INAF-Osservatorio Astrofisico di Torino

Technical Report nr.182

Tools towards the scientific exploitation of the OATo plate archive

Ana Vrancean*, Roberto Morbidelli, Beatrice Bucciarelli, Luciano Lanteri

***** "Alexandru Ioan Cuza" University of Iasi, Romania, Erasmus Traineeship Programme trainee (18/6-18/8 2018)

Pino Torinese, 18/8/2018

Sommario

Abstract

This technical note summarizes the results of the work conducted on the OATo historical photographic plate archive during a 2-month stage of the first author of this document, within the Erasmus Traineeship programme. The subject of investigation is a catalog of circa 3000 archival plates previously recorded and digitized with a commercial scanner at 600-dpi resolution and 8-bit dynamic range. The principal goals were to analyze the content of the catalog in order to check its consistency and fix potential errors, to produce some relevant statistics, and to generate new files compliant with FITS astronomical standards starting from the available images, to be further exploited by means of ad–hoc astronomical reduction s/w.

Introduction

This is a critical review of the available digitized images, and their associated catalog, obtained from a previous program conducted in the early 2000, which was devoted to a systematic recognition of the OATo astronomical plate archive and to highlighting its scientific potential by the analysis of a selected plate sample.

In particular, we have addressed some issues regarding the archived material, i.e. a master table written in EXCEL recording all the available plate information, and the corresponding digitized images, e.g.:

1. Check consistency of the catalogued plate archive and format of digitized images
2. Inspection of the recorded plate info mainly aimed at homogenizing nomenclature
3. Rebuild images in astronomical standard FITS
4. Production of global statistics of archive properties

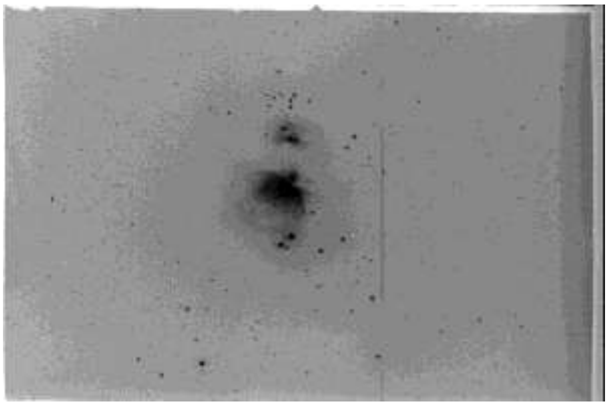
A plate is a medium on which one can detect and store a region of the sky: in this perspective, it is our instrument of observation. The characteristics of the plate, and so the astronomical quality of the recorded field, change as function of the dimension of its photo-sensitive granules; this important parameter is coded in a keyword present in the fits header (see section "Generation of FITS images"). An important concept is that the dimension of granules is related to the resolution of the detected image; from this point of view, the 'goodness' of the plate can be associated to the dimension of granules (the smaller the granules, the better the image details). In any case, the digitization process should be carried out at a resolution comparable with the intrinsic image resolution, which for the material ad hand would correspond to about 2400 dpi (dot-per-inch). All the plates of the present archive have been previously scanned with a commercial scanner at a low resolution of 600 dpi, mainly for preview purposes; nevertheless, a preliminary calibration of such images could be attempted in order to obtain useful information on their astronomical content, such as magnitude limit, photometric properties, detection of selected astrophysical objects.

Assessment of digitized photographic material

The plates are made up of two materials: glass and gelatinous layer. The light-sensitive emulsion (gelatinous layer) of silver salts is coated on a glass plate, typically thinner than common window glass, instead of a clear plastic film. Many famous astronomical surveys were taken using photographic plates. A number of observatories maintain large archives of photographic plates, which are used primarily for historical research on variable stars, but also for dynamical studies of solar system objects, and investigations of different kinds of transient phenomena such as gamma-ray bursts or novae stars.

The digitization of the OATo plates was done during a previous campaign aimed at exploring its large photographic archive that covers about one century of astronomical observations, including also non-scientific plates recording some activities related to the history of the Torino Observatory (see an example below).

The number of digitized scientific images is 2806 in JPEG format, 8-bit resolution. The majority of the exposures were taken with the telescopes of the Astronomical Observatory of Turin (mostly Zeiss and Morais), while a few plates came from the GPO Astrograph at ESO and telescopes of the Harvard Smithsonian collection. A first approach was a visual inspection of the images: such an inspection can immediately tell that some exposures are completely dark and unusable while others are very good. The examples below show two exposures of very different image quality.

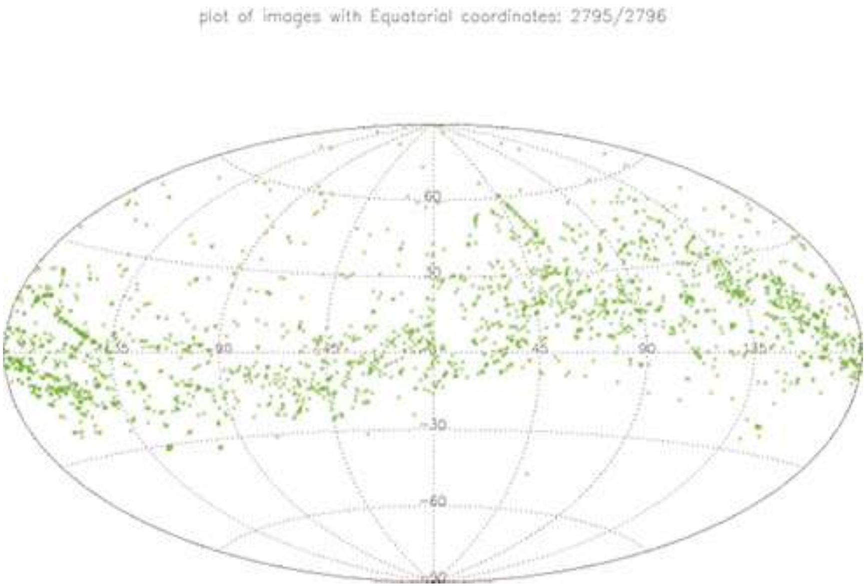


Two examples of plate exposures: the first shows a plate completely saturated, the second case is a part of a plate obtained with the astrograph Zeiss in which the Nebula in the constellation of Orion is clearly recognizable

Two images of non-astronomical subjects. The images are plates obtained probably using a camera of the late nineteen century preserved in the OATo museum

The results of a preliminary attempt to read the photographic material put in evidence that 6 of the digital images were unreadable and we removed them from the list. The remaining images demonstrate, by simply looking at the file system properties, that all scans were produced using a 600-dpi scanning strategy with a 8-bit gray scale (one exception is for image 5550-0001.jpg which was digitized at 24-bit depth). The remaining 2799 images is the complete archive on which the following work was carried out.

If the astronomical coordinates of the field centers are available, the information is reported in the header of the newly generated FITS file; this was the case for all the plates but one, therefore we plot the sky distribution of the 2798 center plate positions in the following Aitoff map projection, in equatorial coordinates.



Distribution of plate centers in equatorial coordinates (Aitoff map projection), as reported in the header fits files obtained by reprocessing the jpeg images (2798 out of the initial 2800 image archive)

We highlight the fact that an important part of this sample lies in the ecliptic plane region, demonstrating that large part of the historical campaigns was devoted to solar system objects, such as minor planets and comets; nevertheless, the distribution of the center points show the presence of other astrophysical objects at higher declinations, representing different scientific investigations carried out in the XX century by the local astronomers.

Standardization of the archived plate images

We have produced some informative statistics of the archive content and at the same time we tried to achieve a minimum standardization of the image files from an astronomical point of view. The existing EXCEL document containing a summary of the scanning operations and of the astronomical information accompanying each plate exposure does not present a one-to-one correspondence between records and plates; in fact, it contains a total of 5546 records: the reason of the discrepancy with the number of image files is that, besides the available digitized material being incomplete, each record is a description of a sky exposure, but some plate contains more than one exposure and the result is that one plate/image is linkable to more than one record in the EXCEL file.

In order to process the entire archive, we built a pipeline of programs written in IDL, which links the image files to the corresponding plate information and converts the JPEG file into a FITS image with the relevant astronomical information included in the FITS header. The choice of IDL was made because a well consolidated astronomical library is available in such an environment, with some routines specifically made to solve some astronomical issues and others oriented at building files in FITS format following the IAU astronomical standards. The resulting pipeline is an interactive and semiautomatic procedure allowing to perform a reprocessing of the archived images, which are checked for readability and completeness of recorded information.

Before doing the reprocessing, the following information, useful for statistical purposes and database queries, was added to each record:

Object code for the different objects	
The objects	**The code**
Comet	C
Galaxy	Gal
Open cluster	OC
Globular cluster	GC
Minor planet	MP

Constellation	CON
Planet	P
Star	S
Satellite	SAT
Calibration	CAL
Miscellanea	M
Radio source	QSO
Ricerca	R
Unknown	UK

Also, in each record was added one more column with the corresponding file name, coded as: ID1-Plate_number.jpg, where:

ID1 is the number under which the plate/image was recorded in the EXCEL table

Plate_number is the plate number. It consists of digits (2447) and letters (J3208); when the plate number is missing, its value in the file name is replaced by ‘NAN’

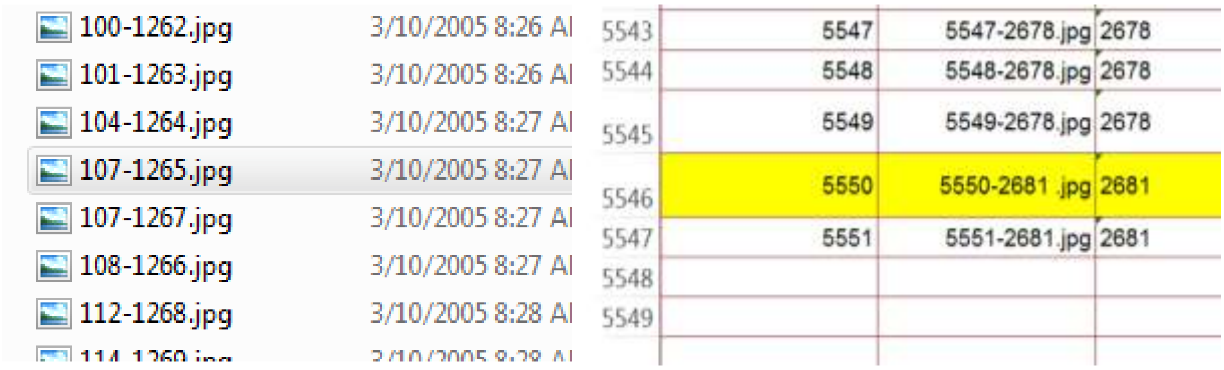
‘jpg’ it is the format of the image

Plate size was introduced in degrees. For different telescopes these data are different. For example: for Zeiss 1^o=2cm, for Morais 1^o=12.3cm, for GPO-ESO 1^o=6.98cm.

We found few inconsistencies in the EXCEL table, which we point out below. One is that the ID1 number is not consecutive, as shown in the top left of the below panel;

48	47	47-1255.jpg	1255
49	48	48-1256.jpg	1256
50	49	49-1257.jpg	1257
51	51	51-1258.jpg	1258
52	52	52-1213.jpg	1213
53	53	53-1217.jpg	1217
54	54	54-1219.jpg	1219

105	105	105-1264.jpg	1264
106	106	106-1264.jpg	1264
107	107	107-1265.jpg	1265
108	108	108-1266j.jpg	1266j
109	109	109-1266.jpg	1266
110	110	110-1267.jpg	1267
111	111	111-1267.jpg	1267

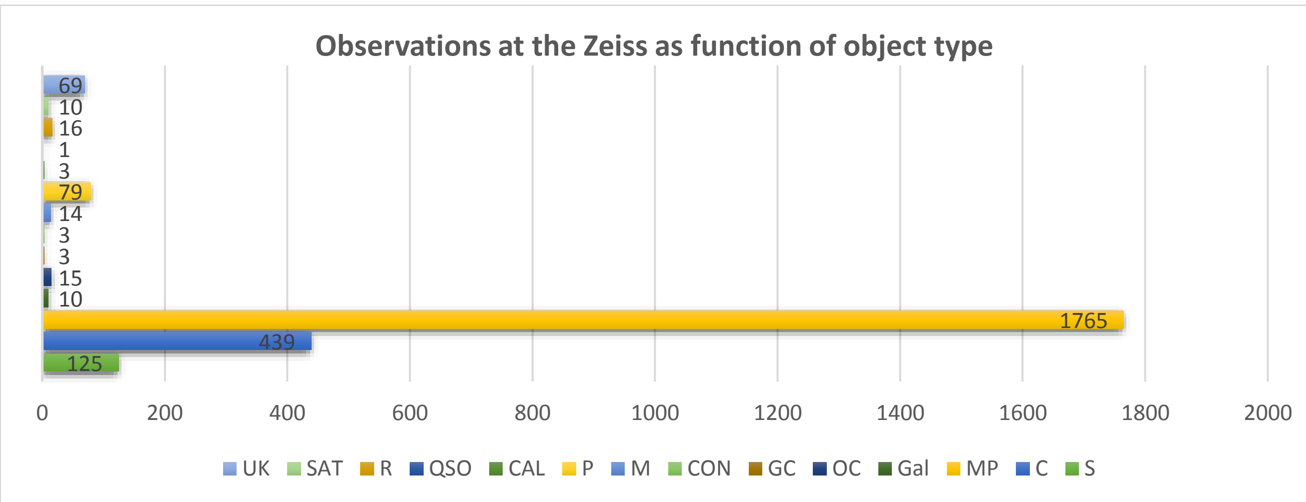


Another inconcistency lies in the difference between the name of the image files and the, presumably, corresponding record in the EXCEL table. For example, two images have the same ID1, but with different plate number, while in the EXCEL archive, these plate numbers have different ID1, see top right and bottom left of the panel; finally, image file 5550-0001.jpg corresponds to ID1 5550 and Plate number 2681 in the EXCEL table, see bottom right of above panel.

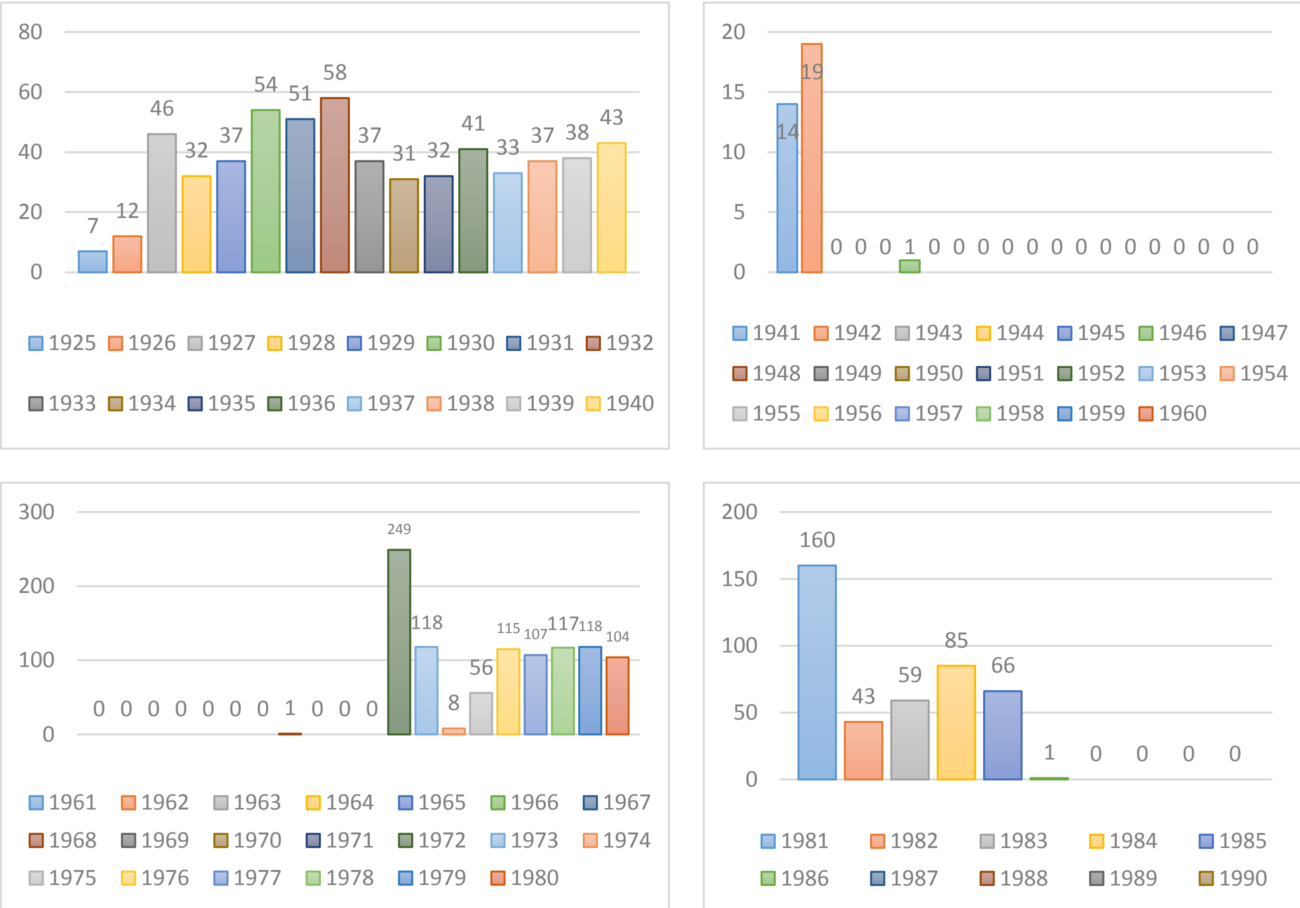
All the empty fields in the EXCEL file corresponding to 'numeric' type were filled in with the value 0, while those which were of 'string' type were filled with 'NAN'.

Graphs of plate exposure frequency distributions

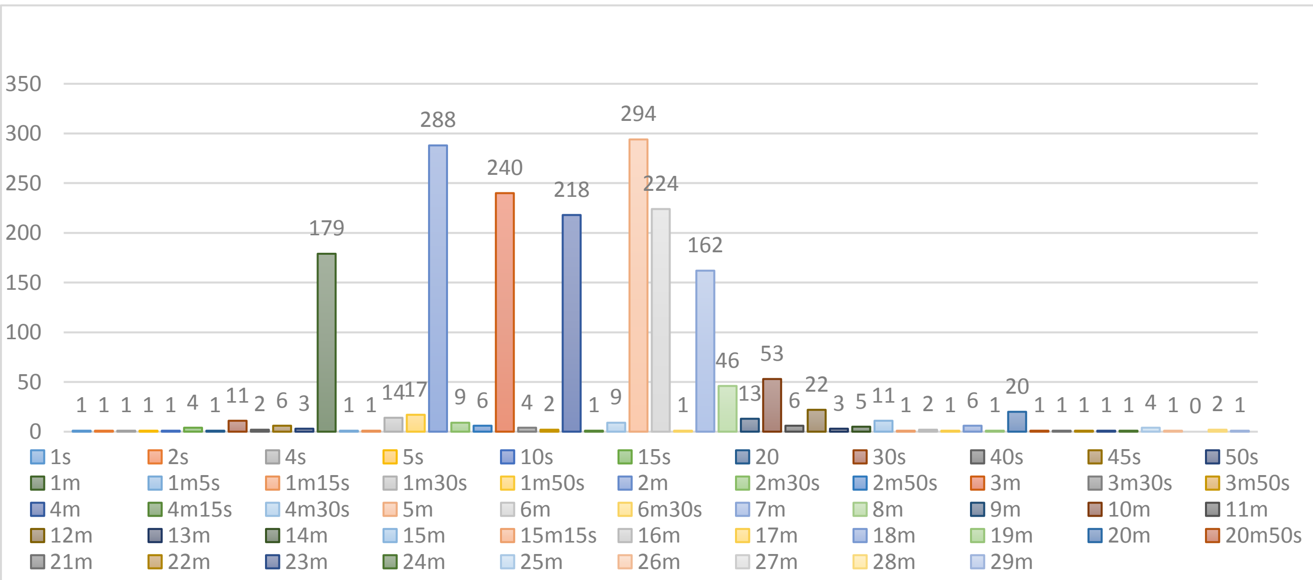
A histogram of the frequency of the type of objects observed at the Zeiss telescope is shown in the following graph. The histogram shows that the minor planets are the most observed objects; furthermore, it points out that the content of the EXCEL file is presumably biased towards the more recent plates as opposed to including a uniform sample of the complete plate material.

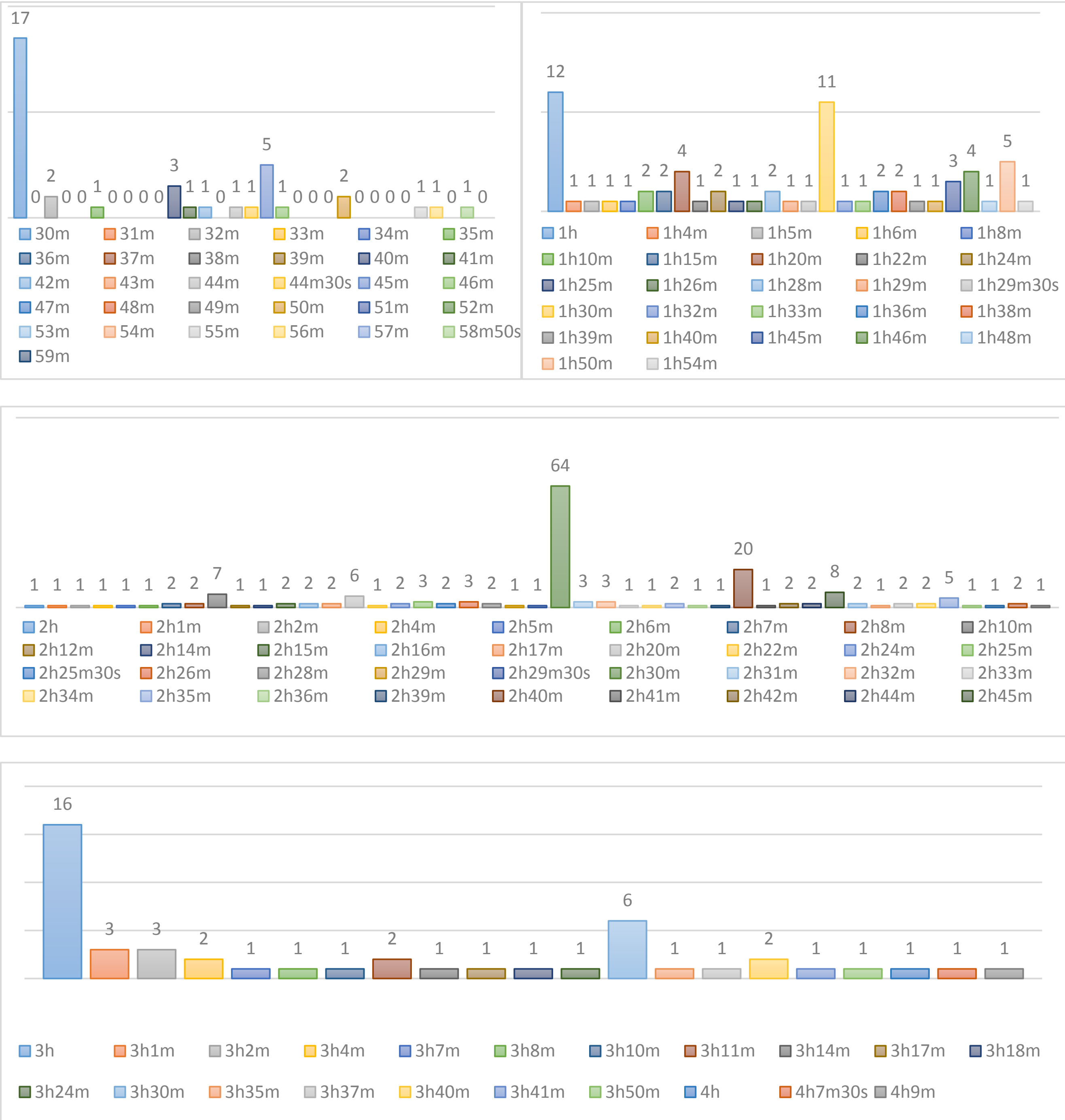


In the following, different statistics are reported for the Zeiss telescope: these histograms, although incomplete because reflecting only part of the photographic work carried out with this telescope, are illustrative of some historical and scientific aspects of astronomical activities at OATo of the last century.

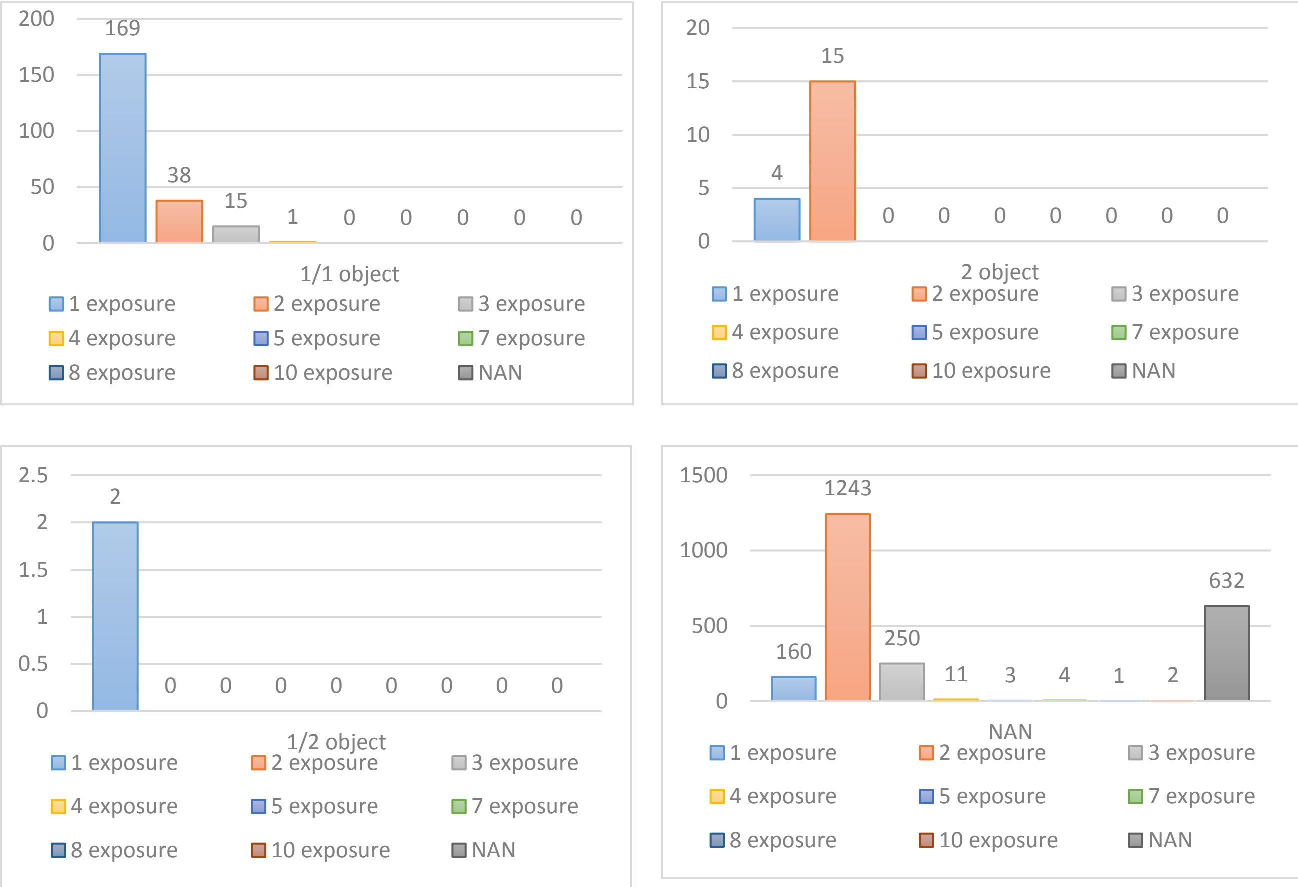


Number of plates exposed at the Zeiss telescope in the years 1925-1990



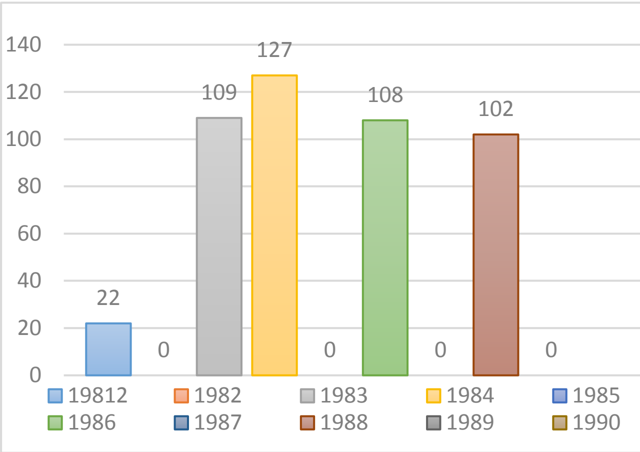


Number of exposures as function of exposure time @Zeiss

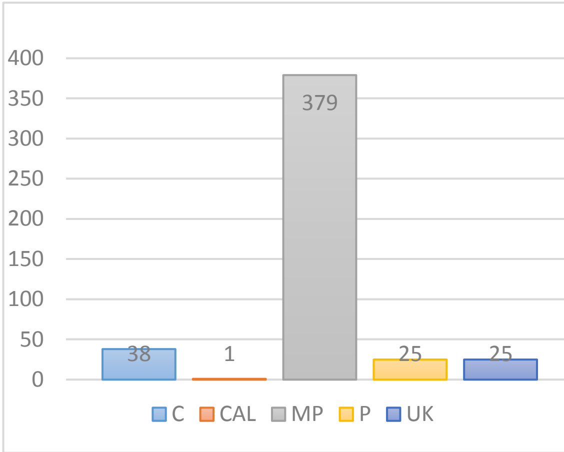


Frequency of single-exposure, multi-exposure, multi-object plates @Zeiss

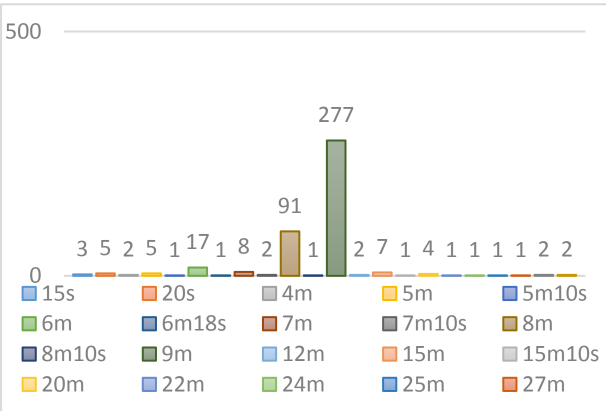
We also report here some informative statistics for the GPO telescope at ESO:



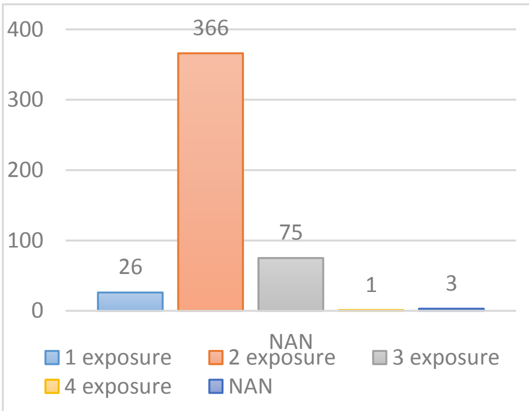
Frequency of observations @GPO-ESO



Observations per object type @GPO-ESO

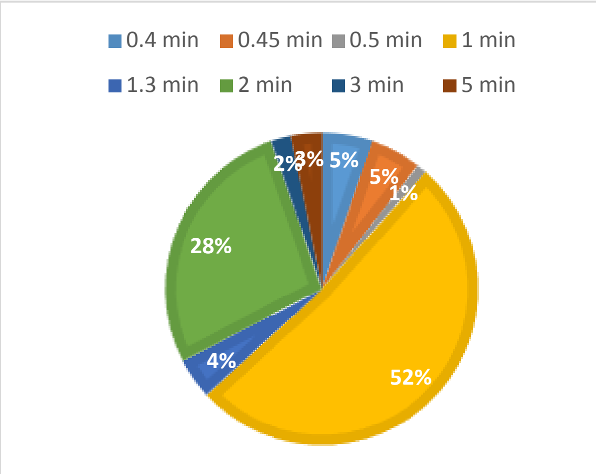


Frequency of exposure times @GPO-ESO

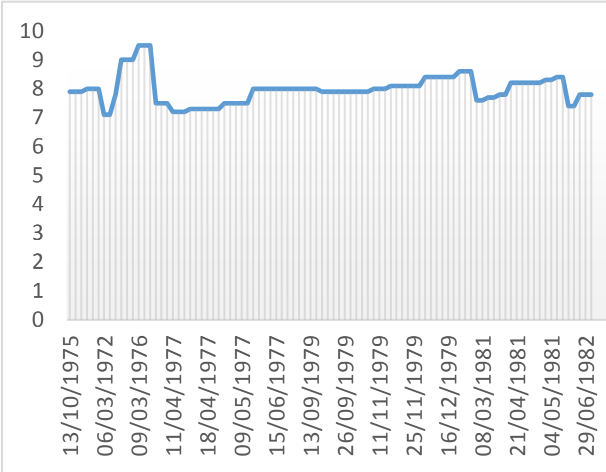


Multi exposures frequency @GPO-ESO

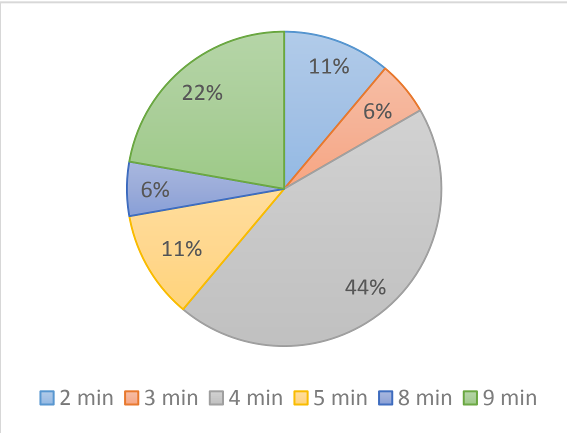
Here are some examples of statistical graphs produced for two different minor planet observations:



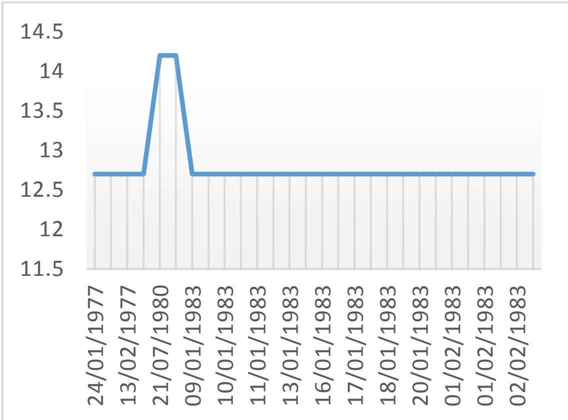
Exposure time @Ceres



Magnitude @Ceres (average is 7.98)



Exposure time @Freia



Magnitude @Freia (average is 12.81)

Generation of FITS images

A correctly generated JPEG image can be conveniently exploited from an astronomical point of view if it contains information that qualifies the physical characteristics of the image, including astronomical qualifiers such as epoch, field coordinates, equinox etc… This has been achieved by writing an ad-hoc IDL code which produces a FITS image and writes in the FITS header all the informations coming from the EXCEL table. Also, an additional table containing some relevant telescopes information is used to compute the scale factor of the images. Such a table is reported in Appendix A.

The condition to build an image is that a JPEG file named after the corresponding EXCEL columns *ID1* and *plate number* exist. In this approach, the astronomical data availability is a sufficient but not necessary condition to create a FITS file.

Some format problems arise when the file is written using *fitsio* NASA routines because the JPEG format has a range of pixel values limited from 0 to 255 and this excursion preserves a poor dynamic range to the image.

The date of observation has been rebuilt to comply with FITS standards, starting from the information contained in the EXCEL table, and written in the form YY/MM/DD. Besides this, other standard and non-standard astronomical FITS keywords are inserted, i.e.:

EPOQUE	the epoch of observation
TELESCOP	the telescope used to acquire the plate image
OBSERVER	the observer who acquired the image
INSTRUME	the instrument used to acquire the data, in our case “plate”
OBJECT	the available information for observed object
FILTER	the filters used or keyword NONE if no filter
PLTSIZE	a non-standard keyword: plate size in cm
M-FIELD	a non-standard keyword: multi-field plate
M-EXPOS	a non-standard keyword: multi-exposure plate
EQUINOX	the reference equinox
DATE-OBS	the date of the observation
RA	Right Ascension of the observed object in degrees
DEC	Declination of the observed object in degrees
DATE	the fits file creation date
AUTHOR	Creator(s) of the given fits file
'REFERENC	OATo-INAF, TN; 180’, (this Technical note)
PLTSCALE	plate scale (“/mm), a function of telescope
XPIXELSZ	X pixel size in microns obtained from scanner properties

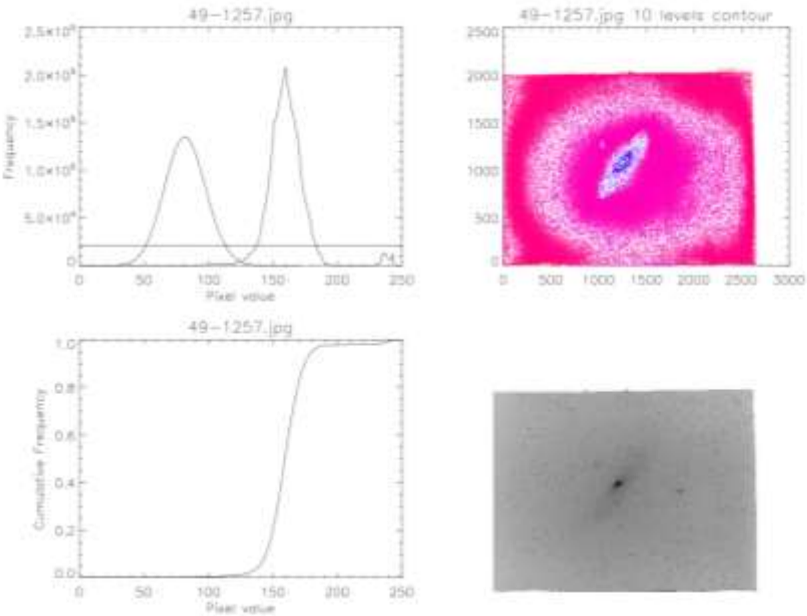
YPIXELSZ	Y pixel size in microns obtained from scanner properties
EMULSION	The type of exposed emulsion
EXPTIME	Exposure time in seconds

Some statistical considerations on image information content

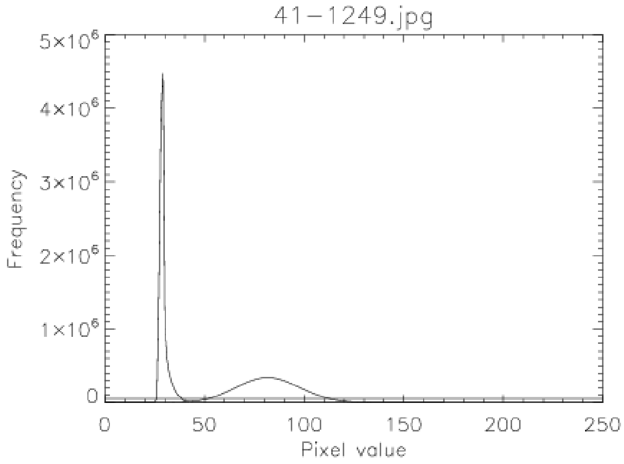
A way to evaluate image quality avoiding the one-by-one visual inspection is to use a statistical approach to image characterization, as detailed below.

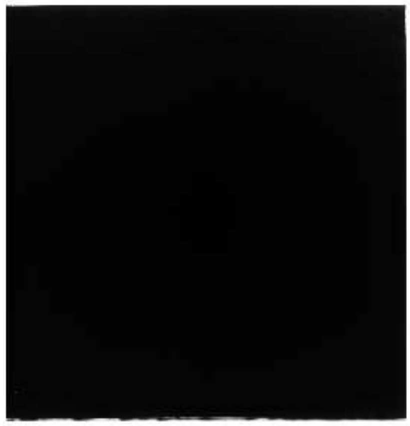
A module written using IDL statistical routines was adopted to generate a report in the form of a postscript file for each fits image. The report produces the following graphics: histogram of pixel-level frequency compared to a standard Poisson distribution of mean λ=256, histogram of cumulative pixel-level frequency and image contour plot.

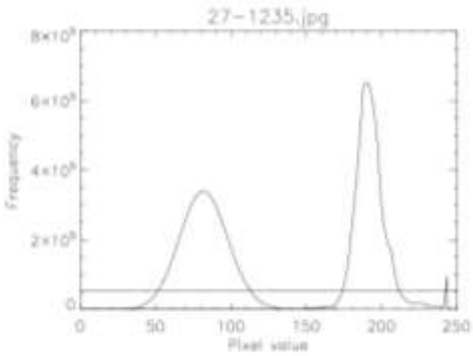
Below we show the reports obtained for 4 selected images, which are representative of different exposure conditions/quality:

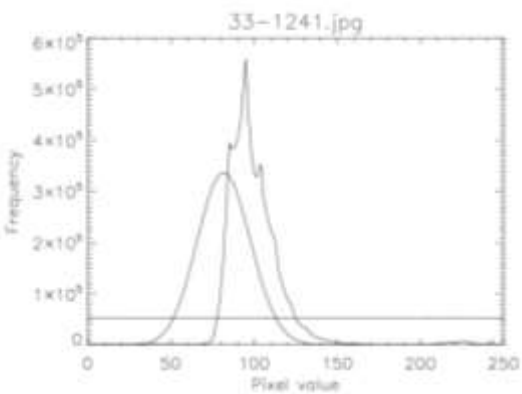


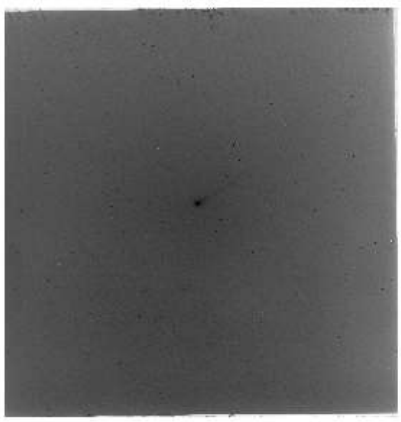
Example of IDL report showing the image, pixel-level frequency and image contour plot


41−1249.jpg
Frequency
Pixel value




27−1235.jpg
Frequency
Pixel value


33−1241.jpg
Frequency
Pixel value



The graphs in the left (top left for the first example), show the frequency distribution of pixel intensities of the actual image, a generic Poisson frequency corresponding to an average pixel level of 256, and a uniform frequency.

Looking at the comparison of the *real* frequency with the Poisson distribution we note the following:

- If the *real* frequency is to the left of the Poisson frequency then the image is dark;
- If the *real* frequency is to the right of the Poisson frequency then the image is lighter;
- If the two frequency distribution intersect then the image is gray

In order to have a better classification, the image is trimmed before computing the statistics. In most cases, the edges of the exposures are too dark or too bright. These defects can be due to the scanning procedure, or to a mix of different causes such as plate damaging, vignetting, fading emulsion. If the plate presents inscriptions which were inserted manually, these 'defects' can create a large error in the classification process, especially when the exposure is generally bright. This is why the inscriptions should be removed before digitizing the image for scientific purposes.

From these findings, we conceived a possible first approach to give a voting for qualifying the images. The classification is done according to the following scale:

0 very bad image

1 bad image, but there is astronomical information

2 Permissible image with astronomical information

3 Image quality is average

4 good image

Comparison of the *maximum* of the frequency distribution of the actual pixel levels with the poissonian one can also be a useful statistical indicator.

Another interesting aspect, which was not addressed by this work, would be to examine the correlation between the distribution of pixel values and the magnitude depth of the plate. Since the plate is a frozen sky image, one could try to use the distribution of pixel "darkness" as an indicator of magnitude "excursion" achieved in the exposure.

This very preliminary analysis has identified a possible first approach to characterizing the plate exposure in terms quality and information content. In any case, this analysis can only have a statistical significance, and the astronomical usefulness of each image must be evaluated case by case, by means of ad-hoc calibration software.

Conclusions

We have presented a critical inspection of some 3000 photographic plates from the OATo photographic archive, previously recorded and digitized at 600-dpi resolution, with the principal aim of fixing practical errors, adding missing information and writing the IDL procedures needed to transform the 8-bit JPEG images into FITS images compliant with the IAU astronomical standard - making them formally ready to be astronomically calibrated. Some statistical graphs have additionally been produced, giving information about the historical and scientific value of the archive. Also, a preliminary approach to an image quality assessment has been conceived which makes use of the real and ideal frequency distributions of pixel densities; this strategy seems promising, although it has been applied to very few images, and would need to be tested on a larger and more diversified plate sample.

Bibliography

R. J. Hanisch, A. Farrs et al.; Definition of the Flexible Image Transport System (FITS); Astronomy and Astrophysics, 376, 359-380 (2001) – DOI: 10.1051/0004-6361:20010923

H. Karttunen, P. Kröger et al.; Fundamental Astronomy; ISBN 978-3-540-34143-7(2003) – DOI: 11685739_2

Renate Budell, Institute of Planetology, Munster, Germany; Technical Bases for Modern Data Reduction on Photographic Plates

D. Ratledge; Digital astrophotography: the state of the art; 1945 – 522.6’3; ISBN 1852337346

https://heasarc.gsfc.nasa.gov/docs/fcg/common_dict.html - Dictionary of Commonly Used FITS Keywords

https://idlastro.gsfc.nasa.gov/contents.html - IDL Astronomy User's Library

Appendix A

Telescope	*Characteristics*	*Number of Plates*	*Dimensions*	*Epochs*	*Programs*
OATo Telescopes					
Astrograph Zeiss	Diameter 200 mm, focal 1140 mm	circa 700	18x24 cm	1920-1935	Asteroids comets; special areas
		circa 1500	9x12 cm	1953-1982	
		circa 500	15x16 cm	1974-1984	
Refractor Morais	Diameter 380 mm, focal 6875 mm	circa 700	16x16 cm	1970-1980	Asteroids; radiosources; open clusters;
		circa 350	16x16 cm 20x20 cm	1983-1990	
Astrometric Reflector REOSC	Diameter 1050mm focal 9942 mm	circa 540	16x16 cm	1972-1983	Double stars; parallaxes
		circa 500	9x12 cm		
		circa 550	16x16 cm 9x12 cm	1983-1990	Radiosources; parallaxes
Other telescopes					
Astrograph ESO/GPO (Cile)	Diameter 400 mm, focal 4000 mm	circa 500	16x16 cm	1981-1988	Asteroids
		circa 400	16x16 cm	1983-1990	Open clusters
Reflector Schmidt (Asiago)	Diameter 920 mm, focal 2150 mm	circa 20	20x20 cm	1960-70	Asteroids; variable stars
Reflector Schmidt at Serra La Nave (Catania)	Diameter 610mm focal 1200 mm	circa 100	9x12 cm	1971	Asteroids
Reflector Cassegrain JKT (La Palma, Canarie)	Diameter 1000mm focal 15000mm	circa 20	18x24 cm	1990	Radio sources
Astrograph Cape Town (Sudafrica)	Diameter 330 mm focal 3438 mm	circa 30	16x16 cm	1992	Open clusters
Miscellanea					
About 1000 mostly old plates, including reproductions obtained from different Observatories, still to be classified.					