180 second exposure per image).

The Very Wide data processed in the "image-by-image" mode consists of 1262 A/B graded images.

The most innovative technical aspect of T0002 as compared to T0001 is SCAMP, for the astrometric and photometric calibration and the field to field photometric rescaling. Since SCAMP is based on a global astrometric calibration method, it can use as much images as possible around each field, leading to a more robust solution and a better accuracy than Astrometrix. For each field we required that all Megacam images encompassed inside a 2.5 deg radius circle around each field be included in the astrometric calibration. The local small-scale weak oscillations observed in Deep fields of T0001 should also vanish since the Wide and Pre-survey data (which have large dither patterns) will be also added, together with deep images, in all astrometric calibrations of deep stacks. The plots and tables presented in the SCAMP section show a sample of output meta-data of a T0002 product for the Wide survey: CFHTLS_W_i_14307+533430_*. The summary table of the SWarp stacking process and the QualityFITS-out summary table of this target are shown below.

6.3 Acces to T0002 information:

In addition to CADC data access points, there are several TERAPIX pages that explain the TERAPIX data processing pipeline, or the data content of each release, or provide statistics on CFHTLS data. The most useful pages are given below. Those with restricted access to CFHTLS registered users need a userid and password.

- General overview of Deep+Wide T0002 data :

  http://terapix.iap.fr/article.php?id_article=506

- Table summary for Deep and Wide stacks:

  http://terapix.iap.fr/rubrique.php?id_rubrique=198

- Explanatory table attached to the summary:

  http://terapix.iap.fr/article.php?id_article=505

- Statistics on CFHTLS images injected in STEP1 of the TERAPIX pipeline (restricted access to CFHTLS registered users):

  http://clix.iap.fr/steeringroup/

- STEP1: Quality assessments for each input CFHTLS Megacam image (restricted access to CFHTLS registered users):

  http://clix.iap.fr/steeringroup/qf

- STEP2: Quality assessments for each Very Wide processed data (restricted access to CFHTLS registered users):

  http://clix.iap.fr/steeringroup/spica/index.php (select "Y" in the SPICA status box)

- STEP2: quality assessments for each Deep, Wide or Pre-W stacks (restricted access to CFHTLS registered users):

  http://clix.iap.fr/steeringroup/spica/survey

- Detailled image selection used for the Deep astrometry process and the Deep stacking process (restricted access to CFHTLS registered users):

  http://terapix.iap.fr/article_intranet.php?id_article=482

- Detailled image selection used for the Wide astrometry process and the Wide stacking process (restricted access to CFHTLS registered users):

  http://terapix.iap.fr/article_intranet.php?id_article=484

- Detailled image selection used for the Pre-W astrometry process and the preW stacking process (restricted access to CFHTLS registered users):

  http://terapix.iap.fr/article_intranet.php?id_article=483

6.4 Next releases

Provided TERAPIX experiences no further delays, T0003 is scheduled to be delivered to CADC by December 15, 2005. TERAPIX will start running Step 1 by July 15, so that new weight and flag map images will be delivered to CADC by September 15. Except for an improvement of source positional accuracy of sources by around 30%, and the possible use of proper motions in the astrometric calibration, we do not foresee innovative development for that release, as compared to T0002.

Beyond T0003, the TERAPIX releases will include two innovative tools:

- Defectix: this is an automated masking software tool based on neural network learning process which classifies and masks most common defects present in images: for example, saturated stars, diffuse stellar halos or satellite tracks. It is currently in preliminary test phase and more work is required before it can be implemented in the TERAPIX pipeline. In principle it should (partly) replace the hand-made masking by providing a FITS mask image that will be easily usable in conjunction with the weight map image. We expect defectix to be ready for T0003 (December 2005).

- PSF homogenization: this is still in study phase. It will use the PSFex analysis software tool to derive a optimal PSF homogenization of all survey data sets and create a PSF-homogenized set of stacked images.

7 TERAPIX software tools

All TERAPIX software are released under the General Public License (GPL). CADC will be able to use all of the software listed as below as the various modules are released. Note that some are still under development and are therefore not yet publicly. The TERAPIX software suite is available at `http://terapix.iap.fr/rubrique.php?id_rubrique=34` . It includes:

- Astrometrix / Photometrix: local and global astrometry and photometry (used for the T0001 release)

- Automask : ASCII DS9 readable mask file.

- Datatransfer : transfer and distribute data on TERAPIX disks;

- Deadix : detect dead CCDs on Megacam images;

- Defectix : detect, classify and mask defects on Megacam images;

- LibVOtable : a VOTable parser;

- MissFits : FITS file management;

- Panorapix : Visualisation tool for very large FITS images;

- PSFEx : constructs a PSF model of a field;

- QualityFITS : input and output quality assessment tools and metadata generation;

- RDBIX : store information in a database;

- SCAMP : new astrometric and photometric software, which will be used for all releases "stacked mode" releases after T0001, and all image-by-image mode releases;

- SExbench : hardware benchmarking software;

- SExtractor : source extraction;

- SkyMaker : image simulation;

- SkyWatcher : monitors the progress of a survey on the celestial sphere;

- SNOOPPIX : monitor data at CADC

- SPICA : image processing pipeline (data management, file and pipeline configuration, process synchronization, data production , user interface)

- Stiff : produces color TIFF images based on FITS file input

- Stuff : simulation of galaxy populations

- SWarp : image re-sampling and co-adding

- WeightWatcher : image weight-map creation

All TERAPIX software is packaged using the GNU autoconf/automake system which allows them to be fully platform-independent. Binary ".rpm" files are also provided which makes installation particularly easy on LINUX systems. Whenever software bugs are found, TERAPIX will notify CADC immediately and will send appropriate patches.

Acknowledgements

Figure 2: A summary panel showing the QualityFITS-in output, as they are displayed in each web page generated for all individual Megacam images processed at TERAPIX binned PNG image and its weight map (top left), background analysis (middle left) and PSF analysis (bottom left). Right: galaxy and star counts (2 first top right panel) and PSF anisotropy mapping (middle right) plots (left). For each analysis, red dots show objects that have been selected in the Rh-magnitude diagramme (extreme right). Bottom right: summary table. Statistics CCD per CCD and averaged over the Megacam field. The first column links to a PNG image of each CCD.

Wide Stack (134)							
Image	Survey	Filter	Ra	Dec	Scamp files	Swarp files	Duration
CFHTLS_W_g_021410-041200	Wide	g	02 h 14 m 10	-04 ° 12 ' 00	105	4	02 h 34
CFHTLS_W_g_021410-050800	Wide	g	02 h 14 m 10	-05 ° 08 ' 00	108	8	03 h 35
CFHTLS_W_g_021800-041200	Wide	g	02 h 18 m 00	-04 ° 12 ' 00	125	4	03 h 49
CFHTLS_W_g_021800-050800	Wide	g	02 h 18 m 00	-05 ° 08 ' 00	126	4	03 h 17
CFHTLS_W_g_021800-060400	Wide	g	02 h 18 m 00	-06 ° 04 ' 00	126	5	03 h 29
CFHTLS_W_g_022150-041200	Wide	g	02 h 21 m 50	-04 ° 12 ' 00	123	5	03 h 13
CFHTLS_W_g_022150-050800	Wide	g	02 h 21 m 50	-05 ° 08 ' 00	124	3	02 h 51
CFHTLS_W_g_022150-060400	Wide	g	02 h 21 m 50	-06 ° 04 ' 00	123	6	03 h 25
CFHTLS_W_g_022539-041200	Wide	g	02 h 25 m 39	-04 ° 12 ' 00	942	5	14 h 57
CFHTLS_W_g_022539-050800	Wide	g	02 h 25 m 39	-05 ° 08 ' 00	124	4	03 h 02
CFHTLS_W_g_022539-060400	Wide	g	02 h 25 m 39	-06 ° 04 ' 00	125	2	02 h 34
CFHTLS_W_g_022539-070000	Wide	g	02 h 25 m 39	-07 ° 00 ' 00	119	4	02 h 51
CFHTLS_W_g_022929-041200	Wide	g	02 h 29 m 29	-04 ° 12 ' 00	962	4	16 h 47
CFHTLS_W_g_022929-050800	Wide	g	02 h 29 m 29	-05 ° 08 ' 00	958	5	15 h 39
CFHTLS_W_g_022929-060400	Wide	g	02 h 29 m 29	-06 ° 04 ' 00	958	5	15 h 39
CFHTLS_W_g_022929-070000	Wide	g	02 h 29 m 29	-07 ° 00 ' 00	607	5	05 h 50
CFHTLS_W_g_023319-041200	Wide	g	02 h 33 m 19	-04 ° 12 ' 00	980	7	18 h 46
CFHTLS_W_g_023319-050800	Wide	g	02 h 33 m 19	-05 ° 08 ' 00	970	10	18 h 48
CFHTLS_W_g_023319-060400	Wide	g	02 h 33 m 19	-06 ° 04 ' 00	970	4	17 h 32
CFHTLS_W_g_023319-070000	Wide	g	02 h 33 m 19	-07 ° 00 ' 00	531	5	05 h 28
CFHTLS_W_g_085011-012700	Wide	g	08 h 50 m 11	-01 ° 27 ' 00	46	5	02 h 26
CFHTLS_W_g_090137-012700	Wide	g	09 h 01 m 37	-01 ° 27 ' 00	108	5	01 h 14
CFHTLS_W_g_090526-012700	Wide	g	09 h 05 m 26	-01 ° 27 ' 00	99	2	00 h 53
CFHTLS_W_g_090526-022300	Wide	g	09 h 05 m 26	-02 ° 23 ' 00	120	4	01 h 06
CFHTLS_W_g_090526-031900	Wide	g	09 h 05 m 26	-03 ° 19 ' 00	120	7	01 h 18
CFHTLS_W_g_090526-041500	Wide	g	09 h 05 m 26	-04 ° 15 ' 00	90	4	02 h 27
CFHTLS_W_g_135933+533431	Wide	g	13 h 59 m 33	+53 ° 34 ' 31	17	4	00 h 34
CFHTLS_W_g_141147+533431	Wide	g	14 h 11 m 47	+53 ° 34 ' 31	689	3	11 h 13
CFHTLS_W_g_141754+514231	Wide	g	14 h 17 m 54	+51 ° 42 ' 31	716	4	11 h 51

Figure 3: User interface generated automatically by SPICA, showing a sub-set of T0002 fields produced by the pipeline. Column 6 indicates the number of images used for the astrometric solution produced by SCAMP and column 7 indicates the number of images used to produce a stack with SWarp. Both columns are clickable areas that link to a detailled list and a description of each image used, as provided by QualityFITS-in. The last column displays the time needed to produce the final stack. It depends on the amount of images SCAMP and SWarp used as well as on the TERAPIX node performances that handled the process contained in the SPICA command line.
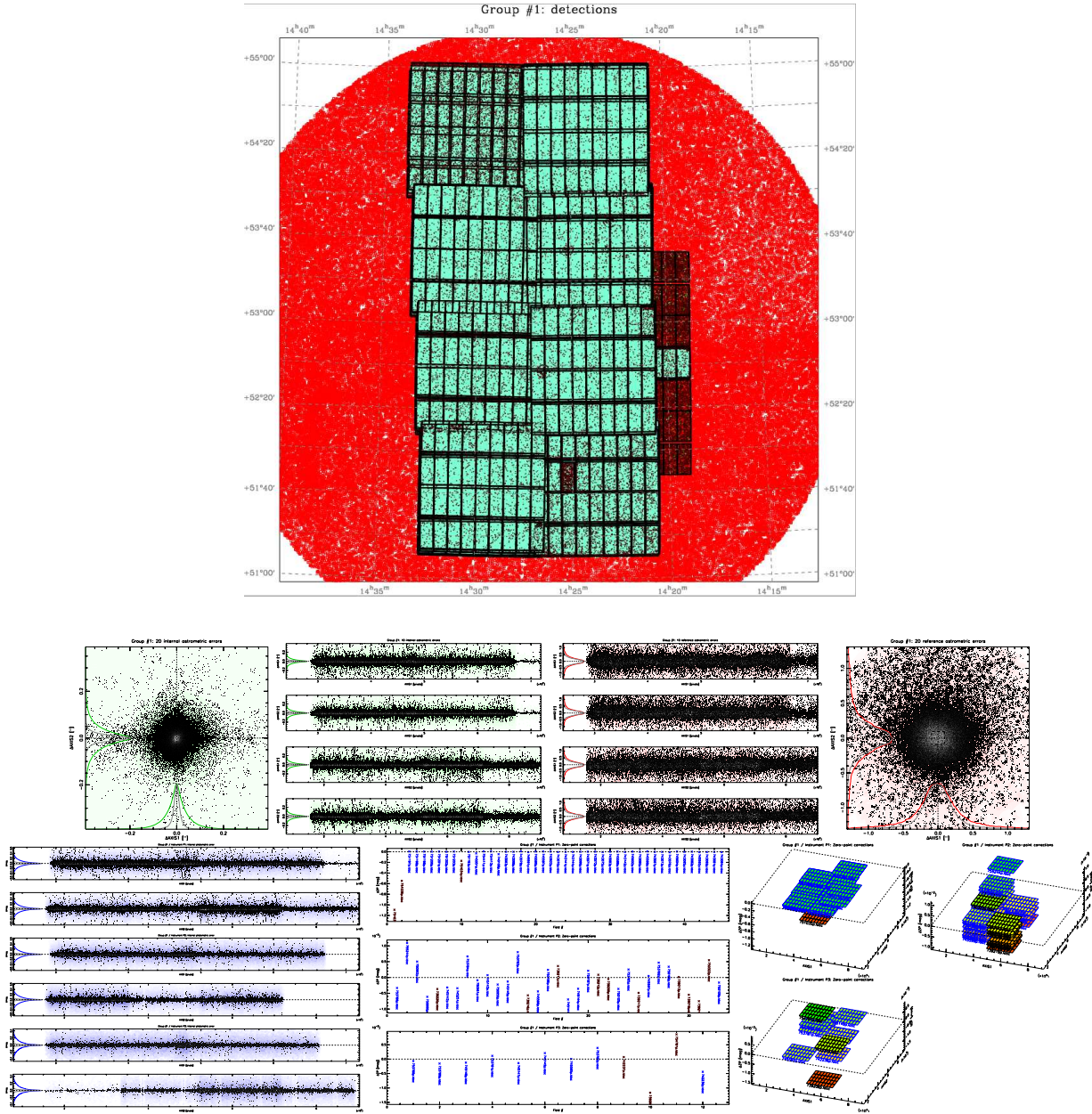
26

Figure 4: A series of quality assessment plots produced by SCAMP for each Deep+Wide stacks and all Very Wide data of the T0002 release. The target is the Wide W3 field in i-band, named w3+2-1 according to the CFHTLS Wide target definition and CFHTLS_W_i_143007+533430_*, according to the SPICA/ICD naming process of released data. The top panel displays all Megacam images that were used for the astrometric calibration. Red dots show the USNO-B1 objects. Green dots are the USNO-B1/Megacam matched objects. The middle panel shows the astrometry quality assessment plots: the left hand panel pairs show the internal errors on 2-dimension or 1-dimension plots. They are based on cross matched pairs located on the overlap regions of Megacam fields. The right hand panel pairs show the external errors. They are based on the rms residual error between the USNO-B1 catalogue and Megacam astrometred positions. The bottom panel shows the internal photometric error and field to field rescaling values derived from SCAMP between each Wide images displayed on top. Megacam images. The bottom panels are photometric calibration plots. Left plot: flux rescaling value derived by SCAMP by matching sources located in overlapping regions. Middel: 1-dimension internal photometric residuals derived by SCAMP after the re-scaling. Left: 3-dimension views of photometrically rescaled relative position of each Megacam image of each group. The z-axis is the zero point offset derived by SCAMP. For all panels, black point are high signal to noise objects, coloured points are low signal to noise objects.
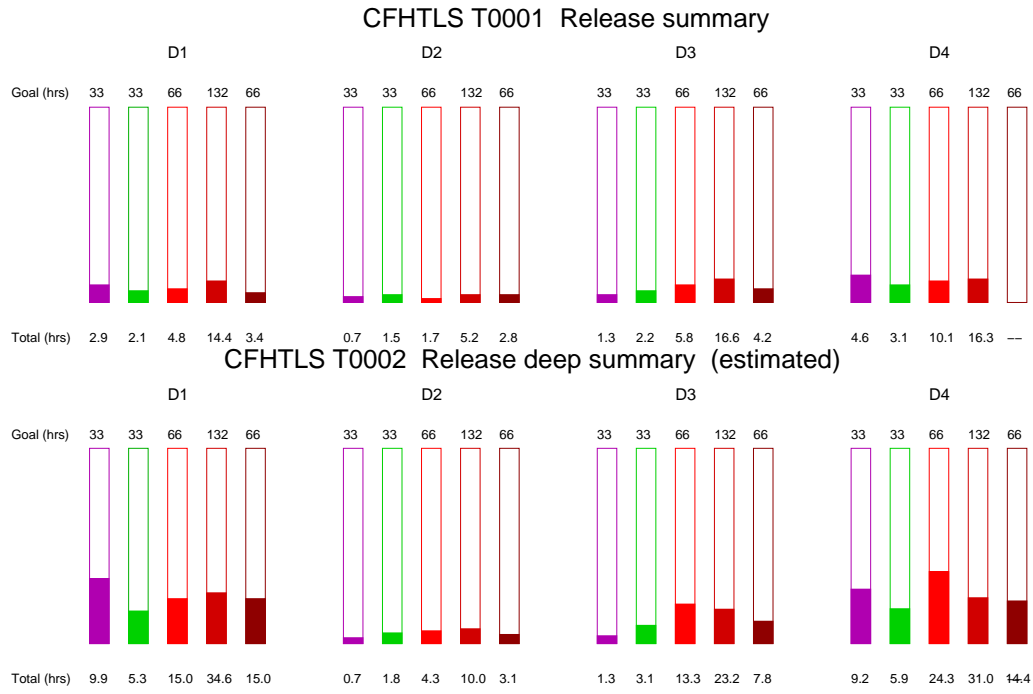
Figure 5: Top panel: CFHTLS T0001 stack summary as compared to the final goals in each filter. T0001 still does not reach 20% of the final release and a significant deficit is observed in D2 and D3 u* with respect to other bands. Bottom panel: anticipated final exposure times of deep T0002 stacked (assuming the on going stacks will be validated).

Figure 6: Summary tables of the Wide w3+2-1 processing produced by QFITS-out and SPICA. A link to QualityFITS-in allow users to view more detailed information regarding individual input images. The information displayed in this panel is available to users for each Deep and Wide stacks and for all Very Wide output data of the T0002 release.