Project PANOPTES: a citizen-scientist exoplanet transit survey using consumer digital cameras

By

Wilfred Gee

A thesis submitted to Macquarie University

for the degree of Master of Research

Department of Physics and Astronomy

July 2017

Examiner's Copy

Typeset in LaTeX 2_ε.

Acknowledgements

A big thank you to my supervisors Dr. Christian Schwab and Dr. David Coutts for helping make the entire thesis happen. It has been an extremely valuable learning experience and a great help to myself and the PANOPTES project. Another giant thanks to the rest of the PANOPTES team, especially Prof. Olivier Guyon, Dr. Nem Jovanovic, and Dr. Josh Walawender for all the help, support, and guidance. Working with Project PANOPTES has allowed me to work with world-class scientists on a fun and engaging project that will not only yield valuable science, but will be a great educational asset to many students and citizen-scientists for years to come.

Thanks to Jen Tong at Google Cloud Platform for supporting the project and providing invaluable resources.

Thanks also to Dr. Lee Spitler and Dr. Anthony Horton for choosing to use POCS for the Huntsman Telephoto Array, which greatly helped to improve the software and also provided me with a chance to work on an additional interesting project.

And, of course, a big old thanks to Krystal for joining me in Australia and helping to keep me sane (while simultaneously driving her a little crazy).

Abstract

Project PANOPTES[1] is aimed at establishing a collaboration between professional astronomers, citizen scientists, and schools in order to discover a large number of exoplanets using the transit technique. An imaging unit based on consumer digital single-lens reflex (DSLR) cameras has been designed to run automatically every night, covering large parts of the sky looking for exoplanet transits. At a cost of $5000 USD, the units are a low-cost and efficient system for wide field astronomical imaging, offering approximately two orders of magnitude better étendue per unit of cost than professional wide-field surveys. Critical to the success of the project, a robust and flexible software control system was developed, as discussed in this work. The properties and capabilities of the DSLR camera as an imaging device are also explored, along with a demonstration of the algorithm required for successful processing and detection. Both science and outreach, our vision is to have dozens of units built by schools and citizen scientists in the next few years, helping make this project the most productive exoplanet discovery machine in the world.

[1] http://www.projectpanoptes.org

Contents

List of Figures

List of Tables

...great and strong Argus [*Panoptes*]*, who with four eyes looks every way. And the goddess stirred in him unwearying strength*: *sleep never fell upon his eyes*; *but he kept sure watch always.*

Aigimios - Hesiod

1 Introduction

PANOPTES Panoptic Astronomical Networked Observatories for a Public Transiting Exoplanets Survey

Ultimately there is one main question that Project PANOPTES aims to answer: what is the most efficient and cost-effective ground-based solution for making a full-sky map of exoplanets that transit nearby stars? Our answer is to deploy a large number of low cost collecting devices to as many geographically disperse locations as possible. While simple on its surface the actual implementation is nevertheless quite difficult to achieve. Advances in consumer-level digital-single-lens-reflect (DSLR) cameras, however, make the solution not only scientifically practical but also affordable, thus opening up exciting possibilities both in terms of scientific output as well as education and outreach.

By creating a reliable and autonomous survey device that is both low-cost and easily reproducible, Project PANOPTES plans to utilize interested citizen-scientists, including high-schools and universities, to help build a world-wide network of imaging units that provide coverage over a large fraction of the sky. These units will come in the form of an easy-to-assemble kit of components that, together with instructions and lesson plans, will allow the participants to construct an autonomous astronomical imaging device (see Figure 1.1).

Project PANOPTES occupies a unique niche in the ongoing search for exoplanets. Previous missions, such as Kepler, have focused primarily on one region of the sky, taking deep images in order to discover transiting exoplanets around mostly distant stars, leading to limited opportunities for imaging follow up as the majority of the targets are far away and dim. Current projects, such as the Hungarian-made Automated Telescopes (HAT), aim to fill this gap by targeting relatively bright and near stars, but continue to utilize "purpose-built heavy-duty mounts and dome enclosures," expensive CCDs, and optimal site locations in order to ensure the success of the project, thus limiting the ability of the project to grow and adapt (Bakos et al., 2004). Similarly, projects like MEarth, which recently

Figure 1.1: PANOPTES unit `PAN001` deployed and operating at the Mauna Loa Observatory (MLO) July 2016.

announced the find of a super-Earth around a nearby M Dwarf, themselves note some of the challenges and limitations associated with single-longitudinal surveys (Dittmann et al., 2017).

Future space-based missions, such as TESS and PLATO, will fill some of these gaps, but still suffer from problems inherent in large-scale dedicated surveys, whether it is a sparser cadence or a smaller field of view, not to mention overall cost (Ricker et al., 2014; Rauer et al., 2013). At approximately $5000 USD per unit and including a $10^\circ \times 15^\circ$ field-of-view (FOV), PANOPTES is at least 100 times more cost-effective (in dollars per square degree per collecting area) than large-scale projects such as the LSST (Guyon & Martinache, 2011).

PANOPTES is attractive because it can complement these missions, doing so cheaply and with the involvement of citizen scientists. And while the primary scientific interest of the PANOPTES project is to aid in the discovery of exoplanets via the transit-detection technique, the unit could also prove useful in the detection and monitoring of other transient phenomena, such as supernovae, near-earth objects (NEO), space debris, and more. It is the hope of the project that schools will come up with novel ways to utilize the unit and the data it gathers.

This thesis aims to prove that it is possible to create an affordable and reliable astronomical imaging unit that can be assembled by citizen scientists and used to aid in the discovery of transiting exoplanets. To accomplish this, a number of specific tasks were outlined:

1. develop a robust software system capable of automated control of an astronomical imaging device, including guaranteed operational safety of the unit,
2. investigate the properties and capabilities of the preferred consumer-level DSLR camera in order to demonstrate exoplanet transit detection, and

3. develop and demonstrate a working transit-detection algorithm capable of overcoming the difficulties associated with the Bayer color filter array in consumer DSLRs.

The result of this thesis work is a well-tested unit that can run autonomously and without interruption, making its own scheduling decisions about what should be observed and when, all while safely and reliably monitoring weather conditions, parking itself when needed and resuming operations without intervention when conditions return to normal. This work has been critical in helping move PANOPTES from the demonstration phase into full-scale beta-testing with third-party volunteers.

This chapter goes on to explain the main motivations that give rise to the project, both technological and social, as well as challenges faced. The complications arising from the use of DSLRs as collecting devices is given particular attention. A brief history of the project is also provided. The exoplanet science behind PANOPTES is given in chapter 2. Chapter 3 describes a PANOPTES unit at the hardware level and outlines performance characteristics of the DSLR cameras. The software architecture and its development is described in chapter 4, including an overview of potential alternatives. Data analysis and algorithm implementation is discussed in chapter 5, again with a particular focus on the complications due to the DSLR. Concluding remarks, including an update on the status of the project and future work, forms the basis of chapter 6.

1.1 Motivations

The desire to watch the entire night sky is one of the primary goals behind PANOPTES. Most information that arrives at earth from distant sources is not captured. Previous generations of astronomy were largely concerned with discovery and classification of astronomical sources as they exist in the spatial domain along with how they evolve over long-term time scales.

With the advent of modern computing and improved technology it has become feasible to start monitoring these sources with a fine-grained time-domain, where events can often be tracked in near real-time for changes. This, in turn, is leading both to valuable new discoveries as well as revisions of existing theories within many subfields of astronomy. Nevertheless, the simple act of capturing large amounts of data is still a significant challenge, in part because of the need to have geographically dispersed collecting devices.

At the same time, citizen-science projects such as Zooniverse have demonstrated that the public has a strong desire to contribute to scientific inquiry (Fortson et al., 2011; Marshall et al., 2015; Jordan Raddick et al., 2013). Coupling the desire of citizens to participate in science with the need

to distribute imaging devices to geographically diverse areas therefore seems natural. The issue then becomes the reliability and cost efficiency of these devices and the ease with which non-technical users and students can assemble, operate, and improve the hardware and software. Critical to this is having a unit that has been well-tested and proven to operate effectively without advanced training or significant time involvement.

The ultimate question Project PANOPTES aims to answer has now become a series of simple questions: Who can participate in exoplanet science? How can citizen-scientists contribute to exoplanet science? What kind of exoplanets can these students and amateurs find? How much of the night sky can be watched? and, finally, how frequently can the whole night sky be watched?

1.1.1 Who can participate in exoplanet science? Engaging citizen scientists

While amateur astronomy has always enjoyed a certain level of popularity, the rise of rapid communication and the internet has allowed for a recent resurgence in amateurs taking part in the scientific process, both in traditional and novel ways. Specifically, the joint interaction of citizen science projects with large and open datasets has created a unique opportunity for the realization of more effective participation on the part of the public in a diverse set of fields, including astronomy (Catlin-Groves, 2012). The Plant Hunters project, and indeed the entire ecosystem of citizen science projects housed under the Zooniverse[1] umbrella, have famously capitalized on this public interest by involving large amounts of interested citizens in the classification of exoplanet light-curves to great success (Wang et al., 2015). PANOPTES aims to utilize this motivated public by involving participants in not only the collection and analysis of the data, but in the actual construction of the data collection devices, the individual PANOPTES units.

The involvement of interested citizen scientists across a variety of disciplines has been studied heavily in recent years, including the mechanics of organization (Franzoni & Sauermann, 2014; Catlin-Groves, 2012; Newman et al., 2011; Wiggins & Crowston, 2011), the types of tasks performed (Catlin-Groves, 2012; Marshall et al., 2015), the motivations and demographics of the participants (Jordan Raddick et al., 2013; Marshall et al., 2015; Tulloch et al., 2013), and the relative effectiveness and scientific value of such projects (including the distrust of publicly collected data among the scientific community (Catlin-Groves, 2012; Szabo et al., 2012)).

[1] https://www.zooniverse.org/

Specifically within astronomy it has been found that the "best" of the projects play to the strengths of the participants while also targeting areas of observation that professional observatories and automated software might miss (Marshall et al., 2015). Of particular note is that a critical motivating feature of citizen scientists in astronomy is the desire to make authentic contributions to the field (Jordan Raddick et al., 2013). Therefore, there are strong indicators that projects that provide opportunities for the citizens that cannot be accomplished by other means, while simultaneously providing a low barrier to entry as well as a pathway to more advanced scientific work, are highly desired (Marshall et al., 2015). For PANOPTES, this can be coupled with the recognition that certain types of projects inherently benefit from the geographic diversity of the participants (Franzoni & Sauermann, 2014).

By its very design, the PANOPTES project seeks to spread the total étendue of the survey over as wide of an area as possible (Guyon & Martinache, 2011). Since étendue is the product of how much light a survey can gather by how much of the sky it sees at one time, the best way to accomplish this task is to have a large number of small collecting devices with a large field of view (FOV) in as diverse a set of global locations as possible. The building of these devices, while inexpensive and relatively non-technical, is nevertheless too time consuming and geographically dispersed to be handled by professional observatories and staff. On the other hand, this requirement provides a perfect opportunity to engage citizen scientists in a task that otherwise would be difficult to accomplish.

1.1.2 How can citizen-scientists help find exoplanets? Utilizing digital cameras and low-cost consumer mounts

Consumer DSLR cameras are designed for taking images that replicate human vision and are thus traditionally not used for scientific research. Despite this, a number of authors have successfully employed DSLRs for astronomical work in recent years. The use of DSLR cameras (see Figure 1.2) provides a low cost per étendue (see section 1.1.4) and serves as the backbone of Project PANOPTES, enabling the answer of how to effectively watch most of the night sky. PANOPTES intends to become the first wide-field survey that relies on DSLR cameras for data capture. The specific advantage that DSLR cameras offer is their low price and usability as compared to traditional CCDs while still offering adequate performance metrics (Hoot, 2007; Kloppenborg et al., 2013; Guyon & Martinache, 2011; Zhang et al., 2016).

In parallel with advances in consumer cameras, advances in consumer-level mounts mean that affordable and reliable equatorial mounts are now available that offer quality tracking along with a

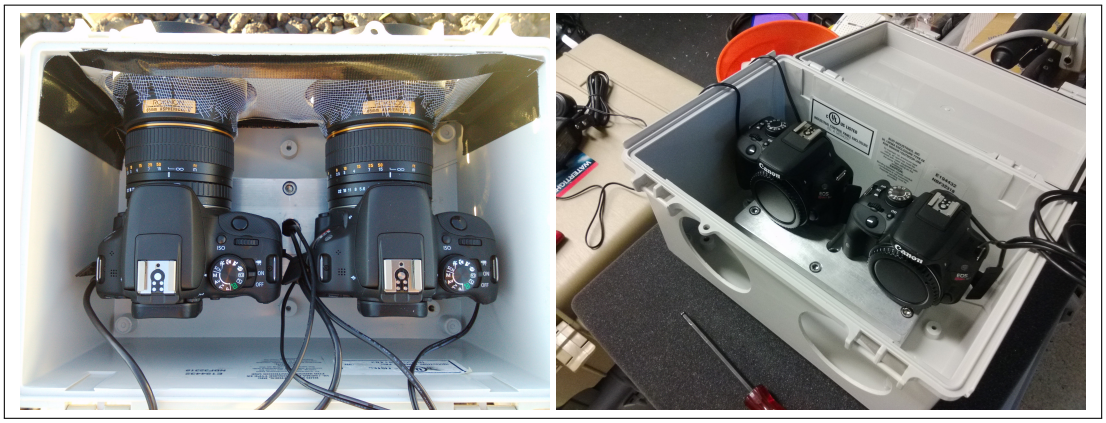

Figure 1.2: *Left:* Two Canon EOS 100D (Rebel SL1) with Rokinon 85mm $f/1.4$ lenses installed inside the `PAN001` camera box. *Right:* Two of the same models being installed inside a camera box for a unit currently under construction.

programmable and customizable control interface that can be operated remotely. While not designed specifically for unattended outdoor and autonomous usage, these consumer mounts nevertheless can be easily customized to withstand drastic weather conditions and provide a low-barrier for repair and replacement should something go wrong, unlike most dedicated astronomical equipment.

DSLRs present a number of complications for use as an astronomical imaging device, which is one key hurdle PANOPTES needed to overcome. These challenges are discussed below in the section on Complications and provide the most significant obstacle in terms of data quality.

1.1.3 What kind of exoplanet science can be performed? Mapping the nearby exoplanets

Exoplanet science has changed dramatically in the last decade. While the first exoplanet detection came in 1995 with the discovery of the gas giant 51 Pegasi b (Mayor & Queloz, 1995), it was the Kepler space mission's launch in 2009 (Borucki et al., 2010), and the subsequent confirmation of hundreds of exoplanets that brought about the largest change in how astronomers thought about extra-solar planetary systems.

However, Kepler was designed to look primarily at dim and therefore distant stars, which made follow-up with ground-based instruments difficult. Since it had been confirmed that many such exoplanets did in fact exist, the desire grew to find exoplanets that were closer to earth, orbiting stars of higher apparent brightness, which enables ground based spectroscopic follow-up, a prerequisite for

the possibility of finding liquid water (Kopparapu et al., 2013).

Yet despite new missions and surveys to map out missing exoplanets, there still remains large sections of the night sky about which we have no information. PANOPTES aims to help fill out the details of this map. Specifically, PANOPTES can cover large portions of the sky in order to fill holes in the map in addition to watching those sections for long periods of time, helping to find those long-period planets that are missed by other surveys.

Details about the types of exoplanets Project PANOPTES searches for are provided in chapter 2.

1.1.4 How much of the night sky can we watch? Increasing the étendue

Wide-field survey projects are measured in étendue, which is simply the product of the telescope aperture and the field-of-view (FOV). Formally,

$$\mathrm{d}G = n^2 \, \mathrm{d}A \cos\theta \, \mathrm{d}\Omega \tag{1.1}$$

where n is the refractive index, $\mathrm{d}A$ is a surface area segment, θ is the angle of incidence, and Ω is the solid angle against the sky. Setting $n = 1$ for air, $\theta = 0$ for a source at infinity and summing across the elements, étendue for wide-field surveys simply becomes:

$$G = A\Omega \tag{1.2}$$

where A is the aperture size and Ω the FOV[2]. Wide-field surveys are measured in meters-squared degrees-squared ($\mathrm{m}^2\mathrm{deg}^2$), which is often denoted by $A\Omega$ (as opposed to G as given in Equation 1.2). One item of note is that étendue is a cumulative property, meaning new systems can be linearly added to effectively increase the overall étendue of a survey. Additional systems observing different fields add to the overall FOV while those observing similar fields add to the total aperture (Law et al., 2015). In this way, additional telescopes and sites can be added to a survey to increase overall étendue.

Previous generation surveys, many of which are still operational, measure étendue values in the single digits (Tyson, 2010). Pan-STARRS, a current-generation survey described as a distributed aperture system, recognizes the cumulative aspect of étendue and makes the important distinction that there is no difference between one large aperture and a collection of smaller apertures (Kaiser et al.,

[2]The formal equation for étendue doesn't take into account issues such as vignetting, which may be present in an optical system and can be wrapped into the *effective* FOV.

2012). Pan-STARRS-1 (PS1), located in Hawai'i, achieves a value of 13 m^2deg^2 with the total Pan-STARRS network of four telescopes expected to achieve an étendue in the range of 40 – 60 m^2deg^2 (Chambers et al., 2016). The Evryscope[3], which employs 27 individual telescopes on one mount in order to simultaneously monitor 8,660 square degrees from one location, has an étendue around 30 m^2deg^2 (Law et al., 2015). The Fly's Eye design, which similarly uses 19 telescopes per mount, is expected to achieve an étendue of 35 – 40 m^2deg^2 (Pál et al., 2013). Kepler, which has a 0.95 meter aperture and a 115 degree FOV, has an étendue of nearly 75 m^2deg^2, the largest of any existing survey.

Next-generation surveys, notably the Large Synoptic Survey Telescope (LSST), will greatly exceed this with a total étendue value of 319 m^2deg^2 cited as a primary science requirement (Ivezić, 2013).

A single PANOPTES unit, equipped with a Rokinon 85mm $f/1.4$ lens, has an effective aperture diameter of 0.06 m per camera. The baseline unit includes two cameras observing the same field, doubling the effective aperture for étendue considerations. Thus, an individual unit is capable of giving an étendue of:

$$2 \times \pi \left(\frac{0.06\,\mathrm{m}}{2} \right)^2 \times 150\,\mathrm{deg}^2 \approx 1.75\,\mathrm{m}^2\mathrm{deg}^2. \tag{1.3}$$

Note that Equation 1.3 does not take into account vignetting on the lens, which would effectively lower the effective aperture. A comparison of a single PANOPTES unit to other wide-field surveys is given in Figure 1.3 and Table 1.1. At approximately 42 units, PANOPTES will have the largest étendue of any dedicated exoplanet survey, while 182 units would provide PANOPTES with the largest étendue of any astronomical facility worldwide.

1.1.5 How often can we watch the entire night sky? Watching all the time

Despite being the accepted metric for wide-field surveys, étendue doesn't give an entirely accurate representation of how effective a survey will be for exoplanet detection. Étendue is primarily an optical metric and is meant to convey the amount of sky coverage that is possible but doesn't say anything about the temporal coverage of the selected areas. More accurate representations for a transit survey include a time fraction that specifies the efficiency of the survey.

[3]One Evryscope is currently in operation and construction of a second unit began in early 2017, which will eventually raise the overall étendue (Law, 2017).

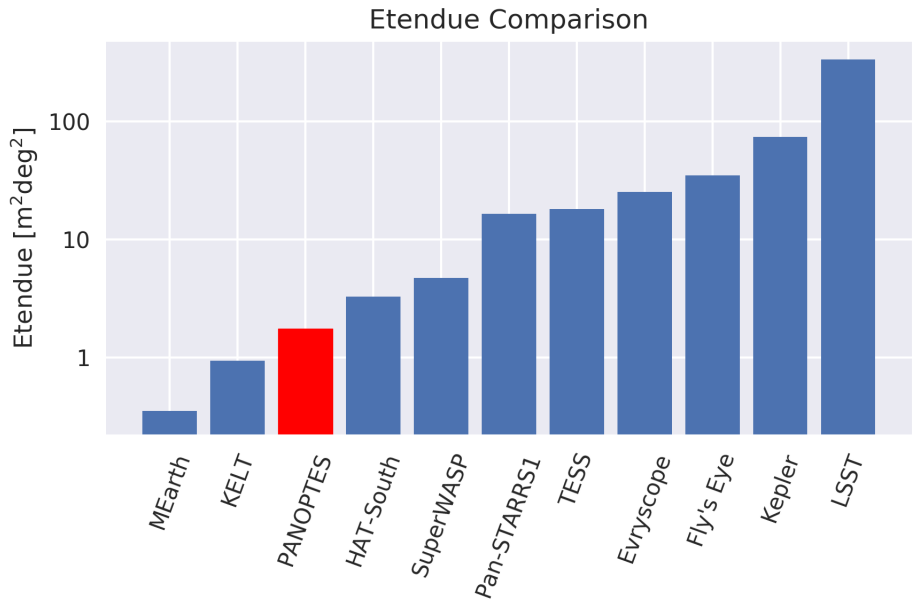

Figure 1.3: Étendue of selected wide-field survey missions per individual units. Many surveys plan to utilize the cumulative effects of étendue in order to increase the overall value, which is not depicted here. See text for a description of the cumulative effects of étendue. See also Table 1.1

Pepper et al. (2002) break a global survey efficiency ($\mathcal{E}$) down into two parameters representing the time fraction available for observing, given by $\mathcal{E}_0$, and the time fraction spent exposing the target, $\mathcal{E}_S$. These parameters come from their equation for optimizing the design of a survey telescope, which comes down to an optimization of the photon collection rate of an imaging device.

This photon collection rate is used by Pepper et al. (2002) to determine the time required by the survey to successfully detect a single hot-Jupiter around stars with certain fiducial parameters. This survey time is ultimately scaled by the efficiency fractions. Since these fractions will always be less than unity, they can have a dramatic effect on the time required to successfully detect an exoplanet.

For instance, in Pepper et al. (2002) the time required to complete a hypothetical survey is $T = 4$ months using values of $\mathcal{E}_0 = 20\%$ and $\mathcal{E}_S = 50\%$, for a total fractional efficiency of $\mathcal{E} = 10\%$.

However, since $\mathcal{E}_S$ represents the fraction of the time spent observing, it can easily be increased by adding additional cameras across multiple units such that the fraction of time spent exposing on a target is much closer to unity. Likewise, in order to mitigate effects of weather and, indeed, the night time sky, one can add additional units at different geographic locations, which can very effectively increase the overall survey efficiency. For a low-cost system like PANOPTES this is not only possible but it actually one of the main motivating factors for the project.

Table 1.1. Étendue of Selected Wide-Field Surveys

Name	FOV [deg^2]	Effective Aperture [m^2]	Étendue† [m^{2}deg^2]	Pixel Size [" / pixel]	Source
MEarth	2.8	0.4	0.35	0.8	Law2015
KELT	676.0	0.04	0.94	23.0	Law2015
PANOPTES	150.0	0.12	1.75	10.3	–
HAT-South	128.0	0.18	3.26	3.7	Law2015
SuperWASP	488.0	0.11	4.72	13.7	Law2015
Pan-STARRS1	7.0	1.73	13^b / 16.42	0.258	Chambers2016
TESS	2300.0	0.1	18.06	21.0	Ricker2014
Evryscope	8660.0	0.06	25.31	13.3	Law2015
Fly's Eye	–	–	35.0^b	22.0	Pal2013
Kepler	105.0	0.95	74.43	4.0	Law2015
LSST	9.6	6.68	319^b / 336.44	0.2	Ivezic2013

†Étendue is calculated from the FOV and aperture unless otherwise noted.

^{b}Reported value, which differs from calculated value due to inclusion of vignetting or other factors.

Examples from Kepler, TESS, and LSST further illustrate the issues concerning efficiency with regard to exoplanet surveys.

Kepler, which dwelt on one field for the duration of the original mission[4], is perhaps the only survey to approach unity for the time fraction. However, the CCD is not constantly integrating and breaks in exposure are encountered between individual images. In addition, 46 hours are lost each quarter due to position adjustments and 26 hours are lost each month while data is downloaded from the spacecraft (Haas et al., 2010). A first-order estimate of the efficiency value for Kepler is thus in the range of $0.8 - 0.9$. Finding exoplanets via the transit method was of course the primary design requirement of the Kepler space mission. In a post-Kepler world, however, it should be noted that Kepler was designed to serve as a proving ground for a space-based exoplanet finding mission. The intent was never to try and map out the entire night sky but instead to map out a small section of sky in intimate detail. This is in contrast to LSST, TESS, and other wide-field surveys that are trying to create a broader, if slightly less detailed, time-resolved map of the entire night sky.

TESS (Transiting Exoplanet Survey Satellite) divides the celestial sphere into 26 observation sectors, 13 per hemisphere, as shown in Figure 1.4 (Ricker et al., 2014). Most fields are viewed for 27 continuous days after which they are never revisited[5]. Considering only the 27-day regions, TESS could be said to have an efficiency of $\mathcal{E}_0 = \frac{27 \text{ days}}{365.25 \text{ days}} \approx 0.074$. Doubling the time doubles the time

[4]K2, the rebooted Kepler mission which operates with only two of the three reaction wheels, has a number of different target fields and therefore does not apply to the following analysis (Howell et al., 2014).

[5]According to the original mission specifications. Like Kepler, the hope is that TESS will secure funding to operate past the original mission life time in which case some fields would be revisited.

fraction such that the 54-day region would have a fraction of ≈ 0.148. A complete account of TESS's time fraction would consider each region according to how much time was spent observing there but it can be seen that when trying to accurately detect all exoplanet transits, as opposed to statistically constraining population estimates, TESS's efficiency can actually be quite low. It should be noted that TESS is specifically looking for exoplanets with periods of $P \approx 10$ days (for the 27-day region) around relatively bright and nearby stars. For each individual sector the exoplanet detection efficiency for planets of $P = 10$ days is relatively high.

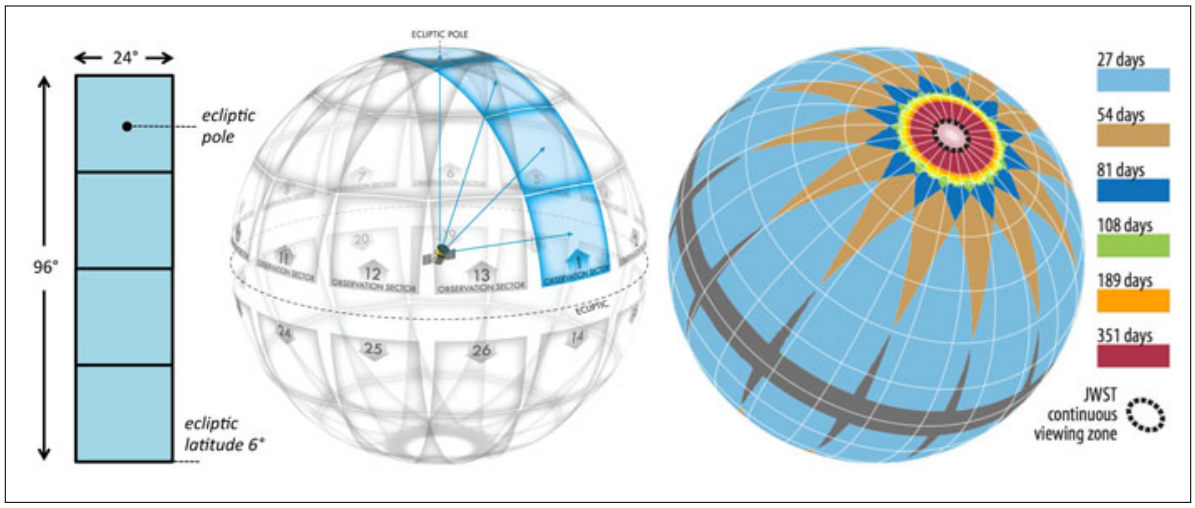

Figure 1.4: A graphical representation of the TESS observing sectors. *Left:* FOV of combined CCDs. *Center:* The 26 observing sectors. *Right:* Time coverage showing total time spent on one area including overlaps. Note that some regions are in a continuous viewing area for their respective hemisphere. Source: Ricker 2014 (Ricker et al., 2014)

LSST, which provides a significantly higher étendue than any other survey, unfortunately has a much lower efficiency when it comes to exoplanet detections. According to the LSST Science Requirements Document (Ivezić, 2013), each sky location will receive 100 visits per year with two 15 sec exposures per visit[6]. For a given year the fraction of time that LSST will spend on a single location will be $\frac{3000\,\text{sec}}{31558150\,\text{sec}} \approx 9.5 \times 10^{-5}$ [7].

Thus, while it has been demonstrated that LSST can provide valuable data for exoplanet science (Lund et al., 2014; Jacklin et al., 2015, 2017), it is clear that the search for exoplanets is not among the

[6]A conservative estimate derived from simulations and including factors such as weather history, etc. See Ivezić 2013 for details.

[7]LSST has been shown to have the ability to detect hot-Jupiter and hot-Neptune class planets with orbital periods of $P = 0.5 - 20$ days within the first year of operation. See (Lund et al., 2014; Jacklin et al., 2015, 2017) for details.

primary requirements for the survey and that high étendue does not necessarily equate to effectiveness with regard to finding exoplanets.

Space-based missions offer the ability to dwell on a single field with high-efficiency but are necessarily limited by cost and practicality in their ability to provide long-term monitoring. Large ground-based surveys, like LSST, are looking to provide some of this long-term monitoring but are also limited by cost such that only a single observatory can be constructed, forcing a cadence that allows the survey to cover as much sky as possible at the expense of dwelling on individual fields.

The solution provided by PANOPTES, enabled by the low-cost of adequate detectors in the form of DSLR cameras, is to distribute a large number of units globally. By providing units in dispersed geographic locations the observing efficiency ($\mathcal{E}_O$) of the survey is increased while multiple units observing the same field serves to increase the shutter efficiency ($\mathcal{E}_S$). In this way the project hopes to accomplish its panoptic goal of watching all of the night sky all of the time.

1.2 Complications

There are obvious complications to building a distributed network of astronomical imaging devices based on consumer products. At the single unit level the two most important considerations are the validity of a DSLR for use as a scientific instrument and the protection of the unit in adverse weather.

1.2.1 DSLR Cameras

While DSLR cameras present a low-cost alternative to traditional astronomical CCDs, they nevertheless present their own subset of problems with regard to gathering scientific-quality data (Holst & Lomheim, 2011). The presence of the Bayer color filter array (CFA), a repeating four-pixel square pattern of red-green-green-blue pixels (see Figure 1.5), makes absolute photometry difficult (Kloppenborg et al., 2013; Guyon & Martinache, 2011). In addition, proprietary processing techniques, low quantum efficiency (QE), small full-well capacity, a built-in infrared (IR) filter, and a smaller dynamic range limit the ability of these cameras (Holst & Lomheim, 2011; Hoot, 2007; Guyon & Martinache, 2011; Guyon et al., 2014; Zhang et al., 2016; Kloppenborg et al., 2013).

ISO settings on a DSLR camera can have particular effects on the image quality. Read noise is actually found to decrease at higher ISO settings (Zhang et al., 2016), making high ISO and defocusing two attractive factors enabling the high quality of photometric results obtained from DSLRs (Littlefield, 2010; Zhang et al., 2016). Higher ISO settings, however, mean shorter exposure times on bright stars

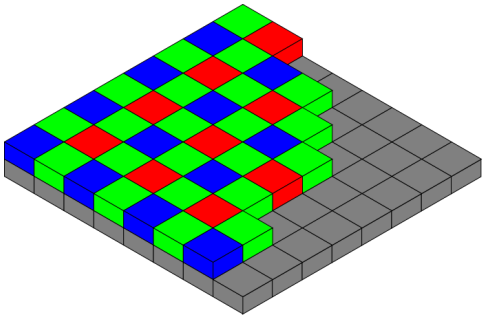

Figure 1.5: A Bayer color filter array present on typical consumer cameras. Represented here is a blue-green-green-red pattern. See also Figure 5.2. Image courtesy of Wikipedia user Cburnett.

which in turn equate to more overall exposures and thus shorter shutter lifetimes.

A Canon EOS 100D has a shutter lifetime rated at 100,000 exposures[8]. Assuming 30 second exposures, a (long) 15 second readout/transition time, and a modest 6-hours of observing per night, shutter failure will occur in just over six months, which is clearly not acceptable for a remote and autonomous imaging device. Even with exposure times of 120 seconds (the PANOPTES default), shutter failure is the first anticipated point of failure for a running PANOPTES unit. With longer exposures times we expect that this will not occur for at least two years on most cameras, at which point the camera body alone can be replaced at low cost.

This leads to a compromise where low ISO settings (and longer exposure times) are recommended for surveys of long duration (Guyon & Martinache, 2011; Zhang et al., 2016). On the positive side, an ISO setting of 100 has been shown to exhibit limited dark current noise, thus removing the necessity of applying dark frame subtraction (Kloppenborg et al., 2013; Zhang et al., 2016).

Focusing is another item of concern. Consumer DSLR cameras tend to have a pixel size comparable to the point-spread-function (PSF) size of the optics, and hence also to the size of a star (just over 10 arcseconds per pixel in our case), leading to an under-sampling of the PSF when the lens is in focus (Guyon & Martinache, 2011). De-focusing of the lens is often used to broaden the stellar PSF (Littlefield, 2010; Kloppenborg et al., 2013). This has the effect of spreading the light across a number of pixels as a means of properly sampling the PSF but at the cost of increased issues with stellar blending and decreased sensitivity for faint stars.

Since the goal of PANOPTES is to monitor as many stars as possible, the stellar blending that results from de-focusing makes this an unviable option (Guyon & Martinache, 2011; Zhang et al., 2016). On the other hand, a strict focusing procedure does not seem to offer any specific advantage

[8] http://www.shutteractuations.com/canon-eos-camera-shutter-lifetime/

to a one-time manual focus as the lenses appear to be thermally stable across multiple months of use (Guyon & Martinache, 2011; Zhang et al., 2016). Testing this thermal stability over longer time scales, an important feature considering the proposed life-span of a PANOPTES unit, is left for future work. However, it should also be noted that the Rokinon lenses that are used on `PAN001` have no focusing motor, meaning that if it is discovered in the long-term that a focusing procedure is required a new lens solution will also be needed.

Anti-blooming features, wherein excess charge from a pixel is discarded rather than leaked onto neighboring pixels, can also lead to non-linear responses as stars begin to saturate the detector (Hoot, 2007). Despite this, is has been found that stars can effectively be monitored across a range of five (5) magnitudes (Kloppenborg et al., 2013). Prior to saturation, the CMOS sensors present in DSLR cameras exhibit near-perfect linearity (Zhang et al., 2016).

The take-away message is that while DSLR cameras might make absolute photometry difficult, they have nevertheless been found to be effective for use with differential photometry where only modest systems are required, which exactly matches the use-case of Project PANOPTES (Hoot, 2007; Guyon & Martinache, 2011; Guyon et al., 2014; Zhang et al., 2016).

Specific image property measurements of the camera used for `PAN001` are discussed along with other hardware in chapter 3. The differential photometry algorithm developed as part of this thesis is described in chapter 5.

One particular note is that all of the reviewed literature used (a variety of) Canon DSLR cameras in their analysis without much consideration or discussion as to the particular advantages Canon cameras seem to offer versus other brand names. One of the advantages of how Project PANOPTES has been designed is that it can act as a platform to test various hardware configurations (see discussion of the hardware agnostic design of the software in section 4.1.1.2 as well as the section on alternative hardware in section 3.3).

1.2.2 Weatherproofing

A PANOPTES unit does not provide a dome or enclosure to protect itself from adverse weather conditions, which would otherwise significantly increase the overall cost of each unit. As such, performing a safety check and maintaining a park position that effectively hides the exposed lenses is one of the primary design goals of the hardware/software interaction. To date this park position (see Figure 1.6) has been successful in weathering snow, rain, and high wind.

Late January 2017 saw sustained winds of over 100 kph for nearly six hours with gusts as high

as 160 kph. Some tearing of the sheet metal that is used to divert rain from the joints on the RA axis resulted, prompting a need for a redesign that includes stiffer material (see Figure 1.7, right and section 3.4).

Operational safety as it relates to weather is discussed along with the software in section 4.1.1.4.

Figure 1.6: PAN001 at MLO in the park position, with cameras facing down, on an otherwise clear night. The unrelated VYSOS-5 dome is on the left. Image courtesy of PANOPTES team member Josh Walawender.

1.3 Project History

PANOPTES began in 2010 with the creation of a one prototype camera unit by Prof. Olivier Guyon, the project's founder. Shortly thereafter, autonomous operation was achieved via a re-wiring of the mount electronics and a second camera was added. In 2013 a third prototype featuring four cameras was designed in order to compare lenses and to test the effectiveness of DSLR cameras with and without an IR filter present. The result of this prototype was the demonstration that 1) cheaper Rokinon lenses were nearly as effective as the more expensive Canon counterparts, 2) the IR-filter inside the camera body did not severely affect photometric performance, and 3) an algorithm could be employed that

Figure 1.7: *Left:* PANOPTES prototype happily parking through the occasional Hawaiian blizzard. Image by Olivier Guyon. *Right:* Tearing of sheet metal due to high winds. Sheet metal designed to divert rain from joint interactions is shown after nearly six hours of sustained high winds ($\sim$ 150 kph). Weatherproofing worked effectively for nearly two years before tear occurred. Image by Josh Walawender.

was effective at achieving a 2% photometric precision (Guyon et al., 2014) despite the presence of the Bayer CFA.

The success of the third prototype led to an official Project PANOPTES with an initial goal of creating a baseline unit that could be easily reproduced for under $5000 USD. Design of this unit initially began in 2014. Whereas the original prototypes had also used a collection of hacked-together C scripts, the educational goals of PANOPTES argued for the adoption of Python and a completely redesigned code-base in order to accommodate the anticipated needs of the PANOPTES project, which is the main aim of this thesis.

2

Science

Project PANOPTES is an efficient wide-field survey, designed to provide comprehensive spatial and temporal coverage of the entire celestial sphere. PANOPTES will primarily target bright and nearby stars with an emphasis on M dwarfs. Jupiter- and Neptune-sized exoplanets causing a 1% reduction in total flux should be attainable. Long-period planets, which are missed by many current and future surveys, will also be targeted. PANOPTES will be well-suited to provide long-term coverage of the same parameter space as that of the forthcoming TESS mission in terms of stellar types and magnitude range. This chapter discusses each of these points in more detail, providing an overview of the particular type of exoplanet science to which a PANOPTES unit will contribute.

2.1 Sky coverage

Sky coverage for PANOPTES will largely be a function of the number of units that can be built. PANOPTES is meant to be an all-sky survey, with the goal of providing a network of individual units that can realize full spatial coverage of the full celestial sphere at all times. To attain this goal a large number of units would need to be built and distributed intelligently across the surface of the globe.

Before the goal is reached, PANOPTES will rely mostly on the availability of volunteer site locations, which cannot easily be controlled. When seeking volunteers to build new units emphasis will be placed on striving for global coverage. New units are currently being planned for Australia[1] and Chilé, which adds important southern hemisphere locations, as well as a number of sites across the USA as part of a NASA/JPL grant that has recently been awarded[2].

[1]Part of this work as well as the future PhD work of the current author.

[2]PANOPTES is part of "NASA's Universe of Learning: An Integrated Astrophysics STEM Learning and Literacy Program" Space Telescope Science Institute - Baltimore, MD. Denise Smith, Principal Investigator, Contract Number: NNX16AC65A

While it would be possible to optimize an individual unit to provide maximum spatial coverage for a given year, this would necessarily cut down on the amount of temporal coverage possible for a given field. Scheduling, therefore, is a balance between coverage in either the spatial or temporal domain. For `PAN001` the goal was to validate the unit against known exoplanet transits and so fields were selected based on ephemeris data[3].

Figure 2.1 shows the fields that were used during an observing run in early 2017, each centered around either a known transit or a "pretty image" field. This serves as an example of the scope of coverage capable by a single unit. Throughout the 32 nights of observations a total area of just over 2000 square degrees was observed. Table 2.1 provides summary details about each of the fields, including the temporal coverage of each target.

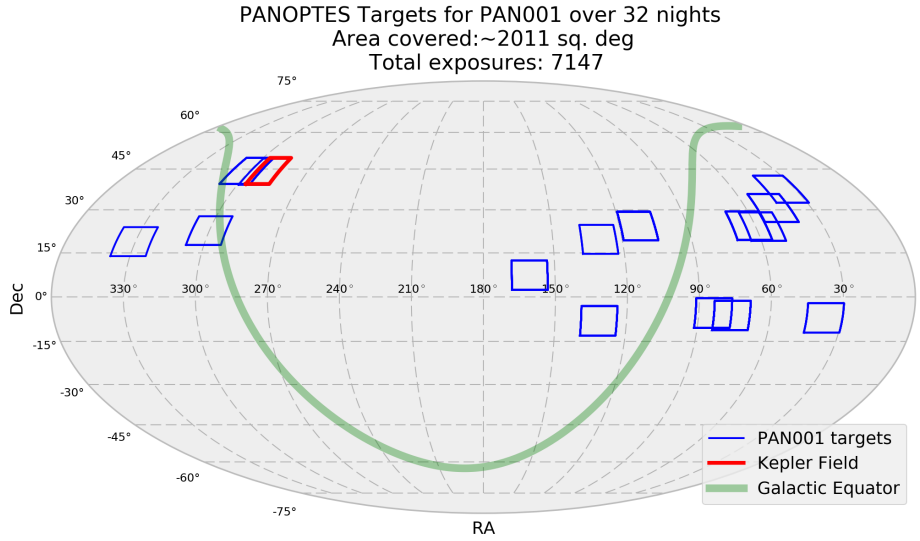

Figure 2.1: Map of over 2000 deg^2 across 15 separate fields captured by `PAN001` over 32 (non-consecutive) nights in late 2016 to 2017 after . The red box shows the Kepler field for reference, which was also one of the fields observed by `PAN001`. For a list of targets included see Table 2.1

With a goal of monitoring as many bright and near stars as possible, emphasis will originally be placed on regions of high stellar density. Galactic latitudes of $\leq \pm 30^\circ$ will be given a higher priority in the PANOPTES scheduler, with the caveat that with nearly 10 $''/_{\text{pixel}}$ blending of stellar sources starts to become a problem.[4] However, because bright stars are located throughout the entire sky (Ricker et al., 2014) there will be no strict requirement about targeting regions near the galactic equator.

[3]Originally provided by the Exoplanet Transit Database (http://var2.astro.cz/ETD/index.php) but later incorporated into utility scripts in the image analysis repository that access the Open Exoplanet Catalogue (www.openexoplanetcatalogue.com/)

[4]Blended eclipsing binaries will continue to remain a “problem” with the PANOPTES pipeline currently unable to

Table 2.1. PAN001 Observing Stats

Field	Number Exposures	Total Exposure Time [sec]
Kepler 1100	12†	1358
HD 209458	32†	3892
HD 189733	35†	3091
M44	121	14453
M45	136	16170
Wasp 33	209	25199
M42	274	32488
Wasp 77	296	35400
KIC 8462852	317	38122
Wasp 36	377	44870
Wasp 11	453	54011
Wasp 104	771	92045
Wasp 35	1328	158347
HAT-P-20	1376	164199
EPIC-211089792	1410	168518
	7147	**236.71 hours**

†A "successful" observation is normally comprised of at least 60 contiguous observations

Accordingly, it should also be noted that each individual unit is fully controllable by the team that built the unit. Teams will be encouraged to design their own parameters or regions of interest, with a default being those regions near the galactic equator at priorities defined by the PANOPTES team. We hope that by giving control to individual teams, serendipitous discoveries might be made in regions outside the galactic equator, and people are encouraged to research into the topic in order to determine their own regions of interest.

2.2 Stellar targets

M dwarf stars offer an attractive target for Project PANOPTES. Besides being relatively abundant, comprising nearly 3/4 of the solar neighborhood within 10 pc (Lurie et al., 2014), they are also thought to harbor a large number of exoplanets (Howard et al., 2013). Estimates of 2.5 ± 0.2 planets per M dwarf are given based off the full Kepler data (Dressing & Charbonneau, 2015). These systems are thought to contain mostly rocky planets in tightly-packed orbits, with many potential habitable zone candidates (Shields et al., 2016). While smaller terrestrial planets would be a significant challenge for a PANOPTES system, the mapping of the larger gas giants in such a system, especially those of

distinguish between these sources and that of a transiting planet. Later work on the pipeline to take place during the PhD will seek to identify these sources via the regular periodicity although they will continue to be detected similar to transits

long-period duration (see section on Period below), would nevertheless provide value to our overall knowledge of these systems.

The PANOPTES Input Catalog (PIC), which formally defines the targets of interest for the project, will be based on the TESS Input Catalog (TIC). The TIC contains over 2 million stellar objects that can be viewed during long-cadence exposures with up to 400,000 selected stars identified specifically for observation during the short-cadence exposures (Stassun et al., 2014; Oelkers et al., 2016).

The bulk of the 400,000 target stars in the TIC are main-sequence and of spectral types F5-M5 (Ricker et al., 2014). The TESS bandpass is centered around the Cousins I_C filter (786.5 nm) with a magnitude range of $I_C \approx 4-13$ whereas a PANOPTES DSLR camera closely approximates a Johnson V filter (540 nm) (Hoot, 2007; Kloppenborg et al., 2013). The magnitude range for exposures of 120 seconds is $V \approx 8.5-12.5$ with an SNR $5-10$ for the dimmer stars (see also discussion of image quality in chapter 3). TESS is therefore more sensitive in the red than the PANOPTES DSLR cameras and can also target brighter stars. Saturation on a PANOPTES camera occurs at $V \approx 8$ with the default exposure time of 120 seconds. Shorter exposure times could potentially be used in order to capture brighter stars, at the expense of losing many of the dimmer stars, and is left as an option for individual teams to explore. Since each unit has two cameras, exposures times could also be varied per camera.

Of particular concern for transit detections with M dwarfs is the amount of stellar activity that is present (Rosenblatt, 1971; Tofflemire et al., 2012; Davenport et al., 2016; Hartman et al., 2011; Hosey et al., 2015). Stellar activity can cause variations in the associated light-curve which interfere with the detection of transiting planets, especially Neptune-class, leading to a high number of false-positives. A long observational baseline can help mitigate these issues, to which PANOPTES is well suited.

2.3 Exoplanet targets

2.3.1 Period

Transit surveys tend to be strongly biased toward planets with short orbits (Lurie et al., 2014; Ricker et al., 2014). The minimum period (P_{min}) that can possibly be observed is ideally placed at the Roche limit for a planet and is on the order of a few hours. Upper limits, on the other hand, are usually determined by the duration of the survey coupled with the fact that potential transits should be observed at least three times for proper confirmation. In addition, complications arising from the combination of multi-day (and multi-observatory) observations sets an upper limit of $P \sim 10$ days on the period that can be observed from the ground (von Braun et al., 2009). For the Kepler mission, an explicit

design goal of detecting earth-size planets at up to a one-year orbit was specified, which in turn led to a mission duration of over three years (Koch et al., 2010). TESS cites the period distribution of Kepler detections and notes it reaches a maximum where the period is near 10 days and thus chooses a P_{max} of 10 days for general observations[5] (Ricker et al., 2014).

The accepted standard for survey sensitivity to long period planets scales with $P^{-5/3}$, meaning that the transit probability is very low for planets on long orbits (Gaudi et al., 2005). As seen with Kepler, TESS, and LSST, this standard is often valid. However, this sensitivity is based upon conditions that don't necessarily apply to PANOPTES as it is envisioned. Specifically, the derivation of this sensitivity assumes that transit surveys have random sampling of the target and limited observational baselines. In fact, the $P^{-5/3}$ scaling is defined specifically for periods longer than the baseline of contiguous observations of the survey (Foreman-Mackey et al., 2016). Since PANOPTES aims to watch fields continuously and for long periods of times, those conditions don't necessarily apply. Similarly, since the

PANOPTES units have a low overall cost and are simple to repair if individual subcomponents fail, making them easy to keep running for long periods of time. This low cost and the longitudinal distribution of individual units help make the project especially well suited to the detection of long-period exoplanets.

2.3.2 Size

Photon-noise is ultimately limited by the number of units PANOPTES will have operating, with the final photometric accuracy of a given target attained by averaging the measurements of all the individual units. As more étendue is added to the network, better results will be obtained.

That said, a noise limit due to inherent systematics is still present. Significant challenges exist when considering a network of systems that are built and installed by non-professionals. The algorithm discussed in chapter 5 seeks to overcome these systematics, leading to an overall noise that is at the level of a few percent. Therefore hot and cold Jupiters around nearby bright stars are well within the limits of the current design whereas Neptune class planets probably represent the lower limit of detectability by the PANOPTES network. The pursuit of this lower limit will be the subject of my PhD work.

[5] A $P_{\text{max}} \gtrsim 40$ days is specified for regions near the ecliptic poles where there is high overlap with the upcoming James Webb Space Telescope (JWST). See subsection 1.1.4 and (Ricker et al., 2014) for details.

3

Hardware

The PANOPTES hardware baseline was iteratively developed by the core PANOPTES team. While the complete design is not within the scope of the current thesis, as part of my work with the PANOPTES core team I participated in the design, build, and installation of the `PAN001` baseline unit (see Figure 1.1 and Figure 3.1), including documentation[1]. Here an overview of the various hardware components is provided followed by more detailed characteristics of the DSLR cameras used. A section on the possible uses of alternative hardware is included at the end of the chapter. More detailed information about each of the hardware components, including descriptions of the weatherproofing of each component, is available online at the project website.

3.1 Baseline Design

The baseline hardware designed consists of three main components: the computer box that houses the running computer and other components related to power distribution; the camera box at the top of the mount that houses the two DSLR cameras; and the mount and pier assembly (see Figure 3.1).

3.1.1 Computer Box

The computer box is a water-tight plastic box (Pelican 1560) with three custom holes drilled on the sides. Two of the holes serve as conduits for the associated wiring; mains power[2] comes into the box from an external source two USB cables (to the camera box and webcam hub), 12 V power (to camera box), and weather cable (to AAG CloudWatcher) leave the box. The mains power is attached to a 12 V Universal Power Supply (UPS) that also has a 12 V battery backup. When mains power is interrupted

[1] Available online at www.projectpanoptes.org

[2] `PAN001` uses 110 V but the system could easily be adapted for 240 V outside the USA.

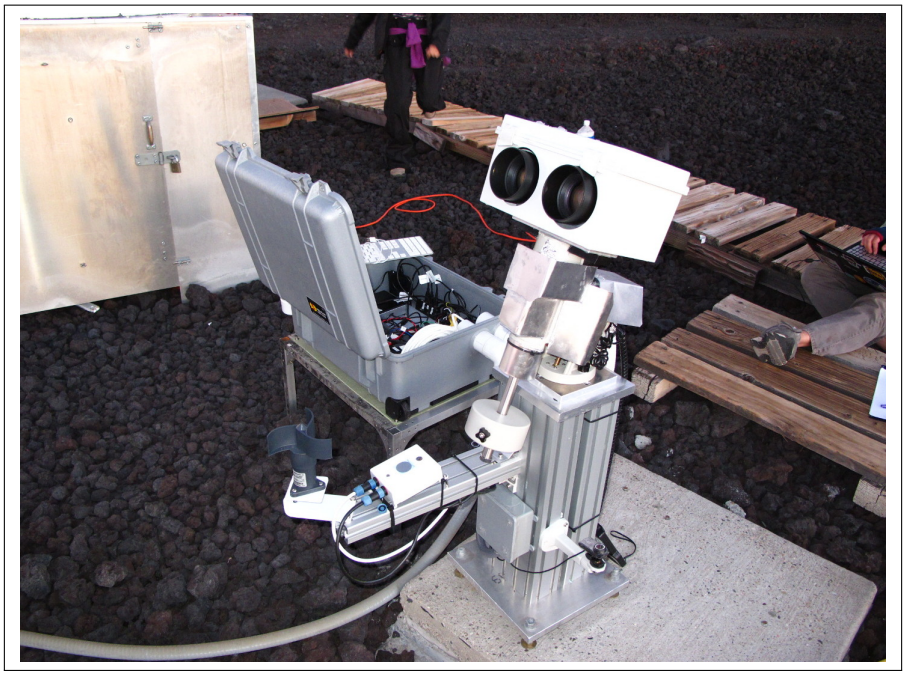

Figure 3.1: An image of PAN001 as installed at the MLO facility. The computer box is open in the background. The mount is currently in the 'Home' position and the two DSLR camera lenses can be seen in the camera box located on top. The pier includes a protrusion for the AAG CloudWatcher weather monitor as well as two webcams that point toward the camera box.

the unit is marked Unsafe and the mount automatically parks itself. The third hole, located on the opposite side of the box, serves as exhaust and features a controllable fan. The holes have plastic PVC pipe conduits which are sealed with silicone. The exhaust conduit is down-facing and covered with a fine-mesh screen and is the only open exposure on the box.

Two custom electronic boards are located inside the computer box, referred to as the power distribution board and the telemetry board. The former takes a 12 V input and splits it into five channels, each with a relay switch and current monitoring. The telemetry board has an Arduino Micro microcontroller responsible for sensor readings inside the computer box (temperature and humidity), control over the relay switches, and aggregation of the current monitoring. The original PAN001 installation did not include the power distribution board, which was added in late March 2017. The need to remotely power-cycle various components of the unit was the primary motivation for this upgrade (see section 3.4) but the board has also been re-designed for electronic safety purposes.

An Intel NUC 5i5RYK running Ubuntu 16.04 LTS serves as the control computer, responsible for running the PANOPTES Observatory Control System (POCS) and other accompanying software (see chapter 4). An external USB hard drive is attached to the unit for storage purposes.

3.1.2 Camera Box

The camera box is a Bud Industries NBF-32318 plastic enclosure with two large holes drilled into the side to accommodate the hoods of each lens, which project outside the box. A hole is also drilled into the bottom of the box through which the USB and 12 V power cables are routed. The camera box is attached to the mount via a dovetail plate secured to the bottom of the box with an additional aluminum plate (to increase rigidity), which is attached to two more aluminum plates on the inside of the box. The two DSLR cameras are attached to the top aluminum plate (see Figure 3.2).

Attached to the lid of the camera box is a USB hub and the camera electronics board. The camera board features an Arduino Micro, which records the values from the temperature/humidity sensor and the 9-axis accelerometer, as well as two voltage regulators that convert the 12V input to a 9V output appropriate for the DSLR cameras[3].

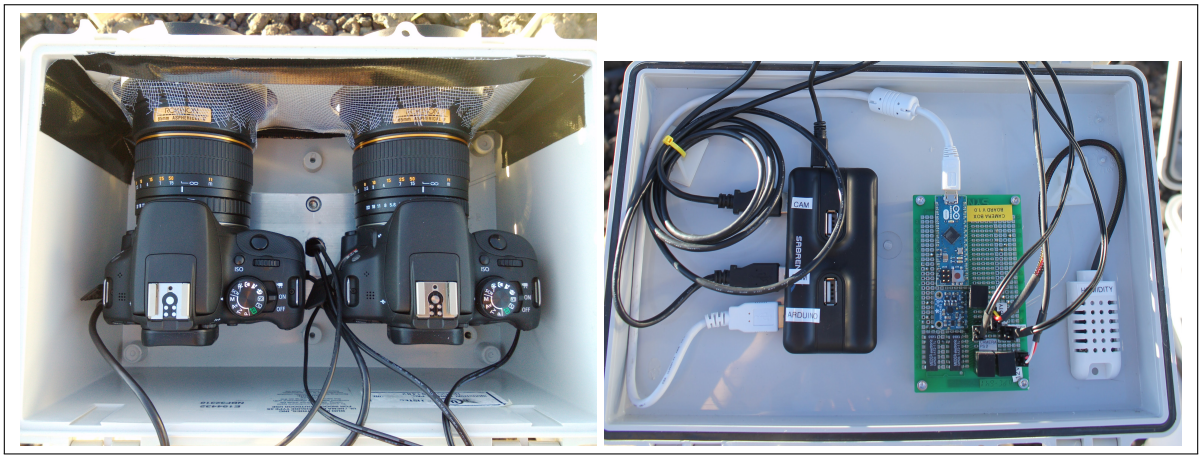

Figure 3.2: Images showing inside of camera box. *Left:* The two DSLR cameras mounted on top of the internal aluminum plates. Power and USB cables are seen entering through the center of the aluminum plates. *Right:* A USB hub and the camera electronics board are attached to the lid of the camera box. A temperature/humidity sensor and a 9-axis accelerometer are also visible.

3.1.3 Mount and Pier

The mount is a consumer iOptron iEQ30 Pro mount that was purchased directly from the factory. The mount is the single largest expense in the project at $1300 USD. While expensive relative to the other PANOPTES components, the mount offers adequate performance and performs well in adverse

[3]The cameras specify a 7.2V 1A input. The next iterative design for the camera board will feature one 8V 2A regulator which will feed both cameras.

Table 3.1. Camera & Lens Properties

Property	Unit	Value
Camera Model		Canon EOS100D
Cost	US $	~ 400
Pixels	#	3476×5208
Sensor Size	mm	14.9×22.3
Pixel Size	μm	4.3
Lens Model		Rokinon 85M-C
Cost	US $	~ 270
Focal Length	mm	85
F-number		$f/1.4$
Pixel Scale	arcsec / pixel	~ 10.31
FOV	degree	~ 10 × 15

weather conditions for which it was not necessarily designed. The mount features a serial interface and is programmable either via an ASCOM driver (which requires a Windows based computer) or a simple subset of commands that can be directly entered via a serial connection.

The mount comes with the ability to alter the tracking rate to a fraction of sidereal rate ($0.99-1.01\times$ sidereal) in order to account for any minor discrepancies in the hardware. However, due to a pre-production error on this particular mount the ability to alter tracking rates for the RA axis was not working for the initial deployment of `PAN001`, leading to significant tracking errors (see section 3.4).

A custom pier is constructed from Mitsumi aluminum bars. This pier features one large protrusion to which an AAG CloudWatcher weather station is mounted as well as two side arms to which webcams have been attached. The two webcams point toward the mount and are used solely for diagnostic purposes (see Figure 3.1).

3.2 DSLR Camera Characteristics

PANOPTES uses two Canon EOS100D (Rebel SL1) DSLR cameras with Rokinon 85mm $f/1.4$ fixed aspherical lenses. Selection of the camera and lens occurred prior to design of the baseline unit with a primary metric being low cost and therefore high étendue per dollar achieved when combining multiple cameras (Guyon & Martinache, 2011; Guyon et al., 2014). Two cameras mounted inside the camera box can be seen in Figure 3.2. Properties of the camera and lens are given in Table 3.1.

3.2.1 Camera & Lens Properties

Each camera takes full resolution images of 5208×3476 pixels on a CMOS sensor with a square pixel width of 4.3μm. With the attached lens the achieved FOV is $14.83° \times 9.928° = 148.409\,\mathrm{deg}^2 \approx 150\,\mathrm{deg}^2$, giving a pixel scale $\approx 10.31''\mathrm{pixel}^{-1}$. As mentioned in the Introduction, the particular complication that arises with DSLR cameras is due to the Bayer color filter array (CFA), which is designed primarily for visible light photography. As such, the CFA is optimized for green light (corresponding to visible sunlight and the human eye) and thus features twice as many green pixels as either red or blue. The EOS100D has a built-in zero offset bias of 2048 counts.

Canon lays out the Bayer array with a red pixel in the top-left corner, which then repeats in a 2×2 RGGB fashion as shown in Figure 3.3. A combination of RGGB pixels is sometimes referred to as a superpixel (Holst & Lomheim, 2011), especially when interpolating a RAW image (although it should be stressed that for PANOPTES processing no interpolation is done).

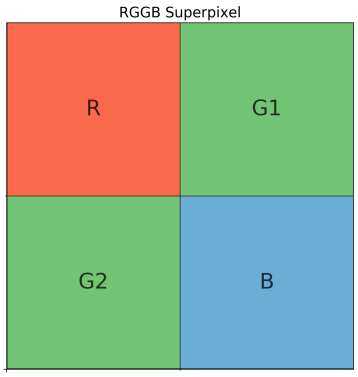

Figure 3.3: Bayer filter representation showing the arrangement of the pixels on the array. Each box corresponds to a pixel with a square side of 4.3 μm. Color channels are referred to by the labels in the image.

Canon writes raw images to a proprietary CR2 format, which our software immediately converts to a FITS image as part of the `Observing` state (states are described in section 4.1.1.4). The CR2 image is written with four color channels corresponding to each of the filter labels and each pixel is represented with 14 bits. The CR2 file is first converted to a 16-bit graymap file via the `dcraw`[4] utility and then to a 16-bit FITS file via a custom-written script. Header information directly from the CR2 file is added to observational header information placed in the FITS file. FITS image sizes are thus $\sim 35MB$ each, resulting in $15 - 20$ GB of raw data generated each night.

[4] https://www.cybercom.net/~dcoffin/dcraw/

Table 3.2. Canon EOS100D Sensor Properties at ISO 100

Property	Unit	Value
Gain	e^- / ADU	1.70
Bias (built-in)	ADU	2048
Quantum Efficiency	%	43[a]
Readout Noise (RMS)	ADU	7.37
Saturation	ADU	13583[b]
Readout Noise (RMS)	e^-	11.5[a]
Saturation	e^-	19568[a]

[a] http://www.sensorgen.info/CanonEOS-100D.html

[b] Includes built-in bias

3.2.2 Sensor Properties

Sensor properties for the EOS100D at ISO 100 are given in Table 3.2. A low ISO setting, corresponding to longer required exposure times, is chosen primarily as a means of increasing the maximum lifetime of the camera shutter, which has a factory rating of 100,000 exposures. Analysis of different ISO settings and the selection of ISO 100 for a wide-field survey has been given by multiple independent parties (Guyon & Martinache, 2011; Guyon et al., 2014; Zhang et al., 2016; Kloppenborg et al., 2013).

Saturation (in electrons), read noise, and quantum efficiency (QE) values reported in Table 3.2 are from publicly available web sites (source provided in table) but correspond closely to values given for different Canon models by (Guyon & Martinache, 2011; Guyon et al., 2014; Zhang et al., 2016). Official product specifications are proprietary and not published by Canon, and overall specifications for the cameras are difficult to attain (Zhang et al., 2016). Saturation counts are measured directly from image data of a bright source and gain values are calculated by taking the ratio of the electron saturation and the ADU saturation (after subtracting bias). That is, gain $= \frac{19568\,e^-}{(13583-2048)\,\mathrm{ADU}} = 1.70\,e^-\mathrm{ADU}^{-1}$. The electron saturation value is obtained from the source listed in Table 3.2 while the ADU saturation count is taken directly from an image of a saturated source.

Linearity for Canon DSLR cameras is almost perfect for all three color channels up to saturation (Guyon & Martinache, 2011; Guyon et al., 2014; Zhang et al., 2016; Hoot, 2007). While this wasn't tested specifically on the EOS100D, linearity is still assumed.

At ISO 100 the camera exhibits negligible dark noise as a result of dark current (Zhang et al., 2016; Kloppenborg et al., 2013), thus obviating the need for dark frames during normal observing routines.

Finally, modern DSLR cameras in general tend to experience very low read noise (Guyon &

Martinache, 2011; Guyon et al., 2014; Hoot, 2007; Zhang et al., 2016). For the PANOPTES prototype this read noise was explored along with the optimal ISO setting in order to determine minimum exposure times for photon-noise limited performance (Guyon & Martinache, 2011; Guyon et al., 2014). The procedure is repeated here for the EOS100D.

Science images were taken on the night of January 28, 2017, when the moon fraction was less than 0.1%. Two cameras each took 60 frames with exposure times of 120 seconds, with approximately 50 seconds between the end of one exposure and the start of another[5]. Background values for each channel were determined with the `photuils` package as described in 5.1.1. Median values (across full frame) for each channel for each frame were obtained and the mean of these medians were used as the per-pixel count values. This in turn leads to an average count per pixel for each of the color channels, as reported in Table 3.3 (see also Figure 3.4).

To get photon-noise limited performance we want the photon noise to match the readout noise (11.5 e^-) based on the minimum counts per second (given by the red and blue channels). Thus, photon noise overcomes readout noise at

$$\frac{11.5^2}{1.70}\left[\frac{e^-}{e^-/\mathrm{ADU}}\right] \approx 78\ \mathrm{ADU}.$$

For a per pixel count of 0.34 ADU per second in the red channel (the lowest count) this leads to a minimum exposure time of $\approx$ 230 seconds for photon-noise limited performance while the green channel, at 0.55 ADU per second, requires $\approx$ 142 seconds. Since PANOPTES is mostly concerned with tracking brighter stars, the photon-noise limit will be reached in 120 second exposures for stars in the $m_V \sim 8.5 - 12.5$ range.

This minimum exposure time is important because long exposure images start to become dominated by systematics and therefore stop providing useful information for a large number of stars in the field. Therefore we want to find the balance where the exposure time is long enough to overcome the photon-noise limit for the majority of the stars in the field while also being short enough such that the systematics don't begin to dominate. Exposure times of 120 seconds were chosen as the initial default time for all observations, regardless of moon conditions, airmass, etc. in order to facilitate initial algorithm development. Varied exposure length will be explored in future work.

[5]Tracking analysis was also being performed so plate-solving and other processing added to the inter-exposure time. Normal operating conditions typically have < 10 seconds between exposures.

Table 3.3. New Moon Background Counts at ISO 100

Property	R Channel	G Channel	B Channel
Mean of Median [ADU]	40.77	65.89	41.83
Mean of RMS Median [ADU]	11.15	12.61	11.57
Per Pixel Value [ADU / sec]	0.34	0.55	0.35
Per Pixel Value [e^-] / sec	0.58	0.93	0.59

Note. — Values reported for one camera only

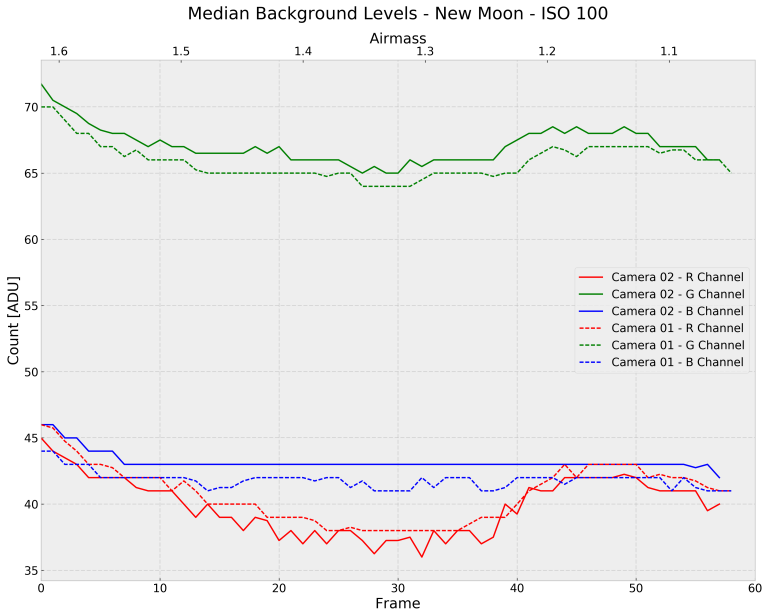

Figure 3.4: Background count values shown per color channel for each of the two cameras across 60 frames. The two cameras generally agree with each other within ±1 ADU for each of the channels.

3.2.3 Image Quality

The FWHM of the PSF on the image is about 2.5 pixels in the center of the image and goes toward 5.0 pixels in the corners (see Figure 3.5). A sample 120 second exposure at ISO 100 can be seen in Figure 3.6 taken from the same observation sequence as used for the background calculations above.

Despite the quality seen in the data coming from the edges of the image, vignetting remains an issue for any image with such a large FOV. The data processing algorithm (described in chapter 5) will automatically select around such stars when determining reference stars for a given target, but the potential reduction of the number of available stars will naturally result. Means of mitigating issues with vignetting will be explored as part of future PhD work.

Saturation at 120 seconds occurs for stars brighter than $m_V \sim 8.0$ while stars at $m_V \sim 12.5$, while

detectable, have a SNR $\sim 5 - 10$.

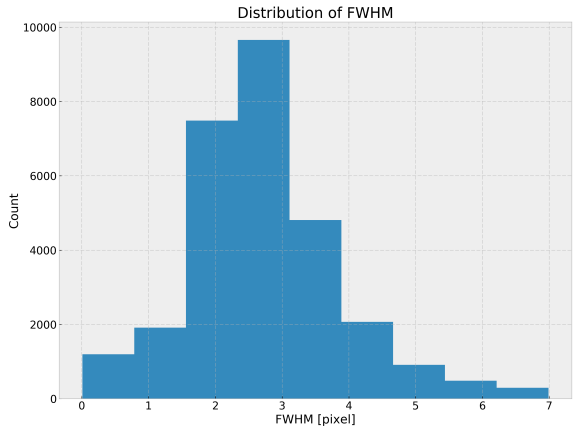

Figure 3.5: The FWHM distribution in a sample image as measured by `sextractor`.

Figure 3.6: Sample image from the Canon EOS100D. Image here is the JPEG extracted from the CR2 file of a 120 second exposure at ISO 100 and thus features interpolation of the pixels but is representative of the field of view and numbers of stars per image. The lower left corner has been zoomed to show aberrations.

3.3 Alternative Hardware Considerations

PANOPTES was designed to be hardware agnostic in terms of the mount and camera used (see chapter 4 for descriptions of the software). While it is important to have a baseline design that is centered around a specific brand and model, it is envisioned that there will be many groups that will want to utilize different hardware configurations. This could be either as a means of putting existing hardware to use, as a means of testing new hardware, or because of a specific science-case need.

While it is envisioned that these hardware differences will be mostly under the purview of PANOPTES, an extreme test of the software flexibility came with the selection of POCS as the primary control software for the Huntsman Telephoto Array[6]. While the underlying software remained the same (with the addition of new hardware drivers), the mount and the cameras were replaced with completely different makes and models to satisfy a completely different science case. POCS will serve as the primary control system when Huntsman is installed at Siding Springs Observatory mid-2017.

For PANOPTES, the different cameras that could potentially be used might pose a problem for the data processing. However, because differential photometry is used this issue can to a large extent be mitigated. See chapter 5 for details.

3.4 Hardware Issues

3.4.1 Tracking Issues

In late 2016 it was discovered that the mount was exhibiting a tracking error that was initially thought to be due to a fault on the RA motor control board. To diagnose the tracking errors a single target was observed for three consecutive nights in December 2016. Tracking errors were measured over the course of observations that were nearly three hours in length. The starting hour angle at which the target was observed each night was comparable and the only difference for each observation was the sidereal offset rate, as described below.

The iOptron mount command set allows for a tracking rate adjustment as a multiple of the sidereal rate in the range $[0.99 - 1.01]$ with increments of 0.0001. Because it was suspected that the offset rates were having no effect, the two extreme offsets (0.99 and 1.01 $\times$ sidereal) were tested against the standard sidereal rate. As suspected, no change was seen in the tracking error despite the different rates. All three observations resulted in a $\Delta_{\text{pixel}} = -30.0$ offset from sidereal after 120 minutes of

[6] https://www.facebook.com/AstroHuntsman/

observation regardless of the rate adjustment tested, as seen in Figure 3.7. Declination experienced negligible drift at the same time.

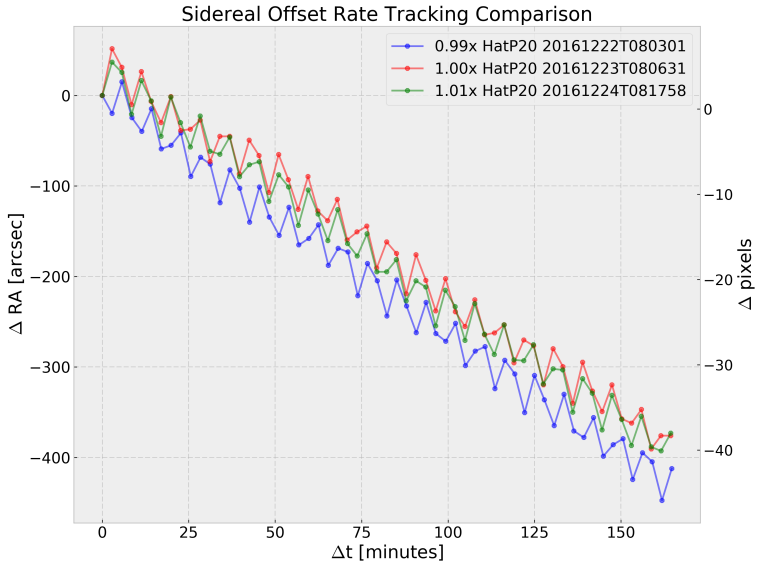

Figure 3.7: RA tracking error across the duration of an observation. Colors indicate offsets from sidereal rate and the legend refers to the start time of observations. High frequency oscillations are due to periodic error in the RA motor, which is almost 100 arcseconds peak-to-peak. Plot is of HAT-P-20 on successive nights at roughly the same hour angle but with different tracking rates. No discernible change in the tracking error is seen despite the different rates, indicating the tracking offsets weren't having any effect. See text for details.

With a pixel scale of 10.3 $''/_{\text{pixel}}$ we should expect a tracking rate of 0.99× sidereal to be approximately -100 pixels behind the true position after an observation of 120 minutes. Similarly, at 1.01× sidereal this value should be over 100 pixels ahead of the true position. The two extremes, therefore, should have an offset that is ∼ 200 pixels from each other, a difference that would be noticeable but was not seen. Offset rates between the two extremes were also tested with no noticeable effect. The conclusion from this analysis was that tracking rate adjustments were having no effect on the mount whatsoever, an issue we continue to diagnosis with iOptron. A closed-loop solution to tracking was later developed, described further below.

3.4.1.1 Drift Alignment Tests

The inability to alter the tracking rate does not indicate the source of the tracking error. To explore the source of drift a set of scripts was created to easily measure the polar misalignment, measured in altitude and azimuth coordinates, using the common drift alignment method. Note that this procedure

is still performed with an equatorial mount, you are simply measuring how far off from the celestial pole your mount is using alt-az directions. To test the altitude misalignment a star near the galactic equator is tracked from a point near the horizon until it begins to drift as measured on the declination axis. To test the azimuthal misalignment a star near the galactic equator is tracked but this time near the meridian and again the drift of the declination axis is monitored. The scripts simply acquired the appropriate stars and then tracked for a length of time, plotting the results for inspection.

The results from two separate tests can be seen in Figure 3.8, both of which show a large and sudden pixel displacement while the mount is tracking with the cameras pointed toward the western horizon. Of note, however, is that this displacement is occurring at different times. Indeed, this test was performed a number of times on different nights with each test showing the same sudden displacement at random times. The best diagnosis for this was that the unit was experiencing some slippage due to gravity and faulty construction. Without additional on-site examination, however, the issue proved difficult to explore further, highlighting some of the inherent complications of PANOPTES.

A site visit to attempt a fix for this issue is described below.

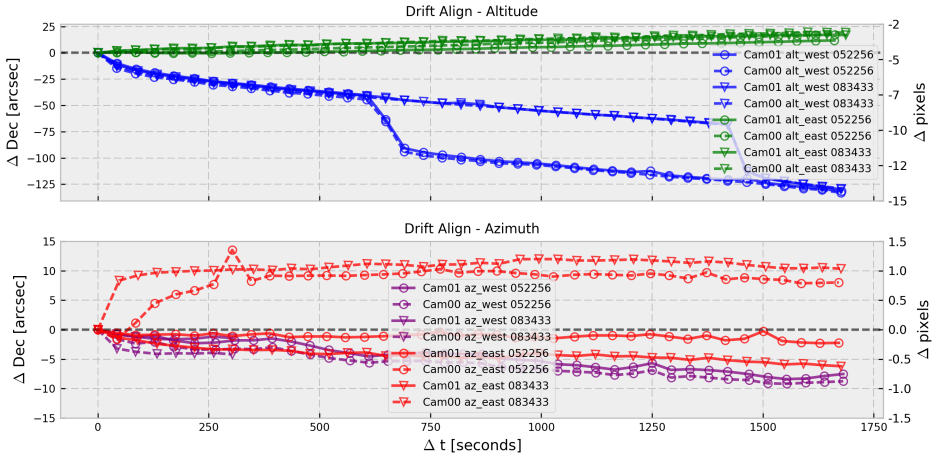

Figure 3.8: An example of a drift alignment analysis performed on a remotely operated unit. The top panel represents a mis-alignment in the altitude of the mount. The bottom panel shows the azimuthal mis-alignment with respect to the pole. Note the different scales on the y-axis. The top panel indicates slippage in the mount, which could not be fixed without an on-site visit. See text for details.

3.4.1.2 Polar Alignment Tests

A comparison of the rotation axis of two images was also performed. For the first image the camera is held fixed while pointing its alignment toward the celestial pole while allowing the celestial sphere

to rotate, whereas in the second image the declination axis is fixed and the mount is rotated on the RA axis. The center of rotation from the first image can be compared to the center of rotation from the second image, giving an altitude offset between the two images (see Figure 3.9).

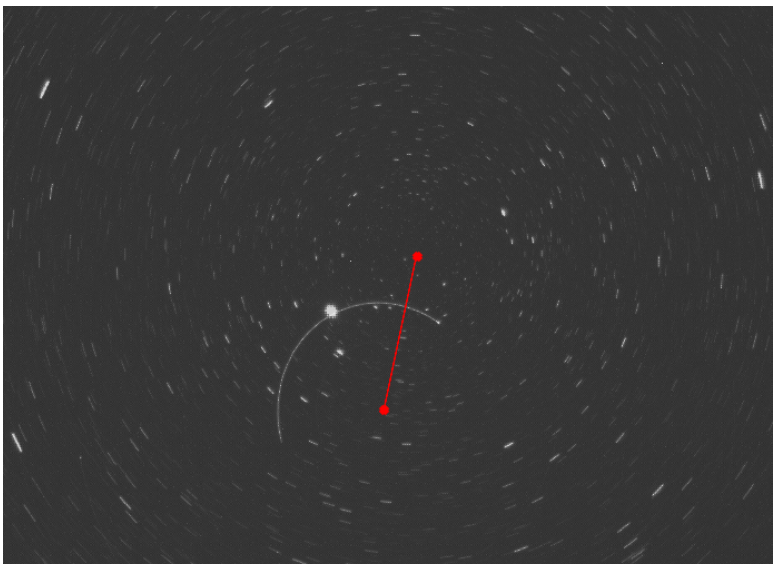

Figure 3.9: An overlay of the rotation around the RA axis and the celestial sphere. Red line indicates the offset between the two centers of rotation, indicative of a mis-alignment in altitude. The image is generated via a script that allows individuals to easily determine the mis-alignment present in their unit. See text for details.

While the misalignment would still require manual intervention in order to fix, scripts were developed in order to automatically determine the mis-alignment in order to make it easier for the user to adjust their unit. While it is envisioned that most PANOPTES units will be set up once and left for a period of time, and therefore not require constant attention paid to the alignment, it is nevertheless important to be able to remotely diagnose alignment issues and to facilitate the alignment process for users who may not experienced amateur astronomers.

The first version of the script required at least a five minute exposure about the celestial sphere in order to create arc-lengths from the star trails that could be automatically measured. After the site visit in March 2017, where further alignment issues presented themselves, this script was improved so that a 30-second image of the celestial pole could be taken, which was then plate-solved in order to determine the pixel location of the actual celestial pole. The later version of the script was incorporated into the administrative website so that a user could push a single button per iteration and obtain a reliable measurement of the misalignment (a screenshot can be seen in the appendix, Figure A.1).

The result of the new script is a procedure that can help diagnose large misalignments, takes

approximately two minutes per iteration and performs reliably each time. In a test of 20 iterations of the polar alignment script, the altitude displacement was found to be measured to within a 10 pixel RMS, which corresponds to about 60 arcseconds.

3.4.1.3 Site Visit - Attempt at a fix

A site visit was made to Mauna Loa in March of 2017 in order to attempt an alignment fix as well as to install new electronic components and upgrade some of the weatherproofing material (described in chapter 3). The initial plan was to replace the entire mount in order to overcome the faulty RA electronics board described above. However, a measurement of the tracking rates in the replacement mount was first made in the labs at the Subaru Telescope base facility using large precision calipers to measure the displacement of the mount using various sidereal offset rates as described above. While the measurements weren't entirely precise, they were accurate enough to cast doubt on the fact that replacing the mount would solve our tracking issues. Since replacement of the mount is not a simple procedure, it was decided to forgo the replacement and instead focus on the closed-loop tracking solution described below. Whether the tracking rate offset complications are due to faulty motor boards, which seems unlikely in two mounts ordered at two separate times, or whether they are due to a combination of software and firmware issues is still being explored.

Unfortunately, while installing the new electronics and weatherproofing material, the tracking on the mount was inadvertently made much worse. After I left Hawai'i it was discovered that the altitude on the mount had been altered, probably due to physical handling of the mount while reinforcing the weatherproofing material. This resulted in an altitude offset of approximately 450 pixels, which is over 1.25 degrees, a significant misalignment. Naturally, this has had drastic effects on the pointing and tracking model on the mount, as shown in Figure 3.10.

3.4.1.4 Closed Loop Tracking

The iOptron iEQ30 command set allows for fixed-duration directional commands to be sent to the mount, which are applied while tracking. These commands tell the mount to move in either the positive or negative direction for both the RA and Dec axis. The commands are given in terms of millisecond pulses in the N-S-E-W directions, with the mount moving at custom rate for the specified number of milliseconds in the commanded direction.

By plate-solving each raw image immediately after it has been taken, an offset in arcseconds from the original position can easily be obtained. Converting this arcsecond offset into the appropriate

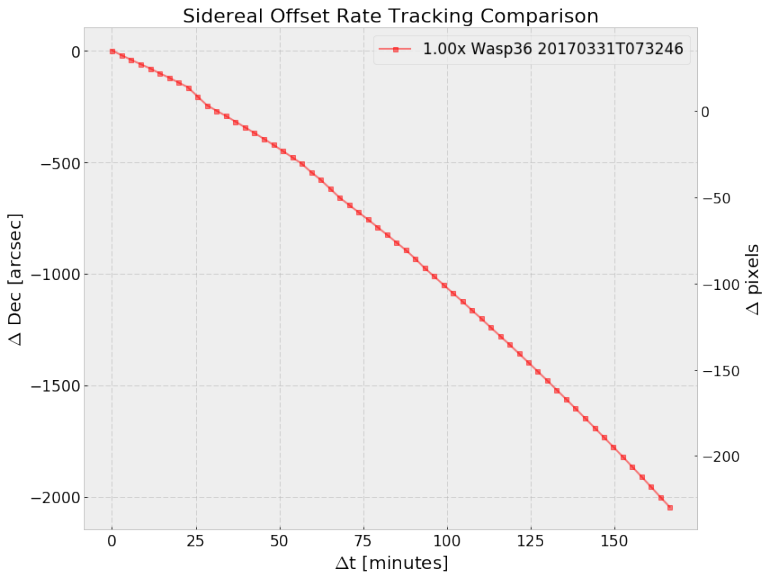

Figure 3.10: A site visit to `PAN001` in March 2017 inadvertently introduced an extremely large polar misalignment, causing extreme drifts in the Declination axis.

number of milliseconds at the given rate allows for the mount to be adjusted even while the next exposure has started. In order not to create visible movement while exposing the specified rate is set at 0.5× sidereal rate.

Implementing this closed-loop solution has led to drastic improvements in the tracking errors despite the extremely large polar misalignment, the results of which can be seen in Figure 3.11. The Declination drift has been eliminated to within one pixel and the RA error has been reduced to that of the periodic error (PE) present on the worm gear, with a peak-to-peak error of approximately 100 arcseconds. Further improvements to the closed-loop tracking model are planned to reduce the PE.

3.4.2 Cable / USB Issues

USB cable connections in the camera box have been found to be unreliable unless firmly secured. Frequent disconnections with each of the cameras was experienced for the first few months of operation, often leading to the cameras being placed in a locked state that required manual power cycling. Significant time was spent trying to remotely power-cycle the cameras via the USB subsystem on the computer, all of which failed.

Fortunately, since the manual power-cycling is a simple process it could be performed by an on-site MLO staff member who was otherwise unrelated to PANOPTES. However the frequency of the disconnects made it obvious that a way to remotely power-cycle each of the subcomponents would be

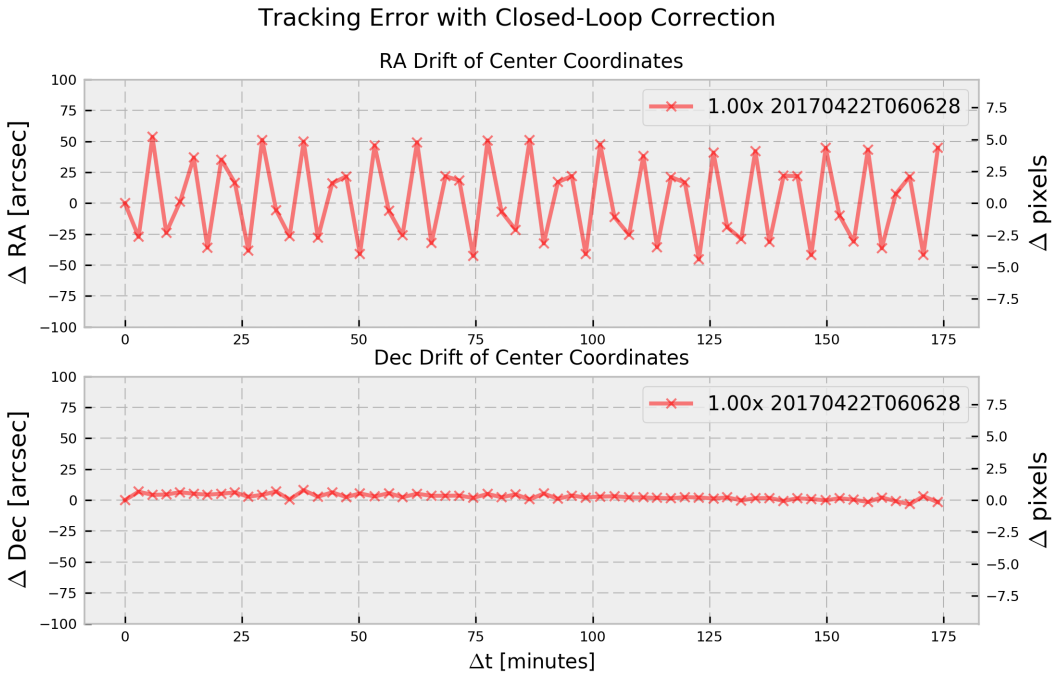

Figure 3.11: The RA and Dec tracking errors with a closed-loop solution implemented using the DSLR cameras. Each image is plate-solved and fixed-length offset positions are sent back to the mount. Periodic error on the RA axis is still present but the overall error for the RA has been reduced to the PE peak-to-peak error. Declination drift has been essentially eliminated. See text for details about the closed-loop solution.

a required upgrade for the electronics boards, leading to the eventual design of the power distribution board described in section 3.1.1.

4

Software

PANOPTES is driven primarily by an open source model with regards to both software and data. In addition, the goal has always been to create an entire "scientific platform" that could be used for educational purposes. Because modern scientific practices usually include some component of software development (Ayer et al., 2014; Wilson et al., 2014), one of the early design goals for the project was to have a software base that was not only completely open source but also easily readable (and modifiable) by individuals just learning how to program. Importantly, the software must also be accurate and robust in order to handle long-term unattended and remote operations and must also be customizable based on potential hardware differences between units (described earlier in section 3.3).

The PANOPTES code repositories[1] can therefore be used in two different contexts: automatically by installed hardware units responsible for collecting data; and interactively by developers, including students and amateurs, who can choose to modify existing operations, add new functionality, or simply use the software as a learning tool for astronomy and software development. The two distinct but equally important uses of the software, one as an automatic observatory control system (OCS) for data collection and the other as a tool for learning, place unique constraints on the decisions made regarding software. I have designed and written all software with these overarching goals in mind.

The design of the software system is first discussed in section 4.1, including an overview of the state machine for overall systems control and the hardware abstraction layer. An overview of the choices made for the software ecosystem is given in section 4.2. A brief section outlining the reasons for not pursuing existing OCS solutions follows in section 4.3.

[1] https://github.com/panoptes

4.1 Design

In order to create a robust OCS a number of specific design features were identified. While running, the chief factor of importance is the ability of the software to always act in an appropriate manner at the appropriate time, especially where safety of the unit (due to weather) is concerned. While developing, a key factor is the flexibility of the software and the ease of development for alternative hardware implementations. The latter condition is satisfied by a well-designed hardware abstraction layer (HAL) while a finite-state machine (FSM) was implemented for the former. The HAL and the FSM are discussed in their appropriate context below.

4.1.1 PANOPTES Observatory Control System (POCS)

The PANOPTES Observatory Control System (POCS)[2] is the primary software responsible for running a PANOPTES unit. `POCS` is implemented as a finite state machine (described below) that has three primary responsibilities: overall control of the unit for taking observations; relaying messages between various components of the system; and determining the operational safety of the unit.

`POCS` is designed such that under normal operating conditions the software is initialized once and left running from day-to-day, with operation moving to a `sleeping` state during daylight hours and observations resuming automatically each night when `POCS` determines conditions are safe.

The POCS repository contains the overall bulk of the software written for Project PANOPTES and for which I have been chief architect, developer, and maintainer. The repository has comprehensive continual integration testing and thorough documentation.

POCS is implemented as four separate logical layers, where increasing levels of abstraction take place between each of the layers. These layers are the low-level Core Layer, the Hardware Abstraction Layer, the Functional Layer, and the high-level Decision Layer (see Figure 4.1).

4.1.1.1 Core Layer

The Core Layer is the lowest level and is responsible for interacting directly with the hardware. For DSLR cameras this is accomplished by providing wrappers around the existing `gphoto2`[3] software package. For PANOPTES, most other attached hardware works via direct RS-232 serial communication through a USB-to-Serial converter. A utility module was written for common `read`/`write`

[2] https://github.com/panoptes/POCS
[3] http://www.gphoto.org/

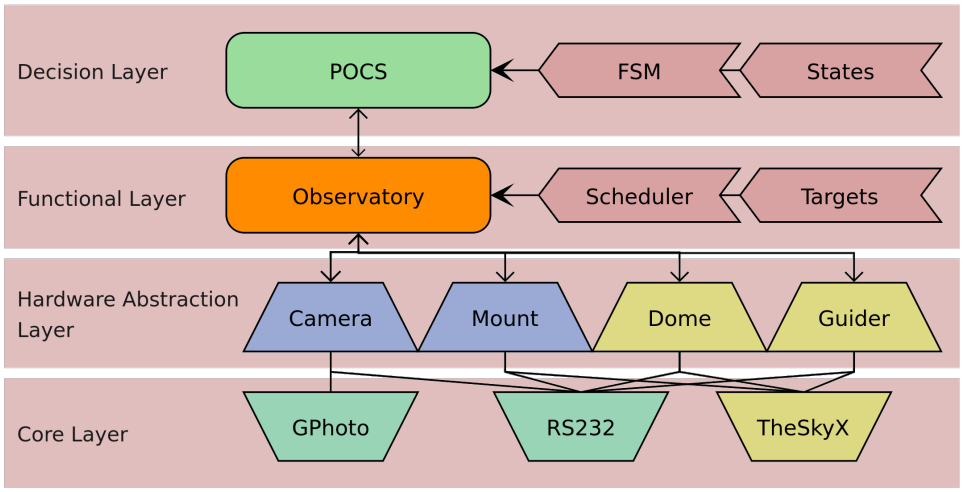

Figure 4.1: Diagram of POCS software layers. Note that the items in yellow (Dome, Guider, and TheSkyX) are not used by PANOPTES but are used by the Huntsman Telephoto Array, which also uses POCS for control (see section 3.3. They are included in the diagram as a means of showing the flexibility of the Functional Layer to interact with components from the HAL.

operations that automatically handles details associated with buffering, connection, etc. Support for TheSkyX was written into POCS for the Huntsman Telephoto Array (see section 3.3). The overall goal of the Core Layer is to provide a consistent interface for modules written at the HAL level.

4.1.1.2 Hardware Abstraction Layer (HAL)

The use of a HAL is widespread both in computing and robotics. In general, a HAL is meant to hide low-level hardware and device specific details from higher level programming (Elkady & Sobh, 2012). Thus, while every camera ultimately needs to support, for instance, a `take_exposure(seconds=120)` command, the details of how a specific camera model is programmed to achieve that may be very different. From the perspective of software at higher levels those details are not important, all that is important is that all attached cameras react appropriately to the `take_exposure` command.

While the Core Layer consists of one module per feature, the HAL implements a Template Pattern (Gamma et al., 1993) wherein a base class provides an interface to be used by higher levels and concrete classes are written for each specific device type. For example, a base `Mount` class dictates an interface that includes methods such as `slew_to_home`, `set_target_coordinates`, `slew_to_target`, `park`, etc. The concrete implementation for the iOptron mount then uses the Core Layer level RS-232 commands to issue the specific serial commands needed to perform those functions. Likewise, a Paramount ME II concrete implementation of the `Mount` class would use

the Core Layer interface to TheSkyX[4] to implement those same methods. Thus, higher levels of the software can make a call to `mount.slew_to_target()` and expect it to work regardless of the particular mount type attached.

Another advantage of this type of setup is that a concrete implementation of a hardware simulator can be created to test higher-level software without actually having physical devices attached, which is how much of the PANOPTES testing framework is implemented[5].

4.1.1.3 Functional Layer

The Functional Layer is analogous to a traditional observatory: an `Observatory` has a `location` from which it operates, attached hardware which it uses to `observe`, a `scheduler`[6] to select from the available `target_list` to form valid `observations`, etc.

The `Observatory` (i.e. the Functional Layer) is thus where most of the operations associated with taking observations actually happen. When the software is used interactively (as opposed to the usual automatic mode) it is with the `Observatory` that an individual would overwhelmingly interact.

The Functional Layer is also responsible for connecting to and initializing the attached hardware, specified by accompanying configuration files. The potential list of targets and the type of scheduler used are also loaded from a configuration file. The particular type of scheduler is agnostic to the `Observatory`, which simply calls `scheduler.get_observation()` such that the external `scheduler` can handle all the logic of choosing a target. In Figure 4.1 this is represented by the "Scheduler" and "Targets" that are input to the "Observatory."

4.1.1.4 Decision Layer

The Decision Layer is the highest level of the system and can be viewed as the “intelligence” layer. When using the software in interactive mode, the human user takes on the role of the Decision Layer while in automatic operations this is accomplished via an event-driven finite state machine (FSM).

A state machine is a simple model of a system where that system can only exist in discrete conditions or modes. Those conditions or modes are called states. Typically states determine how the system reacts to input, either from a user or the environment. A state machine can exist solely in the

[4] http://www.bisque.com/sc/pages/TheSkyX-Professional-Edition.aspx

[5] Writing hardware simulators, while helpful for testing purposes, can also add significant overhead to a project. For major projects such as the LSST or TMT this is obviously a requirement. PANOPTES implements basic hardware simulators for the mount and camera but full-scale hardware simulation of specific components has not yet been achieved.

[6] A modified dispatch scheduler (Denny, 2004) in the case of PANOPTES

software or the software can be representative of a physical model. For PANOPTES, the physical unit is the system and POCS models the condition of the hardware. The "finite" aspect refers to the fact that there are a limited and known number of states in which the system can exist.

Examples of PANOPTES states include: `sleeping`, which occurs in daylight hours, the cameras are facing down, and the mount is unresponsive to slew commands; `observing`, where the cameras are exposing and the mount is tracking; `scheduling`, where the mount is unparked, not slewing or tracking, it is dark, and the software is running through the scheduler, etc. PANOPTES states are named with verbs to represent the action the physical unit is currently performing.

POCS is designed to have a configurable state machine, with the highest level logic written in each state definition file. State definition files are meant to be simple as most of the details of the logic should exist in the functional layer. Students using POCS for educational purposes will most likely start with the state files. An example of the `observing` state can be seen in the section A.3.

State machines are responsible for mapping inputs (e.g. `get_ready`, `schedule`, `start_slewing`, etc.) to outputs, where the particular mapping depends on the current state (Lee & Seshia, 2017). The mappings of input to output are governed by transition events[7]. A diagram of the main observation loop for the current state machine used by POCS is shown in Figure 4.2.

State definitions and their transitions are defined external to `POCS`, allowing for multiple possible state machines that are agnostic to the layers below the Decision Layer. This external definition is similar to the "Scheduler" in the Functional Layer and is represented similarly in Figure 4.1.

`POCS` is responsible for determining operational safety via a query of the weather station, determination of sun position, etc. The transition for each state has a set of conditions that must be satisfied in order for a successful transition to a new state to be accomplished and a requisite check for operational safety occurs before all transitions. If the system is determined to be unsafe the machine either transitions to the `parking` state or remains in the `sleeping` or `ready` state[8]. A unit spends daylight hours in the `sleeping` state and waits in the `ready` state when dark but otherwise unsafe.

4.2 Ecosystem - Python

Python was chosen as the primary language because of its ubiquity within the sciences in general (Millman & Aivazis, 2011) and astronomy in particular (Robitaille et al., 2013). The ease of use

[7]The Python FSM used by POCS is in fact called *transitions* (https://github.com/tyarkoni/transitions). During the course of this thesis I became an open-source collaborator for the *transitions* module.

[8]A special state originally named 'waiting' that was later changed due to confusion as to the state's purpose

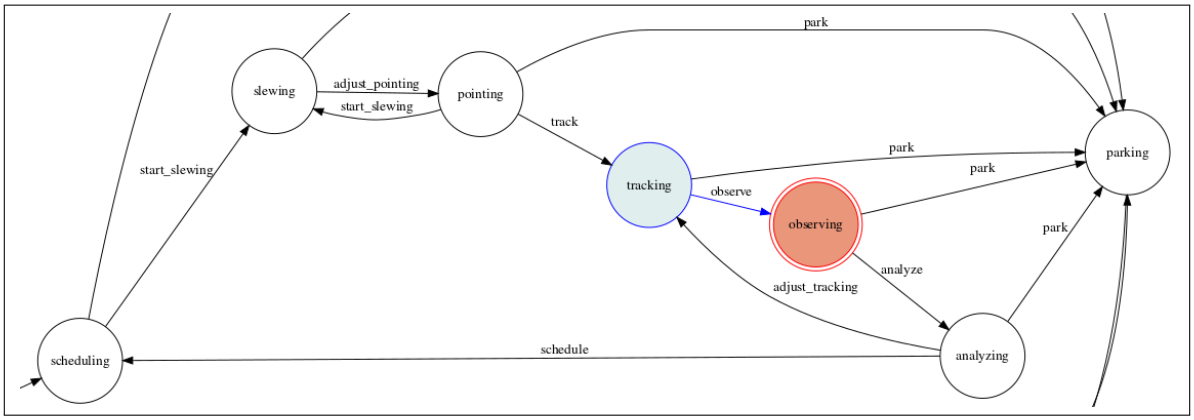

Figure 4.2: A close up of the main loop of POCS State machine showing a currently active `observing` state with a transition from `tracking`. States are represented by circles and the transition input is indicated by the labeled arrows between states. Visual representation of the state machine is updated in real-time and displayed to the user via an administrative web interface. See text for details.

by beginners and non-programmers as well as the availability of powerful open-source libraries for scientific computing (notably the SciPy ecosystem) make it an ideal choice for PANOPTES, which hopes to introduce students to the subject of scientific computing. The free and open-source Astropy library has also matured to the point of becoming a standard tool for astronomers (Robitaille et al., 2013) and has greatly facilitated PANOPTES development.

Along with the decision to use Python came several decisions related to the software ecosystem, such as choice of documentation style, coding standards, and version control. All of the decisions concerning this ecosystem were made with an eye toward increasing openness and boosting participation. As mentioned in the introduction, education and outreach are just as important to the overall project as scientific output and one of the main ways students (and others) will be involved with the project is via the software.

The Anaconda Distribution is used to ensure a consistent Python setup across different machines. Anaconda is described as "an easy-to-install free package manager, environment manager, Python distribution, and collection of over 720 open source packages offering free community support" (Continuum, 2017b). Furthermore, Anaconda promotes the use of Jupyter Notebooks as an effective tool for reproducible research (Continuum, 2017a), an increasingly important concept in modern astronomy (Peng, 2011) (see also section A.2).

4.3 Alternatives

A primary software adage is to avoid "recreating the wheel" and while automated OCS systems are not uncommon, an initial review found that none of the available systems were suitable to the PANOPTES goals outlined above. First, all software that required license fees or was not otherwise free (of cost) and open (to modification) was automatically rejected. Second, software was examined in terms of its ability to handle the hardware and observing requirements of a PANOPTES unit. Third, the ease-of-use of the software was determined, both in terms of installation and usage as well as in ability to serve as a learning tool. Popular alternatives to POCS were evaluated and ultimately dismissed primarily for not conforming to the one of the above three requirements. Alternatives to POCS are described at http://www.projectpanoptes.org/alternative_software.html.

5

Data Processing

The PANOPTES Image Analysis Algorithm (PIAA) performs relative photometry on a target point source by doing a frame-by-frame comparison of the target with an ideal template that is built from other point sources in the field. The algorithm takes advantage of the fact that, given a wide FOV and high stellar density, a significant number of point sources will experience systematic errors in a similar fashion. These errors could include common systematics such as optical differences due to airmass or clouds, detector non-linearity[1], or any other number of known or unknown defects. Furthermore, because DSLR cameras exhibit strong color effects due to the presence of the Bayer array, an additional class of systematic effects must be accounted for at the pixel level before any photometry can be achieved.

By building an ideal reference PSF from an ensemble of similar point sources, differential photometry can be performed for each frame in order to determine if the target differs from the ideal template. Similar schemes aimed at removing systematic errors from wide-field surveys have demonstrated success (Tamuz et al., 2005; Kovacs et al., 2004; Bakos et al., 2010). The examples in the literature, however, have all used traditional astronomical CCDs where the selection of the reference point sources isn't as critical[2]. The careful selection of the reference point sources that demonstrate those same color effects imposed by the Bayer array, discussed below, is therefore a critical step in the processing.

The basic steps of the algorithm can be described as follows, with more specific details given in

[1] DSLR cameras show a highly linear response but local discrepancies may still be present. See section 3.2.2 for details.

[2] Kovacs (2004) picks "almost random" stars from the full field while Tamuz (2005) uses all sources in a simulated field. Bakos (2010) likewise merely states that "certain other stars" must undergo the same variations. Zhang et al. (2016), who similarly process DSLR images, use 11 stars in the field that are manually selected via visual inspection and analysis. Furthermore, Zhang et al. (2016) process at the light-curve level after performing aperture photometry, which is dissimilar from the approach PANOPTES will use.

the sections to follow:

1. Calibrate raw frames by removing the bias, applying the flat-field frames, and performing background subtraction.
2. Identify all point sources in initial frame and match with PANOPTES Input Catalog (PIC).
3. Choose a target point source that will be examined for transit candidates.
4. Create a postage stamp cube (PSC) around each point source that matches the width and height required to encompass the target point source and includes all frames from the observation.
5. Select reference PSCs by comparison of the pixel-level intensity distributions across all frames with those of the target PSC; this yields an ensemble of references that experience systematics similarly to the target and have the same PSF shape and distribution across the colored pixels.
6. Construct an "ideal" PSC template through a linear combination of the reference PSCs.
7. Compare each frame of the target PSC with the corresponding frame from the template PSC in order to determine how target has changed relative to the ensemble.

Note that the full algorithm for PANOPTES will repeat steps 3-7 for all point sources in a field, the challenges of which are non-trivial but are not part of the current work.

5.1 Implementation

5.1.1 Step 1: Image Calibration

Each frame in an observation is initially plate-solved using the open-source `solve-field` program from `astrometry.net` software suite (Lang et al., 2009). Plate-solving is done in an iterative procedure similar to Zhang et al. (2016) and results in a world coordinate system (WCS) for the frame. The WCS of the first frame is used to identify point sources in the image (see below).

Built-in camera bias is subtracted from each raw image before all other calibration. Raw images are stored with the bias and all calibration and analysis is performed on a copy of the raw data. After the initial subtraction of the camera bias from the data and the division of a master flat frame, background estimation is performed on each frame separately. The master flat frame is built from twilight flats that are taken periodically by the unit, rather than during each nightly operation.

Background subtraction is more complicated compared to normal astronomical photometry due to the RGB Bayer array and the strong pixelization that occurs. To account for this, a separate mask

is created for each of the three color channels so that the background can be calculated per channel[3]. Due to the large FOV ($10° \times 15°$) this background is not represented as a single value but instead is created as a 2D map that varies across the image, an example of which can be seen in Figure 5.1. Clearly present in the image is the strong vignetting that exists with the camera and lens combination.

An alternative to the grid pattern would be to perform local background subtraction within each of the PSCs but it was found that this offered no advantage while significantly increasing the computation time (similar to the small box size option described below).

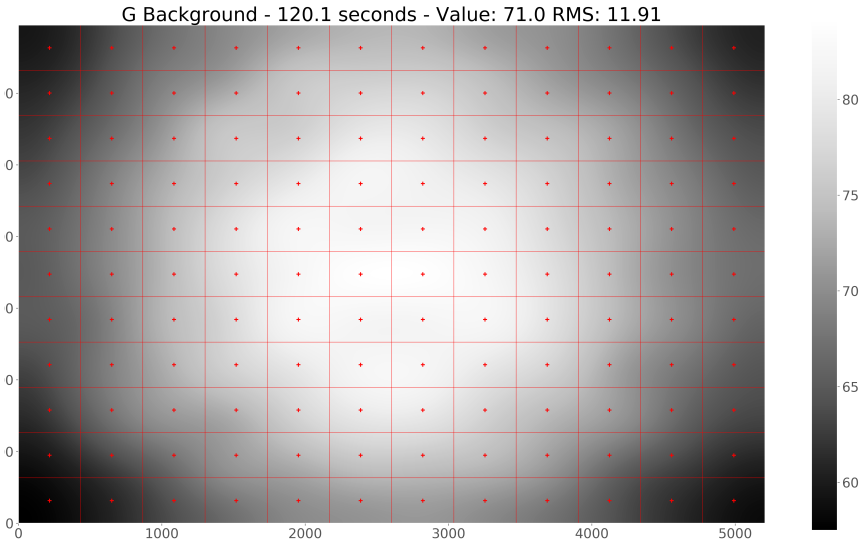

Figure 5.1: A 2D background map for the G channel taken from a 120 second exposure at ISO 100. The map is representative of the other channels although with different values and represents the image before flat-fielding is applied. The red grid corresponds to the mesh grid used to compute the background. See text for details.

To calculate backgrounds, I use the `Background2D` function from the Astropy affiliated *photutils*[4] library and so adopt their terminology. This process creates a grid of boxes across the full frame and calculates the background within each box separately using sigma-clipped data. Sigma-clipping removes point sources from the box by identifying pixel values that are more than 3σ from the median and ultimately masking their values from the calculation of the background value. This clipping procedure is iterated five times per box.

The `Background2D` documentation recommends box sizes large enough to encompass individual point sources while also being small enough to capture local background variations. Boxes should

[3]One color channel is created for all green pixels, unlike Zhang (2016) where the G1 and G2 pixels are processed separately.

[4]https://photutils.readthedocs.io

be an even integer multiple of the size of the image. The Canon EOS 100D has image dimensions of 5208×3476, leading to box sizes of 434×316, creating a grid of 12×11 boxes, as shown in Figure 5.1.

After sigma-clipping of the data the background value for each box is determined using the `MedianBackground` estimator, which uses a simple median to derive the value for the box. The values for each of these boxes therefore creates a low-resolution 2D map of the background which is interpolated to the original size of the frame using the default `BkgZoomInterpolator` option, which performs a spline interpolation in order to generate the full-resolution background. The result is a smoothly varying map for the full frame, which is subtracted from the data after removing the bias and performing the flat-fielding.

Variations on the number of iterations (5, 7, and 10), sigma level (3, 4, and 5), box size (11x12 and 316x434 pixels), moon fraction (0.0 and 0.77), and estimator type (the default SExctractor algorithm, a mean, median, and the DAOPHOT MMM algorithm) were performed in order to determine the correct configuration to use for the `Background2D` function. Computation times between the different combinations varied by up to 20 seconds. Derived background pixel values for all combinations were in agreement to within 5 counts with similar RMS values for all combinations. The box size of 434×316 with a 3σ clip and 5 iterations along with a simple median estimator was therefore chosen in order to optimize computation time.

Each frame takes about 60 seconds to complete background subtraction with standard observations consisting of at least 60 frames. Each observation therefore takes about one hour just to process the background. A typical night can see up to four complete observations resulting in four hours of processing spent on background subtraction. Since it is desired to perform all initial data processing on the individual units, it is important to keep a time budget that allows for full processing of the data during the daylight hours. Background subtraction, and all other steps described here, have not yet been optimized for time.

5.1.2 Step 2: Identify Point Sources

Point sources are extracted from the first frame of each observation using the `sextractor` software package to determine the J2000 RA and Dec coordinates (Bertin & Arnouts, 1996). The RA/Dec coordinates are used during step 4 along with the frame-specific WCS in order to determine `XY` pixel coordinates for the frame. During aperture photometry performed as part of the last step, local peak detection is used in order to determine the centroid around which to place the aperture. The RA/Dec

coordinates, therefore, are merely used to build the stamp and need not have sub-pixel accuracy.

Identification of the point sources is a relatively quick procedure and adds negligible overall time to the algorithm as it must only be performed once per observation.

5.1.3 Step 3: Target Selection

As mentioned in the beginning of the chapter, full image processing will treat each point source as a potential target and thus the algorithm will be run a number of times per observation sequence[5]. This specific step doesn't take any time per se but in normal operations this will be the first step of the iteration over all potential targets.

5.1.4 Step 4: Construct Postage Stamp Cubes (PSC)

A Postage Stamp Cube (PSC) is created for each identified point source in the image. The PSC is constructed after the selection of the target and the target point source dimensions are used to construct the remaining PSCs. The dimensions for each PSC are `width` × `height` × `num_frames`. The location of each stamp is based of the `XY` pixel coordinates derived from the WCS for each frame, with the RA/Dec coordinates being identified as per step 2.

Care must be taken when constructing the reference stamps such that all stamps have identical RGB indexing and size as that of the target. That is, if the target stamp has a red pixel in the upper-left corner, so too must all the reference stamps. This is accomplished by using the same index slicing as for the target source but then shifting the bounds up or left by one pixel depending on the color of the upper-left pixel of the reference stamp. Again, if this causes the point source for the reference to fall on a different pixel index as that of the target these stamps will be automatically discounted below.

The initial target stamp is created via the `XY` pixel locations obtained from `sextractor` during the identification of point sources in step 2. The minimum and maximum locations for both the `X` and `Y` pixels across all frames of the observation are used to determine the size of the stamp. These values are padded with three (3) additional pixels in order to obtain a `width` and `height` for the the stamp that can adequately capture a suitable aperture placed around the point source. Minimum and maximum pixel locations are used in order to capture potential drift of the point source across the detector[6]. A typical stamp size has a width and height on the order of 12×12 pixels. Periodic error

[5]Not all point sources will necessarily be used. The point sources to be treated as targets will be looked up from the PANOPTES Input Catalog. See section 2.2

[6]See tracking related issues in section 3.4.1

on the mount is around 10-pixels peak-to-peak, which is encapsulated within the stamp[7]. Because the stamp size is automatically determined, target sources experiencing drift will automatically have larger postage stamps created to encompass their total drift.

5.1.5 Step 5: Reference Selection

To pick the appropriate reference sources each frame of the PSC is normalized according to the sum of the values within the frame. The normalized value of a pixel is therefore simply the value of the pixel at a certain index divided by the sum of all the pixels within the stamp. This normalized pixel will have a value between $0 - 1$. This normalization of a pixel within a single frame of the PSC can be represented by:

$$\hat{p}_l(t_i) = \frac{p_l(t_i)}{\sum_{k=0}^{P} p_k(t_i)} \tag{5.1}$$

where l and k are the one-dimensional pixel index, i represents the time index for the given frame, and P is the total number of pixels in a stamp. Normalization of a frame means we lose flux information but retain the morphology for each frame. This morphology provides information about the sub-pixel placement of the PSF on the stamp, the implications of which are discussed below.

Note that the PSC is created based on a selected target and the stamp size is built around that target as described in the previous step such that all stamps will have the same index range. The stamp index is represented as a one-dimensional index, as opposed to a row-column indexing scheme, for ease of storage and computation (see Figure 5.2). Stamps are reshaped to their original row-column dimensions for display purposes.

A least squares comparison of the normalized pixel value with the corresponding target pixel is performed for each frame and the sum across all frames is used to rank the reference PSCs according to their overall similarity with the target. This is giving us a ranking of PSCs whose morphology differs the least with respect to the morphology of the target and as such we are trying to minimize:

$$V_j = \sum_{i=1}^{M} \sum_{l=1}^{P} \left(\hat{p}_{0,l}(t_i) - \hat{p}_{j,l}(t_i)\right)^2 \tag{5.2}$$

where V_j represents the variance with respect to the target, M is the total number of frames in the observation, P is the total number of pixels in the stamps, $\hat{p}_{0,l}$ refers to the normalized pixel value at index l for the target and $\hat{p}_{j,l}$ is the corresponding pixel for reference j. Lower values of V_j correspond

[7]Efforts to mitigate the periodic error as part of the closed-loop are underway

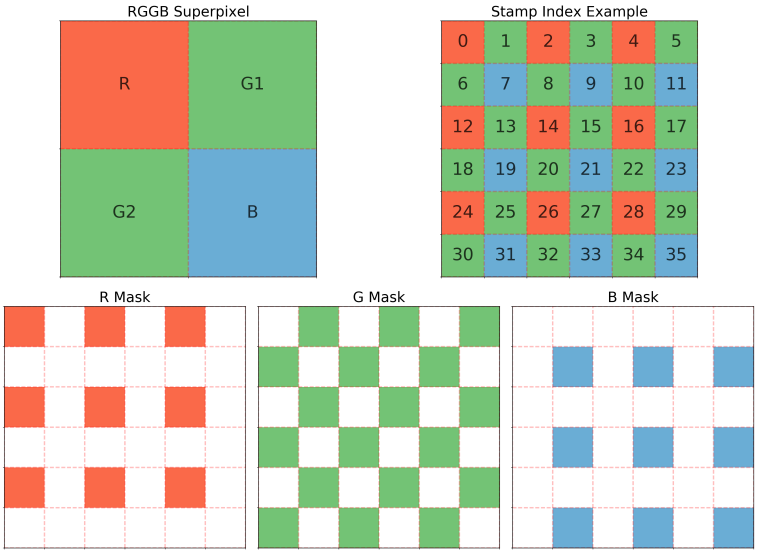

Figure 5.2: *Top left:* An example showing a single superpixel with the individual pixels labeled. *Top right:* A cartoon representation of a 6×6 postage stamp illustrating the one-dimensional indexing scheme. *Bottom:* Example of the array masks that are used to work with the data. Note that when creating the stamps there will not always be a red pixel in the upper left but all stamps will have a similar RGB layout.

to references with similar morphologies as that of the target, with a minimum value of zero given by a comparison of the target with itself (i.e. $\hat{p}_0 = \hat{p}_j$).

In effect, this procedure selects reference stars that have a postage stamp that has the same shape on a pixel-to-pixel level to that of the target star. This means that the reference star lies on the same sub-pixel location with respect to the Bayer array as the target star. It also implies that the PSF shape is similar. This is important because the process is therefore automatically selecting stars that exhibit the same systematic errors. For example, if a target is selected in the corner of the image it will display asymmetric aberrations (coma) that will distort the morphology of the PSF. The minimization above will automatically select those references that undergo the same morphological distortions.

References are sorted according to their V value and the top 100 references (those with the lowest V values) are used to build the template. An automatic way to determine how many references are optimal is left for future work.

5.1.6 Step 6: Template Construction

To construct an "ideal" reference template a linear combination of the selected references is created for each frame of the observation in order to determine a coefficient for each reference PSC. This

coefficient is applied to all pixels within the reference stamp for a given frame (before any aperture summation). The PSC coefficients for each frame are created by minimizing the difference between the linearly combined references stamps and the target stamp for every frame *except* the frame in question. The coefficients created from the other frames are then applied to the references in that frame. When constructed this way, the PSC coefficients for each frame are constructed without any knowledge of the frame itself.

The coefficients are created via a minimization between the target frame and the reference frames using the original (non-normalized) pixel values as follows:

$$C_0(t_i) = \sum_{a=1,a\neq i}^{M} \left(\sum_{l=1}^{P} \left(p_{0,l}(t_a) - \sum_{k=1}^{N} c_k p_{k,l}(t_a) \right)^2 \right). \tag{5.3}$$

where C_0 represents the minimized value for index 0 (the target index), M is the total number of frames, P is the number of pixels, N is the number of references used for comparison, $p_{0,l}(t_a)$ is the value for the lth pixel in the target stamp of frame a, and c_k is the desired coefficient for reference k.

The ideal reference for a given pixel r_l is then constructed for frame i using the set of coefficients obtained from the minimization above:

$$r_l(t_i) = \sum_{k=1}^{N} c_k p_{k,l}(t_i). \tag{5.4}$$

A template PSC can then be constructed using each l pixel in P for each i frame in M.

5.1.7 Step 7: Target Comparison

For each frame i in the set of M frames a simple comparison of the target to the ideal template is made. This results in a depth d that is the percent level by which the target has changed compared to the ideal reference for that target:

$$d_0(t_i) = \frac{\sum_{l=1}^{P} p_{0,l}(t_i)}{\sum_{l=1}^{P} r_l(t_i)} - 1 \tag{5.5}$$

The set of $d_0(t_i)$ values for the target given by index 0 corresponds to the relative flux.

5.2 Results

An image sequence was taken on April 22, 2017 in the Boötes constellation. From this field the star HIP 70617, $m_V = 10.24$ was analyzed. The star was located at XY coordinates `1765.5, 1075.8`,

placing it slightly offset from the center of the image. Due to hardware issues described above this led to a slight rotation, which manifests itself as a slightly larger postage stamp (15×8) required to encompass the star's drift across the detector. An example of the postage stamp for the target and corresponding template frame can be seen in Figure 5.3.

The results of an image taken after the resolution of most tracking errors can be seen in Figure 5.3 and Figure 5.4. The first three frames were removed from the observation sequence due to the delay before the closed-loop tracking corrections take effect. Photon noise for the target is 3.39% with a SNR of 29.5, asserting that sampling errors have been reduced to the level of the systematics.

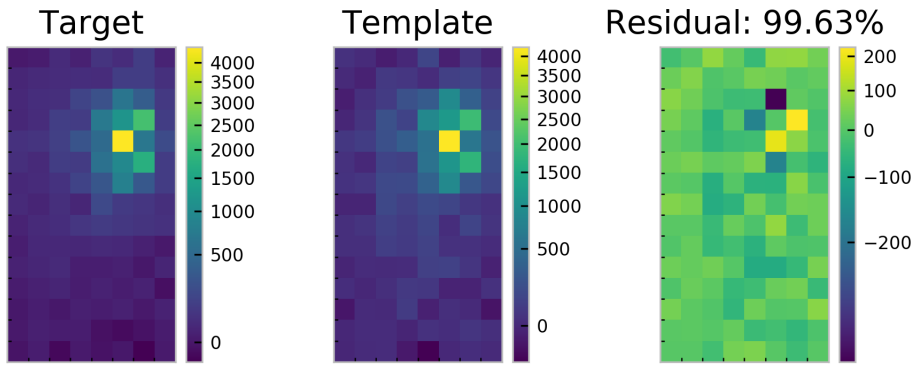

Figure 5.3: *Left:* Target stamp for frame 37. *Middle:* Template stamp built from a linear combination of 75 reference stamps. *Right:* Residual from the target and templates with showing 99.63% match.

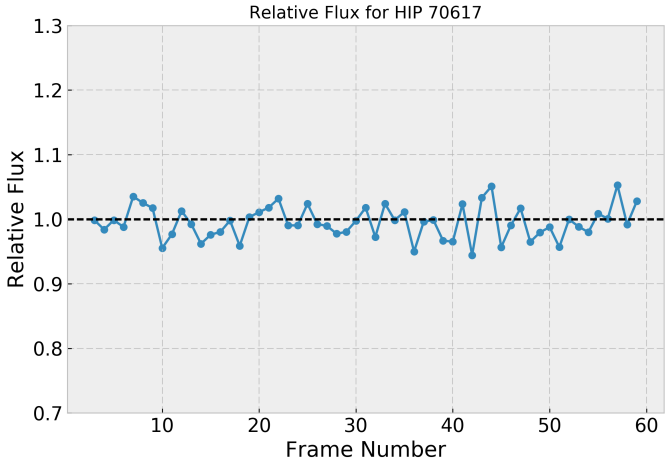

Figure 5.4: Relative flux for HIP 70617 on April 22, 2017. A five pixel aperture (~ 50") was used to derive photometry. The first three frames have been removed due to a delay in tracking correction.

6

Conclusion

6.1 Summary

As part of this thesis work I have designed and built the control software that is responsible for safe operation of the PANOPTES baseline unit. This software has successfully allowed for remote operation and testing of the baseline unit over the course of the thesis work.

Having completed all work related to control of a unit, I have begun building the data processing pipeline that will be used individually on the units to prepare the raw data products and have also begun designing the cloud-based infrastructure that will handle data from multiple units. This initial data processing includes the implementation of the algorithm responsible for accurately processing RGB images from the DSLR camera, which forms the basis of the PANOPTES scientific case.

During this time I have also helped build, install, modify, repair, and test the running hardware on Mauna Loa in Hawai'i. Multiple trips to the mountain were necessary to arrive at a system that takes observations reliably and successfully through adverse conditions with minimal interaction. My work has made a significant contribution toward the success of Project PANOPTES. The project is now able to move past the initial design phases, with the baseline hardware proven and working properly and the software capable of performing autonomously and safely. Funding has been secured for the project, many volunteers have expressed interest in the project, and multiple new units are already being built.

Specific to this thesis, the following tasks were accomplished:

1. a robust software system was developed that is capable of automated control of an astronomical imaging device, including guaranteed operational safety of the unit,
2. the properties and capabilities of the selected consumer-level DSLR camera were investigated as they relate to detecting exoplanet transits, and

3. a working implementation and demonstration of the transit-detection algorithm for DSLR cameras was developed, providing a platform on which to build the complete data pipeline.

The accomplishment of these tasks begins to answer the question of how we can effectively and efficiently capture the most information from the night sky at all times in order to help make a map of transiting exoplanets, all the while involving interested students and citizen-scientists. How much night sky can be watched and who can do the watching is thus answered through the use of the citizen-scientist based Project PANOPTES and the work which I have contributed to the project.

The key findings from this work are that the PANOPTES strategy for exoplanet detection is viable in terms of building and controlling a reliable unit that uses consumer DSLRs to do initial detection and discovery of a large number of Jupiter class planets. This includes confirmation that complications arising from the Bayer color filter array can be overcome at the software level given accurate tracking by the mount. Related to this it was found that closed-loop feedback systems could be developed using no additional hardware that perform reliably and are able to overcome tracking deficiencies, enabling unmodified low-cost consumer mounts to be utilized.

6.2 Future Work

The findings above provide a solid basis upon which more PANOPTES units can be built with confidence that the scientific goals of the project are attainable. PANOPTES itself is moving from baseline development of the individual units to the design and implementation of the wider PANOPTES network. The refinement of the algorithm, especially with respect to computation times, will be an immediate challenge. Exploration of particulars features of the algorithm, such as the construction of the ideal reference frame and the ideal selection of appropriate references, will help improve the detection efficiency of the project. Enhancements to the closed-loop tracking to eliminate the periodic error and achieve sub-pixel tracking will further improve individual results in the different color channels. Particulars related to the camera hardware, such as issues with vignetting, thermal stability of the selected lenses, and evaluation of potential alternatives, among other items, will help refine the exact capabilities of the system. Evaluations of ideal exposure times and cadence, along with the full development of the PANOPTES Input Catalog and stellar selection, will help make PANOPTES a suitable companion to existing and future exoplanet surveys, while at the same time engaging a large number of students and citizen-scientists.

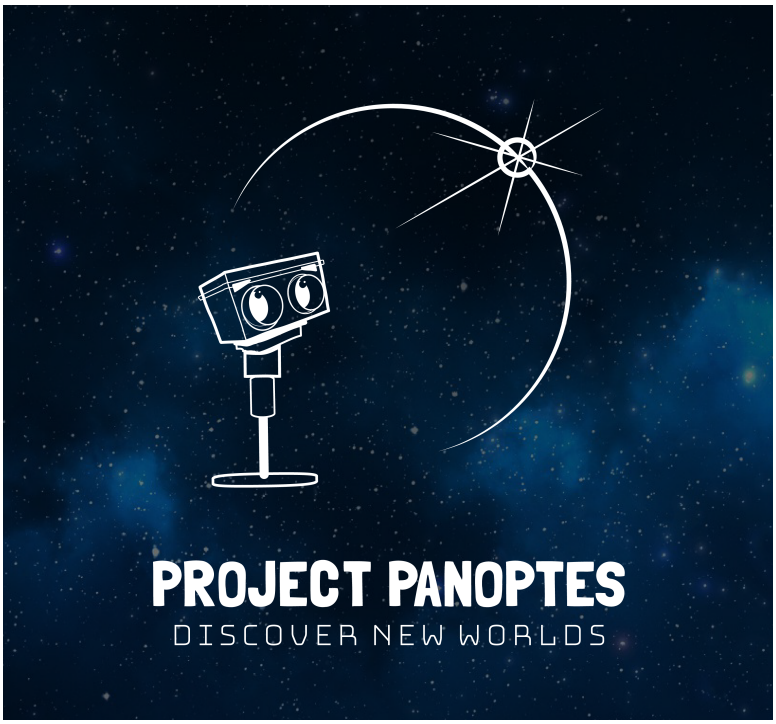

Figure 6.1: The PANOPTES logo, designed by volunteer Laura Summers.

A

Appendix

A.1 Software

The PANOPTES project makes heavy use of the following software libraries:

Python https://www.python.org

SciPy http://www.scipy.org

AstroPy http://www.astropy.org

Anaconda Distribution https://docs.continuum.io/anaconda/

Jupyter http://jupyter.org/

A.2 Jupyter Notebooks

Most of the calculations and plots contained in this thesis are available online in the form of Jupyter Notebooks[1] and are available under an MIT license (along with the rest of the repository). Links to the appropriate file and a short description are given below:

Observing Stats Mollweide Plot of observing runs after tracking fix for a period in early 2017.

Sky Background Calculations Determine the sky background levels for a moonless night. Additional images directories could easily be added to, for instance, compare full moon nights with moonless nights in one plot.

[1] http://jupyter.org/

Background Subtraction Method Code used to analyze which of the background detection methods is most appropriate for use.

Drift Align Test Analysis and plot generation for the drift alignment test.

Tracking Errors - Closed Loop Fix Demonstration of the tracking errors with closed-loop solution.

Superpixels and Stamps Code to create cartoons of superpixels and stamp index.

A.3 Example State Machine

Individual states for the state machine are designed to be simple and written at a high level. An example of the `observing` state can be written as follows[2]:

```
# Called automatically when machine enters state
def on_enter(machine):
    pocs = machine.model        # Main POCS instance
    pocs.next_state = 'parking' # Always the default

    pocs.say("I'm finding exoplanets!")
    try:
        # Start the observing
        camera_events = pocs.observatory.observe()

        # Wait for cameras to finish
        wait_time = 0.
        while not all([event.is_set() for event in camera_events.values()]):
            pocs.check_messages() # Listen for interrupt
            sleep(wait_interval)
            wait_time += wait_interval

    except Exception as e:
        pocs.say("Hmm, I'm not sure what happened with that exposure")
    else:
        # Perform some observing cleanup
        pocs.observatory.finish_observing()

        # Transition to next state
        pocs.next_state = 'analyzing'
```

[2]The code is slightly simplified for presentation with the real version containing more specific error checking.

A.4 Additional Images

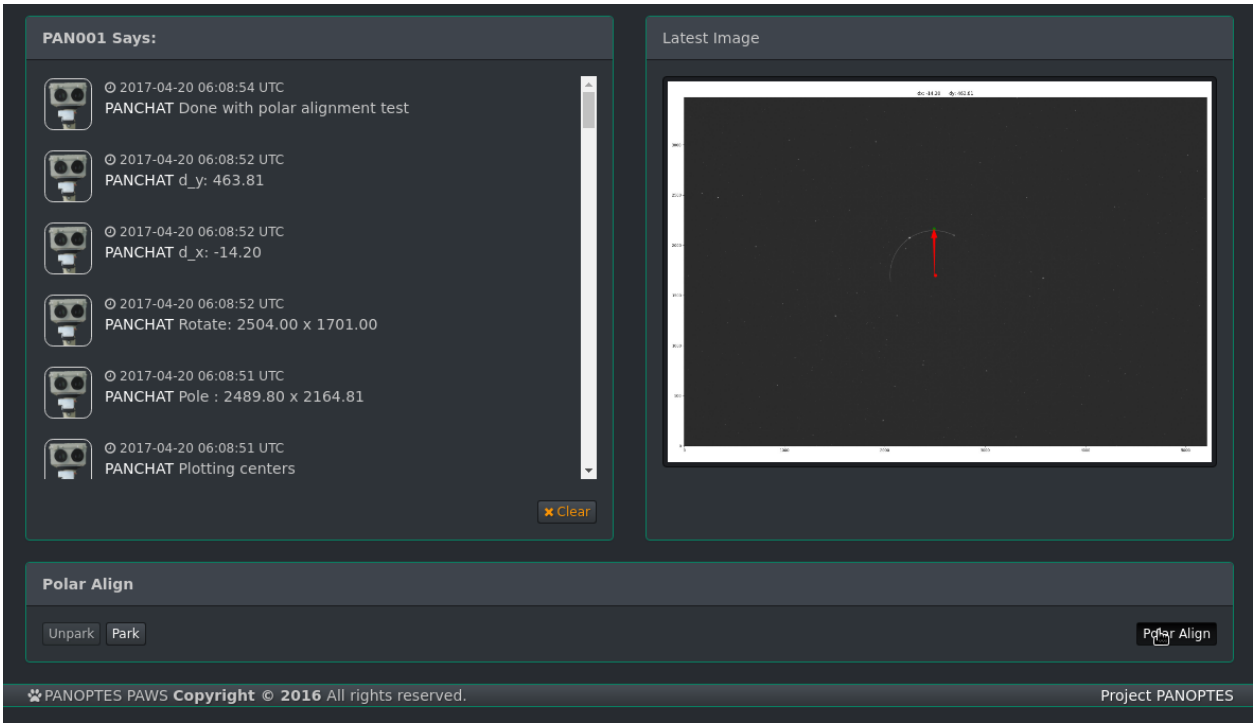

Figure A.1: A screenshot of a portion of the administrative website showing the polar alignment thumbnail, which can be expanded to a full-image, as well as diagnostic output from the mount and the simple one-button interface for iterating the tests.

References

Ayer, V. M., Miguez, S., & Toby, B. H. 2014, Powder Diffraction, 29, S48

Bakos, G., Noyes, R., Kovács, G., et al. 2004, Publications of the Astronomical Society of the Pacific, 116, 266

Bakos, G. Ã., Torres, G., Pál, A., et al. 2010, The Astrophysical Journal, 710, 1724

Bertin, E., & Arnouts, S. 1996, Astronomy and Astrophysics Supplement Series, 117, 393

Borucki, W. J., Koch, D., Basri, G., et al. 2010, Science, 327, 977

Catlin-Groves, C. L. 2012, International Journal of Zoology

Chambers, K. C., Magnier, E. A., Metcalfe, N., et al. 2016, Instrumentation and Methods for Astrophysics

Continuum. 2017a, Anaconda and Docker - Better Together for Reproducible Data Science | Continuum

—. 2017b, Anaconda Distribution | Continuum Analytics: Documentation

Davenport, J. R. A., Kipping, D. M., Sasselov, D., Matthews, J. M., & Cameron, C. 2016, The Astrophysical Journal, 829, L31

Denny, R. B. 2004, Sciences 23rd Annual Symposium on Telescope, 23, 35

Dittmann, J. A., Irwin, J. M., Charbonneau, D., et al. 2017, Nature, 544, 333

Dressing, C. D., & Charbonneau, D. 2015, The Astrophysical Journal, 807, 45

Elkady, A., & Sobh, T. 2012, Journal of Robotics, 2012, 15

Foreman-Mackey, D., Morton, T. D., Hogg, D. W., Agol, E., & Schölkopf, B. 2016, The Astronomical Journal, 152, 206

Fortson, L., Masters, K., Nichol, R., et al. 2011, Database, 11

Franzoni, C., & Sauermann, H. 2014, Research Policy, 43, 1

Gamma, E., Helm, R., Johnson, R., & Vlissides, J. 1993in (Berlin, Heidelberg: Springer Berlin Heidelberg), 406

Gaudi, B. S., Seager, S., & MallenâĂŘOrnelas, G. 2005, The Astrophysical Journal, 623, 472

Guyon, O., & Martinache, F. 2011, Proceedings of the SPIE, 8151, 81511C

Guyon, O., Walawender, J., Jovanovic, N., et al. 2014, Proceedings of the SPIE, 9145, 91453V

Haas, M. R., Batalha, N. M., Bryson, S. T., et al. 2010, The Astrophysical Journal, 713, L115

Hartman, J. D., Bakos, G. Ã., Noyes, R. W., et al. 2011, The Astronomical Journal, 141, 166

Holst, G. C., & Lomheim, T. S. 2011, CMOS/CCD Senors and Camera Systems (JCD Publishing), 75

Hoot, J. E. 2007, The Society for Astronomical Sciences 26th Annual Symposium on Telescope Science, 26, 67

Hosey, A. D., Henry, T. J., Jao, W.-C., et al. 2015, The Astronomical Journal, 150, 6

Howard, A. W., Sanchis-Ojeda, R., Marcy, G. W., et al. 2013, Nature, 503, 381

Howell, S. B., Sobeck, C., Haas, M., et al. 2014, Publications of the Astronomical Society of the Pacific, 126, 398

Ivezić, Å. 2013, LSST Science Requirements Document

Jacklin, S. R., Lund, M. B., Pepper, J., & Stassun, K. G. 2015, The Astronomical Journal, 150, 34

—. 2017, The Astronomical Journal, 153, 186

Jordan Raddick, M., Bracey, G., Gay, P. L., et al. 2013, Astronomy Education Review, 12

Kaiser, N., Burgett, W. S., & Morgan, J. S. 2012, in Advanced Maui Optical and Space Surveillance Technologies Conference No. Dm

Kloppenborg, B. K., Pieri, R., Eggenstein, H.-B., Maravelias, G., & Pearson, T. 2013, Journal of the American Association of Variable Star Observers, 40, 815

Koch, D. G., Borucki, W. J., Basri, G., et al. 2010, The Astrophysical Journal, 713, L79

Kopparapu, R. K., Ramirez, R., Kasting, J. F., et al. 2013, The Astrophysical Journal, 765, 131

Kovacs, G., Bakos, G., & Noyes, R. W. 2004, Monthly Notices of the Royal Astronomical Society, 356, 557

Lang, D., Hogg, D. W., Mierle, K., Blanton, M., & Roweis, S. 2009, The Astronomical Journal, 139, 1782

Law, N. M. 2017, Evryscope-North construction begins - The Evryscope

Law, N. M., Fors, O., Ratzloff, J., et al. 2015, Publications of the astronomical society of the pacific, 127, 234

Lee, E. A., & Seshia, S. A. 2017, Introduction To Embedded Systems, 2nd edn., 537

Littlefield, C. 2010, The Journal of the American Association of Variable Star Observers, 38, 212

Lund, M. B., Pepper, J., & Stassun, K. G. 2014, The Astronomical Journal, 149, 16

Lurie, J. C., Henry, T. J., Jao, W.-C., et al. 2014, The Astronomical Journal, 148, 91

Marshall, P. J., Lintott, C. J., & Fletcher, L. N. 2015, Annual Review of Astronomy and Astrophysics, 53, 247

Mayor, M., & Queloz, D. 1995, Nature, 378, 355

Millman, K. J., & Aivazis, M. 2011, Computing in Science & Engineering, 13, 9

Newman, G., Graham, J., Crall, A., & Laituri, M. 2011, The art and science of multi-scale citizen science support

Oelkers, R. J., Stassun, K. G., Pepper, J., Lee, N. D., & Paegert, M. 2016, in 2016 IEEE International Conference on Big Data (Big Data) (IEEE), 3204

Pál, A., Mészáros, L., Csépány, G., et al. 2013, Astronomische Nachrichten, 334, 932

Peng, R. D. 2011, Science, 334

Pepper, J., Gould, A., & Depoy, D. L. 2002, Acta Astronomica, 53, 213

Rauer, H., Catala, C., Aerts, C., et al. 2013, The PLATO 2.0 Mission No. December (ESA/SRE(2013)5.)

Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2014, Journal of Astronomical Telescopes, Instruments, and Systems, 1, 014003

Robitaille, T. P., Tollerud, E. J., Greenfield, P., et al. 2013, Astronomy & Astrophysics, 558, A33

Rosenblatt, F. 1971, Icarus, 14, 71

Shields, A. L., Ballard, S., & Johnson, J. A. 2016, The habitability of planets orbiting M-dwarf stars

Stassun, K. G., Pepper, J. A., Oelkers, R., et al. 2014, arXiv.org, 1410, 6379

Szabo, J. K., Fuller, R. a., & Possingham, H. P. 2012, Ibis, 154, 468

Tamuz, O., Mazeh, T., & Zucker, S. 2005, Monthly Notices of the Royal Astronomical Society, 356, 1466

Tofflemire, B. M., Wisniewski, J. P., Kowalski, A. F., et al. 2012, The Astronomical Journal, 143, 12

Tulloch, A. I., Possingham, H. P., Joseph, L. N., Szabo, J., & Martin, T. G. 2013, Biological Conservation, 165, 128

Tyson, J. A. 2010, 773303

von Braun, K., Kane, S. R., & Ciardi, D. R. 2009, The Astrophysical Journal, 702, 779

Wang, J., Fischer, D. A., Barclay, T., et al. 2015, The Astrophysical Journal, 815, 127

Wiggins, A., & Crowston, K. 2011, in Proceedings of the Annual Hawaii International Conference on System Sciences, 3426

Wilson, G., Aruliah, D. A., Brown, C. T., et al. 2014, PLoS Biology, 12, e1001745

Zhang, M., Bakos, G. Ã., Penev, K., et al. 2016, Publications of the Astronomical Society of the Pacific, 128, 035001