BRNO UNIVERSITY OF TECHNOLOGY
VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

FACULTY OF MECHANICAL ENGINEERING
FAKULTA STROJNÍHO INŽENÝRSTVÍ

INSTITUTE OF MATHEMATICS
ÚSTAV MATEMATIKY

NUMERICAL METHODS OF IMAGE ANALYSIS IN ASTROMETRY
NUMERICKÉ METODY ANALÝZY OBRAZŮ V ASTROMETRII

DOCTORAL THESIS
DIZERTAČNÍ PRÁCE

AUTHOR — **Mgr. Michaela Honková**
AUTOR PRÁCE

SUPERVISOR — **prof. RNDr. Miloslav Druckmüller, CSc.**
ŠKOLITEL

BRNO 2018

Abstrakt

Velmi přesná následná astrometrie je nezbytným předpokladem sledování blízkozemních objektů, které mohou představovat riziko srážky se Zemí. Tato práce přináší ucelený přehled přesné astrometrie, obsahuje potřebnou matematickou teorii, postup předzpracování snímků v astronomii, a nastiňuje použití filtrů. Navrhuje nové metody pro vyrovnání pozadí snímků před provedením astrometrického měření pro případ, kdy nejsou dostupné kalibrační snímky. Tyto metody jsou založeny na vytvoření syntetického flatfieldu pomocí aplikování filtru na snímek a následné užití tohoto flatfieldu pro odstranění pozadí snímku. Metody byly otestovány na vzorových snímcích a vzápětí použity k získání astrometrických pozic prvního mezihvězdného objektu `1I/2017 U1 ('Oumuamua)`.

Abstract

High-precision follow-up astrometry is essential for tracking Near-Earth Objects, which may pose risk of impact to the Earth. This thesis brings an overview of astrometric process and requirements of precise astrometry, contains the necessary mathematical background, steps of image pre-processing in astronomy, and outlines the use of filters. New methods are proposed to level the background of images prior to performing astrometric measurements for cases when no calibration images were taken. They are based around applying filters on an image to create artificial flatfield, which is then applied back on the image. These methods are tested on sample data and immediately used to obtain astrometry of the first interstellar object `1I/2017 U1 ('Oumuamua)`.

klíčová slova

astrometrie, zpracování obrazu, flatfield, filtry

keywords

astrometry, image processing, flatfield, filters

HONKOVÁ, M. *Numerical Methods of Image Analysis in Astrometry.* Brno: Vysoké učení technické v Brně, Fakulta strojního inženýrství, 2018. 110 s. Vedoucí disertační práce prof. RNDr. Miloslav Druckmüller, CSc..

I declare that I have written the doctoral thesis *Numerical Methods of Image Analysis in Astrometry* on my own according to advice of my supervisor prof. RNDr. Miloslav Druckmüller, CSc., and using the sources listed in references.

March 14, 2018 Michaela Honková

I would like to express my gratitude to Kleť Observatory for allowing me to utilize their images for this work.

I would like to thank my supervisor prof. RNDr. Miloslav Druckmüller, CSc. for supporting me with ideas and imparting unflinching enthusiasm for image processing.

Contents

Chapter 1

Introduction

Asteroids and comets are small, rocky bodies of the Solar System, orbiting the Sun along its planets. Some of them come dangerously close to the Earth and even collide with our planet, occasionally causing mass extinction event. An ability to track and compute position of these so-called Near-Earth Objects (NEOs) ahead of time gives us unprecedented opportunity to avoid this natural hazard to mankind.

To be able to predict an impact event, these objects have to be closely monitored over time. The measuring of their positions on the sky in given time is called astrometry, history and present practice of which is detailed in Chapter 2, following with Chapter 3, which gives overview of the mathematics required for further chapters. The most critical task is obtaining enough of precise positional data of newly discovered objects, called follow-up astrometry, which is essential to compute meaningful orbit of the object and therefore predict its future position.

Kleť Observatory, located on top of Kleť mountain in Blanský forest, operates the biggest telescope in Europe dedicated to follow-up astrometry of NEOs. Being one of the few professional observatories which survived transit from photographic plates to Charge Coupled Device (CCD; essentially a special digital camera), it strives ceaselessly to improve its high precision astrometry, be it by incorporating new hardware, swapping to better stel-

lar catalogs or automatizing processes whenever feasible. Their telescopes, software and work-flow are detailed in Chapter 4, and the process of image acquisition, pre-processing, and the astrometric algorithm itself, in Chapter 5.

The digital images of Kleť Observatory suffer from very uneven background, hindering the performance, or even the possibility, of astrometric measurement. The usual pre-processing methods given in Chapter 5 can not be used in the case of Kleť Observatory, but they inspired newly designed approach detailed in Chapter 6. Six filters are proposed, programmed into software the author is developing for Kleť Observatory's needs, and tested on sample data. Their performance is compared in Chapter 7 and recommendations for their use by the observer are assembled. These filters allow for construction of artificial flatfield (background image), from as little as a single digital image, used to remove the uneven background, eliminating the need to take calibration images during the observations, and allowing re-measurement of vast image archives for which no calibration images are available.

The new filters were immediately put to use to allow astrometry of the first interstellar object `1I/2017 U1 ('Oumuamua)`, as described in Chapter 8. The work is closing with Conclusion in Chapter 9, being followed by the Bibliography, Nomenclature, and finally, Appendices.

Chapter 2

Astrometry of asteroids and comets

2.1 Asteroids and comets

Asteroids and comets are solid bodies of the Solar System orbiting the Sun together with planets and dwarf planets. Their sizes range from several meters to hundreds kilometers big boulders. There is an estimated amount of millions of these bodies in the inner Solar System alone.

To be able to track these bodies, a set of *orbital elements* is computed for each. These can be in turn used to compute *ephemeris*, which is a position of the body on the observer's sky in a given time. Six *orbital elements* are required to unambiguously define an orbit in space to cover all six degrees of freedom of the orbit. There are three parameters to define position and three parameters to define velocity in each position. Due to computational convenience, different set of six parameters is traditionally used in astronomy. These *Keplerian elements*, depicted on Figure 2.1, are an idealized representation of an orbit conforming to Kepler's laws without regard to gravitational perturbations, therefore it is required to state the time *Epoch* for which the approximation is valid. Assuming elliptical orbit,

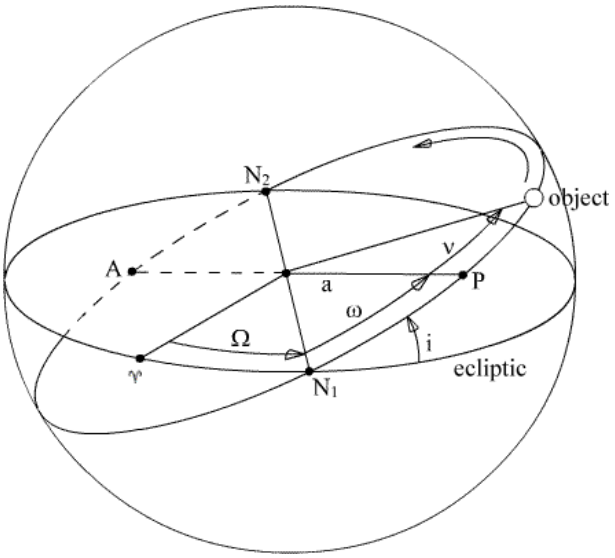


Figure 2.1: Orbital parameters' illustration

e ... *numerical excentricity* gives shape of the elliptical orbit, $0 < e < 1$

a ... *semimajor axis*, the longest distance between two points of the orbit, decides the size of the orbit [AU]

i ... *inclination*, an angle between the orbit and the ecliptic orients the plane vertically [deg]

Ω ... *longitude of the ascending node*, an angle between the ascending node (where the orbit passes upward through ecliptic) and the vernal point Υ orients the orbit horizontally [deg]

ω ... *argument of perihelion*, an angle between the ascending node and perihelion (the point closest to the Sun on the orbit), defines the orientation of the ellipse in the orbital plane

M ... *mean anomaly* gives the position of a mean body on the orbit at the time of Epoch [deg]

There are many subcategories of asteroids and comets sorted by their orbits. Of an immediate interest to mankind are NEOs, Near-Earth Objects. NEOs are asteroids and comets with *perihelion distance* less than 1.3 AU, placing them inside the inner Solar System. There are over 17 700 discovered NEOs

as of January 2018 [1]. Most of NEOs are asteroids, henceforth denoted as NEAs, and they are further divided into four groups according to their *perihelion distance*: Atira (18 asteroids), Aten (1 298 asteroids), Apollo (9 625 asteroids) and Amor (6 666 asteroids), named by the first discovered asteroid of each group.

As of January 2018, 1 886 of these bodies are listed as Potentially Hazardous Asteroids (PHAs). The definition of PHA is based on its ability to threaten the mankind: specifically its size and proximity of its orbit to the orbit of the Earth. A PHA is an asteroid of an absolute magnitude 22.0 or less (corresponding to size at least 150 meters when we assume average albedo of asteroids 0.13) with an Earth MOID (minimal orbit intersection distance, the minimal distance between the orbit of the asteroid and the orbit of the Earth) 0.05 AU or less. About 8 000 of known NEAs have diameter over 140 meters, which makes them capable of causing local disaster upon impact, while about 900 known NEAs have diameter bigger than 1 kilometer, having the potential to cause mass extinction event upon impact to the Earth. To evaluate the risk and predict possible impacts, accurate orbits must be computed, which requires precise astrometry of asteroids and comets to be carried out.

2.2 History of astrometry

Astrometry is a precise measurement of positions and movements of celestial bodies. It includes observational techniques, instrumentation, processing and analysis of observational data, positions and motions of bodies, reference frames, and relies on a number of theoretical aspects, which relate the observations to the laws of physics. Among the most important are celestial mechanics, optics, theory of space and time references, astrophysics, and statistical inference theory.

Astrometry is a fundamental tool for astronomy as it provides data to many

[1] `http://neo.jpl.nasa.gov/stats/`

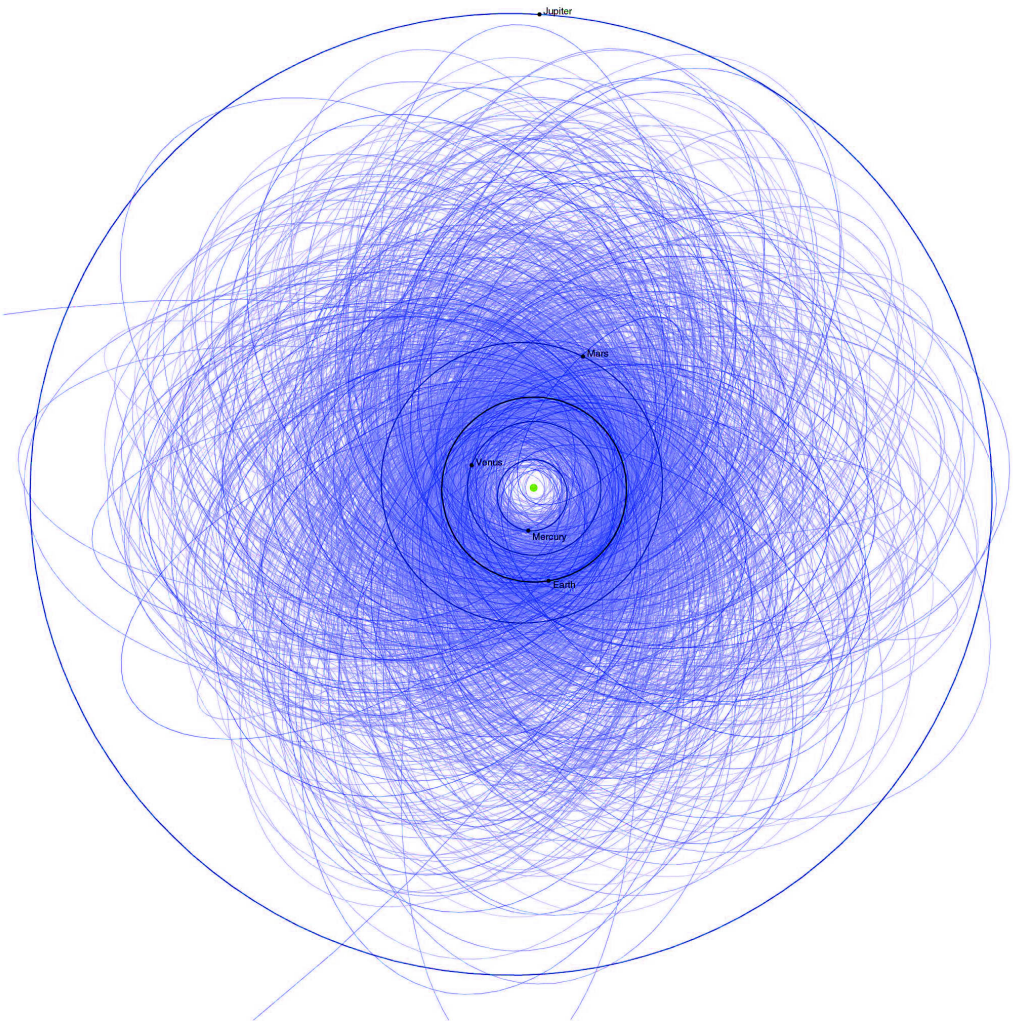


Figure 2.2: The orbits of known PHAs in the inner Solar System.
Source: nasa.gov

of its branches: stellar catalogs assembled with use of astrometry provide reference frame to all other celestial observations and allow us to identify observed objects; stellar dynamics uses stellar catalogs to determine the collective motion of stars in the Galaxy; celestial mechanics uses them to compute motions of celestial objects such as stars, planets including extrasolar planets, asteroids, the Moon and the artificial satellites. Astrometry is also used to determine parallax of the closest stars, thus being an essential step in establishing the scale of cosmic distances. At last, astrometry allows for discovery and tracking of PHAs by detecting these faint, moving bodies against the background of catalogued stars.

Unaided eye

Astrometry emerged with the first naked-eye observations of the sky, leading to discovery of *wanderers* (planets) among the stars, knowledge of the cycles of days and years, prediction of eclipses, calendars and catalogs of observations. The naked eye greatly limited the faintness and positional accuracy of observed objects. Astrometry was based on measuring angles between the objects and drawing these observations down. The tools used were Jacob's staff, also known as cross-staff, triquetrum (parallactic ruler), quadrant, octant, or sextant. The only way to get better measurement was making these devices ever larger, and fixing them to the floor instead of using small, portable versions.

In 150 BC, Hipparchus discovered precession of Earth's axis from observations with positional accuracy 20' [8], using observations of Spica made 160 years earlier, and also assembled first stellar catalog. In 150 AD, Ptolemy issued his *Almagest*, a famous scientific work of his time, which included his own stellar catalog and also theoretical explanation of the planetary motions, placing the Earth into the center of cosmos.

The islamic culture provided the stellar catalogs of Al-Sufi in 960, using armillary sphere, and of Ulugh Beg in 1430, who used a huge sextant, of precision of the order 5–10'. Tycho Brahe (1546–1601) improved quadrant

and sextant by mounting sights and scales on them. His observations with precision of 15–35" were unrivaled in his time, and had a wide impact on astronomy. The observations of planets (particulary of Mars, which was the only planet with observations of the necessary precision) led Kepler to develop his laws of motion, which in turn led to Newton's Universal Law of Gravity. However, it was still impossible to measure stellar parallax, which cast doubts upon Copernican heliocentric theory, since the parallactic displacement of stars due to motion of the Earth around the Sun remained undetected.

Telescope

The discovery of telescope in early 17th century led to a significant jump in accuracy of the measurements due to magnification of the viewed scene. Galileo used this new device to discover the moons of Jupiter and performed on them the first telescopic astrometric observations.

The application of astrometry for navigation and time-keeping led to the founding of national observatories in Paris (1667), Greenwich (1675), Berlin (1701), and St. Petersburg (1725). The telescopes were mounted on quadrants and sextants and the field of view in the telescope was manually drawn and the drawing was later processed. At the Royal Observatory of Greenwich, John Flamsteed produced first great stellar catalog based on telescopic observations *Historiae Coelestis* (1725).

In 1718, Edmond Halley, who would succeed Flamsteed as Astronomer Royal, showed that the bright stars Aldebaran, Sirius, and Arcturus were displaced by many arcminutes from their positions in antiquity, thus discovering proper motions of stars. In 1728, Halley's successor James Bradley discovered stellar aberration, a displacement of a position of a moving object due to the finite speed of light. Later on in 1748 Bradley detected the periodic wobbles of Earth's axis known as nutation, which can amount to 18" difference in position on the sky. The discoverer of planet Uranus, German astronomer William Herschel, detected the motion of the Sun toward constellation Her-

cules from an analysis of stellar proper motions. In 1801, dwarf planet Ceres was discovered, the circumstances of which are detailed in my Bachelor's thesis [7].

Another device widely used in this era was a meridian circle. It was a specialized telescope attached to a circle which was mounted to a fixed horizontal axis, and the time of passage of objects across local meridian was measured, along with their angular distance from the nadir (zenith). With the knowledge of sidereal time (the hour angle of the vernal point), it was possible to deduce right ascension of the observed object, defined in such a way that the sidereal time is equal to the right ascension of the object on the observer's meridian. Meridian circle was first used in 1689 in Copenhagen, and its design was modified in 1806 to allow for declination measurement as well. For over 150 years, the meridian circle was the prime instrument for large-angle observations.

In 1838 Friedrich W. Bessel, well known for his precise observations, used meridian circle to achieve the first measurement of a stellar parallax, measuring parallax 0.315 arcseconds of star 61 Cygni, and thus the truth of the Copernican theory was finally verified and method of directly measuring accurate stellar distances was established. In 1844 Bessel announced discovery of unseen stellar companions of stars Sirius and Procyon, derived from variations of the proper motions of those stars, using positional observations of 0".7 accuracy and proper motions of 0".5 accuracy. These companions were later in the century confirmed by direct telescopic observations.

Photographic plates

While the meridian circle was good for accurate measurements of stars on the hour circle of the instrument all over the night, the photographic plate could observe many stars at once in a small field of view. Starting on the end of 19th century, photographic plates revolutionized astrometry. The precision of observations improved again and the glass panels coated with light-sensitive emulsions removed the disadvantage present through all previous improve-

ments — the necessity to use human eye for recording of the observations. The recorded images also allowed for use of more sophisticated mathematical methods to obtain the objects' coordinates, and allowed for archiving as well as for repeated measurements.

The international Astrophotographic Congress in Paris in 1887 coordinated the first photographic sky survey, the *Carte du Ciel*, resulting in *Astrographic Catalogue* (AC). This survey reached about 12th magnitude and subsequent photographic surveys went as deep as 21st magnitude. Long-focus refractors were being used at the end of nineteenth century in small fields for observations of binary stars, asteroids and satellites.

Charge Coupled Device (CCD)

The last decade of 20th century have experienced a revolution in astrometry due to Charge Coupled Devices (CCDs) replacing use of photographic plates. A CCD works using photoelectric effect. Its image chip consists of a matrix of capacitors covered with a layer of light-sensitive semiconductors which, upon absorbing photons, expel electrons. The electron charge accumulates in the capacitors and at the end of exposition the total charge of electrons on each CCD pixel is amplified, converted to voltage and further to ADU (Analog Digital Units). The relation between the electron count and ADU count is called Electron gain and expressed in electrons/ADU. The number of electrons a pixel can store is called Full Well Depth, and the maximum ADU count available is for example for 16-bit images 65 536 values. Electron gain is usually set to such a value, that Full Well Depth corresponds to the maximum ADU count. At such settings, the Full Well Depth and maximal pixel value are reached nearly simultaneously. The quantum efficiency [2] of a CCD is much higher than photographic plates could ever achieve, reaching over 90% compared to 1% of photographic plate and 0.1% of unaided human eye. The methodology for coordinate determination derived for photographic plates remained principally the same and was applied on the digital images from CCDs. The accuracy of measurements improved to 1 mas (miliarcsec-

[2] the percentage of registered photons

onds), but the image processing became much quicker because of computers, which allowed for advanced and near-automatic image processing.

CCDs can also be fitted to satellites to perform astrometry from space and therefore avoid the distortions caused by the atmosphere. The Hipparcos satellite was a dedicated astrometric satellite launched in 1989 to measure stellar parallaxes, determining positions and proper motions of 120 000 stars, and parallax of 60 000 stars, with 1 mas accuracy. Next astrometric missions could reach microarcsecond accuracies for stars as faint as 20th magnitude. Such precision implies that the light path between the source and the detector has to be modelled with a similar accuracy, making use of general relativity and sophisticated mathematical models.

Guide Star Catalog (GSC) was formed by digitization of Palomar photographic plates for purpose of guiding Hubble Space Telescope. Due to its dense full sky coverage allowing for a number of stars being present at typical CCD's field of view, GSC became the first widely used astrometric catalog. USNO catalogs had been in long and widespread use, progressively as they were issued. Starting with USNO-B1.0, proper motions of stars are included, so the quality of the catalog does not deteriorate that quickly with time. First UCAC catalogs had relatively few stars for most astrometric use and UCAC2 was covering only declinations -90 to +40 degrees, so they were not generally in use until UCAC4 removed these flaws. Although UCAC4 has just one ninth of the stars of USNO-B1.0, due to its much higher accuracy the results are by far superior to USNO-B1.0. UCAC4 catalog is the most precise astrometric catalog to date.

2.3 NEO astrometry in present

Especially for observatories focused on NEOs research, flexibility in observing plan is crucial. Smaller NEOs are often discovered when passing close to the Earth and failure in promptly obtaining enough data to determine their orbit results in inevitable loss of the asteroid. Accordingly, newly discovered NEO should be observed and its astrometric positions sent to the Minor Planet

Center (MPC) as soon as feasible. With the additional data available, the accuracy of the asteroid's orbit increases and there is more time available to obtain further observations before the asteroid would become lost due to its deviation from its predicted position. With enough observations, the computed orbit becomes precise enough to allow for recovering the body at its next favorable apparition. This fast and accurate astrometry of newly discovered bodies is called follow-up astrometry. Examples of its profound effect on body's orbital parameters accuracy can be found in my Bachelor's work [7].

The astrometric observations of minor bodies are collected and processed by Minor Planet Center (MPC), which operates at the Smithsonian Astrophysical Observatory, under International Astronomical Union. The MPC is responsible for the designation of minor bodies (asteroids, comets, moons) in the Solar System, for the efficient collection, computation, checking and dissemination of astrometric observations and orbits of minor planets and comets.

2.3.1 The requirements of precise astrometry

The images entering astrometry are affected by various errors, translating into inaccuracy in obtained position. That leads to inaccuracy in computed orbit, and then of predictions of the body's positions (and possible impacts to the Earth) in future. As there are only a few observations available soon after the discovery, the accuracy of every single one is that more important. Their insufficient quality can result in loss of the asteroid, thus follow-up astrometry has to be not only quick, but also precise. The essential requirements for sufficient astrometric precision are given by the 80 characters long format, in which the observations are sent to Minor Planet Center[3] and are as follows:

- Time measured to accuracy under 1 second
- Topocentric position determined to at least 15 meters
- CCD image with small pixels, high dynamic range, low noise and high quantum efficiency

[3]described at http://www.minorplanetcenter.net/iau/info/OpticalObs.html

- Stellar catalog with high accuracy evenly covering the sky
- Image processing at least partially automatized
- Robust astrometric method

2.3.2 Time determination

As the astrometry is measuring the position of the object in the given time, even the best positional measurement is worthless without well measured time. The astrometric observation format sent to Minor Planet Center allows for time to be recorded with 0.00001 days precision, translating into 0.864 seconds. The computer time itself is not reliable enough. Freeware programs are available to synchronize computer time with chosen time servers, achieving better than one second accuracy.

Global Positioning System is the most widespread supplier of precise time, the receivers providing accuracy down to 0.1 millisecond. Another option is to use DCF77 time signal for Europe or WWVB time signal for North America, which are synchronized with the closest available atomic clocks. Most receivers determine the time to 0.1 second accuracy, which is reasonable for astrometric measurements[2].

The start of exposure is however delayed by CCD shutter, therefore its reaction time must be on sub-second scale or the beginning of the exposure must be corrected for the delay. The errors allowed for astrometry are related to the time it takes for the sky to rotate by 1", which means the shutter delay must be lower than 0.07 second.

2.3.3 Topocentric correction

An ideal image and precise time are by themselves not enough to carry out precise astrometry. The observation is basically a vector in three dimensional space oriented toward the observed object and projected on the two dimensional plane of Earth's sky. To complete the vector, the observer's position is required. The closer the object is, the more accurate position of the observer is required to avoid parallax error. For the Solar System bodies, geocentric

position of observer becomes insufficient approximation and topocentric corrections are introduced. The topocentric coordinates have to be determined to precision better than 0.1" for precision astrometry[32], translating into 15 meters accuracy.

2.3.4 Image distortions

Large reflector telescopes mounted in observatories' domes equipped with CCDs and manipulated by remote controlled motors are the norm of todays professional observational astronomy. Reflector telescopes perform better than refractors, which suffer some loss of incoming light through their lenses and cause color aberrations of the image. The motors of the mount allow for accurate movement across the sky. The CCDs, as described above, have very high quantum efficiency, thus their resolution is limited mainly by seeing of the atmosphere, and are capable of recording the observation in a form of an image file for later, repeated use.

The images are typically taken by a CCD in grayscale, meaning every photon regardless of its wavelength (color) results in an electron gain as given by the camera's spectral sensitivity by its manufacturer, and the information of its color is therefore lost. That can be circumvented by using narrow-band filters, which only let in photons of specific wavelength, to obtain the color components, and assembling the color image in the computer. Astronomical images use special file formats, which allow for appending various data about the conditions the image was taken in in the file's header and saving the image data without compression, which is crucial for subsequent scientific use.

Since the precision of astrometric measurements in modern time is so high, many small sources of inaccuracies have to be dealt with, which were negligible in past. Going from the source to the detector and then to the final image, the quality of the final image may be affected by following effects, depending on the site, telescope and the camera:

Cosmic rays, which are high energy particles coming randomly from the

space. When hitting a CCD detector they appear as bright dots or lines on a picture. The atmosphere protects ground-based telescopes from most of the cosmic rays.

Scintillation, an apparent flickering of position and brightness of the object, caused by the packets of light passing through the atmospheric cells with different refractive index. The change in the refractive index is caused mostly by temperature changes between layers of the atmosphere, in smaller amount by humidity changes and air turbulences. Larger telescopes reduce this effect as they cover more atmospheric cells. Considering the long exposure time of astronomical images, this results in the originally dot sized object being blurred into a fuzzy ball called a *seeing* disc. The seeing is measured in arcseconds at FWHM[4]. Very good observing sites have seeing under 1' and average site has 2–3".

Sky brightness, diffused light which is present on the night sky even in absence of light sources of interest. The light comes from the Moon, nearby cities, Milky Way, air glow, zodiacal light and others, being diffused by the atmosphere and setting an impassable brightness limit on possible observations.

As the telescope is of finite diameter, the image forms an *Airy disc* due to light diffraction. However, for most professional telescopes, the effect is negligible in comparison to seeing.

Telescopes also suffer from inherent image distortions called *aberrations*, caused by nonideal mirrors and lenses. Spherical aberration comes from nonideal shape of the mirror spheroid, so it is better to use parabolic shape for mirrors. Coma causes off–axis point objects to appear elongated. Astigmatism appears when the different areas of the mirror have slightly different focal length. Distortion of field happens when the image scale depends on the distance from the optical axis. Chromatic aberration is caused by refractive

[4]Full Width at Half-Maximum, which is the width of a stars profile at height of half the maximum value.

index dependence on wavelength and is therefore relevant only for refractors. Aberrations can be significantly reduced all at once by lens corrector in front of the CCD camera. Typically, three lenses are required in the corrector's construction.

Vignetting occurs when the arriving wavefront is truncated by objects in different planes obstructing the incoming light. Usually those are the edge of the aperture, the edge of the CCD chip and central obstruction for cameras positioned in the primary focus. Darkening of the CCD image toward the edges can be seen. Larger CCD chips suffer from vignetting more, and focal reducers also introduce vignetting. Vignetting also changes aberration properties.

Dust particles may be present on the chip and shown on the image as dark donuts. Professional CCDs are placed in vacuum chamber to avoid the dust.

Light sensitivity of pixels varies slightly. In addition, *defect pixels* in a form of hot and cold pixels can be present, as well as whole defect sectors or lines. Those always read out either very high (for hot pixels) or very low (for cold pixels) values.

When the CCD accumulates more electrons than a pixel can hold, the electrons leak out to nearby pixels and create *bleeding* streaks around bright stars. Antiblooming was developed to counter those. CCDs with antiblooming use spaces between pixels to drain the excessive electrons. Professional CCDs do not use antiblooming anymore. Devices with antiblooming have larger space between individual pixels for the drainage of the electrons and tend to have nonlinear response once the electron count gets close to full well (the amount of electrons the pixel can hold) as some of the electrons overflow to the drainage already. In addition, we have no way of knowing how much signal was lost to the antiblooming, so interposing an overexposed star with Gaussian curve[5] can yield wrong results.

[5]Common technique to compute position of the star in the image, described in Section 5.4.2

The number of photons which arrived during exposure time fluctuates, which is known as Poisson noise, *photon noise* or shot noise. Photon noise follows Poisson distribution[12].

Dark current is created on the CCD chip by randomly generated electrons, which is known as *thermal noise*. It increases with exposure time but can be reduced by cooling the CCD chip. Because of the limited exposure time it contains in itself also its own shot noise and therefore the scatter can be expressed as square root of the number of thermal electrons generated. Thermal noise of modern cameras is very low, in order of electrons RMS (root mean square)[6].

Read-out noise is the noise of the on-chip amplifier, which converts very small charge of the electrons generated on the CCD into measurable voltage. It is present in all images and does not depend on exposure time, but can be reduced by using lower sampling frequency.

Joining together all of these effects, the quality of the image can be expressed by SNR (Signal to Noise Ratio), which is the ratio of the useful obtained signal versus the total signal including all the noise, as per Definition 5.2.

2.3.5 Image pre-processing

As the asteroids pose moving targets, there may be very short window of opportunity for observations. The observations have to be done in less than ideal conditions, providing noisy images. In addition to the noises mentioned in previous paragraph, we may have to do with images containing a gradient from a nearby bright star or the Moon, dense stellar field of Milky Way, or the asteroid imaged close to the corner of the field of view. It is essential to pre-process the images before reducing them to objects, for which a handful of simple techniques are used.

[6]Also called standard deviation. See Section 3.2.1 for definition.

The first group of pre-processing methods relies on obtaining calibration images and then applying them on the data images. Only their basic explanation is stated in this part, please refer to Section 5.2 for details.

Light Frames

Multiple exposures can be stacked to reduce noise of the image. Enlarging the images before they are stacked allows for subpixel resolution of the result. Taking multiple shorter exposures of the target is meaningful to avoid quick saturation of the CCD chip and for fast moving bodies, where part of the image could be stacked with prior knowledge of the target's velocity and angle of movement. However, astrometry of such a target would be less precise, so this should be avoided unless astrometry could not be obtained otherwise. Additionally, tracking could be implemented to resolve stars from moving objects, causing the stars to appear as streaks and moving objects being a streak of different direction and length on the image. However, the same effect can be achieved by stacking and astrometry of such image would be highly inaccurate.

Bias Frames

These are images taken with closed shutter of the CCD and shortest possible exposure time and serve to remove electrons related noise from the CCD image. Bias still contains read-out noise, therefore it is better to take more bias images and use their median.

Dark Frames

These images are taken with closed shutter of the CCD and the same exposure, temperature and binning as the image we want to remove the noise from. They include bias and thermal noise of the CCD chip, but are affected by read-out noise, therefore it is better to use median of several dark frames. An observatory can create a library of dark frames, each master dark frame being an average of several dark frames of the same exposure and tempera-

ture. If the thermal noise is negligible, bias frames are used instead.

Flatfields

The previous frames were taken with CCD shutter closed. On the other hand, flatfield is taken with shutter open, at the target exposure time and against evenly lit background. It accounts for different pixel sensitivity, dust particles in the optical system and vignetting. It can not counter cosmic ray hits or ambient light on the sky like moonlight or nearby bright stars. Flatfields represent the response of the whole optical system. The pixel values of the flatfield are normalized to lay between 0 and 1 and the final image is then multiplied or divided by the flatfield image. If the optical system changes (the camera is moved, the mirror is re-coated, more dust is accumulated etc.) new flatfields have to be taken. Dark frame should be applied to the flatfields as to any other image. Observatories often create libraries of master flatfields to use later. There are many methods for obtaining a flatfield, for example to image zenith of a twilight sky, an evenly illuminated portion of a dome or cloth pinned to the dome, or to median–combine images of relatively empty star fields.

The second group of pre-processing methods employs computer algorithms to further improve the image. There are a handful free or paid programs available for enhancing images in this way, but the details of their procedures are usually not available, making them unsuitable for scientific work. Methods combining several images into one are sometimes used in astrometry, these are described in Section 5.2.6.

Filtering is a viable option, to which Section 5.3 is devoted. In the first step, the image is converted to frequencies by Fast Fourier Transform. Low pass filter can be then applied, reducing high frequency detail, which results in apparent smoothing of the image. Another filter, which is used often, is Median filter. However any filtering will alter statistical properties of the data, and reducing noise will also inevitably reduce low SNR data, hindering the astrometry itself.

Additional image processing is not used by observatories performing astrometric measurements. Military programmes of asteroid search (LINEAR for example) may use advanced image processing, however these remain unpublished. Subtractive method, which processes the images in such a way that stationary objects i.e. stars are removed from the image, could be employed, leaving only moving objects to be easily spotted by eye or automated procedure, but due to lack of stars astrometry could not be carried out on them.

2.3.6 Stellar Catalog

Stellar catalogs are used for determination of asteroid's position on the sky, therefore the inaccuracy of catalog can contribute largely to the resulting error. Quality catalog is essential for high precision astrometry. Ideal star catalog should be 'dense and deep', offering stars of varying brightness distributed across entire sky, for astrometry to be carried out regardless of asteroid's brightness and position. UCAC4 catalog is the most precise astrometric catalog to date and recommended by Minor Planet Center for observers to employ. As demonstrated in [23], it is possible to determine and account for catalog bias to increase average precision of astrometric positions, however astrometric observations sent to Minor Planet Center have to be unadjusted.

2.3.7 Image astrometry

The CCD images entering the astrometry typically have darkframe or bias and flatfield applied. Sometimes co-adding is used, but further image processing is not employed. Most of the average astrometry is done by one of the available programs capable of astrometric measurements, notably Astrometrica.at, or SAO Image. As mentioned before, the details of their methods are not readily available, thus truly dedicated observatories create their own astrometric software, typically in cooperation with their home universities. This serves two purposes: the observatory can adjust the software to their own specific needs, and the worldwide astrometric measurements avoid cumulating possible systematic errors. In either case, astrometric measurements should consistently have an error lower than 1" for observations using

the same comparison stars, and a night-to-night consistency limited only by the comparison star catalog.

Automated astrometry is so far the unreachable dream in the field. The largest asteroid surveys — Catalina Sky Survey, NEAT and LINEAR — tried to replace the observer by an automated system[35], but even having the possibilities nonmilitary projects are not entitled to, they were forced to go back to using skilled observers again. The automated astrometry in general either lacks completeness (does not recognize part of the objects present on the images) or it lacks reliability (generates a lot of false detections). The human brain is still unmatchable in its ability to recognize faint targets on images and skilled observer is the choice of the observatories.

Typical astrometric procedure is described in detail in Section 5.4. The resulting equatorial coordinates[7] of the measured object are sent to Minor Planet Center with precision 0.01 second in right ascension and 0.1 arcsecond in declination, unless higher precision is warranted. To keep the quality of the included data high, MPC compares the astrometric measurements with the expected positions and inconsistent measurements are rejected by the automatic system.

[7]Defined in Section 3.1.3.

Chapter 3

Mathematical background

Required mathematical background is given in this chapter. In no way complete, as the text would become enormously long, only basics are written about the topic in this text. Required coordinate systems are defined and their transformations stated. A few essential statistics concepts are described, including normal probability distribution, leading to derivation of linear least squares method. Fourier transform, necessary for image processing, is described, including its discrete version.

3.1 Celestial Sphere

Astronomy uses two kinds of coordinate systems for describing positions in 3D, which are Cartesian and Spherical coordinate systems. The most used form of spherical coordinate system in astronomy is the equatorial coordinate system. Gnomonic projection then allows us to pass between equatorial coordinates and 2D cartesian coordinates on the CCD image.

3.1.1 Cartesian coordinate system in 3D

The cartesian coordinate system in three-dimensional case is formed by an ordered triplet of vectors **i**, **j**, and **k**, which are pair-wise perpendicular, have a single unit of length and begin at the same point, called origin of the coordinate system. The lines, on which these vectors lay, are usually called axes x, y and z.

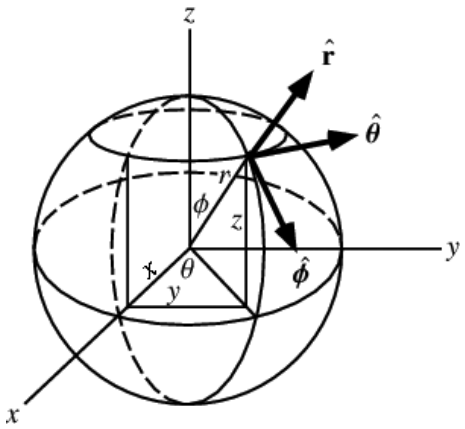


Figure 3.1:
Spherical Coordinate System.
Source: Wolfram MathWorld

The coordinates of a point R are represented by a position vector $\mathbf{r}$ going from the origin of the coordinate system to the point R. The vector can be written as linear combination of unit vectors $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$,

$$\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k},$$

where the unit vectors $\mathbf{i}^\top = (1, 0, 0)$, $\mathbf{j}^\top = (0, 1, 0)$, $\mathbf{k}^\top = (0, 0, 1)$ form a standard basis.

The coordinates of a point R are written as $R(x, y, z)$. The corresponding angles α, β, γ lay between the vector $\mathbf{r}$ and the coordinate axes. The coordinate system can have two orientations depending on the direction of axis z, either right-handed, or left-handed.

3.1.2 Spherical coordinate system

Definition 3.1. (Spherical coordinate system). Spherical coordinate system is formed by vector $\mathbf{r}$ and angles ϕ, θ, where $\mathbf{r}$ is *radial distance* (radius) from a point to a fixed origin of the coordinate system, ϕ is *polar angle* or *zenith angle* measured from the positive fixed zenith direction (z-axis) with $0 \leq \phi \leq \pi$, and θ is *azimuthal angle* of its orthogonal projection on a

reference plane xy that passes through the origin and is orthogonal to the zenith direction, measured from a fixed reference direction (x-axis) on that plane, with $0 \leq \theta < 2\pi$.

Another term we will meet is *great circle* of a sphere, also known as orthodrome or Riemannian circle, which is the intersection of the sphere and a plane that passes through the center point of the sphere, therefore it contains the diameter of the sphere.

The spherical coordinates (r, θ, ϕ), are related to the Cartesian coordinates (x, y, z) by

$$\begin{aligned} r &= \sqrt{x^2 + y^2 + z^2} \\ \theta &= \arctan \frac{y}{x} \\ \phi &= \arccos \frac{z}{r}, \end{aligned}$$

where $r \in [0, \infty)$, $\theta \in [0, 2\pi)$, $\phi \in [0, \pi]$, and the arctan must be suitably defined to output correct quadrant.

Reversely, the Cartesian coordinates (x, y, z) are related to the spherical coordinates (r, θ, ϕ) by

$$\begin{aligned} x &= r \cos\theta \sin\phi \\ y &= r \sin\theta \cos\phi \\ z &= r \cos\phi \end{aligned}$$

3.1.3 Celestial equatorial coordinate system

Is a spherical coordinate system used in astronomy to specify positions of celestial objects, and is based on the concept of the *celestial sphere*.

Definition 3.2. (Celestial Sphere.) *Celestial sphere* is an imaginary sphere of a single unit radius surrounding the planet Earth. The celestial sphere rotates about the axis passing through the North and South poles of the Earth, which intersects the celestial sphere at the *North celestial pole* and *South celestial pole*, respectively.

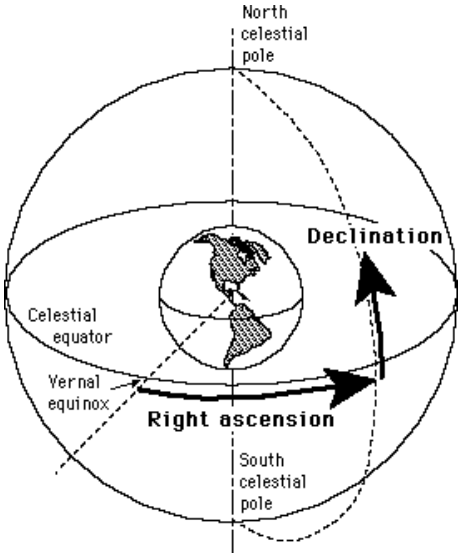


Figure 3.2:
The equatorial coordinate system.
Source: mwit.ac.th

Definition 3.3. (The equatorial coordinate system.) The equatorial coordinate system has its origin at the center of the Earth, its fundamental plane consist of the projection of Earth's equator onto the celestial sphere, forming *celestial equator*, and primary direction points towards the vernal point Υ, which is the apparent location of the Sun at the spring equinox. The system uses right-handed convention.

Its coordinates are right ascension and declination, denoted (RA, Dec) or (α, δ). *Declination* δ is measured in degrees north and south of the celestial equator, $\delta \in \langle +90, -90 \rangle$. *Right Ascension* α is measured eastward along the celestial equator from the vernal point Υ to its intersection with the object's hour circle (the great circle passing through the object and through the north celestial pole), and is measured in hours, minutes and seconds[1], where 24 hours of right ascension are equivalent to 360 degrees; $\alpha \in \langle 0, 24 \rangle$.

[1]The rationale for this unit is based on 24 hours (sidereal time) required for the stellar object to pass through the observer's local meridian — the great circle passing through the celestial poles, the zenith and the nadir of the observer's location.

Since the equatorial system is aligned with Earth's equator and poles, declination δ is the equivalent of latitude, and right ascension α is analogous to longitude. The right-hand convention means that the coordinates increase northward from and eastward around the fundamental plane. The equatorial coordinate system is based on the celestial equator and the vernal point Υ, therefore it does not rotate with the Earth and remains relatively fixed[2] against the background stars. Changes in the latitude and longitude of the observer do not affect equatorial coordinates of the object, and neither does the annual motion of the Earth.

3.1.4 Gnomonic projection

Gnomonic projection allows us to compute spherical coordinates from CCD image by describing how these spherical coordinates project onto the tangent plane of the CCD chip.

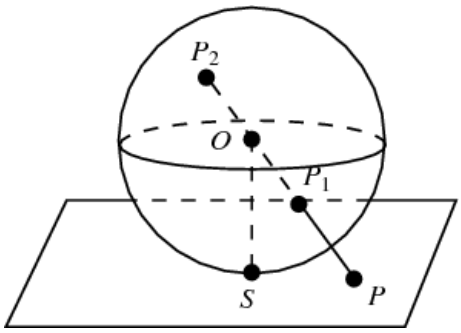


Figure 3.3:
Gnomonic projection.
Source: Wolfram MathWorld

Definition 3.4. (Gnomonic Projection). Gnomonic projection is a nonconformal projection obtained by projecting point P_2, (or P_1) on the surface of sphere from a sphere's center O to point P in a plane that is tangent to a point S [9].

[2]The precession and nutation of Earth's axis cause a slow drift of the vernal point Υ. In order to fix the primary direction, it is necessary to specify the date, known as Epoch, for which the coordinate system is valid.

In the Figure 3.3, S is the south pole, but it can in general be any point on the sphere. Since this projection obviously maps antipodal points P_1 and P_2 to the same point P in the plane, it can only be used to project one hemisphere at a time. In a gnomonic projection, great circles (sections of the sphere which contain the diameter of the sphere) are mapped to straight lines. The gnomonic projection represents the image formed by a spherical lens, and is sometimes known as the *rectilinear projection*.

The gnomonic projection is well known in geodesy [10] , using (λ, ϕ), where λ is longitude and ϕ is colatitude, computed from latitude by $\phi = 90° - \delta$. Let's have celestial equatorial coordinates (α, ϕ), where α is right ascension and $\phi = 90° - \delta$ is computed from declination δ, and the center of the image $S = (\alpha, \phi) = (0, 0)$ is laying in known equatorial coordinates (α_0, ϕ_1). Then the transformation equations of the plane tangent at the point S are given by [24]

$$x = \frac{\cos\phi \sin(\alpha - \alpha_0)}{\cos c}$$
$$y = \frac{\cos\phi_1 \sin\phi - \sin\phi_1 \cos\phi \cos(\alpha - \alpha_0)}{\cos c},$$

where c is the angular distance of the point (x, y) from the center of the projection, given by $\cos c = \sin\phi_1 \sin\phi + \cos\phi_1 \cos\phi \cos(\alpha - \alpha_0)$. The inverse transformation equations are,

$$\phi = \arcsin(\cos c \sin\phi_1 + y \sin c \cos c \cos\phi_1)$$
$$\alpha = \alpha_0 + \arctan \frac{x}{\cos\phi_1 - y \sin\phi_1}.$$

3.2 Theory of errors

Every observation includes an error, a difference between the measured and true value. The error is not known (otherwise we could just add it to the measurement to obtain the true value), but using an associated statistics, it can be evaluated and quantified into *uncertainty*, which characterizes the dispersion of measured values.

Let $\hat{x}$ be the estimated value of a parameter x, and σ_x be the quantity such, that for the normal distribution, the true value x_0 has a probability of 0.683 or being in the interval $\hat{x} - \sigma_x < x_0 < \hat{x} + \sigma_x$. Then we write

$$x_0 = \hat{x} \pm \sigma_x.$$

The quantity σ_x is called the *standard deviation*, also standard uncertainty or root mean square (RMS) and is defined in following chapter.

3.2.1 Random errors

Generally, the best available estimate of the expected value of x that varies randomly, and for which n independent values x_k were measured, is the arithmetic mean or average, μ,

$$x = \mu = \frac{1}{n} \sum_{k=1}^{n} x_k.$$

The residual of each measurement is

$$r_k = x_k - \mu.$$

We introduce the *variance* of the measurements,

$$\sigma^2(x_k) = \frac{1}{n-1} \sum_{k=1}^{n} r_k^2, \tag{3.1}$$

and the *standard deviation*, σ_k, which is the square root of the variance.

3.2.2 Combination of random quantities

Let's have a quantity to be determined, z, which depends upon N different random parameters x_i, and a model that links these parameters

$$z = f(x_1, x_2 \dots x_N).$$

Let us assume that each of the parameters x_i is random, with an expected value μ_i and standard deviation σ_i. For small deviations from these expected values, the first-order Taylor expansion gives

$$z - \mu_z = \sum_{i=1}^{N} \frac{\partial f}{\partial x_i} (x_i - \mu_i)$$

The expected value of z then is

$$\mu_z = f(\mu_1, \mu_2 \dots \mu_N),$$

and

$$(z-\mu_z)^2 = \sum_{i=1}^{N} \left(\frac{\partial f}{\partial x_i}\right)^2 (x_i-\mu_i)^2 + 2\sum_{i=1}^{N}\sum_{j=i+1}^{N} \frac{\partial f}{\partial x_i}\frac{\partial f}{\partial x_j}(x_i-\mu_i)(x_j-\mu_j). \quad (3.2)$$

The expected value of $(z-\mu_z)^2$ is the variance σ^2 of z, and for each of the parameters, the expected values of $(x_i-\mu_i)^2$ are σ_i^2.

Now in the Equation 3.2, the expected value of $(x_i-\mu_i)(x_j-\mu_j)$ is written as $\sigma_i\sigma_j\rho_{ij}$, where ρ_{ij} is by definition the *correlation coefficient* of x_i and x_j. The product $\sigma_{ij} = \sigma_i\sigma_j\rho_{ij}$ is called the covariance of x_i and x_j. If x_i and x_j are independent variables, meaning that their variations do not depend one on another or on a common cause, the correlation coefficient is null. With these notations, Equation 3.2 becomes

$$\sigma_z^2 = \sum_{i=1}^{N} \left(\frac{\partial f}{\partial x_i}\right)^2 \sigma_{x_i}^2 + 2\sum_{i=1}^{N}\sum_{j=i+1}^{N} \frac{\partial f}{\partial x_i}\frac{\partial f}{\partial x_j}\sigma_i\sigma_j\rho_{ij}.$$

We can finally introduce *variance-covariance matrix*, the symmetric matrix $|\sigma_{ij}|$, whose elements are $\sigma_{ii} = \sigma_i^2$ in row i, column i, and σ_{ij} both in row i, column j and row j, column i.

3.2.3 Probability density function

A probability density function (PDF) is a function of a continuous random variable whose integral across an interval gives the probability that the value of the variable lies within the same interval. A PDF is nonnegative everywhere and its integral over the entire space is one.

Central limit theorem states, that the sum of large number of mutually independent random variables tends to a probability density function called Gaussian, or *normal*, density function, given by

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\, \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}},$$

where μ is the mean of the probability distribution and σ is the standard deviation. The normal PDF is represented on Figure 3.4.

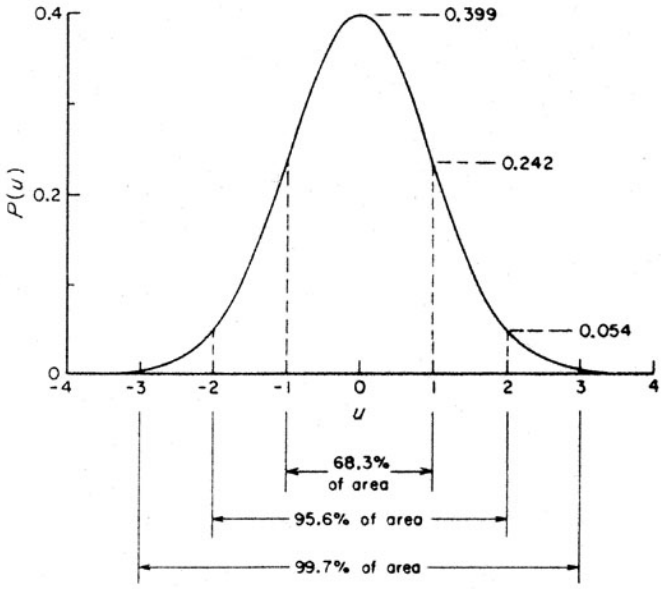


Figure 3.4: Normal probability density function.
Source: ned.ipac.caltech.edu

Level of confidence C is then the probability of a value to be within an interval $\mu - k_p\sigma < x < \mu + k_p\sigma$, where k_p is the *coverage factor*,

$$C = \int_{\mu - k_p\sigma}^{\mu + k_p\sigma} f(x)\mathrm{d}x.$$

An important property of normal PDF is that any random variable, formed by taking a linear combination of independent normally distributed random variables, is itself also normally distributed.

A sample of N observations, that obeys the normal PDF, with variance σ^2, is described by *Chi-square* PDF. The average of the sample is

$$\overline{x} = \sum_{i=1}^{N} \frac{x_i}{N},$$

and the variance of the sample

$$s^2 = \frac{1}{N-1}\left(\sum_{i=1}^{N} x_i^2 - N\overline{x}^2\right).$$

We then consider the random variable

$$y = \frac{\nu s^2}{\sigma^2},$$

where $\nu = N - 1$ is the *degree of freedom*. This random variable has Chi-square probability density function, $\chi^2_\nu(y)$,

$$\chi^2_\nu(y) = K y^{(\nu-2)/2} \mathrm{e}^{y/2},$$

where the constant K is chosen in such a way, that $\int_{-\infty}^{+\infty} \chi^2(\nu) = 1$. We have

$$K = \frac{1}{2^{\nu/2}\Gamma(\nu/2)},$$

in which Γ is the *gamma function*

$$\Gamma(x) = \int_0^\infty \mathrm{e}^{-t} t^{x-1} \mathrm{d}t.$$

The Chi-square PDF is usually tabulated as a function of degree of freedom ν and is used to find the probability that s^2 exceeds some given value, which shows how well is the sample representative of the distribution.

3.3 Least squares method

Although this method is not the only one used to estimate an unknown parameters of a measurement model, its special properties justify its wide use. Description of the method from [18] was adapted for this text, as it relates well to our use.

Let $\mathbf{X}(x_1, x_2 \ldots x_n)$ be the vector of unknowns, $\mathbf{A}(a_1, a_2 \ldots a_m)$ the vector of measurements, and $\mathbf{P}(p_1, p_2 \ldots p_q)$ the vector of the parameters entering the model which relates $\mathbf{A}$ to $\mathbf{X}$. We have

$$\mathbf{A} + \mathbf{E} = \mathbf{F}(\mathbf{X}, \mathbf{P}), \tag{3.3}$$

with $n + q$ unknowns and $\mathbf{E}(\varepsilon_1, \varepsilon_2 \ldots \varepsilon_m)$ is an unknown error vector of the measurements A.

3.3.1 Linearization of the equations

The method of least squares is designed to find the most probable estimation $\hat{a}_i$ of the components a_i of $\mathbf{A}$. The necessary assumptions are, that the unknowns

- have normal probability density function
- are sufficiently small, so that we can neglect their squares

The first assumption is met if all the parameters, which may cause a bias, are taken into account by Equation 3.3, so that their uncertainties may be considered as random. The second assumption implies that, unless the unknowns are indeed small, we know an approximate value $\mathbf{X}_0(x_1^0, x_2^0 \dots x_n^0)$ and $\mathbf{P}_0(p_1^0, p_2^0 \dots p_q^0)$, so that we can develop Equation 3.3 as follows:

$$\mathbf{A} = \mathbf{F}(\mathbf{X}_0, \mathbf{P}_0) + \sum_{i=1}^{n} \left(\frac{\partial \mathbf{F}(\mathbf{X}_0, \mathbf{P}_0)}{\partial x_i} \right) \delta x_i \tag{3.4}$$

$$+ \sum_{j=1}^{q} \left(\frac{\partial \mathbf{F}(\mathbf{X}_0, \mathbf{P}_0)}{\partial p_j} \right) \delta p_j - \mathbf{E}, \tag{3.5}$$

where ∂x_i and ∂p_j are the new unknowns and, following our assumption, the second order of development is negligible and can be added to $\mathbf{E}$ without changing its randomness. So we shall include in $\mathbf{X}$ all unknowns, whether they are of astronomical interest or are additional parameters. Writing Equation 3.4 for each component of $\mathbf{A}$, we get a set of m equations

$$a_j - F_j(x_1 \dots x_k) = \sum_{i=1}^{k} \left(\frac{\partial F_j(x_1 \dots x_k)}{\partial x_i} \right) \delta x_i - \varepsilon_j, \tag{3.6}$$

where $1 \leq j \leq m$, and $k = n + q$ is the number of unknown quantities, while ε_j describes the statistical properties of uncertainties of a_j. This set of linear equations in δx_i can not be solved if $k < j$. Therefore it is necessary to increase the number of equations by adding more observations obeying the same model. For example during plate reduction, described in the part of Section 5.4.3 where plate parameters are computed, we measure a sufficient amount of reference stars, each providing two equations. The least squares method allows any number m of Equations 3.6 as long as m is strictly larger

than the number k of the unknowns. A general experience is that is it desirable that m be of the order of two to three times larger than k. The quantity $m - k = v$ is the degree of freedom as introduced in Section 3.2.3.

A set of Equations 3.6 is called the *equations of condition*, or *design equations*, and the coefficients of the second members form the design matrix $\mathcal{D}$. Equations 3.6 can be written as

$$\mathbf{C} = \mathcal{D}\mathbf{Y} + \mathbf{E}, \tag{3.7}$$

where $\mathbf{Y}$ is the vector of the new unknowns $(\delta x_1 \dots \delta x_k)$, and $\mathbf{C}$ is the vector of the left-hand members of Equations 3.6

$$\mathbf{C} = (a_j - F_j(x_1 \dots x_k)), \quad j = 1 \dots m.$$

3.3.2 Principle of the least squares method

If the conditions stated in the beginning of Section 3.3.1 are met, the expected value of $\mathbf{E}$ is zero and its variance is

$$\frac{1}{m-1} \sum_{j=1}^{m} \varepsilon_j^2.$$

In reality, this of course is not the case. The objective of the least squares method is to determine the vector $\mathbf{Y}$ that minimizes the sum of the squares of the components of $\mathbf{E}$,

$$\sum_{j=1}^{m} \varepsilon_j^2 = \mathbf{E}^\top \cdot \mathbf{E} = (\mathbf{C} - \mathcal{D}\mathbf{Y})^\top \cdot (\mathbf{C} - \mathcal{D}\mathbf{Y}), \tag{3.8}$$

where the superscript $\top$ denotes matrix transposition. The quantity in Equation 3.8 is minimum if its derivative with respect to $\mathbf{Y}$ is zero. The right-hand member of Equation 3.8 can also be written as

$$\mathbf{C}^\top \cdot \mathbf{C} - (\mathcal{D}\mathbf{Y})^\top \cdot \mathbf{C} - \mathbf{C} \cdot \mathcal{D}\mathbf{Y} + (\mathcal{D}\mathbf{Y})^\top \cdot \mathcal{D}\mathbf{Y}.$$

Its derivative with respect to $\mathbf{Y}$ is

$$-\mathcal{D}\mathbf{C} + \mathcal{D}^\top\mathcal{D}\mathbf{X} + (\mathcal{D}\mathbf{Y})^\top\mathcal{D} - \mathbf{C}^\top\mathcal{D} = 0,$$

which is equivalent to

$$\mathcal{D}^\top \mathcal{D}\mathbf{Y} = \mathcal{D}^\top \mathbf{C}. \tag{3.9}$$

The product of the transposed design matrix by itself is a square matrix of dimension k. Hence we can solve Equation 3.9 and obtain the solution

$$\hat{\mathbf{Y}} = (\mathcal{D}^\top \mathcal{D})^{-1} \mathcal{D}^\top \mathbf{C}, \tag{3.10}$$

which is the estimate of $\mathbf{Y}$ by the least squares procedure.

3.3.3 Weighted least squares solution

An important result concerning this solution is the Gauss–Markov theorem, which states that the least squares estimate $\hat{\mathbf{Y}}$ from Equation 3.9 is better estimator than a solution given by any other possible linear estimator, provided that all of the components of $\mathbf{E}$ have the same variance and are uncorrelated. We shall assume the estimate of the variance of an observation ε_j was done.

The absence of correlation usually implies that the unknowns are well chosen, in a sense that they correspond to the minimum needed to describe the model. For example, a rotation around an axis should be represented by an angle θ, and not by $\sin\theta$ and $\cos\theta$, which would be correlated.

A simple way to equalize the variances in Equation 3.3 is to multiply each equation by the inverse of ε_j. Let $\mathcal{G}$ be a square matrix of order m, whose elements are all zeroes except for the main diagonal components, which are $g_j = \varepsilon_j^{-1}$. This operation is performed on Equation 3.7 by multiplying both sides of the equation by $\mathcal{G}$,

$$\mathcal{G}\mathcal{D}\mathbf{Y} + \mathcal{G}\mathbf{E} = \mathcal{G}\mathbf{C},$$

so that the right-hand member of Equation 3.8 becomes

$$(\mathcal{G}\mathbf{C} - \mathcal{G}\mathcal{D}\mathbf{Y})^\top \cdot (\mathcal{G}\mathbf{C} - \mathcal{G}\mathcal{D}\mathbf{Y}).$$

Let $\mathcal{W}$ be the square matrix of order k

$$\mathcal{W} = \mathcal{G}^\top \mathcal{G}.$$

After some calculations, Equation 3.10 becomes

$$\hat{\mathbf{Y}} = (\mathcal{D}^\top \mathcal{W} \mathcal{D})^{-1} \mathcal{D}^\top \mathcal{W} \mathbf{C}. \tag{3.11}$$

The matrix $\mathcal{W}$ is the weight matrix and all its terms are zeroes, except on the main diagonal, where they are $w_{jj} = \varepsilon_j^{-2}$. The procedure described by Equation 3.11 is *weighted least squares method* to which the Gauss–Markov theorem applies.

The weight matrix can also be generalized to include cases when correlation between the unknowns is present. Then the non-diagonal terms are not all zeroes.

Remark 3.5. The definition of $\mathcal{G}$, as adopted here, implies that the variances, and hence the standard deviations, are unity. We could use $g_{ii} = s\varepsilon_j^{-1}$, where s is standard deviation, and verify whether the assumptions concerning the variances are significant.

Let's compute the variance of the residuals by applying Equation 3.8 to the estimated unknowns, and extending Equation 3.1 to $m-k$ degrees of freedom. We get

$$s_0^2 = \frac{1}{m-k} (\mathbf{C} - \mathcal{D}\hat{\mathbf{Y}})^\top \mathcal{W} (\mathbf{C} - \mathcal{D}\mathbf{Y}).$$

Or, if we call $\mathbf{R}(r_i)$ the vector of residuals,

$$\begin{aligned} \mathbf{R} &= \mathbf{C} - \mathcal{D}\hat{\mathbf{Y}}, \\ s_0^2 &= \frac{1}{m-k} \mathbf{R}^\top \mathcal{W} \mathbf{R}. \end{aligned} \tag{3.12}$$

It can be shown that s_0^2 is an unbiased estimate of the *unit weight variance*.

3.3.4 Variance–covariance matrix of the estimation

An important result is that the variance–covariance matrix for the estimation $\hat{\mathbf{Y}}$ is

$$\mathcal{V} = (\mathcal{D}^\top \mathcal{W} \mathcal{D})^{-1},$$

proof of which is above the scope of this work and can be found in literature on statistics. Then Equation 3.11 can be rewritten as

$$\hat{\mathbf{Y}} = \mathcal{V}\mathcal{D}^{\top}\mathcal{W}\mathbf{C}.$$

The diagonal terms of matrix $\mathcal{V}$, v_{ii} are the *formal* variances of $\hat{y}_i$. They are computed as if the model described by Equation 3.6 is exact, which means the true value is actually underestimated. To get more realistic estimations, we should multiply them by the unit weight variance s_0^2. The best estimate of standard deviation of $\hat{y}_i$ then is

$$\hat{\sigma}_i = s_0\sqrt{v_{ii}}.$$

The best estimates of correlation coefficients are not affected by s_0, and so we have

$$\hat{\rho}_{ij} = \frac{\hat{\sigma}_{ij}}{\hat{\sigma}_i\hat{\sigma}_j} = \frac{v_{ij}}{\sqrt{v_{ii}}\sqrt{v_{jj}}}.$$

3.3.5 Chi-square test

It is generally good idea to check whether the model used is indeed representative of the reality described by the observations $\mathbf{C}$, which can be simply achieved by Chi-square test. Let us introduce the random variable u

$$u = \sum_{i=1}^{n}\left(\frac{r_i}{s_0}\right)^2,$$

and apply the probability distribution of u, as given in Section 3.2.3,

$$\chi^2(\nu, u) = \left(2^{\nu/2}\Gamma\left(\frac{\nu}{2}\right)\right)^{-1} u^{(\nu-2)/2}\mathrm{e}^{-u/2},$$

where ν is the degree of freedom. When ν increases, the Chi-square PDF tends to normal distribution, whose mean is ν and variance 2ν. The functions (or their integrals) giving the probability of having $u < u_0$ are tabulated in most books on probability or in statistical software.

3.3.6 Goodness of fit

The computation of χ^2, corresponding to the residuals of least squares solution, and the determination of the probability that it exceeds the computed

value, gives an indication of the fit to the adopted model.

If the value obtained is such that the probability of exceeding it is small, some systematic effects may remain in the residuals. That may be caused by small number of observations with large residuals increasing χ^2. For example during astrometric reduction of the CCD image to stars, a star could have been mistaken for another, or its position extremely inaccurately measured (perhaps a satellite dash crossing the star confusing the software). These cases should be investigated by considering the distribution of residuals, which should obey the normal law. The small number of the observations with large residuals (outliers) can then be omitted and the least squares solution recomputed. If the new solution is still insufficient, some systematic cause remains and the linearized model used to describe the observations is not adequate.

The drawback of Chi-square test is that the values to compare χ^2 to χ_0^2 are dependent on the degree of freedom ν. For $\nu > 20$ or 30, we can employ another test independent of ν, derived from $F2$ statistics, which is a transformation of χ^2 statistics of the least square fit,

$$F2 = \left(\frac{9\nu}{2}\right)^{1/2}\left(\left(\frac{\chi^2}{\nu}\right)^{1/3} + \frac{2}{9\nu} - 1\right).$$

If χ^2 follows the Chi-square distribution with a degree of freedom ν, then $F2$ is approximately normal with a mean value zero and with a unit standard deviation. The value of $F2$ is the *goodness of fit*. A value larger than 3 is a clear indication of modelling error (probability < 0.0027), especially if the outliers were already removed.

3.4 Fourier Transform

The Fourier transform is an integral transform, which transforms functions of one or more variables into another function of the same number of variables. Since CCD image is a function of two spatial variables and Fourier transform of a function is in general case a function with complex image, we

will deal in Section 3.4.1 only with functions $f : \mathbb{R}^2 \to \mathbb{C}$. The following section deals with the Fourier transform of functions in $\mathbb{R} \times \mathbb{R}$ discrete domain. In general, CCD images are rectangles, however the images from KLENOT telescope are solely square, which allows us to deal here with square images only. Although the discrete nature of CCD images calls for using discrete Fourier transform, some derivations of image processing methods are better done with the Fourier transform of functions with the domain $\mathbb{R}^2$, since operations like rotation and scaling are easily modelled on them. The connection between the Fourier transform of functions defined on $\mathbb{R}^2$ and the discrete Fourier transform is shown at the end of this chapter. Many of the definitions and theorems are adapted from [31].

3.4.1 Fourier Transform of function $f : \mathbb{R}^2 \to \mathbb{C}$

The standard definition of the Fourier transform of a function of two real variables is as follows.

Definition 3.6. (Fourier Transform). Let $f(x, y) : \mathbb{R}^2 \to \mathbb{C}$ be a function such that

$$\iint_{\mathbb{R}^2} \mid f(x, y) \mid \, \mathrm{d}x \, \mathrm{d}y$$

exists and is finite. The *Fourier transform* of f is function $\mathcal{F}\{f\} = F(\xi, \eta)$ defined as

$$F(\xi, \eta) = \iint_{\mathbb{R}^2} f(x, y) \mathrm{e}^{-i(x\xi + y\eta)} \, \mathrm{d}x \, \mathrm{d}y.$$

Function F if also called *Fourier spectrum* of function f.

Other definitions of the Fourier transform can be found in literature. They differ in multiplicative constant before the integral or in the exponent, and in the meaning of the integral over $\mathbb{R}^2$ (for our definition treated in [31]). Different theorems can be proven with different definitions, therefore one must be cautious when taking information about Fourier transform from individual sources.

The domain of $f(x, y)$ is commonly called *spatial domain* or *time domain*,

depending on the type of data recorded, while the domain of $F(\xi, \eta)$ is called *frequency domain.*

Definition 3.7. (Inverse Fourier Transform). Let $F(\xi, \eta) : \mathbb{R}^2 \to \mathbb{C}$ be a function such that

$$\iint_{\mathbb{R}^2} \mid F(\xi, \eta) \mid \mathrm{d}\xi \, \mathrm{d}\eta$$

exists and is finite. The *Inverse Fourier Transform* of function F is function $\mathcal{F}^{-1}\{F\}(x, y) = f(x, y) : \mathbb{R}^2 \to \mathbb{C}$ defined as

$$f(x, y) = \frac{1}{4\pi^2} \iint_{\mathbb{R}^2} F(\xi, \eta) \mathrm{e}^{\mathrm{i}(x\xi + y\eta)} \, \mathrm{d}\xi \, \mathrm{d}\eta.$$

It is not generally true that an inverse Fourier transform of a Fourier transform of a function is the function itself, $\mathcal{F}^{-1}\{\mathcal{F}\{f\}\} \neq f$. However, the assumption would be correct under certain conditions, as proven by Fourier Inversion Theorem[13][14].

Theorem 3.8. (Fourier Inversion Theorem). Let $f(x, y) : \mathbb{R}^2 \to \mathbb{C}$ be a continuous function such that

$$\iint_{\mathbb{R}^2} \mid f(x, y) \mid \mathrm{d}x \, \mathrm{d}y$$

exists and is finite. Let $F(\xi, \eta)$ be the Fourier transform of f. Then for every $(\xi, \eta) \in \mathbb{R}^2$

$$f(x, y) = \lim_{\epsilon \to 0} \frac{1}{4\pi^2} \iint_{\mathbb{R}^2} F(\xi, \eta) \mathrm{e}^{\mathrm{i}(x\xi + y\eta)} \mathrm{e}^{-\epsilon^2 \frac{\xi^2 + \eta^2}{2}} \, \mathrm{d}\xi \, \mathrm{d}\eta.$$

If also the integral

$$\iint_{\mathbb{R}^2} \mid F(\xi, \eta) \mid \mathrm{d}\xi \, \mathrm{d}\eta$$

exists and is finite, then

$$\mathcal{F}^{-1}\{\mathcal{F}\{f(x, y)\}\} = \frac{1}{4\pi^2} \iint_{\mathbb{R}^2} F(\xi, \eta) \mathrm{e}^{\mathrm{i}(x\xi + y\eta)} \, \mathrm{d}\xi \, \mathrm{d}\eta = f(x, y).$$

Digital image is a matrix of finite values, which can be approximated by fitting smooth functions defined on $\mathbb{R}^2$ through the pixel values and therefore will have finite integrals and fulfill the conditions for *Fourier Inversion Theorem*. That is very important step as in image processing, *Fourier Inversion Theorem* plays critical role.

The basic strategy in many applications of image processing is to apply the Fourier transform, perform operations on the data set, and then apply the inverse Fourier Transform. This seemingly more complicated way is in fact advantageous due to some operations being simplified if done in *frequency domain*, as opposed to doing them in the *spatial domain*. The notions of *amplitude spectrum* and *phase spectrum* are commonly referred to in image processing. They are defined as follows.

Definition 3.9. (Amplitude Spectrum, Phase Spectrum). Let function $f(x, y) \in \mathcal{L}(\mathbb{R}^2)$ have Fourier spectra $F(\xi, \eta)$. The *amplitude spectrum* [31] of function f is function $A(\xi, \eta) : \mathbb{R}^2 \to \mathbb{R}_0^+$ defined as

$$A(\xi, \eta) = |\mathcal{F}\{f(x, y)\}| = |F(\xi, \eta)|.$$

The *phase spectrum* [31] of function f is function $\Phi(\xi, \eta) : \mathbb{R}^2 \to \langle 0, 2\pi)$ defined as

$$\Re F(\xi, \eta) = A(\xi, \eta) \cos \Phi(\xi, \eta),$$

$$\Im F(\xi, \eta) = A(\xi, \eta) \sin \Phi(\xi, \eta).$$

If $A(\xi, \eta*) = 0$ for some (ξ, η), we define that $\Phi(\xi, \eta) = 0$.
One of the operations frequently used in this regard is convolution, which becomes multiplication in *frequency domain*, as per following *Convolution Theorem*.

Definition 3.10. (Convolution). Let functions $f_1(x, y)$, $f_2(x, y)$ fulfill the assumptions of Definition 3.6. The *convolution* [16] $f_1 * f_2$ of functions f_1, f_2 is a function

$$f(x, y) = \iint_{\mathbb{R}^2} f_1(s, t) f_2(x - s, y - t) \mathrm{d}s \, \mathrm{d}t.$$

Theorem 3.11. (Convolution Theorem). Let functions $f_1(x,y), f_2(x,y)$ fulfill the assumptions of Definition 3.6. Let $F_1(\xi,\eta)$, $F_2(\xi,\eta)$ be their Fourier spectra, and let $\mathcal{F}$ denote the Fourier transform operator. Then

$$\mathcal{F}\{f_1(x,y) * f_2(x,y)\} = F_1(\xi,\eta) \cdot F_2(\xi,\eta)$$
$$\mathcal{F}\{f_1(x,y) \cdot f_2(x,y)\} = \frac{1}{4\pi^2} F_1(\xi,\eta) * F_2(\xi,\eta).$$

For proofs see [31].

3.4.2 Discrete Fourier Transform

In image processing, we deal with two-dimensional images composed of pixels. Consequently, the Fourier transform changes into discrete Fourier transform as the integrals change to sums to handle the discrete data. The images this work deals with are square datasets, so we will define the discrete Fourier transform for function $f(x,y)$ defined on $\{0,1,\ldots,N-1\}\times\{0,1,\ldots,N-1\}$, where N is natural number.

Definition 3.12. (Discrete Fourier Transform). Let $f(x,y)$ be a function $\{0,1,\ldots,N-1\} \times \{0,1,\ldots,N-1\} = \{0,1,\ldots,N-1\}^2 \to \mathbb{C}, N \in \mathbb{N}$. The *discrete Fourier transform* of function $f(x,y)$ is function $\mathcal{D}\{f\}(\xi,\eta) = F(\xi,\eta) : \{0,1,\ldots,N-1\}^2 \to \mathbb{C}$ defined as

$$F(\xi,\eta) = \sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y)\mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}(x\xi+y\eta)}. \tag{3.13}$$

Function F is also called the *Fourier spectrum* of function f.

Definition 3.13. (Inverse Discrete Fourier Transform). Let $f(x,y)$ be a function $\{0,1,\ldots,N-1\}^2 \to \mathbb{C}, N \in \mathbb{N}$ and let $F(\xi,\eta)$ be its discrete Fourier transform. The *inverse discrete Fourier transform* of function $F(\xi,\eta)$ is function $\mathcal{D}^{-1}\{F\}(x,y) : \{0,1,\ldots,N-1\}^2 \to \mathbb{C}$ defined as

$$\mathcal{D}^{-1}\{F\}(x,y) = \frac{1}{N^2}\sum_{\xi=0}^{N-1}\sum_{\eta=0}^{N-1} F(\xi,\eta)\mathrm{e}^{\frac{2\pi\mathrm{i}}{N}(x\xi+y\eta)}.$$

Digital images have finite size, therefore N is finite and the condition of *Fourier Inversion Theorem* is fulfilled. For proof of the following theorem see [31].

Theorem 3.14. (Fourier Inversion Theorem). Let $f(x, y)$ be a function of $\{0, 1, \ldots, N-1\}^2 \to \mathbb{C}, N \in \mathbb{N}$ and let $F(\xi, \eta)$ be its discrete Fourier transform. Then the inverse discrete Fourier transform of function $F(\xi, \eta)$ is function $f(x, y)$,

$$\mathcal{D}^{-1}\{\mathcal{D}\{f(x, y)\}\} = f(x, y).$$

When talking about digital image processing, the notions of *high frequency* and *low frequency* are commonly used, referring to the opposite parts of the domain of the Fourier spectrum of a function. Low frequencies lie near the origin of the Fourier spectrum and represent the underlying trend in the examined data, while high frequencies lie far from the origin and represent the minute differences on small scale in the data, which is usually noise where sensitive scientific measurements are taken. In the case of Discrete Fourier Transform in $\mathbb{R}^2$, low frequencies correspond to the regions close to the corners $([0, 0], [N-1, 0], [0, N-1], [N-1, N-1])$ of the spectrum, while the high frequencies are in the center of the domain of the spectrum.

The definition of the amplitude and the phase spectrum of functions defined on $\{0, 1, \ldots, N-1\}^2$ is analogous to the definition for functions defined on $\mathbb{R}^2$.

Definition 3.15. (Amplitude Spectrum, Phase Spectrum). Let function $f(x, y)$ be a function $\{0, 1, \ldots, N-1\}^2 \to \mathbb{C}, N$ with Fourier spectrum $F(\xi, \eta)$. The *amplitude spectrum* [16] of function f is function $A(\xi, \eta) : \{0, 1, \ldots, N-1\}^2 \to \mathbb{R}$ defined as

$$A(\xi, \eta) = |\mathcal{D}\{f(x, y)\}| = |F(\xi, \eta)|.$$

The *phase spectrum* [16] of function f is function $\Phi(\xi, \eta) : \{0, 1, \ldots, N-1\}^2 \to \langle 0, 2\pi)$ defined as

$$\Re F(\xi, \eta) = A(\xi, \eta) \cos \Phi(\xi, \eta),$$
$$\Im F(\xi, \eta) = A(\xi, \eta) \sin \Phi(\xi, \eta).$$

If $A(\xi, \eta) = 0$ for some (ξ, η), we define that $\Phi(\xi, \eta) = 0$.

In the previous section, we discussed the connection of convolution to the Fourier transform. Here we do the same with the discrete Fourier transform

and the discrete periodic convolution. The periodic extension of a function defined on $\{0, 1, \ldots, N-1\}^2$ is necessary for this to be defined. Since the Equation 3.13 can be evaluated for any $(\xi, \eta) \in \mathbb{Z}^2$, this allows us to define the periodic extension as follows:

Definition 3.16. (Periodic extension of function and its Fourier spectrum). Let $f(x, y)$ be a function $\{0, 1, \ldots, N-1\}^2 \to \mathbb{C}, N \in \mathbb{N}$ and let $F(\xi, \eta)$ be its Fourier spectrum. The *periodic extension of Fourier spectrum* F [31] is function $\widetilde{F}(\xi, \eta) : \mathbb{Z}^2 \to \mathbb{C}$ defined as

$$\widetilde{F}(\xi, \eta) = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \mathrm{e}^{-\frac{2\pi \mathrm{i}}{N}(x\xi + y\eta)}.$$

The *periodic extension of function* f is function $\widetilde{f}(x, y) : \mathbb{Z}^2 \to \mathbb{C}$ defined as

$$\widetilde{f}(x, y) = \frac{1}{N^2} \sum_{\xi=0}^{N-1} \sum_{\eta=0}^{N-1} F(\xi, \eta) \mathrm{e}^{\frac{2\pi \mathrm{i}}{N}(x\xi + y\eta)}.$$

The definition has following consequences (for details see [31]):

Corollary 3.17. Let $f(x, y)$ be a function $\{0, 1, \ldots, N-1\}^2 \to \mathbb{C}, N \in \mathbb{N}$. Then for every $(x, y), (\xi, \eta) \in \{0, 1, \ldots, N-1\}^2$ and $k, l \in \mathbb{Z}$ it holds:

$$f(x, y) = \widetilde{f}(x + kN, y + lN),$$

$$F(\xi, \eta) = \widetilde{F}(\xi + kN, \eta + lN).$$

In particular,

$$\begin{aligned} \widetilde{f}(x, y) &= f(x, y), \qquad & \widetilde{f}(-x, -y) &= f(N - x, N - y), \\ \widetilde{F}(\xi, \eta) &= F(\xi, \eta), \qquad & \widetilde{F}(-\xi, -\eta) &= F(N - \xi, N - \eta). \end{aligned}$$

Definition 3.18. (Discrete Fourier transform of function's periodic extension.) Let $f(x, y)$ be a function $\{0, 1, \ldots, N-1\}^2 \to \mathbb{C}, N \in \mathbb{N}$. The discrete Fourier transform of the periodic extension of function f, $\widetilde{f}(x, y) : \mathbb{Z}^2 \to \mathbb{C}$ is function $\mathcal{D}\{\widetilde{f}\}(\xi, \eta) = F(\xi, \eta) : \{0, 1, \ldots, N-1\}^2 \to \mathbb{C}$ defined as

$$F(\xi, \eta) = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} \widetilde{f}(x, y) \mathrm{e}^{-\frac{2\pi \mathrm{i}}{N}(x\xi + y\eta)}.$$

Definition 3.19. (Inverse discrete Fourier transform of function's periodic extension.) Let $f(x,y)$ be a function $\{0,1,\ldots,N-1\}^2 \to \mathbb{C}, N \in \mathbb{N}$ and let $F(\xi,\eta)$ be its discrete Fourier transform with periodic extension $\widetilde{F}(\xi,\eta) : \mathbb{Z}^2 \to \mathbb{C}$. The inverse discrete Fourier transform of function $\widetilde{F}(\xi,\eta)$ is function $\mathcal{D}^{-1}\{\widetilde{F}\}(x,y) : \{0,1,\ldots,N-1\}^2 \to \mathbb{C}$ defined as

$$\mathcal{D}^{-1}\big\{\widetilde{F}\big\}(x,y) = \frac{1}{N^2}\sum_{\xi=0}^{N-1}\sum_{\eta=0}^{N-1}\widetilde{F}(\xi,\eta)\mathrm{e}^{\frac{2\pi\mathrm{i}}{N}(x\xi+y\eta)}.$$

Corollary 3.20. Let $f(x,y)$ be a function $\{0,1,\ldots,N-1\}^2 \to \mathbb{C}, N \in \mathbb{N}$ with Fourier spectrum $F(\xi,\eta)$. For every $(x,y) \in \{0,1,\ldots,N-1\}^2$, it holds:

$$\mathcal{D}\{f(x,y)\} = \mathcal{D}\big\{\widetilde{f}(x,y)\big\},$$

$$\mathcal{D}^{-1}\Big\{\mathcal{D}\big\{\widetilde{f}(x,y)\big\}\Big\} = \mathcal{D}^{-1}\big\{\widetilde{F}(\xi,\eta)\big\} = f(x,y).$$

We can proceed to the definition of the discrete periodic convolution and its connection with the discrete Fourier transform.

Definition 3.21. (Discrete periodic convolution). Let $f_1(x,y)$, $f_2(x,y)$ be functions of $\{0,1,\ldots,N-1\}^2 \to \mathbb{C}$, $N \in \mathbb{N}$. Functions $f(x,y) : \{0,1,\ldots,N-1\}^2 \to \mathbb{C}$ is called the *discrete periodic convolution* of functions f_1, f_2 denoted by $f(x,y) = f_1(x,y) * f_2(x,y)$, if

$$f(x,y) = \sum_{s=0}^{N-1}\sum_{t=0}^{N-1} f_1(s,t)\widetilde{f}_2(x-s,y-t).$$

Theorem 3.22. Let functions $f_1(x,y)$, $f_2(x,y) : \{0,1,\ldots,N-1\}^2 \to \mathbb{C}$, $N \in \mathbb{N}$ have Fourier spectra $F_1(\xi,\eta)$, $F_2(\xi,\eta)$. Then

$$\mathcal{D}\{f_1(x,y) * f_2(x,y)\} = F_1(\xi,\eta)\cdot F_2(\xi,\eta),$$
$$\mathcal{D}\{f_1(x,y)\cdot f_2(x,y)\} = \frac{1}{N^2}F_1(\xi,\eta) * F_2(\xi,\eta).$$

For proofs see [31].

Remark 3.23. This approach using periodic extension can be found for example in [16], while in [15] uses different approach employing modulo arithmetics.

Connection between the Fourier series and the discrete Fourier transform, sampling

The connection between the discrete Fourier transform and the Fourier series is described in [16]. Let f be a $p-$periodic function $\mathbb{R} \to \mathbb{R}$ and $\Phi_f(x)$ its Fourier series. The Fourier series is in the exponential form as per [13],

$$\Phi_f(x) = \sum_{k=-\infty}^{\infty} c_k \mathrm{e}^{2\pi \mathrm{i} \frac{kx}{p}} = c_0 \mathrm{e}^0 + \sum_{k=1}^{\infty} \left(c_k \mathrm{e}^{2\pi \mathrm{i} \frac{kx}{p}} + c_{-k} \mathrm{e}^{-2\pi \mathrm{i} \frac{kx}{p}} \right) = \tag{3.14}$$

$$= c_0 + \sum_{k=1}^{\infty} \left(c_k \cos \frac{2\pi kx}{p} + c_k \mathrm{i} \sin \frac{2\pi kx}{p} + c_{-k} \cos \frac{2\pi kx}{p} - c_{-k} \mathrm{i} \sin \frac{2\pi kx}{p} \right) =$$

$$= c_0 + \sum_{k=1}^{\infty} \left((c_k + c_{-k}) \cos \frac{2\pi kx}{p} + \mathrm{i}(c_k - c_{-k}) \sin \frac{2\pi kx}{p} \right), \tag{3.15}$$

where

$$c_k = \frac{1}{p} \int_0^p f(x) \mathrm{e}^{-2\pi \mathrm{i} \frac{kx}{p}} \mathrm{d}x$$

are complex numbers. The Fourier series can also be written in the real form described in [13],

$$\Phi_f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} \left(a_k \cos \frac{2\pi kx}{p} + b_k \sin \frac{2\pi kx}{p} \right), \tag{3.16}$$

where a_k, b_k are real numbers. Comparing equations 3.15 and 3.16, we obtain the connections between c_k and a_k, b_k,

$$c_0 = \frac{a_0}{2} \tag{3.17}$$

$$c_k + c_{-k} = a_k \tag{3.18}$$

$$\mathrm{i}(c_k - c_{-k}) = b_k. \tag{3.19}$$

At this point, a_k, b_k are real, while c_k is complex. Let us split the complex numbers c_k into their real and imaginary parts $c_k = \alpha_k + \mathrm{i}\beta_k$ in the equations 3.18 and 3.19. Noting that the left-hand sides of these equations are real,

$$c_k + c_{-k} = (\alpha_k + \alpha_{-k}) + \mathrm{i}(\beta_k + \beta_{-k}) \text{ is real } \implies \beta_k = -\beta_{-k},$$
$$c_k - c_{-k} = (\alpha_k - \alpha_{-k}) + \mathrm{i}(\beta_k - \beta_{-k}) \text{ is imaginary } \implies \alpha_k = \alpha_{-k}.$$

Hence, $c_k = c^*_{-k}$.

Let us now sample the function f with an equidistant sampling using N steps of length $\frac{p}{N}$ in each interval of length p. Denoting

$$\overset{\sqcap}{f}(n) = f\Big(n\frac{p}{N}\Big), \qquad n \in \mathbf{Z} \tag{3.20}$$

and plugging formally Equation 3.14 into Equation 3.20 in a point where Φ_f converges to f, we obtain

$$\overset{\sqcap}{f}(n) = \sum_{k=-\infty}^{\infty} c_k \mathrm{e}^{2\pi\mathrm{i}\frac{k}{p}n\frac{p}{N}} = \sum_{k=-\infty}^{\infty} c_k \mathrm{e}^{2\pi\mathrm{i}\frac{kn}{N}}.$$

Function $\mathrm{e}^{2\pi\mathrm{i}\frac{kn}{N}}$ is periodic in interval k with period N, therefore we can separate k into $k = l + Nm$, where $m \in \mathbb{Z}$, $l \in \{0, 1, \ldots, N-1\}$, and rewrite the Equation 3.20 as

$$\overset{\sqcap}{f}(n) = \sum_{m=-\infty}^{\infty} \sum_{l=0}^{N-1} c_{l+Nm} \mathrm{e}^{\frac{2\pi\mathrm{i}}{N}n(l+Nm)} = \sum_{m=-\infty}^{\infty} \sum_{l=0}^{N-1} c_{l+Nm} \mathrm{e}^{2\pi\mathrm{i}\frac{nl}{N}}.$$

Due to the series being absolutely convergent, we can swap the order of summation,

$$\overset{\sqcap}{f}(n) = \sum_{l=0}^{N-1} \sum_{m=-\infty}^{\infty} c_{l+Nm} \mathrm{e}^{2\pi\mathrm{i}\frac{nl}{N}}. \tag{3.21}$$

Denoting

$$\overset{\sqcap}{F}(l) = \sum_{m=-\infty}^{\infty} c_{l+Nm},$$

we can rewrite Equation 3.21 as

$$\overset{\sqcap}{f}(n) = \sum_{l=0}^{N-1} \overset{\sqcap}{F}(l) \mathrm{e}^{2\pi\mathrm{i}\frac{nl}{N}},$$

which is almost exactly the formula for one dimensional inverse discrete Fourier transform. We can use its formula to express $\overset{\sqcap}{F}(l)$ from the last equation as

$$\overset{\sqcap}{F}(l) = \frac{1}{N} \sum_{n=0}^{N-1} \overset{\sqcap}{f}(n) \mathrm{e}^{-2\pi\mathrm{i}\frac{nl}{N}},$$

which yields the connection between the Fourier series and the discrete Fourier transform as

$$N\overset{\sqcap}{F}(l) = N \sum_{m=-\infty}^{\infty} c_{l+Nm} = \sum_{n=0}^{N-1} \overset{\sqcap}{f}(n) \mathrm{e}^{\frac{-2\pi \mathrm{i}}{N} nl}.$$

Note: If a sequence c_k has a subsequence of length N or less, each $\overset{\sqcap}{F}(l)$ equals to zero or to one particular c_k. This means that the Fourier series of the original function f has a finite number of terms and can be fully reconstructed from $\overset{\sqcap}{F}(l)$. Such functions are called *bandlimited* and the Fourier series is equivalent notion with the discrete Fourier transform for them. According to *Nyquist-Shannon sampling theorem*, a bandlimited signal can be fully reconstructed from its samples provided the sampling rate exceeds twice the maximum frequency in the bandlimited signal.

Chapter 4

Kleť Observatory

Kleť Observatory is located 1068 meters above sea level in southwestern part of the Czech Republic in a protected landscape area with Bortle's dark sky class 1 to 2. Typical seeing[1] on the site is 1.5–2". The observatory ranks among the worlds most prolific professional NEO follow-up programmes. Kleť Observatory's team consists currently of just two members, but handles the largest telescope in continental Europe used exclusively for astrometric observation of asteroids and comets. The program includes confirmation of newly discovered NEO candidates, early follow-up of newly discovered NEOs, long-arc follow-up of NEOs in need of further astrometric data, recovery of lost NEOs and detection of cometary features. The highest priority is given to Virtual Impactors and Potentially Hazardous Asteroids.

Kleť Observatory has been carrying out astrometric measurements since 1969 using photographic plates, and since 1994 using CCD. Starting in 2002, it operates 1.06-m KLENOT telescope, the largest telescope in Europe used exclusively for follow-up astrometry of asteroids and comets. Kleť Observatory developed a workflow allowing for astrometric positions to be sent within few minutes of images' acquisition. The KLENOT project contributed 52 658 astrometric measurements of 5 867 bodies to Minor Planet Center between 2002 and 2008. Since 2014, KLENOT project cooperates with Space System Awareness Program of European Space Agency and obtained total

[1] See Section 2.3.4 for seeing explanation.

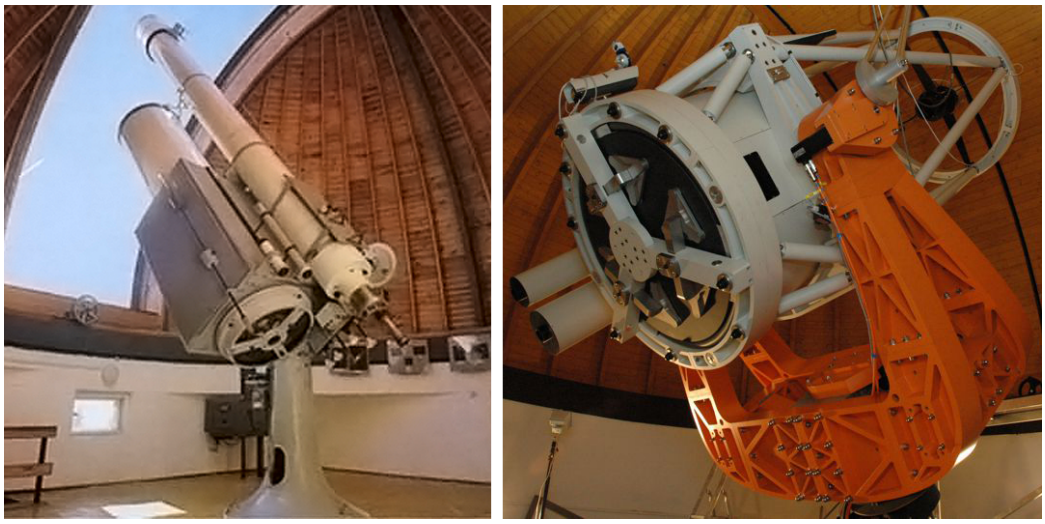


Figure 4.1: The two telescopes of Kleť Observatory.
Left: 0.57-m telescope. Right: 1.06-m KLENOT telescope.
Source: Kleť Observatory

27 426 astrometric measurements of 3 455 bodies between October 2014 and February 2017.

4.1 Hardware

Kleť Observatory uses two reflectors, 1.06-m KLENOT telescope and 0.57-m telescope, both equipped with CCDs.

```
KLENOT Telescope
1.06-m f/3 main mirror + 4-lenses primary focus corrector
f/2.7 optical system
CCD camera FLI ProLine PL-230, thermoelectric cooling (since 2013)
chip e2v 2048x2048 pix, pixel size 15 microns, scale 1.1''/pix
FOV 37'x37', limiting magnitude 22 mag for 120s exposure time


0.57-m Telescope
0.57-m f/5.2 main mirror
CCD camera SBIG ST-8, thermoelectric cooling, antiblooming
chip 1530x1020 pix, pixel size 9 microns, scale 0.9''/pix
FOV 16'x10', limiting magnitude 19.5 mag for 120s exposure time
```

The 1.06-m KLENOT telescope is equipped with computer controlled parallactic mount. Calibration images are not used. The whole setup, meaning the camera, mount and the dome, is operated from a control room.

The mount of the 0.57-m telescope is manually operated and its motors accuracy allows up to few minutes exposures. The focusing of the CCD is done manually as well. Dark frames are taken before each image set and immediately applied. Either 2x2 or 3x3 binning is used to avoid oversampling. Flatfields are not used.

4.2 Software and workflow

The observing process can be divided into several steps, each of which has its own software developed directly by Kleť Observatory to keep control over all of the aspects of the processing and of future development. The observatory's workflow is depicted on Figure 4.2. The observing time for both telescopes is taken from freeware program *AboutTime*, which provides the CCD-controlling computer with averaged time, including correction from chosen servers, to a subsecond precision.

The *Ephem* tool lists the observable objects of desired category and brightness in right ascension order. Observational priority is given to NEOCPs[2], V.I.s[3], minor bodies scheduled for observation by radio telescopes, and NEOs with high orbit uncertainty [1].

A set of the target object's images is taken and its up to the observer, supplemented with custom tool *Blink*, to identify the object. *Blink* is developed by the author of this work and can adjust the contrast, smooth the images,

[2]NEO Confirmation Page of Minor Planet Center with newly discovered asteroids in acute need of confirmatory observations

[3]Virtual Impactors, asteroids with nonzero collision probability within the next 100 years

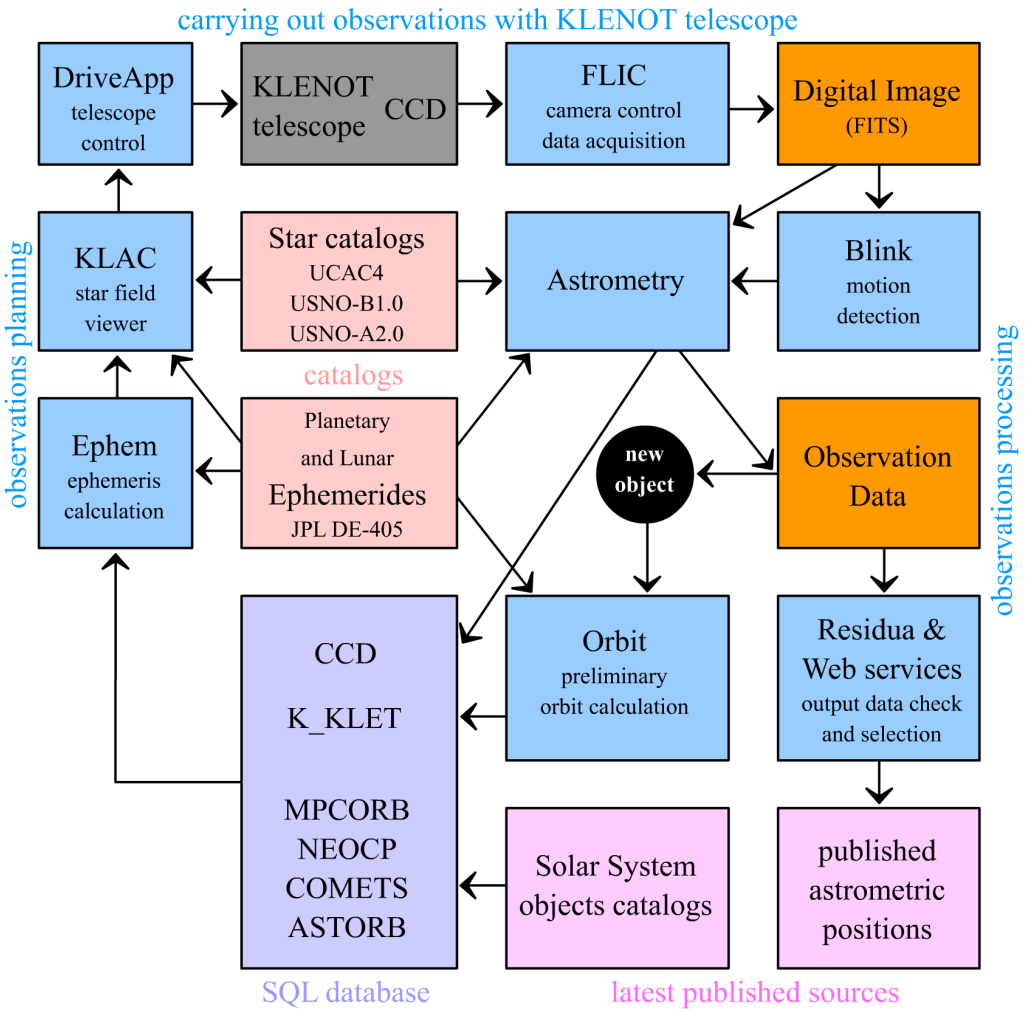


Figure 4.2: Workflow of KLENOT project's observing night.

and automatically align them. Printscreen of the running program can be found in Appendix E. *Blink* can animate the images of the set, making it easier for the observer to spot moving object, and stack the images to effectively lengthen their exposure time and therefore raise SNR, and even stack the images on a moving target. The last technique was immediately used to obtain astrometry of dwarf planet (136472) Makemake[4] [3]. In addition, KLENOT team developed an algorithm to mark spots where known asteroids and comets are expected on the image, which was also implemented into *Blink* to further simplify the work of the observer.

The *Astrometry* tool is used to reduce the image to objects and carry out the astrometry utilizing a stellar catalog. Although the core of *Astrometry* was written in 1993, new astrometric catalogs are implemented when available and the software is being continually upgraded. Printscreen of the running program can be found in Appendix F. The identification of stars is carried out employing similarity of triangles. The newest UCAC4 stellar catalog is used for identification of stars[33]. Plate parameters are computed, stars' positional errors are determined and standard 80 character output is produced for the targeted object. Typical SNR of observed targets, as measured by the program, is around 5, though it can be as low as 0.1 or as high as 20. The measured magnitude would therefore contain high errors, but an estimated value can be provided for other observatories to assess observability of the object on their own telescopes. The measured data is also automatically added to local MySQL databases.

The *Residua* tool is ran and the observed positions of the object are compared with ephemerides computed using local orbital elements database, which is updated daily with new Minor Planet Center data. The observer can therefore verify the correctness of the identification and possibly drop the observations with higher error before sending them to Minor Planet Center. The

[4] A precovery program by the author of the work determined it should be possible to find the new dwarf planet on old archived images. The images were stacked and the astrometry carried out, which lengthened the observing arc years prior to the dwarf planets discovery, leading directly to its orbit determination.

		Residuals		
Years	N Obs	RA	Decl	System Setup
2002-2007	47326	0.13 ± 0.53	0.25 ± 0.50	1.06-m + S300 + USNO A2.0
2008-2012	8134	0.05 ± 0.49	0.11 ± 0.49	1.06-m + S300 + USNO B1.0
2013-2017	28355	0.05 ± 0.44	-0.03 ± 0.45	1.06-m + PL-230 + UCAC4

Table 4.1: Kleť Observatory's average astrometric error.

whole process from taking the image set to sending out the objects observations takes up to several minutes.

The usual astrometric accuracy of professional observatories is under 0.5–0.7", depending on the star catalog used. The observatory's average astrometric error improved over the years [32], as can be seen in Table 4.1.

Chapter 5

Image acquisition and processing

The details of the described observing workflow are laid in this chapter. For every chosen target asteroid, a set of images is taken. That is, however, only the start of a process to obtain the desired positional data of the asteroid.

5.1 CCD Image

Let's describe and specify the parameters of a grayscale CCD image we deal with in this text. The image is taken by a CCD, a Charge Coupled Device, which constitutes of usually square elements arranged into flat matrix. The elements, called pixels, release electrons upon impact of photons on their photosensitive material. The electrons pile up for the integration time of the image's exposure and then are read-out by the device, assembling an image. Therefore a grayscale CCD image is a discrete function of two variables in Cartesian coordinates defined as,

Definition 5.1. (Grayscale CCD image.) Let $M = \{0, 1, 2, \ldots, \omega - 1\} \times \{0, 1, 2, \ldots, h - 1\}$ where $\omega, h \in \mathbb{N}$, let $R \in \mathbb{R}$. Function

$$f : M \to R$$

is called a grayscale CCD image. The set M is called field of view. The natural number ω is called image width and the natural number h is called image height. Elements of $[x, y] \in M$ are called pixels and the value of

function f in pixel $[x, y]$, $f(x, y) = I_{x,y}$ is called pixel value or intensity on pixel $[x, y]$.

Images are often taken in sets of $n \in \mathbb{N}$ images of the same area of sky, with the same exposure and taken in quick succession, called *image sets*.

Signal-to-Noise Ratio (S/N or SNR) is a single most important property involved with regard to astrometry. It defines the ratio between the useful signal from the observed target versus the the total measured signal which includes sky background and noises of any origin, and therefore determines the accuracy, or even feasibility, of the observation.

Definition 5.2. (Signal-to-Noise Ratio.) Let $(I_T + I_\xi)$ be a sum of pixel values over $n \in \mathbb{N}$ pixels on which the target is present. Let I_ξ be a sum of pixel values over $n \in \mathbb{N}$ pixels without the target present (intensity of noise). Then SNR of the target is

$$\text{SNR} = (I_T + I_\xi)/I_\xi.$$

The larger the SNR, the less uncertain are the data measured. However, this definition does not allow for easy measurement of SNR in practice. Instead, image's background must be determined at first, along with its standard deviation σ, which is a measure of the noise in the image. Then $(I_T + I_\xi)$ is measured over the $n \in \mathbb{N}$ pixels covered by the target, and divided by $\sqrt{n}$. The result is divided by σ to obtain SNR.

5.2 Image Calibration

In an idealized case, the image taken by the CCD accurately captures the function of light intensities in the field of view. In reality, that is not the case and the image is deformed by the used optical and imaging system, and also contains additional noise which was not present in the incoming light. The purpose of image calibration is to remove or reduce these instrumental influences without altering the attained light. The theory described in this chapter is based on [34].

5.2.1 CCD camera response model

Pixel values on a CCD image can be expressed as follows

$$s(x, y) = B(x, y) + tD(x, y) + tG(x, y)I(x, y) + \xi, \tag{5.1}$$

where $B(x, y)$ is the bias value of each pixel, t is the integration time (also called exposure time), $D(x, y)$ is the dark current, $G(x, y)$ is the instrument's sensitivity and $I(x, y)$ is the light flux reaching the pixel $[x, y]$. The remaining component, noise denoted ξ, is not dealt with in image calibration.

To solve the Equation 5.1 for I, we have to express B, D, and G, which is done by taking calibration images (also called calibration frames) under controlled conditions: bias, dark and flat frames respectively.

5.2.2 Bias Frame

An image taken with the shortest possible exposure time $t = 0$ and closed camera shutter $I(x, y) = 0$ is called bias frame. The bias frame contains only readout noise[1]. The Equation 5.1 becomes

$$b(x, y) = B(x, y) + \xi.$$

Since the noise is random, we use b to obtain the estimate of B, using a tilde to denote the estimate

$$\widetilde{B}(x, y) = b(x, y).$$

We can further refine the estimate by averaging several bias frames. Due to the random nature of the included noise, its contribution decreases with the square root of the number of frames averaged. The refined bias frame is called *master bias frame*.

$$\widetilde{B}(x, y) = \frac{1}{N} \sum_{i=1}^{N} b_i(x, y)$$

[1] See [6] for details on various instrumental noise

5.2.3 Dark Frame

An image taken with the exposure time t_d and closed camera shutter $I(x, y) = 0$ is called dark frame. The dark frame contains thermal noise and readout noise, which means it includes bias frame in itself.

$$d(x, y) = B(x, y) + t_d D(x, y) + \xi$$

We subtract the bias and divide by the exposure time t_d to obtain dark current estimate

$$\widetilde{D}(x, y) = \frac{d(x, y) - \widetilde{B}(x, y)}{t_d}$$

and refine the dark current estimate by averaging M dark current estimates

$$\widetilde{D}(x, y) = \frac{1}{t_d}\frac{1}{M}\sum_{i=1}^{M} d_i(x, y) - \widetilde{B}(x, y).$$

In practice, we do not work with dark current frames. Instead, dark frame is taken with the same exposure time as data images and the included bias is left as is. Several such dark frames are averaged to reduce the noise and the resulting image is called *master dark frame*.

$$\widetilde{D}_M(x, y) = \frac{1}{M}\sum_{i=1}^{M} d_i(x, y) = \frac{1}{M}\sum_{i=1}^{M} B_i(x, y) + tD_i(x, y) + \xi \tag{5.2}$$

This *master dark frame* also includes master bias frame and therefore solves first two terms in Equation 5.1.

5.2.4 Flat Frame

Flat frame, generally called flatfield frame, is an image taken with the exposure time t_f of an evenly illuminated field, where $I(x, y)$ is constant. Flatfield image represents the response of the whole optical system to the arriving light. It accounts for different pixel sensitivity, dust particles in the optical system, vignetting, and also contains dark frame (and therefore also bias frame) in itself. The flat frame is

$$f(x, y) = B(x, y) + t_f D(x, y) + t_f G(x, y)L + \xi, \tag{5.3}$$

where L is the light flux reaching each pixel, assumed equal across the frame. We take a set of the flat frames and subtract master dark frame $D_M^F(x,y)$ from each, obtaining

$$f'(x,y) = f(x,y) - D_M^F(x,y) = t_f G(x,y)L + \xi$$

and average the images to reduce the noise contribution, arriving to *master flat frame*

$$\widetilde{F}_M(x,y) = \frac{1}{N}\sum_{i=1}^{N} f'(x,y) = \frac{1}{N}\sum_{i=1}^{N} f(x,y) - D_M^F(x,y). \tag{5.4}$$

If we knew the light flux L, we could solve the Equation 5.3 for $G(x,y)$. However, the value of L is not known and so instead we settle with the assumption that L is constant, which allows us to remove the light flux variations across the frame. We introduce

$$G(x,y) = \overline{G}g(x,y),$$

where $\overline{G}$ is the average of $G(x,y)$ across the image and the average of $g(x,y)$ across the image is 1. The Equation 5.4 becomes

$$\widetilde{F}_M(x,y) = t_f \overline{G} g(x,y)L + \xi.$$

We take the average of $\widetilde{F}_M(x,y)$ across the frame

$$\overline{F} = \sum_x \sum_y t_f \overline{G} g(x,y)L = t_f \overline{G} L \sum_x \sum_y g(x,y) = t_f \overline{G} L$$

and then by dividing the Equation 5.4 by $\overline{F}$ we obtain the *normalized master flat frame*

$$\frac{\widetilde{F}_M(x,y)}{\overline{F}} = g(x,y) \tag{5.5}$$

5.2.5 Reducing data frames

Now we can proceed to reduce our data frame

$$s(x,y) = B(x,y) + tD(x,y) + t\overline{G}g(x,y)I(x,y) + \xi.$$

We subtract the *master dark frame* from Equation 5.2 and divide by the *normalized master flat frame* from Equation 5.5

$$\frac{\overline{F}}{\widetilde{F}_M(x,y)}[s(x,y) - \widetilde{D}_M(x,y)] = t\overline{G}I(x,y) + \xi$$

$$\widetilde{I}(x,y) = \frac{1}{\overline{G} \cdot t}\frac{\overline{F}}{\widetilde{F}_M(x,y)}[s(x,y) - \widetilde{D}_M(x,y)]$$

to obtain an estimate of the arrived light flux, which is the best we can do without a reference source calibrated in absolute units.

5.2.6 Frame combining methods

Before we can combine the frames, they need to have the same range of intensity values, so we have to normalize the images. They also need to be aligned for stacking, which could be done by astrometric measurement of the image as described in Section 5.4 to obtain the *plate parameters*, then using their knowledge to achieve the desired translation, rotation and scaling of the images. In practice, mostly translation is used for images in the set. The astrometric measurement provides coordinates of the center of each image, which is sufficient to translate the images appropriately for averaging them. Although that may mean the arithmetic average, there are also several other notable methods of frame combining worth mentioning.

Arithmetic mean

Arithmetic mean, or simply mean, or average, defined as

$$\overline{I}(x,y) = \frac{1}{N}\sum_{i=1}^{N} I_i(x,y)$$

is computed for each pixel $[x, y]$ of the N data images. Arithmetic average reduces the random noise in the image and is suitable for gaussian-distributed data. The result is not an integer value, unlike the data, therefore it smoothes out the quantization effect. On the downside, the arithmetic mean is very sensitive to outliers (data points far from the average of the most other data

points). These deviant values, caused by for example cosmic ray hits, easily skew the average value.

Median

Median is the middle value of a sorted set of numbers, which makes it more suitable for handling data with outliers than arithmetic mean. It makes median useful to remove cosmic ray hits from the images. There is no widely accepted standard notation for the median, hence the following definition is based on [11].

Definition 5.3. (Median.) Let's have a set of measured pixel intensities $(I_1, I_2, \ldots, I_N)$, where $N \in \mathbb{N}$, and arrange them from smallest to greatest into a sorted list

$$I_{(1)} \leq I_{(2)} \leq \cdots \leq I_{(N)},$$

then *median* $\widetilde{I}$ is defined as the middle number of this sorted list,

$$\widetilde{I} = \begin{cases} I_{(\frac{N+1}{2})} & N \text{ is odd} \\ \frac{1}{2}(I_{(\frac{N}{2})} + I_{(\frac{2}{2}+1)}) & N \text{ is even.} \end{cases}$$

In our case, median is computed $x \cdot y$ times, once for each pixel $[x, y]$ of the N data images $I(x, y)$ and the $x \cdot y$ medians construct the resulting image. The outliers do not affect the result in a major way. On the downside, as the result is also an integer value, the median does not smooth the quantization effect. Median also reduces gaussian noise less than the arithmetic mean does.

Mean–median

Mean–median counters the disadvantages of median by improving the statistical efficiency of median and bypassing the quantization problem. Denoting $J_0 = \{1, .., N\}$, where N is the number of images in the set, we start by computing the median $e(x, y)$, and then the standard deviation $\sigma(x, y)$ of the pixel values around the median $e(x, y)$,

$$\sigma(x, y) = \sqrt{\frac{1}{N} \sum_{i=J_0} (I_i(x, y) - e(x, y))^2}.$$

We discard the values outside a specified number of standard deviations κ (usually 1.5 or 2)

$$J_1(x,y) = \{i \in J_0 \mid I_i(x,y) - e(x,y) \mid \leq \kappa\sigma\}.$$

The remaining values are averaged

$$e_1(x,y) = \frac{1}{\mid J_1 \mid} \sum_{i=J_1} I_i(x,y).$$

The computation of mean–median $e_1(x,y)$ is repeated for every pixel $[x,y]$ of the image, and the computed mean–medians produce the final image. The mean–median works well on large image sets, where the outliers do not increase the computed standard deviation drastically.

Kappa–sigma clipping

Kappa–sigma clipping is more elaborate version of mean–median. Denoting $J_1 = \{1,..,N\}$, where N is the number of images in the set, and setting the iteration count $k = 1$ for the first iteration, we calculate the median $e_k(x,y)$ and the standard deviation $\sigma_k(x,y)$ of the pixel values around the median $e_k(x,y)$,

$$\sigma_k(x,y) = \sqrt{\frac{1}{N} \sum_{i=J_k} (I_i(x,y) - e_k(x,y))^2}.$$

The values with large deviations (usually κ is set to 1.5 or 2) relative to the standard deviation are excluded

$$J_{k+1}(x,y) = \{i \in J_k \mid I_i(x,y) - e_k(x,y) \mid \leq \kappa\sigma\}$$

and the remaining values are averaged using arithmetic mean

$$e_{k+1}(x,y) = \frac{1}{\mid J_{k+1} \mid} \sum_{i=J_{k+1}} I_i(x,y).$$

Second iteration $k = 2$ is ran to compute anew the standard deviation $\sigma_k(x,y)$ and so on, repeating the process until there is no change in the arithmetic mean or the iteration limit is reached. The arithmetic mean $e_{k+1}(x,y)$ then becomes the final value for the examined pixel $[x,y]$. The process is repeated for each pixel of the image.

5.3 Filters

Filters re-evaluate the value of every pixel in an image. The basic notion is that the image contains signal we want to enhance or suppress. Mathematically, filters are example of use of convolution, as per Definition 3.10. Convolution is a mathematical operation on two functions (in our case, that is the image and the filter) which produces third function (the adjusted image). Operation reverse to convolution is deconvolution. Depending on the process of re-evaluation, we can divide filters into different categories:

- *Linear* Filters, whose output values are linear combinations of the pixels in the original image, are well understood and fast to compute, but incapable of smoothing the image without simultaneously blurring it. *Nonlinear* filters lack this property.

- Filters applied in *spatial* or *frequency* domain. Making use of *Fourier transform*, it is possible to decompose the image into frequencies that make it up, called *frequency domain representation*, as described in Section 3.4. Filters can then be applied either in the *spatial* domain, or in the *frequency* domain. Some operations are advantageous to perform in one of the domains, for example convolution in spatial domain corresponds to ordinary multiplication in frequency domain. After performing the operations, transformation of the result back to the spatial domain is made.

- *High-pass* or *Low-pass* filters. Taking advantage of *Fourier transform* again to represent the image in the frequency domain, *low-pass* filter passes signal with low frequency and suppresses signal with high frequency, effectively blurring the image. *High-pass* filter does the opposite, passing signal with high frequency and suppressing signal with low frequency, effectively sharpening the image. When a low-pass and high-pass filter is combined, it results in *band-pass* filter which removes both low and high frequencies.

5.3.1 Linear low-pass filters

Low-pass filter passes signal with low frequency and suppresses signal with high frequency. Also called blur or smooth filter, it reduces noise by removing small fluctuations on the image. The filter consists of a set of weights, and each filtered pixel is obtained as the weighted mean of itself and small number of its neighbors.

Moving average, also called Box filter, is an example of a simple linear, low-pass filter. It replaces each pixel value by the average of pixel values in a square centered at that pixel. Other linear filters work in a similar manner, just making use of a weighted average instead.

Let's have a discrete function f_{ij}, where $i, j = 1, \ldots, n$, denoting the pixel values of the CCD image. A linear filter of size $(2m+1) \times (2m+1)$, with specified weight w_{kl} for $k, l = -m, \ldots, m$, gives

$$g_{ij} = \sum_{k=-m}^{m} \sum_{l=-m}^{m} w_{kl} f_{i+k,j+l}, \tag{5.6}$$

for $i, j = (m+1), \ldots, (n-m)$, meaning excluding the edges m of the image. There are various ways to deal with the border edges of g:

- Discarding, resulting in g smaller than f
- Setting their pixels to zero or to f
- Modifying the filter, for example to
  - ignore the pixels laying outside the image f
  - reflect the image f, so that $f_{i,n+1} = f_{i,n-1}$ etc.
  - wrap the image f around so that $f_{i,n+1} = f_{i,1}$ etc.

Additionally, all considered filters in this work have odd number sizes for simplicity.

If all the elements in w are *positive*, the effect of the filter is smoothing. The smoothing blurs the image, but we can avoid severe loss of resolution by choosing m smaller than one half of the width of the expected features.

If some of the weights are *negative*, the filter outputs a difference between pixel values, which emphasizes edges.

The most commonly used smoothing filter is moving average, for which the weights from the Equation 5.6 are $w_{kl} = 1/(2m+1)^2$. That means $(2m+1)^2$ additions and multiplications per pixel involved, which may make the filtering slow for huge images. However, the filter is *separable*.

A filter is said to be separable, if it can be performed by first filtering the image inside a $(2m+1) \times 1$ window (column), and then inside a $1 \times (2m+1)$ window (row). We obtain

$$h_{ij} = \sum_{k=-m}^{m} w_k^c f_{i+k,j}$$

for $i = (m+1), \ldots, (n-m)$, $j = 1, \ldots, n$ while using column weights $w_{-m}^c, \ldots, w_m^c$, followed by a row operation

$$g_{ij} = \sum_{l=-m}^{m} w_l^r f_{i,j+l}$$

for $i, j = (m+1), \ldots, (n-m)$ and row weights $w_{-m}^r, \ldots, w_m^r$. In order for his to be possible, the array of weights w_{kl} must be expressible as the product of the column and row weights,

$$w_{kl} = w_k^c w_l^r \qquad \text{for } k, l = -m, \ldots, m.$$

This reduces the number of operations per pixel from $(2m+1)^2$ to $2(2m+1)$.

Even quicker computation is possible using *recursive* implementation. The filter's window moves along pixels, but most of the pixels remain the same between two steps. Therefore instead of computing each pixel intensity from scratch, it is possible to take the output from location (i, j) and update it to obtain the output at location $(i+1, j)$. Specifically, the first $(2m+1)$ pixel values in column j are averaged,

$$h_{m+1,j} = \frac{1}{2m+1} \sum_{k=1}^{2m+1} f_{kj}.$$

Then, the pixel value in the first row $f_{1,j}$ is dropped from the average and the pixel in row $(2m+2)$ is added. The operation is repeated for every value in column j,

$$h_{ij} = h_{i-1,j} + \frac{f_{i+m,j} - f_{i-m-1,j}}{2m+1}$$

for $i = (m+2)\ldots(n-m)$. This procedure is repeated for each column $j = 1, \ldots, n$ to obtain h. Then the same algorithm is applied along each row of h, to obtain g. The number of operations per pixel has been reduced to 4 *irrespective of the filter size* m.

Although the moving average filter is simple and fast, it has two drawbacks: it is not *isotropic* (circularly symmetric), but smoothes further along diagonals, and the weights have abrupt cut-off instead of gradually dropping to zero, which leaves discontinuities in the smoothed image. These disadvantages are removed by employing a *Gaussian* filter, which has weights determined by the probability density function of bivariate Normal distribution with variance σ^2. With circular Gaussian function, where $\sigma_i = \sigma_j = \sigma$, the weights are

$$w_{ij} = \frac{1}{2\pi\sigma^2}\ \mathrm{e}^{-\frac{i^2+j^2}{2\sigma^2}},$$

for whole numbers $i, j = -3\sigma, \ldots, 3\sigma$. Note that the divisor of $2\pi\sigma^2$ ensures that the weights sum to unity, which is a common convention for smoothing filters. As to its computational speed, Gaussian filter can be implemented efficiently by approximating it with several passes of moving average filters.

5.3.2 Linear high-pass filters

High-pass filter passes signal with high frequency and suppresses signal with low frequency. Also called sharpen, brings out fine details but also the noise in the image. The process consists of baseline or continuum estimation and subsequent subtraction. That can be done for example by least squares fitting of a polynomial.

Unsharp masking is an example of a linear, high-pass filter. It is often available in image processing software and commonly used in astronomy to bring

out the details and enhance the contrast of images.

Let's have a CCD image $f(x, y)$. First, heavy blur using a low-pass filter is applied to $f(x, y)$, resulting into $f_B(x, y)$. The blurred image is subtracted from the CCD image, creating a mask of edges and other high frequency components of the image. This mask is then added to the CCD image.

$$g(x, y) = f(x, y) + \alpha(f(x, y) - f_B(x, y))$$

Scaling factor α controls the strength of the unsharp masking effect. The degree to which $f_B(x, y)$ is blurred controls what size structures the mask functions on most effectively: strongly blurred image enhances large structures but will miss small structures, while small subtle blurs enhance tiny structures but miss larger structures.

It is also possible to work in Laplacian sharpening, which blurs the CCD image at several different scales, producing mask for each scale. Laplacian transform is a true wavelet transform, which is why this method of sharpening is sometimes referred to as wavelet sharpening. There are many various wavelet transforms which can be used for image enhancement in various ways.

5.3.3 Nonlinear filters

Linear, low-pass (smoothing) filters inevitably blur the edges, because both noise and edges are high frequency components of the image. Nonlinear filters are able to reduce noise while preserving the edges. However, there is a vast number of filters to choose from, each to be used for different noise conditions and image types, they may be computationally demanding and may also distort the features in the image, so they should be used with caution, especially in astrometric research.

Many of these filters are based on the histogram of pixel values around the pixel to be adjusted, or on spatial distribution of pixel values in a neighborhood.

An example of a low-pass, but nonlinear filter is *moving median* filter. It uses median, described in Section 5.2.6, instead of average and is favored for images with high intensity changes (edges), where average would result in high blurring. Median can also effectively remove salt-and-pepper noise, which in our case are hot and cold pixels of CCD image.

In general sense, other more *robust estimators* can be used in the place of median, for example *trimmed mean*, which is obtained by averaging the 50% of pixel values nearest to the median in each filter window [30].

5.3.4 Adaptive filters

An adaptive filter is a system containing filter that has a transfer function controlled by variable parameters and means to adjust those parameters according to an optimization algorithm.

Minimum variance filter is an elegant and fast adaptive filter. For this filter, mean $\overline{f}$ and variance S are evaluated in five $(2m+1) \times (2m+1)$ subwindows within a $(4m+1) \times (4m+1)$ window, and the filter output is the mean of the subwindow with the smallest variance.

5.4 Astrometry algorithm by Kleť

This chapter mathematically covers the method in which Kleť Observatory obtains its astrometrical measurements. As the point of this work is to eliminate background variations from CCD images to raise both the quality and quantity of the acquired astrometrical positions of asteroids and comets, it is essential to understand how these are computed from the CCD images. The incentive for this chapter comes from the source code of *Astrometry* program, routinely used on Kleť Observatory for astrometric measurements of small bodies of the Solar system, and was also published in [4].

5.4.1 Gray level remapping

After pre-processing and filtering of the data image is done, an observer should be able to identify the object to be measured on the image by his sight. Since the CCD has many more intensity values than human eye is capable of perceiving, and screen is capable of displaying, it is necessary to remap the intensity values to 256 intensities intended to be viewed on screen.

Definition 5.4. (CCD image). The image has dimension $n \times m$, where $n, m \in \mathbb{N}$, and contains $N = n \cdot m$ pixels with intensities $I \in \mathbb{N}$. [2] The pixels are arranged in a two-dimensional matrix, I_{xy}, where $x = \{0, \ldots, n-1\}$, $y = \{0, \ldots, m-1\}$.

Assuming the image has level background, the program constructs a histogram m_i from the N image pixel intensities I,

$$N = \sum_i m_i. \tag{5.7}$$

The *Background* value is determined as a mean μ according to the following equation, because $\sum_i i m_i = \sum I$. Standard deviation σ is also computed.

$$\mu = \frac{1}{N} \sum_i i \cdot m_i \tag{5.8}$$

$$\sigma = \sqrt{\frac{1}{N} \sum_i m_i \cdot (\mu - i)^2}$$

The histogram is limited to 3σ around the current *Background* value μ and the computation is iterated. This is cutting off the wings of the numerosity graph until the *Background* value stabilizes or an arbitrary 20 iterations are reached.

$$N = \sum_{i=\mu-3\sigma+1}^{\mu+3\sigma} m_i. \tag{5.9}$$

Since Kleť Observatory focuses on observing faint objects, the image intensities are then remapped to show to the observer 2σ intensity levels around the *Background* value,

$$J_{xy} = \frac{I_{xy} - \mu}{4\sigma} \cdot 255,$$

[2] Since the camera uses 16-bit A/D electronics, the available values for $I \in \langle 0; 65535 \rangle$

while lower values are assigned `RGB[0,0,0]` and higher values are assigned `RGB[255,255,255]`. This remapping is of interim nature for the sake of the observer only and the following astrometry is done with the full intensity scale of the data.

5.4.2 Image reduction

The astrometry starts with reduction of CCD image to objects. Objects on the image are discerned by an algorithm, which scans the image for intensity peaks. Detection is successful if the peak intensity is above the SNR threshold chosen by the user. Usually, as in Definition 5.2, SNR is defined as $\mathrm{SNR} = I_{xy}/\sigma$, however Kleť algorithm uses $\mathrm{SNR} = (I_{xy} - \mu)/\sigma$, therefore the program is unable to detect objects with peak intensity under the estimated *Background* value μ.

The position of each detected object is refined to subpixel accuracy by fitting the object with a two-dimensional centroid. The calculation is analogous to a center of mass calculation of an object, with the local mass density replaced by pixel intensity values and allowing a discrimination level at the average background μ. The computation is performed over a box size $[w, h]$ preset by the user, starting at coordinates $[X_0, Y_0]$. P_{ij} represents the intensity value of the pixel at coordinates $[X_0 + i, Y_0 + j]$.

$$X_{\mathrm{centroid}} = X_0 + \frac{\sum_{i=0}^{w-1}\sum_{j=0}^{h-1} i \times P_{ij} - \mu}{\sum_{i=0}^{w-1}\sum_{j=0}^{h-1} P_{ij} - \mu}$$

$$Y_{\mathrm{centroid}} = Y_0 + \frac{\sum_{i=0}^{w-1}\sum_{j=0}^{h-1} j \times P_{ij} - \mu}{\sum_{i=0}^{w-1}\sum_{j=0}^{h-1} P_{ij} - \mu}$$

Remark 5.5. The positional accuracy of astrometry is in no small amount affected by the choice of the CCD chip. Optimally, the CCD chip should have pixel size half the best seeing on the site, and use binning (pairing of

pixels) for worse observing nights. If the pixels are too big, the image of a star becomes undersampled — the image does not cover enough pixels to interpose Gaussian profile and determine the center of the star accurately. In extreme case, having the star on one pixel means its position can be determined only to one pixel accuracy. On the other hand, too small pixels will result in oversampled image. The useful signal is spread over more pixels than necessary, each pixel containing noise as well, and SNR lowers.

5.4.3 Catalog stars' identification

The identification of the star field is near automatic. The user chooses a key star and the software, using known target coordinates of the telescope, lists in its memory catalog stars in the field of view, starting with the key star's coordinates. The following parts of identification algorithm is repeated until the user approves the result.

Triangle identification

At this moment we have the CCD image reduced to list of objects with positions $[x_O, y_O]$ on the image, and list of catalog stars in the field of view with positions $[x_H, y_H]$. Let us match three of the objects, which we will mark L, M, N, with three of the catalog stars, which we will mark i, j, k.

Starting with the key star i chosen by the user, the algorithm compares distances between each two catalog stars with distances between every pair of reduced objects, until a match is found (within some marginal error). Let us mark the matching pair L, M and i, j. A third star N, such that the resulting triangle has every corner angle greater than 10 degrees, is chosen, and distances of the stars in the triangle L, M, N are computed and compared with distances between objects i, j and any another object, until matching object k is found. As similarity of triangles is used to match the triangle of stars to reduced objects, hence even rotated and flipped images will be flawlessly identified. If there happens to be no match at any of the steps, the algorithm repeats with next suitable choice on the list until a match is found.

The three objects L, M, N and their matching three catalog stars i, j, k are highlighted in magenta in user's interface (see Appendix F).

Matching catalog stars to reduced objects

The CCD image does not overlap with the catalog field exactly: it may be moved, rotated, flipped or scaled. To automatically match the rest of the stars, the transformation covering these changes has to be found. The triangle of the three reduced objects $L, M, N = \{[x_{OL}, y_{OL}], [x_{OM}, y_{OM}], [x_{ON}, y_{ON}]\}$ and the corresponding triangle of the three catalog stars $i, j, k = \{[x_{Hi}, y_{Hi}], [x_{Hj}, y_{Hj}], [x_{Hk}, y_{Hk}]\}$ are used to determine field parameters, a similarity transformation between the two triangles covering translation, rotation and scaling of the triangle.

$$A_2 = \frac{(x_{ON} - x_{OL}) \cdot (x_{Hj} - x_{Hi}) - (x_{OM} - x_{OL}) \cdot (x_{Hk} - x_{Hi})}{(y_{Hk} - y_{Hi}) \cdot (x_{Hj} - x_{Hi}) - (y_{Hj} - y_{Hi}) \cdot (x_{Hk} - x_{Hi})}$$

$$A_1 = \frac{x_{OM} - x_{OL} - A_2 \cdot (y_{Hj} - y_{Hi})}{x_{Hj} - x_{Hi}}$$

$$A_0 = x_{OL} - A_1 \cdot x_{Hi} - A_2 \cdot y_{Hi}$$

$$B_2 = \frac{(y_{ON} - y_{OL}) \cdot (x_{Hj} - xHi) - (y_{OM} - y_{OL}) \cdot (x_{Hk} - x_{Hi})}{(y_{Hk} - y_{Hi}) \cdot (x_{Hj} - x_{Hi}) - (y_{Hj} - y_{Hi}) \cdot (x_{Hk} - x_{Hi})}$$

$$B_1 = \frac{y_{OM} - y_{OL} - B_2 \cdot (y_{Hj} - y_{Hi})}{x_{Hj} - x_{Hi}}$$

$$B_0 = y_{OL} - B_1 \cdot x_{Hi} - B_2 \cdot y_{Hi}$$

The transformation is applied on position $[x_H, y_H]$ of every catalog star to obtain their expected positions on the image $[A_0 + A_1 \cdot x_H + A_2 \cdot y_H, B_0 + B_1 \cdot x_H + B_2 \cdot y_H]$, which are then compared with coordinates of reduced objects. The catalog stars, which were matched with objects, are highlighted in lime color.

The result is presented to the user as per the image in Appendix F. On the left side of the program is the stellar catalog field, with the triangle of the three stars used for field parameters determination highlighted. On the right is the reduced CCD image taken by the telescope. Objects entering astrometry algorithm are marked in yellow, the triangle of stars used in field parameters determination is marked in magenta and objects which were

matched with catalog stars are marked in lime color. The user either rejects the identification, causing the algorithm to start anew in search for the next suitable triangle, or confirms the correctness of the identification.

Gnomonic projection

At this point, the cartesian coordinates of the reduced objects are matched to the cartesian coordinates of the catalog stars. However, the celestial equatorial coordinates $[\alpha, \delta]$ of the stars can not be used right away. The catalog celestial equatorial coordinates $[\alpha, \delta]$ of stars need to be projected from spherical plane to the flat CCD image. Given that the center of the image $[\alpha_C, \delta_C]$ is known, gnomonic projection, described in Section 3.1.4, can be used to obtain the projected coordinates $[\xi, \eta]$ of each identified star.

Plate parameters

The resulting projected position $[\xi, \eta]$ of any object is a function of position on the image $[x_0, y_0]$. Least squares method, described in Section 3.3, is used on the two-dimensional data of identified stellar positions. For the number of stars $N < 16$ linear method is used, otherwise quadratic fitting gives more accurate results. Known positions $[x_0, y_0]$ of the N identified stars are used to build vector $\mathbf{s}$ according to chosen method and matrices A and B are calculated. Gauss elimination is used to solve matrix A to obtain plate parameter $\mathbf{a}$ and to solve matrix B to obtain plate parameter $\mathbf{b}$.

$$\text{Linear Method} : \mathbf{s} = \left(\sum_{i=1}^{N} 1, \sum_{i=1}^{N} x_i, \sum_{i=1}^{N} y_i \right)$$

$$\text{Quadratic Method} : \mathbf{s} = \left(\sum_{i=1}^{N} 1, \sum_{i=1}^{N} x_i, \sum_{i=1}^{N} y_i, \sum_{i=1}^{N} x_i^2, \sum_{i=1}^{N} x_i y_i, \sum_{i=1}^{N} y_i^2 \right)$$

$$\text{Matrix } A \text{ for } \mathbf{a} : A = \left(\mathbf{s} \cdot \mathbf{s}^\top | \sum_{i=1}^{N} \xi_i \cdot \mathbf{s}^\top \right)$$

$$\text{Matrix } B \text{ for } \mathbf{b} : B = \left(\mathbf{s} \cdot \mathbf{s}^\top | \sum_{i=1}^{N} \eta_i \cdot \mathbf{s}^\top \right)$$

5.4.4 Object's astrometry

The user marks desired unknown object on the image and plate parameters are used on its cartesian position $[x, y]$ to determine its projected celestial coordinates $[\xi, \eta]$. Vector $\mathbf{s}$ is computed for the method used and $N = 1$ objects, then the projected coordinates $[\xi, \eta]$ are

$$\xi = \sum_{i=1}^{N} a_i s_i, \qquad \eta = \sum_{i=1}^{N} b_i s_i.$$

Inverse gnomonic projection is applied to obtain object's celestial coordinates $[\alpha, \delta]$. In the same manner position $[\alpha_0, \delta_0]$ of each identified star is determined. The differences against the catalog values along with variance from Chi-square test are written into a table and presented to the user. The user can reject identified stars based on their higher astrometry error, upon which the astrometry repeats the computations from determination of plate parameters onward. When the measurement is accepted by the user, the program writes output in standard MPC line[3] and saves basic CCD image data to local MySQL server database for later use.

5.4.5 Limitations

The algorithm described performs very well on an image taken by KLENOT telescope in ideal conditions, requiring dark, moonless nights with low humidity, the measured object not being close to the edges of the CCD and without bright stars being in the field of view, and shorter exposure times being used. However, such ideal conditions are often not possible.

As Kleť Observatory does not use calibration images, the background is very uneven due to many contributing factors, including fringing, vignetting, and background gradient from various sources, and also other random noise, to the point of being unable to carry out astrometric measurements at all. Additionally, starting to use calibration images as of now (2018) helps very little with mining information from the vast archive of Kleť, containing over 150 000 digital images.

[3]Which can be found on http://www.minorplanetcenter.net/iau/info/OpticalObs.html

The weak point of the astrometric process is assuming level background on the image. SNR threshold set by the observer is used to reduce the image to objects, their position refined by fitting through a Gaussian profile. Whenever the strict assumption of level background is not locally satisfied, the astrometry is burdened by error. Even worse situation occurs for objects in darker parts of the image, where the intensity values are under the average background computed by Equation 5.8, making the object literally undetectable by *Astrometry* even when the observer can see it. On the other end of the problem are brighter parts of the image, where even random noise is bright enough to be confused for objects by *Astrometry* program. An example of a problematic image is included in Appendix A, and its counterpart from *Astrometry* program right after it in Appendix B. Constructing an artificial flatfield to level the background of the images will have an undisputable impact on astrometry of KLENOT project.

Chapter 6

Chosen background flattening methods

Since KLENOT project observes very faint objects, usually only several pixels across in size, tampering with high frequency noise on the image will inevitably lead to corruption of the astrometric data. On the other hand, the low frequency noise (background) could be removed.

While many image processing methods are known, most are not suitable for our special case. These methods were carefully considered and the most promising ones were chosen. Some were adjusted to be suitable for our use as appropriate, and new methods were developed to deal with our images.

Six methods were developed for this work, where the first three deal with a set of consecutively taken images, which is the usual way of taking images for astrometric purposes, while the last three deal with a single image. They were programmed into *Blink*, software for image processing and viewing which the author is developing for Kleť Observatory's needs, and tested for performance. *Blink* is routinely used for astrometric observations by Kleť Observatory, and newly carries out the new flatfield processes automatically when the user chooses the desired method.

6.1 Median flatfield

Median filter, described in Section 5.2.6, is commonly used to reduce gaussian noise in the image, for which the images have to be precisely aligned with each other. Such use of median does not achieve background removal through artificial flatfield creation. However, photographers use median to remove moving objects from an image to take scenery photos at frequented places without the walking people being present on the resulting image. The median of values on each pixel of the set of images is used for the result, therefore a person walking across photographed scene present only in less than half of the images on given pixels is removed as the result.

The KLENOT telescope takes a set of images of each observed field of view. Its imperfect tracking causes the stars to 'walk' across the field of view, which led me to the idea of using median to create an artificial flatfield from the image set.

The images have to be normalized due to each image of the set having slightly different background. For each image of the set $I_i(x, y)$, where $i = (1, .., N)$, *Background* values β_i are computed according to equations 5.7 to 5.9. The images are normalized by

$$F_i(x, y) = \frac{\beta_i}{I_i(x, y)}. \tag{6.1}$$

This effectively turns the images into 'flatfields', although still containing stars and other objects. For each pair of (x, y), their median is computed according to Definition 5.3, and the resulting flatfield $\widetilde{F}(x, y)$ comprising of the $x \cdot y$ medians is applied back on the image set as usual.

6.2 Kappa–sigma clipping flatfield

Kappa–sigma clipping, described in Section 5.2.6, is an elaborate version of median filter from Section 6.1, employing mean and iterative process to get around the most profound weaknesses of median filter. The filter was programmed into *Blink*. As with the median filter, it was not used to stack an

image set, but to construct an artificial flatfield.

Taking up the normalized images $F_i(x, y)$ and their median image $\widetilde{F}(x, y)$ from the previous section, for each pixel $[x, y]$ we proceed by the equations of Section 5.2.6, where the median is $e_k(x, y) = \widetilde{F}(x, y)$ and $\kappa = 2$. The iteration is stopped after 5 repetitions or when the mean's 4th decimal place does not change between the iterations, and the resulting mean is $e_{k+1}(x, y) = D(x, y)$.

The flatfield $D(x, y)$, comprised of the iterated means, is then applied on the image set as usual.

6.3 SMin flatfield

The previous two methods tend to leave 'shadows' of the brightest stars in the flatfield data. A possible solution is to construct the flatfield not from a median or its variations, but take the minimal values instead. This filter creates an image from minimal values on each pixel in the image set, which is sometimes used to achieve defringing of digital images. The correction image created by SMIN filter has included noise equivalent to one image of the series. The method was adapted into *Blink* as follows.

As with Median flatfield in Section 6.1, the images of the set $I_i(x, y)$, where $i = (1, .., N)$, have to be normalized into $F_i(x, y)$. *Background* values β_i are computed according to equations 5.7 to 5.9 and the image is normalized using Equation 6.1.

Again, this effectively turns the images into 'flatfields' containing stars. The final flatfield is created from maximal values on each pixel (x, y) of the $F_i(x, y)$ set, corresponding to the minimal values of $I_i(x, y)$, by equation

$$D(x, y) = \max_{\forall i} F_i(x, y).$$

The flatfield $D(x, y)$ is applied back on the image set as usual.

6.4 α-quantile flatfield

This is a filter the author came up with to create an artificial flatfield for any single CCD image of mostly stellar field. It is not needed to work a set of images or take any calibration images. Originally, the intent was to improve background estimate for *Astrometry* program, by working with local background value instead of global background value as per Section 5.4.1. We could compute background in a given window, and stack these windows beside each other to create an image representation of the computed backgrounds, which in turn could be used as an artificial flatfield. The smaller the background windows, the finer details it can preserve, but the more affected by the stars it becomes. There is also going to be intensity jumps at the edges between the windows. All this can be resolved by designing a filter instead, and computing background value for each pixel from a handful of its neighbors.

Let $I(x, y)$ be a CCD image with measured intensities. The filter's mask the size of $m \times m$, where m is odd, is centered on a pixel $I(x, y)$ and determines the part of image from $I(x - \frac{m-1}{2}, y - \frac{m-1}{2})$ to $I(x + \frac{m-1}{2}, y + \frac{m-1}{2})$ containing $N = m \cdot m$ image pixels. These are arranged from smallest to greatest into a sorted list $I_{(n)}$, where $n = (1, 2, \ldots, N)$,

$$I_{(1)} \leq I_{(2)} \leq \cdots \leq I_{(N)}.$$

The α-quantile $\widetilde{I}(x, y)$ lies between 0 and 1 and is defined as such a value, that $\alpha \cdot 100$ percent values in the sorted list $I_{(n)}$ falls under $\widetilde{I}(x, y)$. It is determined as

$$\widetilde{I}(x, y) = I_{(\alpha(N+1))},$$

or its closest index. For $\alpha = 0.5$, the $\widetilde{I}(x, y)$ is median, for $\alpha = 0.25$, the $\widetilde{I}(x, y)$ is first quartile. The user chooses α, for a stellar field it should be suitable to set $\alpha \in \langle 0.25, 0.5 \rangle$.

The computation is carried out for all pixels (x, y), and the resulting α-quantile intensities $\widetilde{I}(x, y)$ form an artificial background image, which is then turned into flatfield by Equation 6.1 and applied on the original image $I(x, y)$.

The second controllable parameter of this method is the size of the filter $m \times m$. It should be big enough for the stars to constitute less than half (in case of median when α=0.5) of the filter's pixels, but not too large, since smaller background variations will get omitted with too big a mask.

The method was programmed into *Blink* and various α and m were tested for performance.

6.5 Unsharp masking filter

This image sharpening technique is commonly used in astronomy to bring out the details in the image. Unsharp masking, described in Section 5.3.2 was programmed into *Blink*. The low-pass filters chosen were Gaussian Blur (kernel 7×7) and Box Blur (kernel 7×7), since Gaussian kernel converges to Box kernel after a few passes. The filter strength parameter α was set to $\alpha = 1$.

6.6 Savitzky–Golay filter

This filter, called *Savitzky–Golay*[25], but also *least squares*[20] or *DISPO* (Digital Smoothing Polynomials)[29], is widely used to smooth the data in analytical chemistry. Even though it has an exceptional features, it is rarely used in image processing. The smoothing increases SNR without greatly distorting the signal. Savitzky–Golay filter works with single image.

The filter locally fits a low degree polynomial through a subset of the data by linear least squares method, one data point at time: each data point is replaced with a new value obtained from a polynomial fit to $2n+1$ neighboring points (including the point to be smoothed), with $n >=$ than the order of the polynomial. In general case, this would be achieved by using NURBS (NonUniform Rational B-Spline) fitting, which would be computationally burdensome with over 16 million pixels of our CCD images.

Savitzky–Golay filter derives from a particular formulation of the data smoothing problem. When the data points are equally spaced, which holds true for a CCD image, an analytical solution to the least squares equations can be found[19] in a form of a set of *convolution coefficients*, which enormously simplifies its application. The convolution coefficients can be computed ahead of time and then applied on the data in a form of a *weighted moving average* filter, where the weights are the precomputed convolution coefficients.

The method was popularized by Savitzky and Golay in [25], where they published tables of the convolution coefficients[1], and was later extended for the treatment of 2D and 3D data. The filter got into widespread use after being included in [22]. Although it is originally used for smoothing, the filter's second derivative can be used for baseline flattening.

Let us first derive the analytical solution. A digital filter is applied to a series of N equally spaced data values $f_i \equiv f(t_i)$, where $t_i \equiv t_0 + i\Delta$ for a constant sample spacing Δ, $i \in \{1, \ldots, N\}$ and $n = \ldots -2, -1, 0, 1, 2, \ldots$. The simplest type of digital filter (the nonrecursive or finite impulse response filter) replaces each data value f_i by a linear combination of itself and a number of nearby neighbors,

$$g_i = \sum_{n=-n_L}^{n_R} c_n f_{i+n}, \tag{6.2}$$

where n_L is the number of points used 'to the left' of a data point i, meaning earlier in case of time series, while the number n_R is the number used to the right, meaning later for time series.

Let us consider a filter with fixed $n_L = n_R = \frac{m-1}{2}$, where m is the number of fitted points, and where we also exclude the $\frac{m-1}{2}$ edges[2] of the dataset by taking only $\{i \mid n_L < i < N - n_R\}$. Let's take each g_i as an average of the data points from f_{i-n_L} to f_{i+n_R} and the constant $c_n = 1/m$. This filter would

[1] Some errors of the original paper are corrected in [26]

[2] these excluded edges usually are either cropped off the result or the data can be extended by assuming an edge of constant value around the dataset or 'mirroring' the data beyond the edge.

be called *moving average filter* and is also a special case of Savitzky–Golay filter, where the data is fitted with a straight line. If the underlying function is constant, or linearly changing (increasing or decreasing), then no bias is introduced into the result. Higher points on one end of the averaging interval are balanced by lower points on the other end in the resulting average. However, if the underlying function has nonzero second derivative, a bias is introduced.

The idea of Savitzky–Golay filtering is to find filter coefficients c_n, that approximate the underlying function not by a constant (whose estimate is the average), but by a polynomial of higher order: for each point f_i we do a least squares fit with a polynomial to all $m = n_L + n_R + 1$ points in the moving window, and then set g_i to be the value of that polynomial at position i. Then we move on to a point f_{i+1} and do anew the least squares fitting using the shifted window.

Since the process of least squares fitting involves only a linear matrix inversion, the coefficients of a fitted polynomial are themselves linear in the values of the data. That means we can do all the fitting in advance for a given polynomial degree and window size, using a fictional data consisting of all zeros except for a single 1, and then do all the fitting with real data just by taking linear combinations of the result. That is the key advantage of Savitzky–Golay filter: there exists a particular set of filter coefficients c_n for which Equation 6.2 'automatically' accomplishes the process of least squares fitting inside a moving window.

To derive coefficients c_n, consider how g_0 might be obtained: we want to fit a polynomial of degree M in z, namely $\mathbf{a} = a_0 + a_1 z + \cdots + a_M z^M$ to the values $f_{-n_L}, \ldots, f_{-n_R}$. Then g_0 will be the value of that polynomial at $z = 0$, namely a_0. The design matrix for his problem as per Section 3.3 is

$$A_{zj} = z^j \qquad z = -n_L, \ldots, n_R, \quad j = 0, \ldots, M,$$

and the normal equations for the vector of a_j's in terms of the vector of f_z's

is in matrix notation

$$(\mathbf{A}^\top \cdot \mathbf{A}) \cdot \mathbf{a} = \mathbf{A}^\top \cdot \mathbf{f} \qquad \text{or} \qquad \mathbf{a} = (\mathbf{A}^\top \cdot \mathbf{A})^{-1} \cdot (\mathbf{A}^\top \cdot \mathbf{f})$$

We also have the specific forms

$$\left\{\mathbf{A}^\top \cdot \mathbf{A}\right\}_{zj} = \sum_{n=-n_L}^{n_R} A_{nz} A_{nj} = \sum_{n=-n_L}^{n_R} n^{z+j}$$

and

$$\left\{\mathbf{A}^\top \cdot \mathbf{f}\right\}_{j} = \sum_{n=-n_L}^{n_R} A_{nj} f_n = \sum_{n=-n_L}^{n_R} n^{j} f_n.$$

Since the coefficient c_n is the component a_0 when $\mathbf{f}$ is replaced by the unit vector $\mathbf{e}_n$, $-n_L \leq n < n_R$, we have

$$c_n = \left\{(\mathbf{A}^\top \cdot \mathbf{A})^{-1} \cdot (\mathbf{A}^\top \cdot \mathbf{e}_n)\right\}_0 = \sum_{m=0}^{M} \left\{(\mathbf{A}^\top \cdot \mathbf{A})^{-1}\right\}_{0m} n^m. \tag{6.3}$$

Note that the Equation 6.3 says that we need only one row of the inverse matrix. Numerically we can obtain it by LU decomposition with only a single backsubstitution.

Let's choose a 5-point cubic polynomial defined $g = a_0 + a_1 z + a_2 z^2 + a_3 z^3$, where $z = (-2, -1, 0, 1, 2)$. Its coefficients are obtained by solving the normal equations of a matrix $\mathbf{a} = (\mathbf{J}^\top \mathbf{J})^{-1} \mathbf{J}^\top \mathbf{f}$ and are

$$\begin{aligned}
a_{0,i} &= \frac{1}{35}(-3f_{i-2} + 12f_{i-1} + 17f_i + 12f_{i+1} - 3f_{i+2}) \\
a_{1,i} &= \frac{1}{12}(1f_{i-2} - 8f_{i-1} + 0f_i + 8f_{i+1} - 1f_{i+2}) \\
a_{2,i} &= \frac{1}{14}(2f_{i-2} - 1f_{i-1} - 2f_i - 1f_{i+1} + 2f_{i+2}) \\
a_{3,i} &= \frac{1}{12}(-1f_{i-2} + 2f_{i-1} + 0f_i - 2f_{i+1} + 1f_{i+2})
\end{aligned} \tag{6.4}$$

The coefficients of f are known as the *convolution coefficients* and are elements of the matrix $\mathbf{C} = (\mathbf{J}^\top \mathbf{J})^{-1} \mathbf{J}^\top$. They were tabulated and published by Savitzky and Golay in [25]. The fraction in a form $\frac{1}{h}$ in the equations is a scaling factor with h being a sum of used coefficients[3] c_n.

[3]Some sources call convolution coefficients directly c_n, while others mean $\frac{c_n}{h}$. Care must be taken to check the author's notation.

In general, adding to Equation 6.2, we get

$$(C \otimes f)_i = g_i = \sum_{n=-n_L}^{n_R} c_n f_{i+n}. \tag{6.5}$$

The numerical derivatives are obtained by differentiating the fitted polynomial g, which means that the derivatives are calculated for the smoothed data curve.

$$\begin{aligned}
g &= a_0 + a_1 z + a_2 z^2 + a_3 z^3 = a_0 && \text{at } z = 0 \\
\frac{\mathrm{d}g}{\mathrm{d}z} &= \frac{1}{h}(a_1 + 2a_2 z + 3a_3 z^2) = \frac{1}{h}a_1 && \text{at } z = 0 \\
\frac{\mathrm{d}^2 g}{\mathrm{d}z^2} &= \frac{1}{h^2}(2a_2 + 6a_3 z) = \frac{2}{h^2}a_2 && \text{at } z = 0 \\
\frac{\mathrm{d}^3 g}{\mathrm{d}z^3} &= \frac{6}{h^3}a_3 && \text{at } z = 0
\end{aligned}$$

and h is a scaling factor.

Filling into these equations the convolution coefficients computed in Equations 6.4, the i-th data point is computed as follows for smoothing, 1st derivative and 2nd derivative:

$$\begin{aligned}
g_i &= \frac{1}{35}(-3f_{i-2} + 12f_{i-1} + 17f_i + 12f_{i+1} - 3f_{i+2}) \\
g'_i &= \frac{1}{12h}(1f_{i-2} - 8f_{i-1} + 0f_i + 8f_{i+1} - 1f_{i+2}) \\
g''_i &= \frac{1}{7h^2}(2f_{i-2} - 1f_{i-1} - 2f_i - 1f_{i+1} + 2f_{i+2})
\end{aligned}$$

We compute g_i (or its derivatives) of every data point f_i in the dataset to obtain the result.

The specific nature of Savitzky–Golay filter assumes that the underlying function can be locally well-fitted by a polynomial. When that is true, as for smooth line profiles not too much narrower than the filter window, Savitzky–Golay filtering is very useful. While the polynomial itself performs smoothing on the dataset, the 1st (and 3rd) derivative locates maxima and minima in

the data, the 2nd derivative can be used for baseline flattening, and 2nd or 4th derivative can enhance the resolution of the data.

Extrapolation into 2D

Savitzky–Golay filter can be extended for use on two-dimensional datasets, such as intensity values in a CCD image, which is composed of rectangular set of pixels. The trick is to transform the rectangular grid into a single row by ordering the indices of the data points. Where the one-dimensional filter coefficients are found by fitting a polynomial in variable z to a set of m data points, the two-dimensional filter coefficients are found by fitting a polynomial in variables v, w to a set of $m \times n$ data points. The process parallels the 1D-fitting as above.

Let's choose a bi-cubic polynomial, with moving window $m \times n$ having $m = 7$, $n = 5$, $v = (-3, -2, \ldots, 2, 3)$, $w = (-2, -1, \ldots, 1, 2)$ and the fitted polynomial is $g = a_{00} + a_{10}v + a_{01w} + a_{20}v^2 + a_{11}vw + a_{02}w^2 + a_{30}v^3 + a_{21}v^2w + a_{12}vw^2 + a_{03}w^3$. There are $m \times n = 35$ data points in the moving window, indexed $d_1, \ldots, d_{35}$,

w \ v	-3	-2	-1	0	1	2	3
-2	d_1	d_2	d_3		...		d_7
-1	d_8	d_9	d_{10}		...		d_{14}
0			...				
1			...				
2	d_{29}	d_{30}	d_{31}		...		d_{35}

which become vector $\mathbf{d} = (d_1 \ldots d_{35})^\top$. The Jacobian has 10 columns, one for each of a_{00} to a_{03}, and 35 rows, one for each pair of w and v values. Each row has the form,

$$J_{row} = 1 \quad v \quad w \quad v^2 \quad vw \quad w^2 \quad v^3 \quad v^2w \quad vw^2 \quad w^3.$$

The convolution coefficients are calculated as

$$\mathbf{C} = (\mathbf{J}^\top \mathbf{J})^{-1} \mathbf{J}^\top$$

The first row of $\mathbf{C}$ contains 35 convolution coefficients, which are multiplied with the 35 data values, respectively, to obtain the polynomial coefficient a_{00}, which is the smoothed value. The convolution coefficients are applied by

$$g = \sum_{i=0}^{p} \sum_{j=0}^{q} a_{i,j} v^i w^j,$$

where $p < m$ and $q < n$, proposed in [28].

Application

Savitzky–Golay filtering for two-dimensional dataset was programmed into *Blink*. Two-dimensional convolution coefficients computed by Chandra Shekhar in [27] were used. Possible practical application of this filter for Kleť Observatory's needs include image smoothing to raise SNR, use of 2nd derivative of the filter for baseline flattening to remove background irregularities, and use of 4th derivative for resolution enhancement in case the observed asteroid happens to be too close to a nearby star to resolve the two apart. The last implementation, although appealing in case of dense stellar fields in the region of milky way on the sky, goes beyond the scope of this work and will not be carried out for now.

Upon choosing the Savitzky–Golay filter in the menu of *Blink*, the user can specify the filter parameters within boundaries of the included coefficient table: fitted polynomial degree M, the size of square filter window $m \times m$ and the order of the derivative. The algorithm finds the corresponding $m \times m$ filter coefficients, a_{00} for smoothing or a_{20} and a_{02} for the second derivative, in included file and applies the filter on the image by equation

$$g_{(x,y)} = \sum_{i=-n_L}^{n_R} \sum_{j=-n_L}^{n_R} C_{(i,j)} f_{(x+i,y+j)},$$

for each (x, y) of the image excluding the $n_L = n_R$ edges to compute the result. For smoothing (0th derivative), the $C = a_{00}$, and for 2nd derivative, partial derivative with respect to x is applied by setting $C = 2a_{20}$, then partial derivative with respect to y is applied on the result by setting $C =$

$2a_{02}$. Since the second derivative flattens the baseline, an extra intensity buffer has to be added to avoid the intensity values oscillating around zero.

Chapter 7

Comparison of performance

The six methods detailed in previous chapter were programmed into *Blink*, an image processing and viewing software routinely used by Kleť Observatory, which the author of this work is developing. Image set of a typical stellar field was processed by these methods and the results were compared at first by human eye in *Blink*, as the observations rely on experienced human observer to identify the observed body, and then in Kleť Observatory's astrometric program *Astrometry*, where an algorithm described in Section 5.4 is employed to process the images for astrometric positions of the observed body.

7.1 Object detection in Blink

The images were taken by 1.06-m KLENOT telescope equipped with CCD camera on 8th June 2014. The set contains 13 images. On the Figure 7.6 we can see section of the the raw image 3 of the set on the top and corresponding sections of the corrected images under it in the order the six methods are described. The full raw image can be found in Appendix C. The settings for *Background* and *Range*, which recomputes the images from the 65 536 intensities recorded by the camera to 256 intensities intended to be viewed by human eye, were kept the same between the processed images for clarity of comparison. Results of each method is discussed, and their side-by-side comparison is included on the Figure 7.6.

Median flatfield

The flatfield was created as a median of the set, as described in Section 6.1, and applied back on the images. We can see dramatic increase in the image's quality on Figure 7.6, however the stars cast shadows on the corrected image. That is caused by the stars walking too slowly (or alternatively, the image set being too short) in the set of images used for the flatfield's creation.

Kappa–sigma clipping flatfield

The flatfield was created by kappa–sigma clipping method described in Section 6.2 from the image set, removing outliers of standard deviation > 2, with maximum 5 iterations allowed, and the resulting flatfield was applied back on the images. On the Figure 7.6 we can see dramatic increase in the image's quality very similar to using median filter, but the stars cast slightly less pronounced shadows on the corrected image than they did with an image corrected by median filter.

SMIN flatfield

The flatfield was created using SMIN method from the image set as per Section 6.3 and applied back on the images. In the Figure 7.6, we can see the stars no longer cast visible shadows on the result. Images treated with SMIN filter retain slightly more gaussian noise. The corrected image is slightly lighter because the minimal intensity chosen to represent a pixel is used as a denominator while creating a flatfield, resulting in lighter image. Usually, the image's contrast would be recomputed for human eye, which was omitted for the sake of more clear comparison between the methods.

α-quantile flatfield

The flatfield was created from a single image using the new method described in Section 6.4, and applied back on the image. The processed image offers very sharp, even scene, however slight remnants of the background noise can

be seen. The stars do not cast any shadow, but brighter stars developed a dark center. As these are too overexposed to be used in final astrometry computation anyway, this effect can be safely disregarded.

The filter's performance is controlled by two parameters, size of the filter $m \times m$, and α, which specifies the quantile of the result. The effect of varying these parameters can be seen on a small cut-out of the image in Figures 7.1 and 7.2. On Figure 7.1, we can see that high m omits smaller background variations in the image, while small m would not cover enough background pixels close to bigger stars and in dense stellar fields. On Figure 7.2, α is varied. High α leaves the stars intact, therefore they are present in the flatfield by which the image is divided and cause dark stains over the stars, while very low α leaves some of the background noise intact.

It was concluded that for the usual stellar field image, the best performance from the observer's point of view are reached for $\alpha = 0.5$, which is median, and $m = 13$.

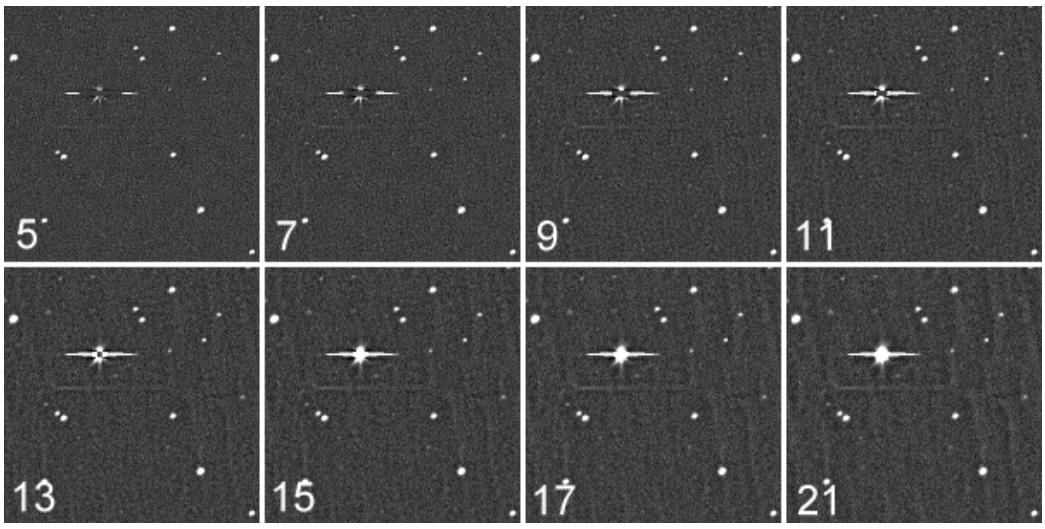


Figure 7.1: Results of parameters $\alpha = 0.5$ and varying m as denoted in the cut-outs.

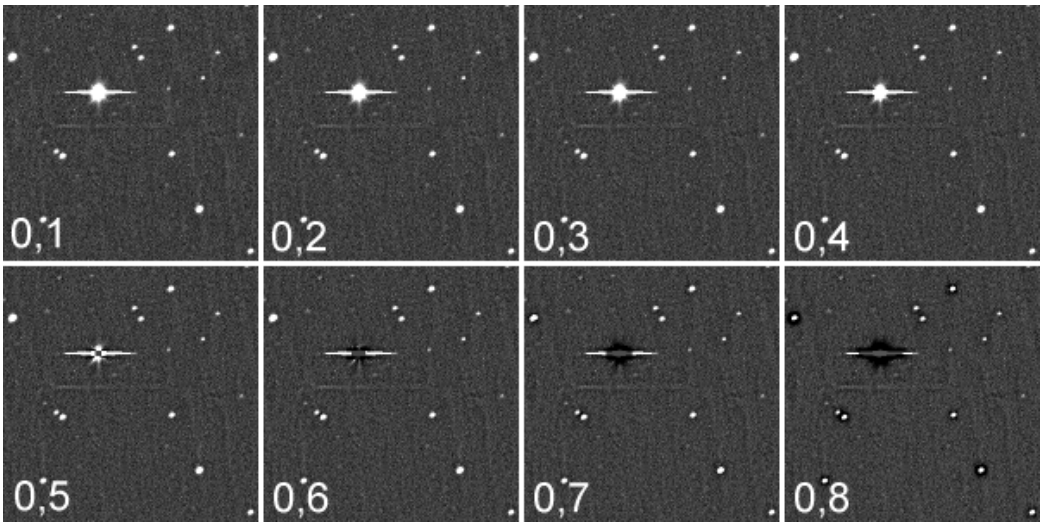


Figure 7.2: Results of varying α as denoted in the cut-outs and $m = 13$.

Unsharp masking filter

Unsharp masking method as described in Section 6.5 was applied on the image. The method works directly with an image without an assistance of flatfield creation. Strength of the filtering is controlled by scaling factor α and the size of affected structures is controlled by the degree of blurring. There could be two possible ways for us to use this filter, which were tested.

1) Heavy blur: A very heavy blur in the first step should make stars disappear, which basically creates an approximate background image, which could be then used to create flatfield instead of carrying out the rest of the unsharp masking procedure. Since the image is blurred, small sized features will not be affected by the image processing.

It shows that not even heavy blur removes the stars due to the CCDs high dynamic range resulting in high difference between the background and the peaks of the stars. In astrophotography, stars would often be removed by hand, which is unacceptable for our purposes.

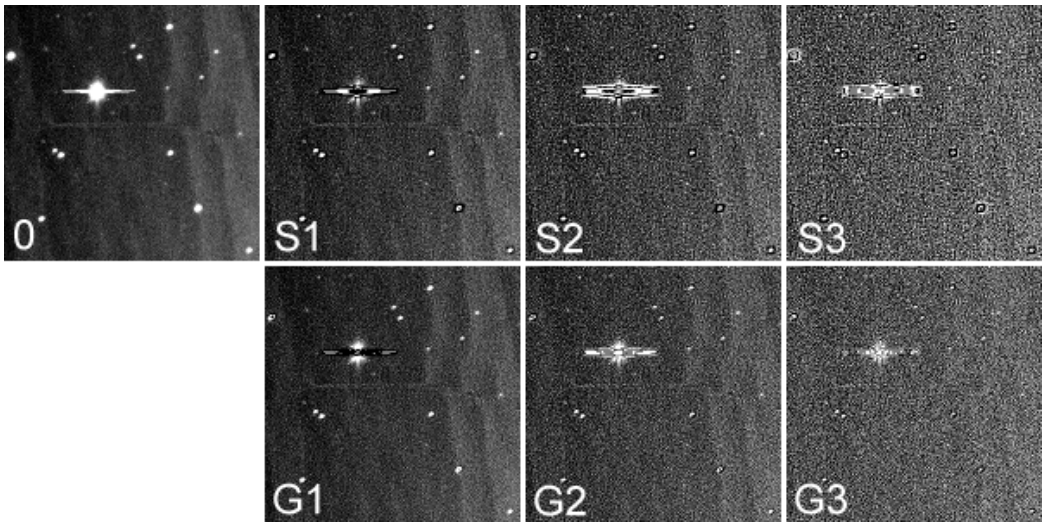


Figure 7.3: Cut-outs of unsharp masking results. O: Original, S1,S2,S3: One to three passes with Box Blur $7x7$ kernel, G1,G2,G3: One to three passes with Gaussian Blur $7x7$ kernel.

2) Small filter kernel: Unsharp masking filter with small kernel blur should affect small structures on the image, effectively enhancing their contrast and thus raising SNR, making the faint asteroids easier to measure.

While the processed image itself may look sharper to human eye overall, closer inspection reveals that it is substantially noisier. The filter enhanced contrast in noise just as well as in the real features. It is a common misuse of this very popular and widely used filter to try to bring faint objects out of the sky's background. Although unsharp masking is globally flux conserving, it does not conserve flux locally on the scale of a few times the smoothing width, which is why, when trying to enhance faint features like asteroids and faint stars out of the background, we might just as well elevate local noise to the status of star-hood.

With a few passes of the filter, as seen in Figure 7.3, it is obvious that while the total flux in the image is conserved, the SNR decreases with each pass. Images suitable for this filter should have high SNR and low contrast scenes with fine details we wish to enhance, which is the direct opposite of

our images. For low SNR and high contrast stellar field, this method is highly unsuitable.

These results illustrate that a popular, widely used method which is recommended often does not guarantee good results. Understanding of the mathematics behind the filter's method is essential to picking correct filter.

Savitzky–Golay filter

This filter works directly on single image without an assistance of flatfield creation. User chooses the fitted polynomial degree M and the size of square filter window $m \times m$. *Blink* finds the corresponding $m \times m$ filter coefficients in included file and applies the filter on the image.

The smoothing of the image with Savitzky–Golay filter looks promising on Figure 7.4. Although it does not remove the background, the image is noticeably less noisy without visible loss of detail. The 2nd derivative, depicted on Figure 7.5, easily removes the background, but has adverse effect on stellar objects, which would interfere with astrometric measurements, due to steep intensity increases in stellar profiles. Additional adjustments may be implemented, for example to apply logarithmic scaling before and after the filter, which are left to future research.

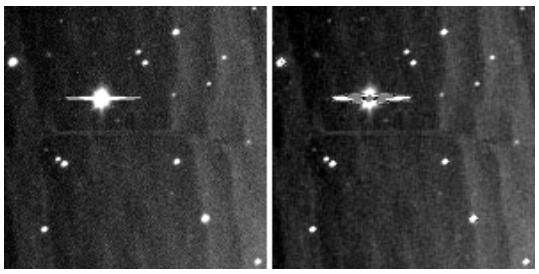


Figure 7.4: Cut-outs of a CCD image and its equivalent smoothed by Savitzky–Golay filter.

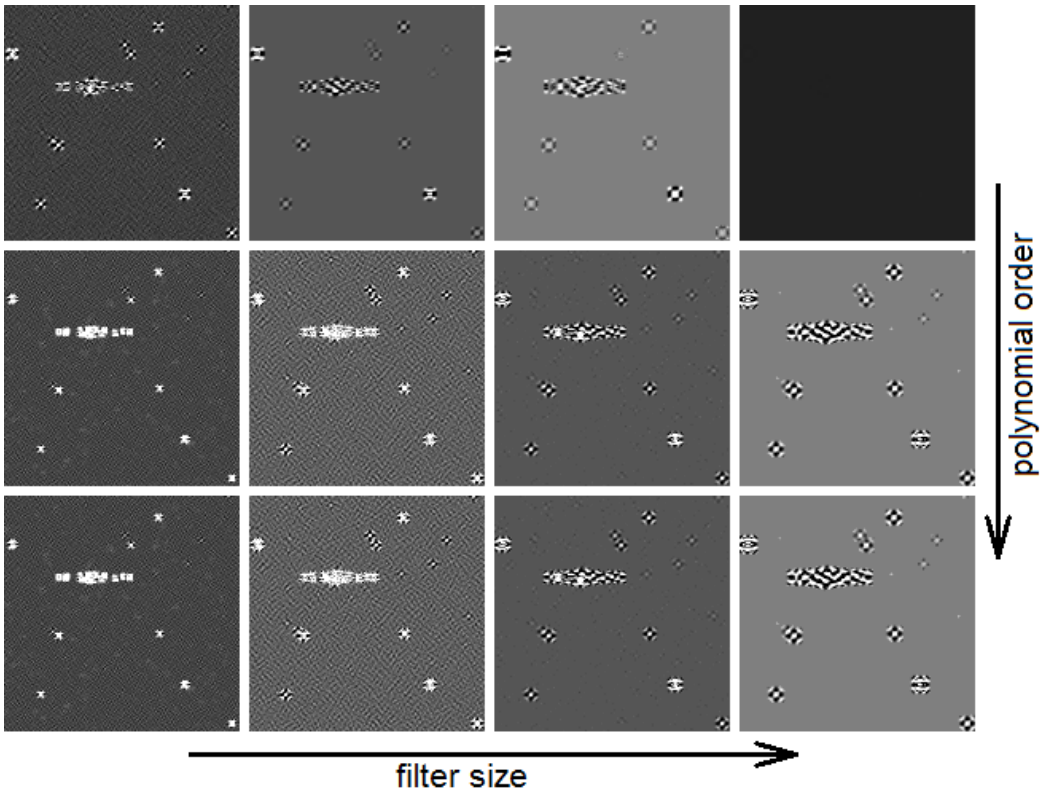


Figure 7.5: Cut-outs of images with applied 2nd derivative of Savitzky–Golay filter. The filter sizes used are 5, 7, 9, and 11 pixels, the fitted polynomials were of 3rd, 4th and 5th order.

Side-by-Side comparison

The adjusted images were set side by side for better comparison on Figure 7.6. The first three filters work with an image set and offer good background removal, although SMIN filter seems the best.α-quantile filter is slightly worse, but can be used with a single image. Unsharp masking filter only adds more noise, and Savitzky–Golay filter evens the background sacrificing the ability to perform the astrometric measurement altogether.

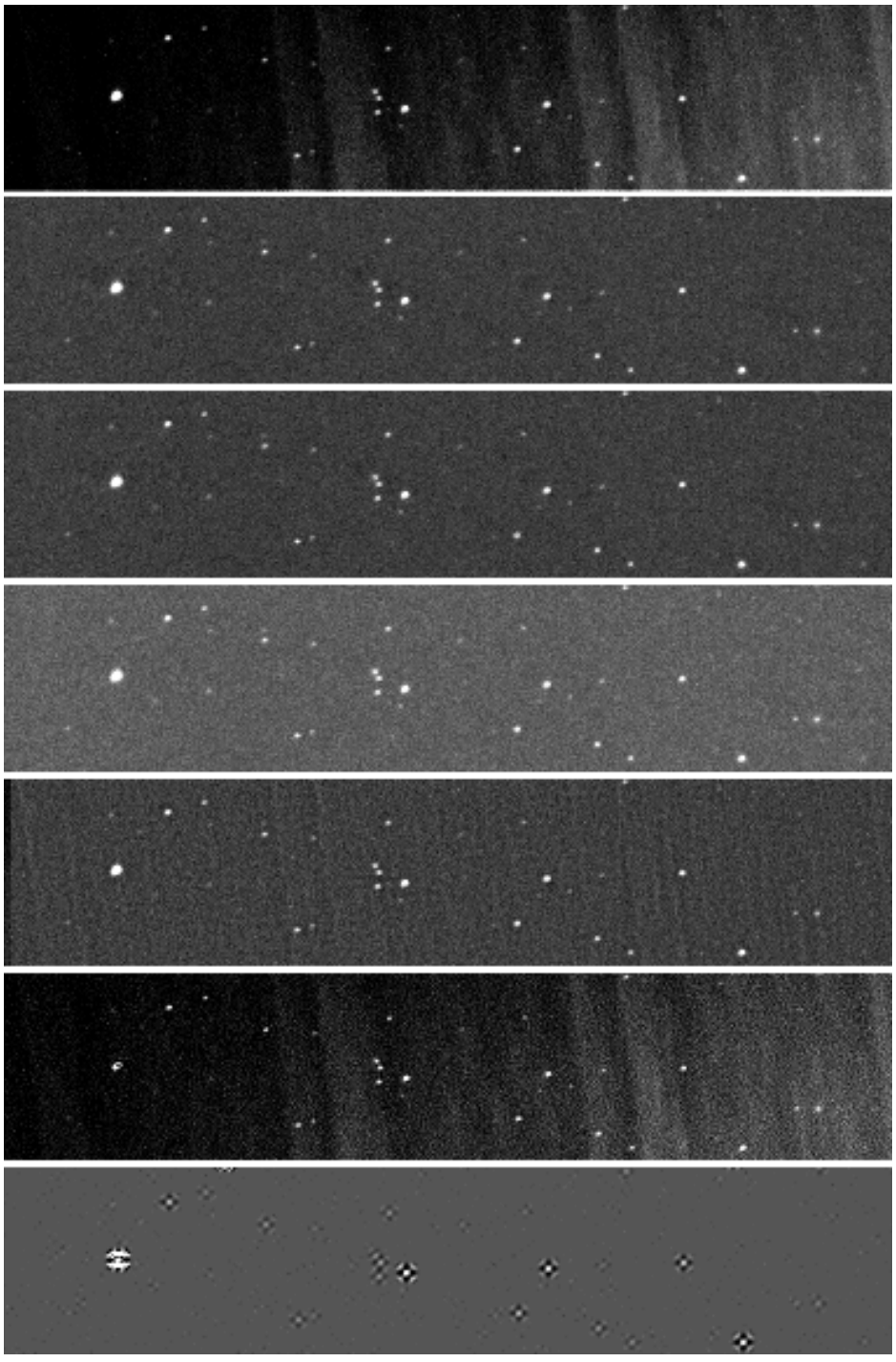


Figure 7.6: Comparison, from top to bottom: Raw image, Median flatfield, Kappa–sigma clipping flatfield, SMIN flatfield, α-quantile flatfield ($\alpha = 0.5$, $m = 13$), Unsharp Masking filter (Gaussian kernel $7x7$), Savitzky–Golay 2nd derivative filter ($M = 4$, $m = 9$)

7.2 Object measurement in Astrometry

The adjusted images were saved and opened in *Astrometry* software, described in Section 5.4, for carrying out the astrometric measurements. Using the same parameters settings (BoxSize 5, SNR for Display/Reduction 2.5, SNR for Identification 6.0), images were reduced to detected objects and displayed.

The reduction of original image to objects, as per Section 5.4.2, along with the reduction of the images processed by the chosen methods, are displayed in the Figure 7.7. Since only four of the six methods were able to provide enhancement of the image, only these four are depicted. The yellow dots represent objects which will enter identification algorithm, the gray dots represent the rest of the detected objects. The red dots stand for overexposed objects and the blue dots for diffuse objects.

Let us compare the first three methods, which employ a set of images for background flattening. They all offer drastically better object detection than the unprocessed image, as well as enhancing their SNR and, consequently, positional accuracy. **Kappa–sigma clipping flatfield** seems to perform slightly better than **Median flatfield**, and **SMIN flatfield** performs slightly worse. Choosing a very faint object at random, it shows its SNR is undetectable at the raw image (SNR < 0.7 results in range check error due to too many possible objects), while with median filter its SNR is 4.9, with kappa–sigma clipping flatfield it is 5.3 and with SMIN flatfield 4.4.

Let's inspect astrometric performance of **α-quantile flatfield**, which requires only single image. α-quantile flatfield with α=0.5, $m = 13$ offers a very even result across the whole field of view, as seen in Figure 7.7, although not as good as previous methods, which can be attributed to the fact we are dealing only with one image instead of a whole set. The same faint object as before now has SNR 3.5. If we vary the parameters of the method as in Table 7.1, we can see the results are more dependent on the size of the mask $m \times m$ than on α.

$\alpha = 0.5$, m is varied								
m	5	7	9	11	13	15	17	21
SNR	2.5	3.0	3.3	3.5	3.5	3.6	3.6	3.6
$m = 13$, α is varied								
α	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8
SNR	3.4	3.5	3.5	3.5	3.5	3.6	3.5	3.3

Table 7.1: An effect of varying parameters m and α on SNR of the chosen faint object.

The detection interface of *Astrometry* for varying m can be found in Appendix D.

Now for the last two methods, which are not included in Figure 7.7. **Unsharp masking filter** enhanced contrast of the noise, making *Astrometry* run into 'Range check error' already at 1.2 SNR, and the chosen faint object is impossible to measure on the image. **Savitzky–Golay filter**'s 2nd derivative is out of question due to inability to discern faint objects altogether.

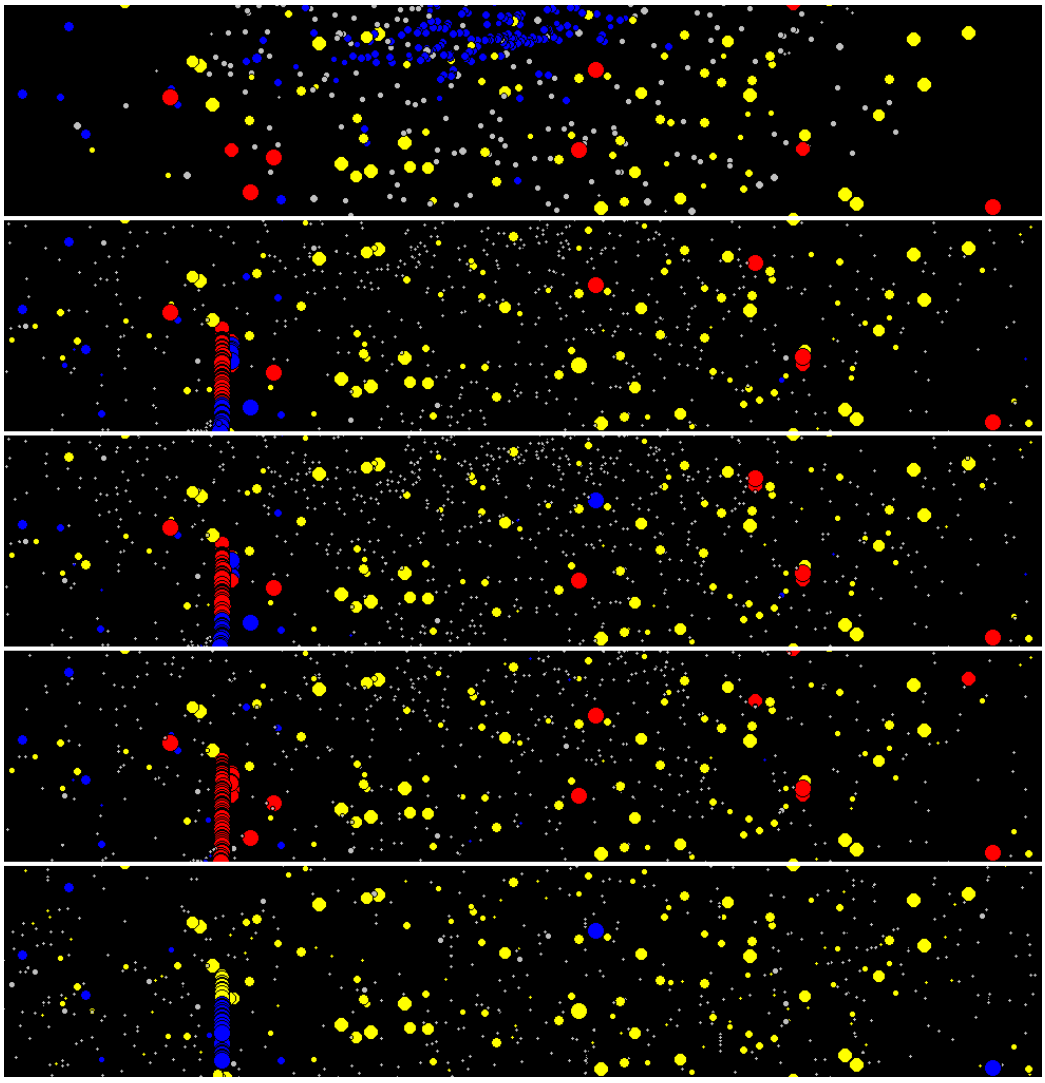


Figure 7.7: Comparison of images' astrometry. Image cut-outs from top to bottom: Raw image, Median flatfield, Kappa–sigma clipping flatfield, SMIN flatfield, α-quantile flatfield ($\alpha = 0.5$, $m = 13$).

7.3 Recommendations for use

The use of these methods newly programmed into *Blink* is discussed in this section to serve as guidance for the observers using the program. Four of the tested methods mean a huge leap in quality — or even possibility — of the astrometric measurement for KLENOT project, although there are conditions which make one or another more useful.

It is suggested to employ **Savitzky–Golay** smoothing implemented in 2D for smoothing the image intended for object detection by human sight, as it preserves the contrast differences better than the simple blur used in its place now.

For an image set where the stellar field shifts slightly between the images, it is recommended to use my customized **kappa–sigma clipping flatfield** to create an artificial flatfield for the set. In the case when the stellar 'shadows' are too pronounced while using this method, it is advised to use **SMIN flatfield** instead.

For a single image or stars moving too slowly across the field of view, it is recommended to use the newly devised **α-quantile flatfield** with default $\alpha = 0.5$, $m = 13$ for creation of artificial flatfield. Its superior advantage is the ability to generate artificial flatfield from a single image.

The rest of the examined methods are not suitable for our use in astrometry of faint asteroids. **Median flatfield** is replaced by better kappa–sigma clipping flatfield, and neither **unsharp masking filter** or 2nd derivative of **Savitzky–Golay filter** were able to provide utilizable results.

Blink also provides an option to save the modified images, which are then in turn loaded into *Astrometry* to carry out the astrometric measurement itself.

Chapter 8

First results: Interstellar body 1I/2017 U1 'Oumuamua

The small object with extreme speed was discovered on 19th October 2017 by Pan-STARRS and originally designated as comet `C/2017 U1` due to its atypical orbit. No cometary features were detected and the designation was changed to inactive comet `A/2017 U1`. Excentricity of its orbit was $e = 1.18$, which corresponds to hyperbolic orbit, placing the origin of the object outside the Solar System. The object's designation was changed again and `1I/2017 U1`, named `'Oumuamua` — meaning 'Scout' in Hawaiian, became the first member of newly created category of Interstellar objects.

The body is elongated, with size estimated to be about 230 by 35 meters, and composed of metal-rich rock, with dark red hue at the surface associated with longterm exposition to cosmic radiation. It flew past the Sun and left the Solar System in the direction of Pegasus constellation.

1.06-m KLENOT telescope observed the object on 19th October 2017, when ten 5-second exposures were taken. The experienced measurer M. Tichý was unable to process the images of the fast, faint body for astrometric positions, as it was 20.4 mag (V band) in the time of observation.

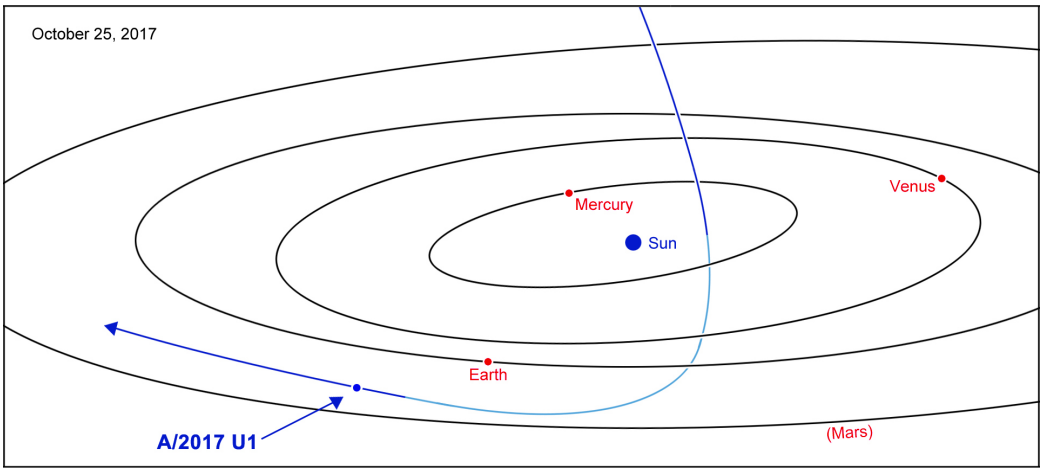


Figure 8.1: Orbit of the interstellar object `1I/2017 U1 ('Oumuamua)`.
Source: nasa.gov

The images were treated with SMIN method the author of this work programmed into *Blink*. It was then possible to measure the object's position in *Astrometry*. The astrometric measurements of 1.06-m KLENOT telescope were published in [5] together with the object's astrometry from 2.4-m telescope in New Mexico, USA, and 5.0-m telescope on Mt. Palomar, USA.

Chapter 9

Conclusion

The doctoral thesis *Numerical Methods of Image Analysis in Astrometry* gives a systematic overview of astrometric measurement of small Solar System bodies. The history of astrometry is briefly described, being followed by the astrometric process as done in present, the considerations and requirements it demands. The necessary mathematical background for following parts is given.

As the aim of this work is related to Kleť Observatory's astrometric needs, the overview of the hardware, software and the observing process of the observatory is given. Image processing and calibration is described in detail, followed by the usual image combining methods and basic filters in use. The astrometric algorithm of the Kleť Observatory is detailed.

The aim of the thesis is 'to find a suitable mathematical method to even out the background of a CCD image without affecting the small irregularities containing combined noise and data to achieve increase in both quality and quantity of astrometric measurements on Kleť Observatory'[6]. Total six promising methods were chosen to be programmed into *Blink*, software for image processing and viewing in routine use by Kleť Observatory, which the author of this work is developing. They were tested, compared, and recommendations for their use were assembled for the observer to follow. The new image processing routines were immediately used and al-

lowed astrometry of the first, and so far the only known, interstellar object 1I/2017 U1 ('Oumuamua)[5].

The doctoral thesis fulfills the goals set in the treatise as written above. Four from the six methods were successful in evening out the background of the CCD images without tampering with its high frequency components. Median flatfield and Kappa–sigma clipping flatfield make use of the stellar field's drift between images of the set, while SMIN flatfield can be applied even on static set. α-quantile flatfield requires only single image to perform its task. These methods allow for astrometric measurements to be carried out even in unfavorable conditions, e.g. at areas where the raw image was darker or brighter than the average background for the image, or pronounced background intensity gradient was present, where astrometry would be impossible before. Additionally, flattening the background by these methods raises the SNR of the objects, thus it is recommended to apply these methods even in the cases when the astrometry can be performed on the raw images, to raise the precision of the result. Not only these methods can be applied on new observations, but they can be also utilized on an extensive archive of CCD images for precovery use.[3]

Furthermore, new options of enhancing the astrometry of Kleť Observatory were revealed and will be subject to additional research. Namely, the artificial flatfield creation using the new filters could be automatized, and the astrometric algorithm itself could be partially re-coded for better performance. Savitzky–Golay filter can be implemented for more effective smoothing of the image for the observer's sight. Its 2nd derivative could be further developed for background removal, and its 4th derivative may have promising application for resolution enhancement to resolve asteroid nearby a star or in a dense stellar field.

Bibliography

Author's publications

[1] TICHÁ, J., TICHÝ, M., KOČER, M., HONKOVÁ, M., KLENOT PROJECT 2002–2008 contribution to NEO astrometric follow-up, *Meteorit. Planet. Sci.*, 2009, vol. 44, Issue 12, pp. 1889–1895. doi: 10.1111/j.1945-5100.2009.tb01998.x

[2] TICHÝ, M., HONKOVÁ, M., TICHÁ, J., KOČER, M., *Inaccuracies Affecting the Calculation of Orbital Elements*, presented at 2013 IAA Planetary Defense Conference, Flagstaff, AZ, USA, 2013. (submitted to Experimental Astronomy).

[3] TICHÝ, M., TICHÁ, J., KOČER, M., HONKOVÁ, M., *TNO Precovery Survey Using the KLENOT Telescope Archive*, presented at Asteroids, Comets, Meteors 2008 held July 14–18, 2008 in Baltimore, Maryland, USA. LPI Contribution No. 1405, paper id. 8067. Bibliographic Code: 2008LPICo1405.8067T

[4] HONKOVÁ, M., TICHÝ, M., TICHÁ, J., KOČER, M., *The Astrometry Algorithm Used at the Kleť Observatory*, presented at 2013 IAA Planetary Defense Conference,Flagstaff, AZ, USA, 2013.

[5] TICHÝ, M., TICHÁ, J., HONKOVÁ, M., BLAGOROD-NOVA, N., YE, Q.Z., ZHANG, Q., PEFFER, K., RYAN, W.H., RYAN, E.V., WILLIAMS, G.V., *A/2017 U1*, Minor Planet Electronic Circ., No. 2017-U265, 2017. Available from `https://minorplanetcenter.net/mpec/K17/K17UQ5.html`

[6] HONKOVÁ, M., *Numerical Methods of Image Analysis in Astrometry*, Doctoral thesis statement, Faculty of Mechanical Engineering, Brno University of Technology, 2014.

[7] HONKOVÁ, M., *Následná astrometrie blízkozemních planetek a její vliv na přesnost určení dráhových elementů a efemerid*, Bachelor's thesis, Faculty of Science, Masaryk University, Brno, 2006. Available from `https://is.muni.cz/th/106774/prif_b/Nasledna_Astrometrie_-_Honkova.pdf`

Books

[8] DICK, S.J., *History of Astronomy, An Encyclopaedia*, John Lankford (Ed.), Garland Publishing Inc., New York & London, pp. 47–60, 1997. ISBN-13: 978-0815303220.

[9] COEXTER, H.S.M., *Introduction to Geometry*. 2nd ed, New York: Wiley, p. 93 and pp. 289–290, 1969.

[10] COEXTER, H.S.M., GREITZER, S.L., *Geometry Revisited*. Washington, DC: Math. Assoc. Amer., pp. 150–153, 1967.

[11] ZVÁRA, K., ŠTĚPÁN, J., *Pravděpodobnost a matematická statistika*. MatfyzPress, 2006. ISBN: 978-80-86732-71-8.

[12] WALL, J.V., JENKINS, C.R., *Practical Statistics for Astronomers*, Cambridge University Press, 2003. ISBN 0-521-45416-9.

[13] FOLLAND, G.B., *Fourier analysis and its applications*. American Mathematical Society, 1992, 433 pp. ISBN 978-0-8218-4790-9.

[14] STEIN, E.M., GUIDO, L.W., *Introduction to Fourier analysis on Euclidean spaces*. Princeton, N.J.: Princeton University Press, 1975, 297 pp. ISBN 06-910-8078-X.

[15] ČÍŽEK, V., *Diskrétní Fourierova transformace a její použití*. 1st ed, Praha, SNTL — Nakladatelství technické literatury, n.p., 1981, 160 pp. Matematický seminář SNTL. ISBN 04-019-81.

[16] BEZVODA, V., JEŽEK, J., SAIC, S., and SEGETCH, K., *Dvojrozměrná diskrétní Fourierova transformace a její použití: I. Teorie a obecné užití.* 1st edition, Praha, 1988, 191 pp. ISBN 17-135-88.

[17] VAN DE KAMP, *Principles of Astrometry*, W. H. Freeman and Company, San Francisco and London, 1967.

[18] KOVALEVSKY, J., SEIDELMANN, P.K., *Fundamentals of Astrometry*, Cambridge University Press, 2004. ISBN 0-521-64216-7.

[19] GUEST, P.G., *Ch. 7: Estimation of Polynomial Coefficients. Numerical Methods of Curve Fitting.* Cambridge University Press, 1961, re-edition in 2012. pp. 147—. ISBN 978-1-107-64695-7.

[20] HAMMING, D.F., *Digital Filters.* 2nd. ed., (Englewood Cliffs, NJ: Prentice Hall), 1983.

[21] GLASBEY, C.A, HORGAN, G.W., *Image Analysis for the biological sciences*, Wiley, 1995. ISBN-13: 978-0471937265.

[22] PRESS, W.H., TEUKOLSKY, S.A., VETTERLING, W.T., FLANNERY, B.P., *Numerical Recipes in C: The Art of Scientific Computing* 2nd ed., Cambridge University Press, ISBN 0-521-43108-5.

Articles

[23] CHESLEY, S.R., BAER, J., MONET, D.G., Treatment of star catalog biases in asteroid astrometric observations. *Icarus.* 210 (2010), pp. 158–181. doi: 10.1016/j.icarus.2010.06.003

[24] SNYDER, J.P., Map Projections - A Working Manual. *U. S. Geological Survey Professional Paper 1395.* Washington, DC: U. S. Government Printing Office, pp. 164–168, 1987. Available at: `https://pubs.usgs.gov/pp/1395/report.pdf`

[25] SAVITZKY, A., GOLAY, M.J.E., Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry.* 36 (8): 1627—39. 1964. doi: 10.1021/ac60214a047.

[26] STEINIER, J., TERMONIA, Y., and DELTOUR, J., Comments on Smoothing and Differentiation of Data by Simplified Least Square Procedure. *Analytical Chemistry*. 44 (1972), pp. 1906–1909. Available at: https://www.people.iup.edu/jford/courses/CHEM421/Resources/CommentsOnSmoothingAndDifferentiationOfDataBySimplifiedLeastSquareProcedure.pdf

[27] SHEKHAR, C., On Simplified Application of Multidimensional Savitzky–Golay Filters and Differentiators, *Progress in Applied Mathematics in Science and Engineering*, 2015, Bali, Indonesia. doi: 10.1063/1.4940262

[28] NIKITAS, P., PAPPA-LUISI, A., Comments on the two-dimensional smoothing of data. *Analytica Chimica Acta*. 415: 117–125, 2000. doi: 10.1016/S0003-2670(00)00861-8

[29] ZIEGLER, H., Properties of Digital Smoothing Polynomial (DISPO) Filters. *Applied Spectroscopy*, vol. 35, pp. 88–92, 1981. doi: 10.1366/0003702814731798

[30] MITCHELL, H.B., MASHKIT, N., Noise smoothing by a fast K-nearest neighbour algorithm. *Signal Processing: Image Communication*, vol. 4, issue 3, pp. 227–232. Signal Processing Computing Division, Elta Electronics Industries Ltd, Ashdod, Israel, 1992. doi: 10.1016/0923-5965(92)90028-E

Other

[31] DRUCKMÜLLEROVÁ, H., *Phase-correlation based image registration*. Master thesis, Faculty of Mechanical Engineering, Brno University of Technology, 2010. Available from https://www.vutbr.cz/www_base/zav_prace_soubor_verejne.php?file_id=27760

[32] *Minor Planet Center*, http://www.minorplanetcenter.net/iau/mpc.html

[33] *U.S. Naval Observatory (USNO)*,
`http://www.usno.navy.mil/USNO/astrometry/optical-IR-prod`

[34] *GCX User's Manual*,
`http://astro.corlan.net/gcx/html/node7.html`

[35] *Tichý, M., personal communication.* Eleanor Helin, NEAT, 1996; Jennifer Evans, LINEAR, 1998; Eric Christensen, CSS, 2013.

Nomenclature

The following tables contain abbreviations and symbols used thorough the thesis. Locally used nomenclature is not included.

Abbreviations

AU	Astronomical Unit
CCD	Charge Coupled Device
FWHM	Full-Width Half Maximum
MPC	Minor Planet Center
NEA	Near-Earth Asteroid
PDF	Probability Density Function
RMS	Root Mean Square
SNR	Signal-to-Noise Ratio

Symbols

$\Im z$	Imaginary part of a complex number z
$\mathbb{C}$	The set of complex numbers
$\mathbb{N}$	The set of natural numbers
$\mathbb{R}$	The set of real numbers
$\mathbb{Z}$	The set of integer numbers
$\mathcal{F}$	Fourier transform
$\mathcal{F}^{-1}$	Inverse Fourier transform
μ	Arithmetic mean or average
$\Re z$	Real part of a complex number z
σ	Standard deviation, or root mean square
$f * g$	Convolution of functions f and g

Appendices

Appendix A
Image from KLENOT telescope with CCD Photometrics S300.

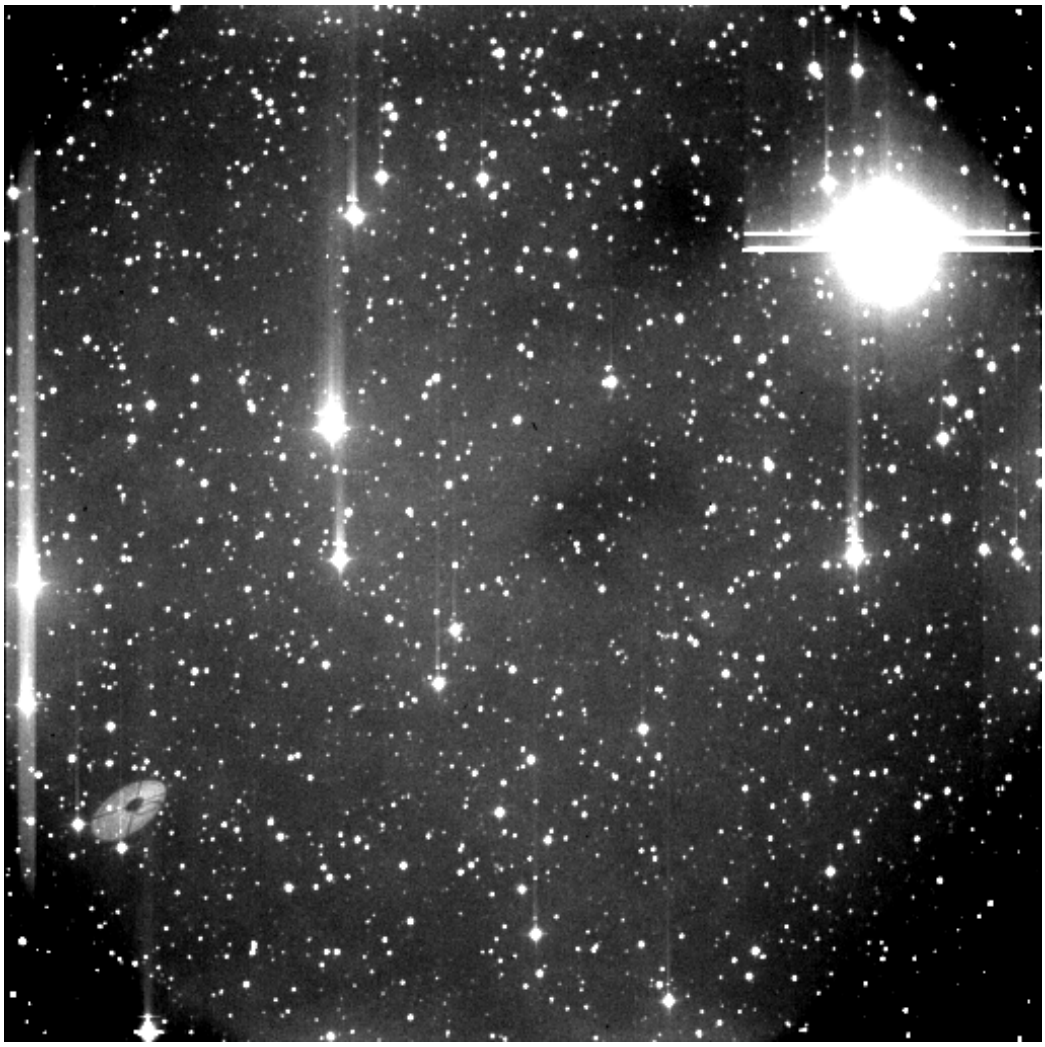
Appendix B
Object detection in *Astrometry* program.

Appendix C
Image from KLENOT telescope with CCD ProLine PL-230.

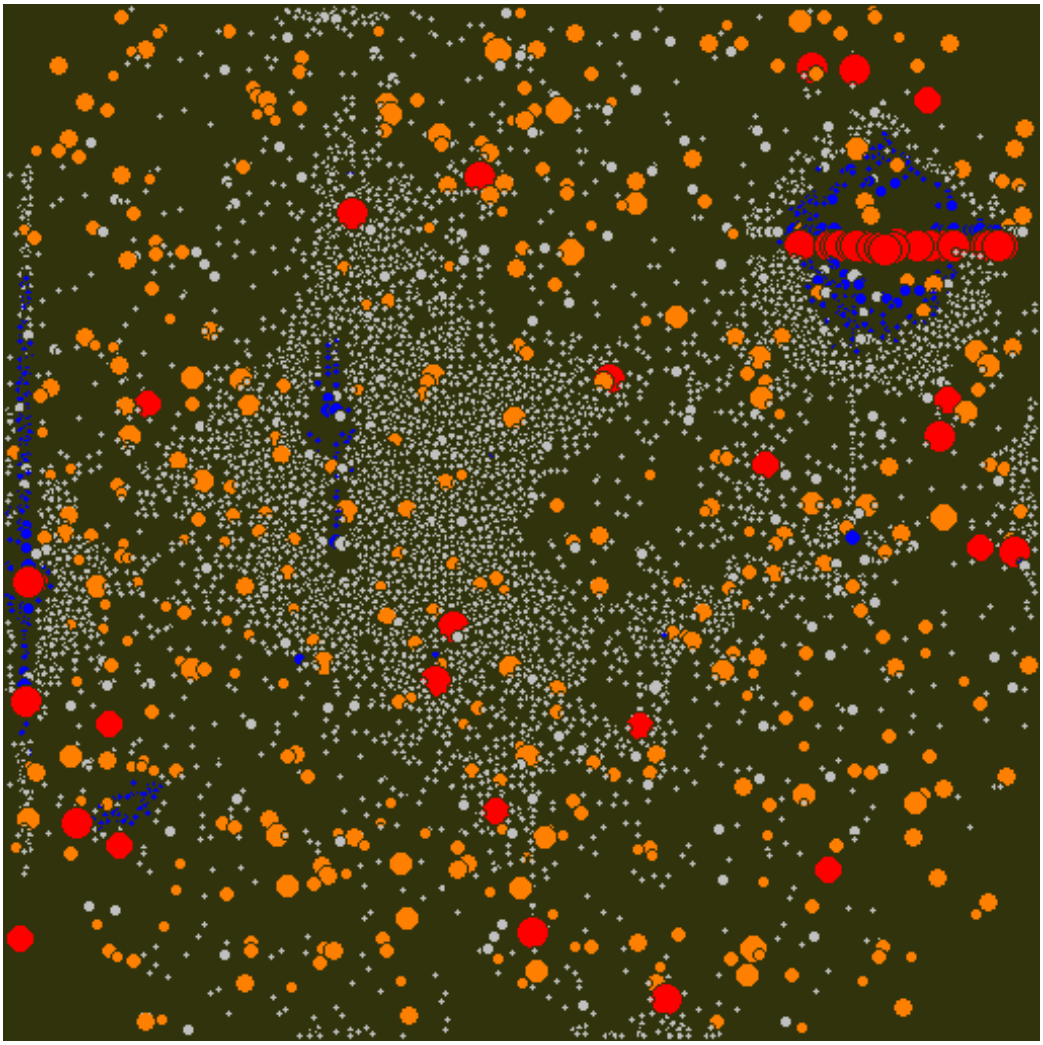
Appendix D
Astrometry of images processed with α-quantile flatfield with different window sizes m.

Appendix E
Interface of *Blink* program developed by the author of this work.

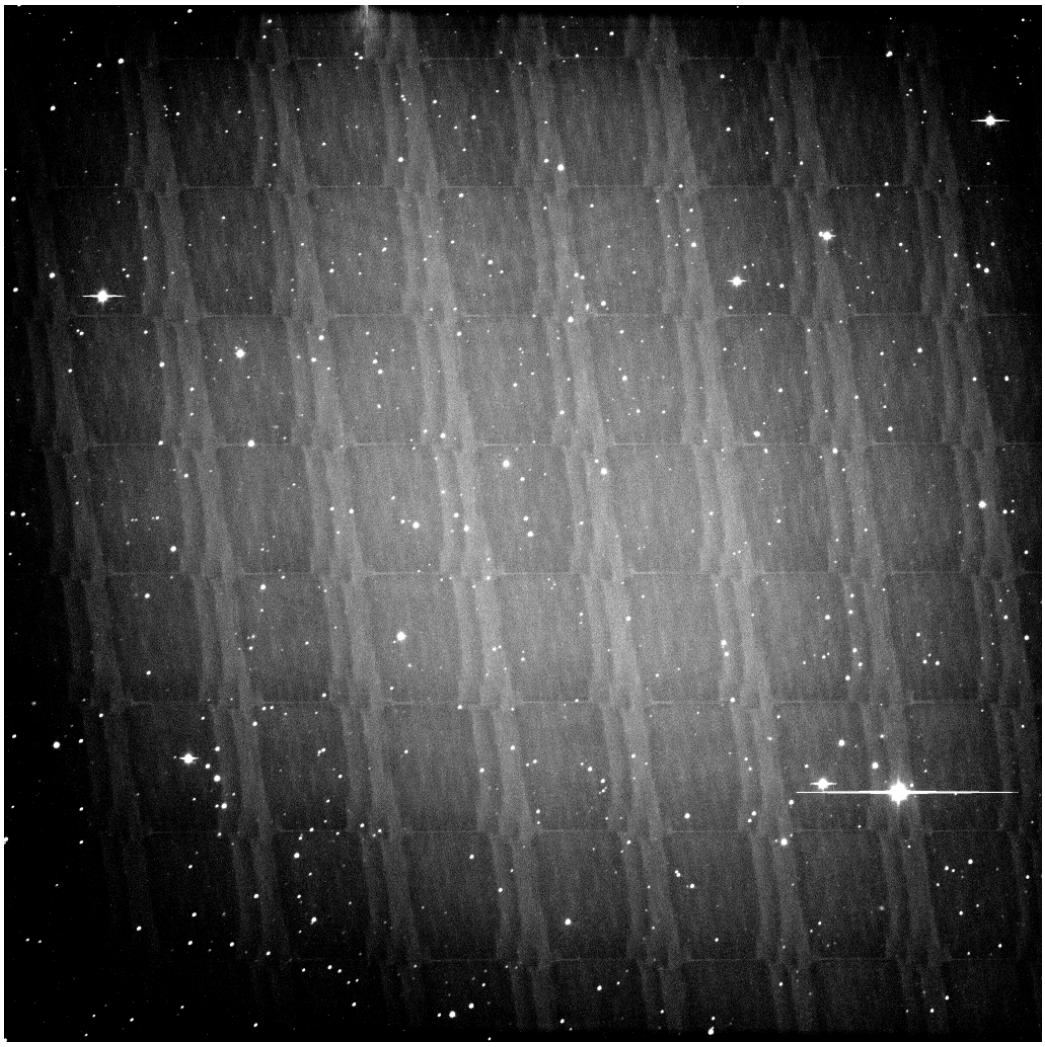
Appendix F
Stellar field identification in *Astrometry* program.



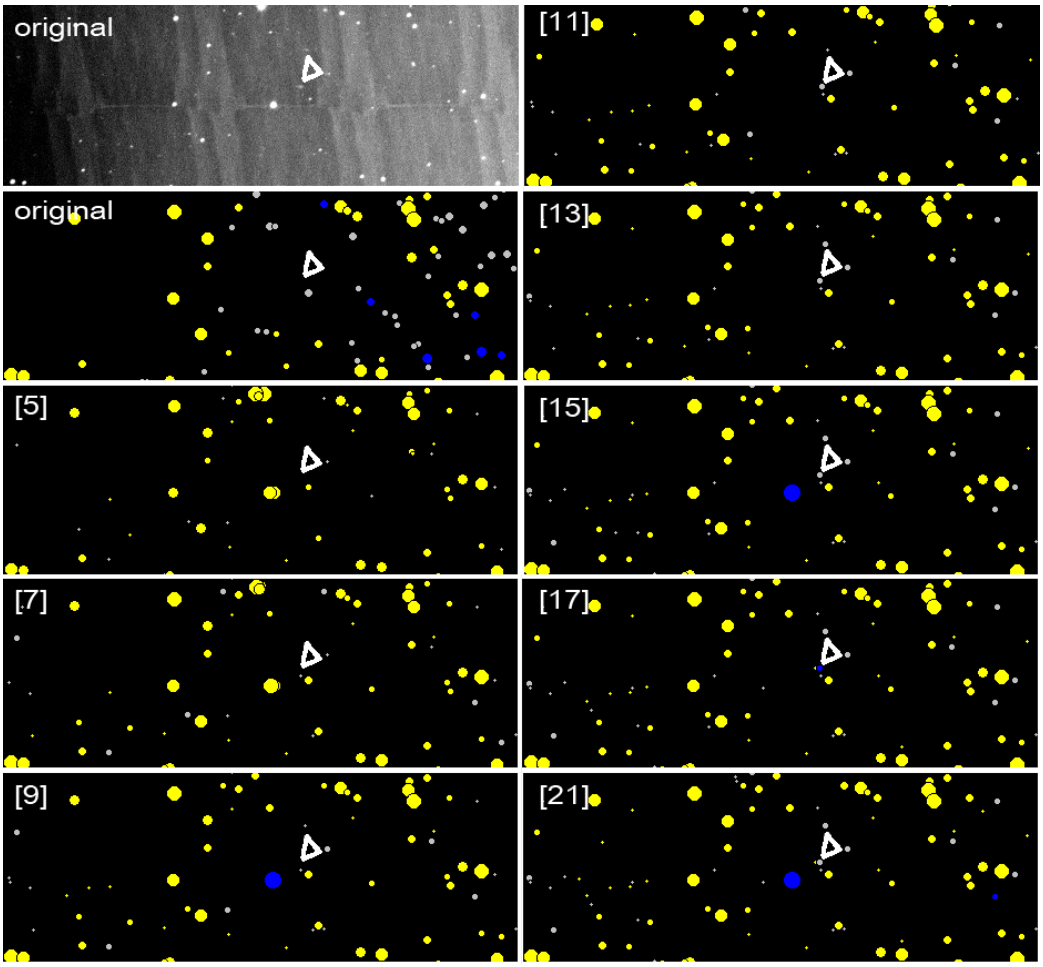
Appendix A: Image taken by KLENOT telescope with CCD Photometrics S300, in use 2002–2012.
Source: Kleť Observatory



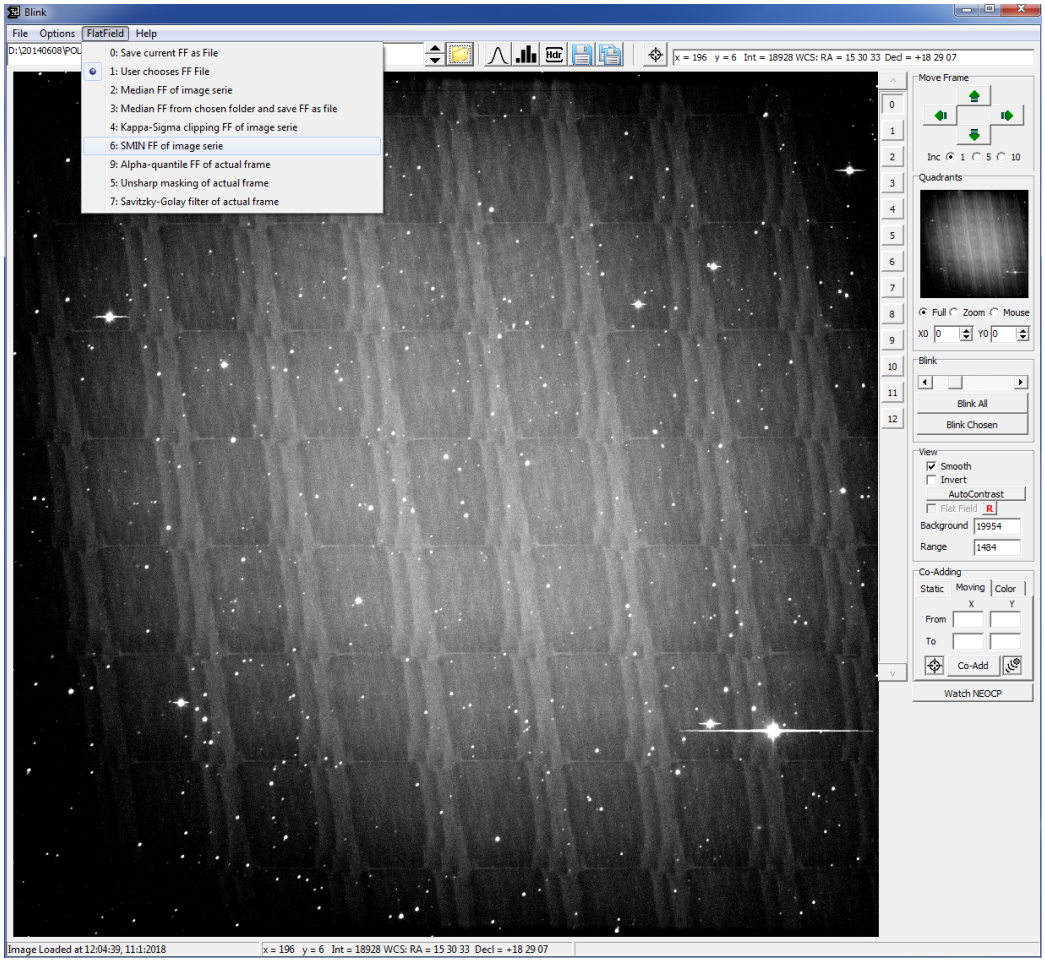
Appendix B: The object detection from the image in previous figure performed by *Astrometry* program. Notice the lack of detections at darker regions (notably, corners) and high number of false detections in brighter regions.
Source: Kleť Observatory



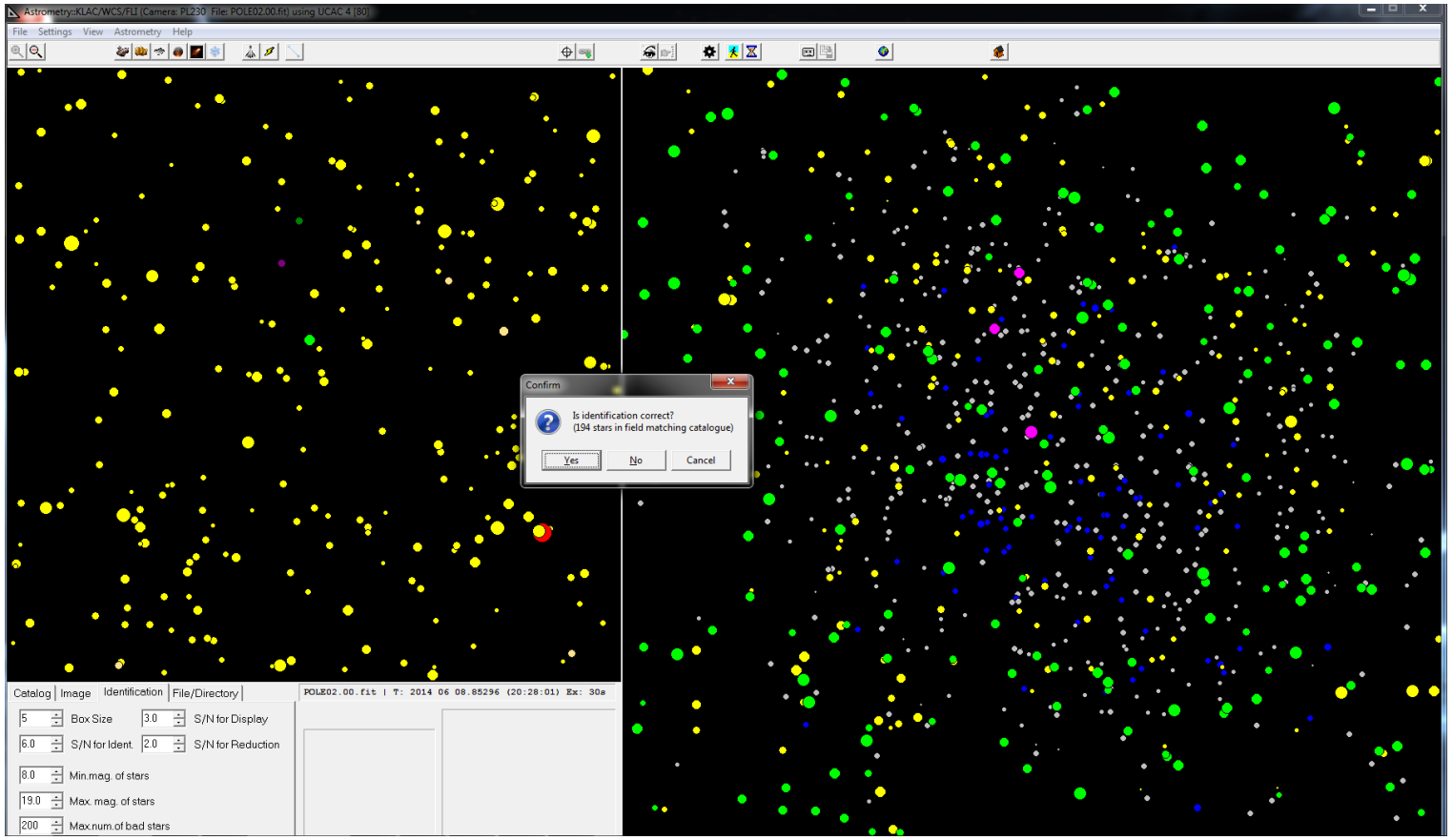
Appendix C: Image taken on 8th June 2014 by KLENOT telescope with CCD ProLine PL-230, used for filter's performance comparison in Chapter 7.
Source: Kleť Observatory



Appendix D: Comparison of astrometry with an image processed by α-quantile flatfield with different filter's sizes [m]. Notice the marked triangle. No dot means no detection by *Astrometry*, dot implies low SNR, small circle better SNR. Blue circle means the program assesses the detection as diffuse object. The best results are given by m 13 and 15.



Appendix E: *Blink* program, developed by the author of this work for Kleť Observatory. The program automatically centers the images (manual centering is also possible), allows animating the image set or its part, zooming to part of the image, smoothing, inverting, stacking the images including stacking on moving object, and other useful functions for the observer. The methods from Chapter 6 were programmed in and can be chosen by the observer to process the images.



Appendix F: Stellar field identification in *Astrometry* program. Left: catalog stars. Left Bottom: parameters of objects' reduction. Right: CCD image reduced to objects. The identification triangle stars are magenta, the identified stars are lime. Yellow are other objects above 'S/N for Ident' threshold which were not matched to stars, gray are objects under that threshold. Blue are diffuse objects. A lot of gray and blue objects are in fact noise, while in the corners stars go undetected. Source: Klet' Observatory