

# Identification of Impact-hazardous Asteroids

## Literature Study

J. G. P. Vermeulen



# Identification of Impact-hazardous Asteroids

## Literature Study

by

J. G. P. Vermeulen

In partial fulfillment of the requirements  
to obtain the degree of Master of Science  
at the Delft University of Technology,

Student number: 4382889  
Supervisor: Dr. J. Guo, TU Delft  
Dr. M. J. Heiligers, TU Delft

Cover image: P. C. Budassi "Asteroid belt landscape", 9 August 2020, CC BY-SA 4.0. Retrieved from:  
[https://commons.wikimedia.org/wiki/File:Asteroid\\_belt\\_landscape.png](https://commons.wikimedia.org/wiki/File:Asteroid_belt_landscape.png)



# Abstract

Ever since the Tunguska event of 1908 and the discovery linking the extinction of the non-avian dinosaurs to the impact of a meteor, humanity has been searching the skies in an attempt to discover asteroids on an impact trajectory before their collision with Earth. In this literature review, comprehensive background research is provided on proposed research contributing to the problem of impact-hazardous asteroid identification.

Through discussion of the current knowledge on the near-Earth asteroid population, optics, and currently existing asteroid surveys, an autonomous space-based survey system is proposed. This system has to be capable of highly autonomously assessing the risk posed by objects it detects. Then, the state-of-the-art in image processing and trajectory determination techniques is discussed: firstly the image has to be processed to obtain a precise measurement of the target's position relative to the spacecraft. Then, a variety of methods can be used to assess its trajectory and associated risk.

The field of trajectory determination is deemed the most appropriate direction for further research based on its current state of the art, and the possibilities for application of established methods. Several techniques have been selected for evaluation: an approach based on the current state-of-the-art in orbit determination, and two systems applying knowledge from other fields: a Kalman Filter, and a novel approach based on artificial neural networks. Based on this decision, an outline is presented for the subsequent work, aiming to deliver its final results near the end of January 2022.



# Preface

What started as a concept in researching the feasibility of deflecting asteroids using hypervelocity impacts, has evolved into a proposal of using simulations to calculate trajectories and assess the risk of those asteroids in the first place. Undoubtedly a challenge, and one that has cost me plenty of thought on whether these methods are actually possible; a neural network or a Kalman filter is a cool concept, but *how* can we apply it to this problem, given the many constraints? Add to that the competition of observational beasts such as the upcoming Large Synoptic Survey Telescope, and the challenge has been set. Nevertheless, I am excited about the next steps in testing my thoughts, and my gratitude goes out to Dr. Jian Guo, who through his supervision has frequently challenged my ideas and assertions, feeding the thought processes leading to these conclusion. On to the next step?

*Arjan Vermeulen, April 2021*



# Contents

<b>Abstract</b>	<b>iii</b>
<b>Preface</b>	<b>v</b>
<b>List of Figures</b>	<b>ix</b>
<b>List of Tables</b>	<b>xi</b>
<b>List of Symbols and Abbreviations</b>	<b>xiii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Motivation . . . . .	1
1.2 Research Topics . . . . .	1
<b>2 Population of Near-Earth Asteroids</b>	<b>3</b>
2.1 Asteroid Classification . . . . .	3
2.2 Size-Frequency Distribution . . . . .	5
2.3 Orbital Parameters . . . . .	6
2.4 Major Near-Earth Asteroid Groups . . . . .	9
<b>3 Optical Detection Technology</b>	<b>17</b>
3.1 Optical Systems . . . . .	17
3.2 Mathematics of Optics . . . . .	19
3.3 Diffraction Limitation . . . . .	23
3.4 Noise and Bias . . . . .	24
<b>4 Near-Earth Asteroid Surveys</b>	<b>27</b>
4.1 Earth-based Surveys . . . . .	27
4.2 Space-based Surveys . . . . .	28
<b>5 Image Processing Techniques</b>	<b>31</b>
5.1 Background Removal . . . . .	31
5.2 Centerfinding . . . . .	33
5.3 Applications of Neural Networks . . . . .	35
<b>6 Trajectory Determination Techniques</b>	<b>39</b>
6.1 Gauss' Method . . . . .	39
6.2 The Method of Very Short Arcs . . . . .	41
6.3 Kalman Filters . . . . .	43
6.4 Artificial Neural Networks . . . . .	45
<b>7 Conclusion and Thesis Research Questions</b>	<b>49</b>
7.1 Conclusion on Literature and Knowledge Gap . . . . .	49
7.2 Thesis Research Questions . . . . .	50
7.3 Project Plan . . . . .	51



# List of Figures

2.1	Photographs of 4 Vesta and 67P/Churyumov-Gerasimenko taken by the Dawn and Rosetta spacecraft, respectively (not to scale). . . . .	4
2.2	Cumulative population, potential impact energy and impact interval for Near Earth Asteroids. . . . .	7
2.3	Population of Near Earth Asteroids and population of those currently detected. . . . .	7
2.4	Distribution of orbital elements of known NEO's (as of 2002) and the largest 10.000 main-belt asteroids. ( <b>originandevolution</b> ) . . . . .	9
2.5	Expected (black line) and known as of 2001 (grey histogram) distribution of orbital elements of all NEA's with H<18. ( <b>debiased</b> ) . . . . .	10
2.6	Diagram showing the characteristic orbits of the NEA groups. . . . .	11
2.7	Expected (black line) and known as of 2001 (grey histogram) distribution of orbital elements of all Amor asteroids with H<18. ( <b>debiased</b> ) . . . . .	12
2.8	Expected (black line) and known as of 2001 (grey histogram) distribution of orbital elements of all Apollo asteroids with H<18. ( <b>debiased</b> ) . . . . .	13
2.9	Expected (black line) and known as of 2001 (grey histogram) distribution of orbital elements of all Aten asteroids with H<18. ( <b>debiased</b> ) . . . . .	14
2.10	Expected (black line) and known as of 2020 (black histogram) distribution of orbital elements of all Atira asteroids with H<18. ( <b>debiased</b> ). The black histogram was not presented in the original paper, but was added by the author of this report. This was done because there were zero known Atira asteroids in 2001. Note that there is still a lot of uncertainty in this population. . . . .	15
3.1	True color image of dwarf planet Pluto taken by the New Horizons space probe in 2015. . . . .	17
3.2	CCD sensor from a SONY Alpha commercial camera. . . . .	18
3.3	Overview of common mirror and lens setups. ( <b>SMAD</b> ) . . . . .	19
3.4	Illustration of coordinate frames used in the camera attitude determination . . . . .	20
3.5	Description of elements of an optical system. The presented figure is for a refractive system, although mathematically a reflective system functions similarly. ( <b>SMAD</b> ) . . . . .	22
3.6	An Airy diffraction disk created by shining a laser on a 90-micrometer hole. . . . .	24
3.7	Simulated intensity of airy disk in 2D in dimensionless units $u$ . . . . .	25
4.1	Artist rendition of the NEOSSat satellite, showing the aperture of the 0.15m Maksutov-Cassegrain telescope. (Image Courtesy: Canadian Space Agency) . . . . .	29
5.1	Overview of the process of image processing (images from <b>ATLASBlog</b> ). Note that these images were not obtained using the mentioned methods, but merely serve as illustration. . . . .	32
5.2	Demonstration of the image subtraction process from a series of stacked images. The path of the transient is shown in red ( <b>processingclassic</b> ). . . . .	33
5.3	Some images from ATLAS data. The first two rows show actual detections of asteroids, whereas the latter two rows show visual artifacts. Note how the white streaks in the bottom images might be mistakes for asteroid tracks in a more simple algorithm ( <b>AtlasDL</b> ). . . . .	35
5.4	Schematic of the network used by <b>processingDLtwo</b> for classification into seven classes of observations. . . . .	36
5.5	Comparison of real and simulated image sequences from <b>processingDLtwo</b> . . . . .	37
5.6	The system trained by <b>autoencoder</b> and its results. . . . .	38
6.1	Actual orbit and orbit as determined by Gauss' method for a short arc . . . . .	41
6.2	Sketch of the construction of the admissible region ( <b>classicmodernorbits</b> ). . . . .	42

6.3 An example of an admissible region, for asteroid 2003BH <sub>84</sub> using exclusively data from the day of first observation. The vertical axis is $r$ [AU/day], the horizontal axis is $f(r) = 1 - e^{-r^2/(2s^2)}$ where $s$ is a normalization term. The real position of the asteroid is marked near the top left with a $\oplus$ . ( <b>classicmodernorbits</b> ) . . . . .	43
6.4 Comparison of a biological neuron with an artificial neuron. The Neuron takes inputs, combines those, and forwards them to produce a set of outputs. Note that this is where the analogy ends: biological neurons behave very differently to ANN's. CC BY-SA 4.0 Prof. Vu-Quoc. . . . .	45
6.5 Structure of a feedforward neural network. CC BY-SA 3.0 Glosser.ca . . . . .	46
7.1 Gantt chart for the thesis work. Shown in red is literature research work, blue is theoretical research work and pink the programming work. Milestones are indicated as red diamonds. Each rectangle represents a week. . . . .	53

# List of Tables

2.1	Summary of asteroid taxonomic classes. . . . .	5
2.2	Mean albedo for asteroid taxonomic classes. . . . .	5
2.3	Estimated population and fraction discovered of NEA's by absolute magnitude . . . . .	8
2.4	Definitions of the four main near-Earth asteroid groups. . . . .	9
3.1	Examples of diffraction target resolution limit for several apertures and distances for $\lambda = 0.5\mu m$ visible light. . . . .	23
3.2	Typical values for spacecraft optical system . . . . .	25
7.1	Thesis timeframe planning. . . . .	51



# List of Symbols and Abbreviations

## Latin symbols

Symbol	Meaning	Unit
$A$	Aperture	$\text{m}^2$
$B_e$	Background signal	$\text{sr}^{-1} \text{s}^{-1}$
$D$	Diameter	$\text{m}$
$DN$	Data number	-
$F\#$	F-stop	-
$H$	Absolute magnitude	-
$I$	Light intensity	$\text{m}^{-2}\text{sr}^{-1}\text{Hz}^{-1}\text{s}^{-1}$
$M$	Magnification	-
$N$	Number	-
$N_d$	Dark current	$\text{s}^{-1}$
$N_r$	Read noise	-
$Q$	Apohelion	AU
$Q$	Quantum efficiency	-
$X'$	Resolution limit	$\text{m}$
$a$	Semi-major axis	AU
$b$	Bias term	-
$c$	Speed of light	$\text{ms}^{-1}$
$e$	Eccentricity	-
$f$	Focal length	$\text{m}$
$g$	Camera sensor gain	$\text{DN/e}^-$
$h$	Distance to target	$\text{m}$
$h$	Pixel magnitude	DN
$h$	Apparent visual magnitude	-
$h^*$	Average apparent visual magnitude	-
$i$	Inclination	deg
$k$	Palomar-Leiden slope	-
$p$	Geometric albedo	-
$p_v$	Assumed geometric albedo	-
$q$	Perihelion	AU
$r$	Position	$\text{m}$

## Greek symbols

Symbol	Meaning	Unit
$\alpha$	Right ascension	deg
$\delta$	Declination	deg
$\eta$	Learning rate	-
$\lambda$	Wavelength	m
$\mu$	Gravitational parameter	
$\phi$	Angle of twist	deg
$\sigma$	Standard deviation	-

## Other symbols

Symbol	Meaning	Unit
$\oplus$	Earth	-
$\odot$	Sun	-
$\mathcal{E}$	Energy	J
$\mathcal{N}$	Normal distribution	-

## Abbreviations

Abbreviation	Meaning
ANN	Artificial neural network
CCD	Charge-coupled device
CMOS	Complementary metal-oxide substrate
CNN	Convolutional neural network
DEC	Declination
ELU	Exponential linear unit
FOV	Field of view
LSST	Large Synoptic Survey Telescope
LSTM	Long-short term memory
MOID	Minimum orbital intersection distance
NEO	Near-Earth object
PHA	Potentially hazardous asteroid
PSF	Point-spread function
RA	Right ascension
ReLU	Rectified linear unit
RNN	Recurring neural network
SNR	Signal-to-noise ratio
ZTF	Zwicky transient facility

# 1

# Introduction

## 1.1. Motivation

Astronomical phenomena have long fascinated humanity, and few are as awe-inspiring or destructive as a large asteroid impacting Earth. The most well-known example is probably the Chixculub impactor, believed to be the cause of the extinction of the non-avian dinosaurs 66 million years ago. However also more recent examples of impactors destructive potential are known: consider the Tunguska event, flattening over 2000 km<sup>2</sup> of forest in 1908, or the 2013 Chelyabinsk meteor, filmed on many dashcams.

Considering the seemingly random nature of these events, and their potential for global catastrophe, it is only natural that humanity began its search for these hazardous asteroids. By now, these methods have evolved into automatic systems scanning the entire visible sky every night, resolving targets less than 100 meters in diameter. However, humanity's knowledge of smaller asteroids is still very much incomplete, and observations from Earth are hindered by weather, atmospheric interference, and glare from our Sun. Thus, the need for a space-based system becomes apparent. Of course, astronomy is facing the same juxtaposed challenge and opportunity of massive data streams highlighted in all areas of modern technical research. With the lakes of data that can be generated daily, downlinking the data to Earth and processing it by means of human observers has become impossible; therefore an integral part of our surveys to understand the asteroid population has become its automation.

The aim of this literature study, and the proposed follow-up research, is to provide a contribution to our capabilities of autonomous spacecraft assessing the hazard of asteroids on Earth; hopefully warning us, before it is too late.

## 1.2. Research Topics

As outlined above, the research will focus on several topics. Firstly, it is important to understand the population of asteroids that can actually pose a threat to Earth. This will be discussed in chapter 2. Then, a short description of how the optical systems used for imaging work will be given in chapter 3. This is primarily important for the practical side of the research work, as the current state of the field of optics is far beyond our expertise. Before continuing to discussing the state of the research in the relevant fields, a small aside is made to give some examples of the current survey capabilities in chapter 4. It is believed that this is essential to place the system in context, and to ensure its feasibility.

After the earlier, more context-oriented, chapters, the later chapters will focus more on determining the content of the research. Firstly, the state-of-the-art in image processing techniques, both classical and using machine learning applications, will be discussed in chapter 5. Then, the current and hypothesized methods for determining trajectories of objects, and of asteroids in specific, will be described in chapter 6. This review will conclude with a derivation of the applicable knowledge gap, thesis questions, and work breakdown in chapter 7.



# 2

## Population of Near-Earth Asteroids

Before examination of the methods available for detecting asteroids, it is imperative to study the intended targets. Asteroids are minor planets orbiting our Sun, and are possibly the most diverse object class in our solar system: ranging from tiny chips of rock and dust to objects nearing 1000km in diameter such as Vesta or the dwarf planet Ceres. Additionally, the shape, material and appearance can differ significantly from asteroid to asteroid, such as can be seen in Figure 2.1. Vesta is characterized as a spherical object littered with craters, whereas the comet 67P/Churyumov-Gerasimenko, target of ESA's Rosetta mission, is a very erratic shape featuring a ragged icy surface with relatively few craters.



Figure 2.1: Photographs of 4 Vesta and 67P/Churyumov-Gerasimenko taken by the Dawn and Rosetta spacecraft, respectively (not to scale).

In this chapter, firstly in section 2.1, an overview of the various classes of asteroids will be given, based on their physical and optical properties. Afterwards, section 2.2 will present the relation between asteroid size and frequency, as this is an important factor in determining the threat posed by asteroids. Next, in section 2.3 the distribution of orbital parameters of asteroids will be presented and lastly in section 2.4 some major asteroid families will be discussed.

### 2.1. Asteroid Classification

The main taxonomy of asteroids is based on the work of **taxonomy**, upon which several subsequent studies have expanded. Most highly regarded, the work of **smaSSII** aims to more precisely categorize asteroids through newly attained spectroscopic survey data, and several others such as **newClassification** have further improved this process. Because the spectroscopic properties of asteroids are only a minor part of the research, the following exposition will focus on the taxonomy as proposed by Tholen. As it would be infeasible to visit every known asteroid, large-scale asteroid classifica-

tion is done through spectroscopic surveys. Thus, the main classification is performed based on spectroscopic properties. Table 2.1 gives a summary of the taxonomic classes (**spectroscopicProperties**).

Table 2.1: Summary of asteroid taxonomic classes.

<b>Tholen Class</b>	<b>Albedo</b>	<b>Spectral Features</b>
A	Moderate	Very steep red slope shortward of $0.75 \mu\text{m}$ ; moderately deep absorption feature longward of $0.75 \mu\text{m}$ .
B, C, F, G	Low	Linear, generally featureless spectra. Differences in UV absorption features and presence/absence of narrow absorption feature near $0.7 \mu\text{m}$ .
D	Low	Relatively featureless spectrum with very steep red slope.
E, M, P	From low (P) to very high(E)	Generally featureless spectrum with reddish slope; differences in subtle absorption features and/or spectral curvature and/or peak relative reflectance.
Q	Moderate	Reddish slope shortward of $0.7 \mu\text{m}$ ; deep, rounded absorption feature longward of $0.75 \mu\text{m}$ .
R	Moderate	Moderate reddish slope downward of $0.7 \mu\text{m}$ ; deep absorption longward of $0.75 \mu\text{m}$ .
S	Moderate	Moderately steep reddish slope downward of $0.7 \mu\text{m}$ ; moderate to steep absorption longward of $0.75 \mu\text{m}$ ; peak of reflectance at $0.73 \mu\text{m}$ .
T	Low	Moderately reddish shortward of $0.75 \mu\text{m}$ ; flat afterwards.
V	Moderate	Reddish shortward of $0.7 \mu\text{m}$ ; extremely deep absorption longward of $0.75 \mu\text{m}$ .

Out of all the spectroscopic properties, the albedo is the most important parameter for the further research. Next to determining the visibility of the asteroid in sunlight, the albedo is also the main source of error in estimating the size of the target body (**quantifyingrisk**). The fundamental relation which allowed for determining an asteroid's diameter from its brightness and albedo was first developed by **sizealbedo** and later used by many others (e.g. **populationofnea** and **quantifyingrisk**). This relation is shown in Equation 2.1. Here, the brightness is given in terms of absolute magnitude  $H$ , diameter  $D$  in km and  $p_v$  representing the assumed albedo of the asteroid. For an albedo of  $p_v = 0.15$ , this implies a diameter of  $1kM$  for an object with  $H = 17.6$  and a diameter of  $100\text{m}$  for an object with  $H = 22.6$ .

$$D = \frac{1329\text{km}}{\sqrt{p_v}} 10^{-H/5} \quad (2.1)$$

Because of the relation between diameter and albedo, it is important to get a reasonably accurate determination of the albedo of asteroids. Of course, because of the dependence on several unknown factors, primary among which geometry, a general estimate is to suffice. **albedovalues** provides an estimate of albedo values per taxonomic class based on a numerical analysis of 84 known asteroids. This overview is shown in Table 2.2.

Out of the presented classes, the C, S and V classes are most interesting for their commonality (**smaassiitwo**). The carbonaceous C-types make up approximately 75% of the known asteroids, the stony S-types are the second most common at 17% of all known asteroids. The V-type asteroid or Vestoid makes up approximately 6% of known asteroids although they are of special interest due to their proximity to Earth and their large size, making impacts particularly dangerous (**vestoid**).

Table 2.2: Mean albedo for asteroid taxonomic classes.

Asteroid classes	Mean albedo
C, G, B, F, P, T, D	$0.058 \pm 0.004$
M	$0.124 \pm 0.009$
S, Q	$0.184 \pm 0.011$
E, V, R	$0.403 \pm 0.032$

## 2.2. Size-Frequency Distribution

Possibly the most important relation in assessing the threat of asteroid impacts is the relationship between asteroid size and frequency (**firstpaper**). Whereas events such as the 2013 Chelyabinsk meteor might represent a once-in-a-decade event, globally destructive events such as the Chixculub asteroid<sup>1</sup> might only occur once every millions of years. At this point, as Earth impacts are discussed, it is important to further constrain the selection of asteroids to those actually threatening the Earth. To this end, firstly the definition of a Near Earth Objects as used by ESA<sup>2</sup> and NASA<sup>3</sup> is given: a Near Earth Object (NEO) is an asteroid or comet with perihelion distance smaller than 1.3AU. The vast majority of these objects are asteroids, with comets only accounting for around 1% of the population. For simplicity, the description will be constrained to Near Earth Asteroids (NEA's). Among NEA's, an additional distinction exists for Potentially Hazardous Asteroids (PHA's): asteroids with a minimum orbital intersection distance of less than 0.05AU and an absolute magnitude of 22 or less, corresponding to objects of approximately  $D \geq 140\text{m}$ . This last limit was determined by **neosizelimit** as a balance between the difficulty of detection of smaller objects, their higher frequency, and the lower impact energy of smaller bodies. At this size, they estimated that by detecting 90% of NEA's above this threshold, 90% of the remaining risk over all size categories would be alleviated (**subpopulations**). However, this might not be the minimal size required to properly defend planet Earth against asteroids. As **smallneos** conclude, smaller objects are also worthy of study, with the Tunguska event of 1908 as an example. Objects like this, in the  $30\text{m} > D > 50\text{m}$  range, "are also capable of causing significant damage to Earth... with much of the damage [of the Tunguska asteroid] caused due to shock waves from the explosion of the object in the Earth's atmosphere." They estimate such an event to occur once every 300 years, and advise "...to detect as many 30- to 50-meter objects as possible."

Currently, it is estimated that all NEA's with absolute magnitude  $H \leq 15$  are known, corresponding to all objects of  $D \geq 3.5\text{km}$  at  $p_v = 0.15$ . To estimate the frequency of smaller asteroids, a power law was proposed by **asteroidpowerlaw** which is shown in Equation 2.2. The exponent  $k$ , the so-called Palomar-Leiden slope, is in the range of 2.95 to 3.5 according to **impactrate**. Although not entirely accurate, this relation provides a good estimate for the population of asteroids. An overview of the results of this law is given in Figure 2.2 (**populationofnea**). As the impact frequency is logically an inverse function of population, and the impact energy is a linear function of mass (and therefore cubically related to diameter), these are shown as an additional reference.

$$\frac{dN}{dD_p} \propto D_p^{-k} \quad (2.2)$$

Furthermore, when assessing the feasibility of a system for asteroid detection, it should be considered which asteroids still need to be detected. As of March 2021, no objects score above a 0 ("NO HAZARD" on the Torino scale<sup>4</sup>, therefore a potential threat will come from a currently unknown asteroid. It follows from this that the ratio of known to unknown asteroids becomes an important parameter. Frustratingly, this becomes a paradoxical problem: models of the population of asteroids are based on

<sup>1</sup>The asteroid popularly seen as the cause of the extinction of the non-avian dinosaurs.

<sup>2</sup>see: <http://neo.ssa.esa.int/definitions-and-assumptions>

<sup>3</sup>see: [https://cneos.jpl.nasa.gov/about/neo\\_groups.html](https://cneos.jpl.nasa.gov/about/neo_groups.html)

<sup>4</sup>see: <http://neo.ssa.esa.int/risk-page> for an overview of known NEA's with non-zero chance of impact in the next 100 years.

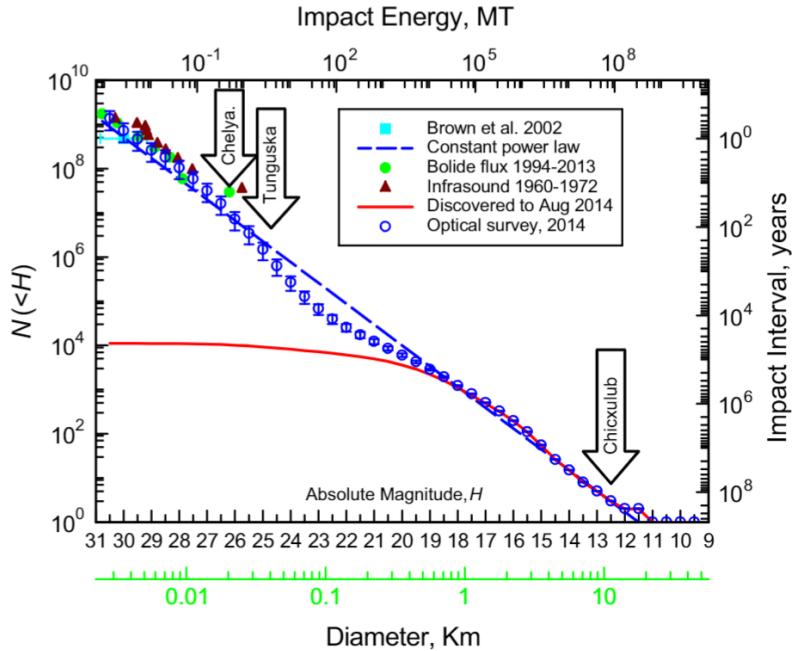


Figure 2.2: Cumulative population, potential impact energy and impar interval for Near Earth Asteroids.

the asteroids we have presently detected, therefore there is no guarantee these models provide adequate information on the population of unknown asteroids. To resolve this problem, **populationofnea** recently performed a numerical simulation of various distributions of orbital elements to find a distribution of orbital elements resulting in a distribution of detected objects as observed from Earth. The results are shown in Figure 2.3. As mentioned previously, the population of asteroids of  $H \leq 15$  is assumed to be near 100% known. However, they also calculate that on the lower bound of NEA sizes, at  $21.5 < H < 22.0$ , only 6.64% of NEA's is currently discovered. The full results are presented in Table 2.3. This fraction will get smaller as the absolute magnitude of the target body decreases, further strengthening the case for attempting to detect asteroids below the 140m threshold. As is apparent, there are still a large number of PHA's to be discovered.

## 2.3. Orbital Parameters

When discussing the hazard of asteroids, the characteristic orbits should be discussed. As can be seen in Figure 2.4, the orbital elements are varied, but a few general conclusions can be drawn. In this section, a general overview of all NEA orbits will be given, and in section 2.4 the classification will be specified into the different asteroid families.

The distribution of orbits can be understood better in the context of the lifecycle of near-Earth asteroids. Most near-Earth asteroid orbits are unstable on megayear timescales **neonieuw**, and will eventually have their orbits altered by orbital resonances. This will lead to either impact with a planet, being launched out of the solar system by a large planet like Jupiter, or coming too close to the Sun. Because of this decay, the population has to be maintained by sources. It follows logically that over astronomical timescales a steady-state solution should be reached, which can be used as a model for the population.

The list of accepted primary sources for NEO's is the one presented by **debiased**, further expanded upon by **originandevolution**, in rough order of importance:

- The  $\nu_6$  secular resonance between asteroids in the outer solar system and Saturn. This slows asteroids down and lowers their orbit.
- The 3:1 mean motion orbital resonance with Jupiter in the asteroid belt.

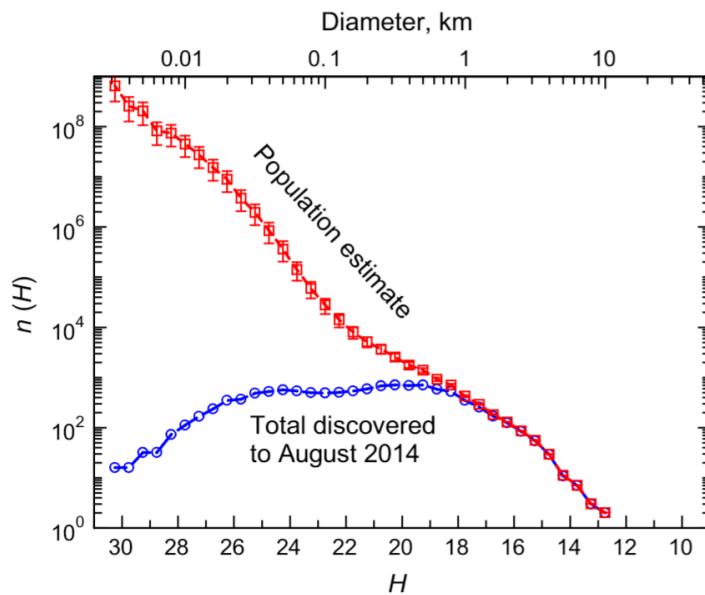


Figure 2.3: Population of Near Earth Asteroids and population of those currently detected.

Table 2.3: Estimated population and fraction discovered of NEA's by absolute magnitude

H range	Estimated population	Fraction discovered
(15.0; 15.5]	112	0.978
(15.5; 16.0]	199	0.969
(16.0; 16.5]	329	0.953
(16.5; 17.0]	512	0.927
(17.0; 17.5]	802	0.886
(17.5; 18.0]	0.123E4	0.826
(18.0; 18.5]	0.194E4	0.740
(18.5; 19.0]	0.286E4	0.631
(19.0; 19.5]	0.425E4	0.509
(19.5; 20.0]	0.603E4	0.387
(20.0; 20.5]	0.859E4	0.276
(20.5; 21.0]	0.123E5	0.185
(21.0; 21.5]	0.174E5	0.115
(21.5; 22.0]	0.255E5	0.664E-1

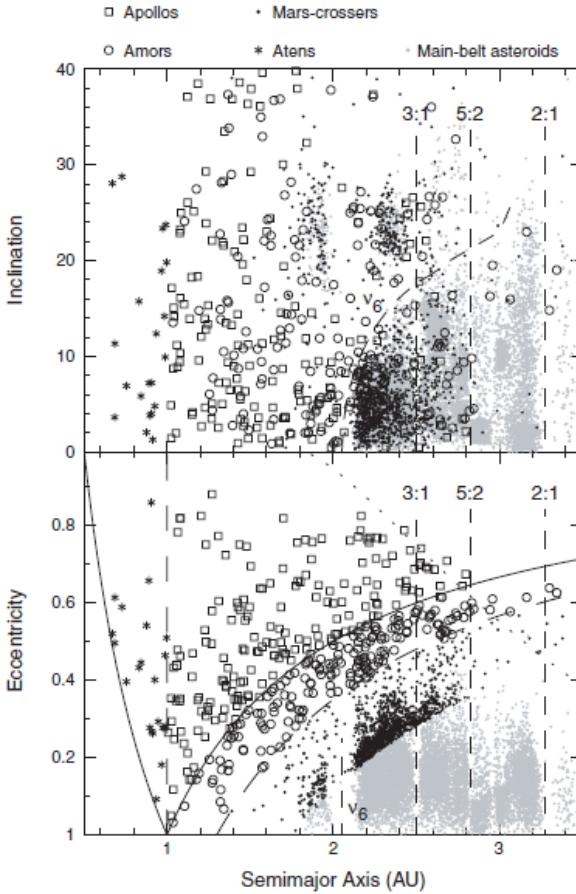


Figure 2.4: Distribution of orbital elements of known NEO's (as of 2002) and the largest 10.000 main-belt asteroids. ([originandevolution](#))

- Asteroids coming from the Intermediate Mars crosser population, being placed in near-Earth orbits by mean motion resonances with Mars, three body mean motion resonances like with Jupiter and Saturn, or other slower resonances similar to the  $v_6$  resonance.
- Jupiter family comets, due to various gravitational interactions. However, due to their large semi-major axes and therefore long periods, this is a slow process.
- Outer main belt asteroids on unstable orbits in various two and three body resonances.

As might be intuitive from the variety of sources, and the evolution being driven by slow gravitational effects, the distribution of orbits is varied. An overview of known NEO's orbital distribution is shown in Figure 2.4 and Figure 2.5. Several attempts have been made at modelling the populations of NEO's based on these sources, most notably the work of **debiased**, the work of **subpopulations** based on the NEOWISE mission and more recently the work of **neonieuw**. Some general conclusions can be drawn from this, but with some exceptions, most combinations of orbital elements are a possibility. This is important information for the further trajectory determination process.

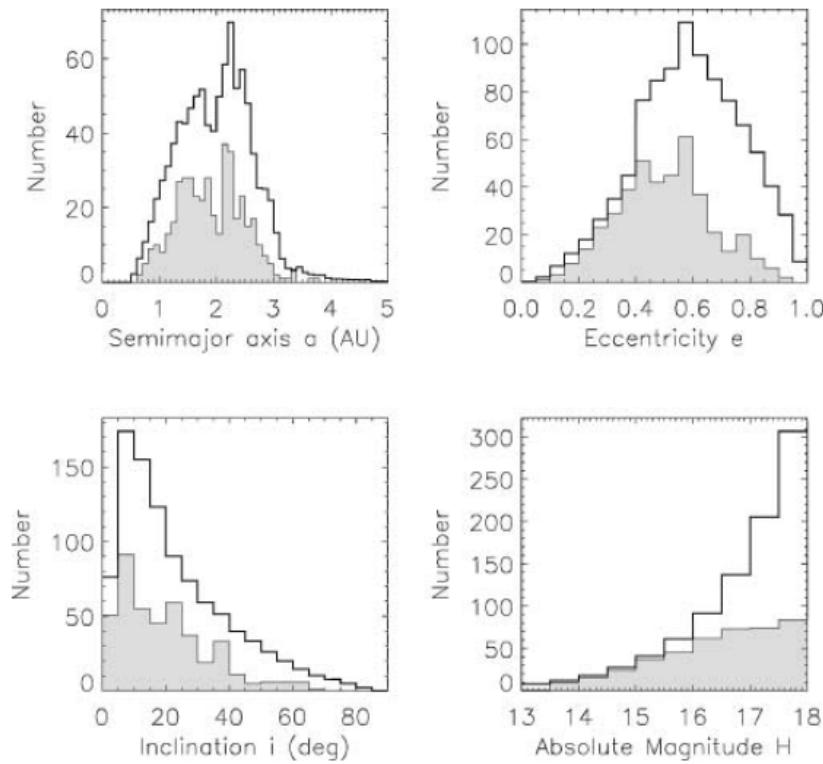


Figure 2.5: Expected (black line) and known as of 2001 (grey histogram) distribution of orbital elements of all NEA's with  $H < 18$ . (debiased)

## 2.4. Major Near-Earth Asteroid Groups

Table 2.4: Definitions of the four main near-Earth asteroid groups.

Group	$q$	$a$	$Q$	Earth-crossing
Amors	$>1.017$	$>1.0$	-	No
Apollos	$<1.017$	$>1.0$	-	Yes
Atens	-	$<1.0$	$>0.983$	Yes
Atiras	-	$<1.0$	$<0.983$	No

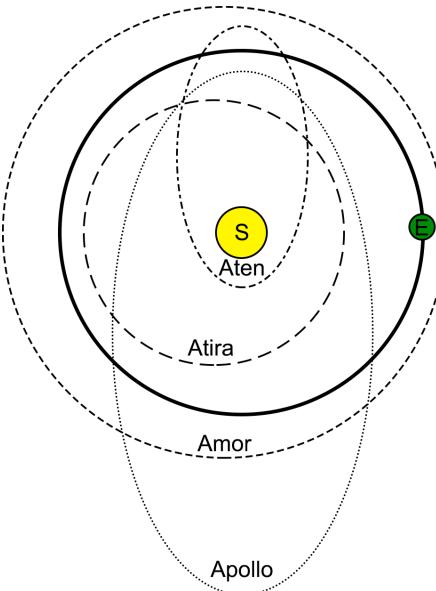


Figure 2.6: Diagram showing the characteristic orbits of the NEA groups.

To conclude the discussion of near-Earth and potentially hazardous asteroids, a cursory overview of the four main dynamical asteroid groups in the near-Earth asteroid population is presented. These groups are based on the perihelion  $q$ , aphelion  $Q$  and semi-major axis  $a$  of the asteroid's orbit. The definitions of the groups as used by NASA/JPL<sup>5</sup> is shown in Table 2.4 and schematically in Figure 2.6. The boundary values for  $q$  and  $Q$  were chosen to coincide with the perihelion and aphelion of Earth. Although useful for reference, the Earth-crossing column does not necessarily indicate anything about the danger of impact. As explained by **don't do ECA**, the calculations involved in determining whether the asteroid will cross the orbit of Earth are laborious and error-prone. The dimensions of the orbit do not suffice: an object with a perihelion  $<1.0$  AU and aphelion  $>1.0$  AU is not guaranteed to intersect Earth's orbit at all; after all the orbits of asteroids can be highly inclined.

<sup>5</sup>see: [https://cneos.jpl.nasa.gov/about/neo\\_groups.html](https://cneos.jpl.nasa.gov/about/neo_groups.html)

### 2.4.1. Amor Asteroids

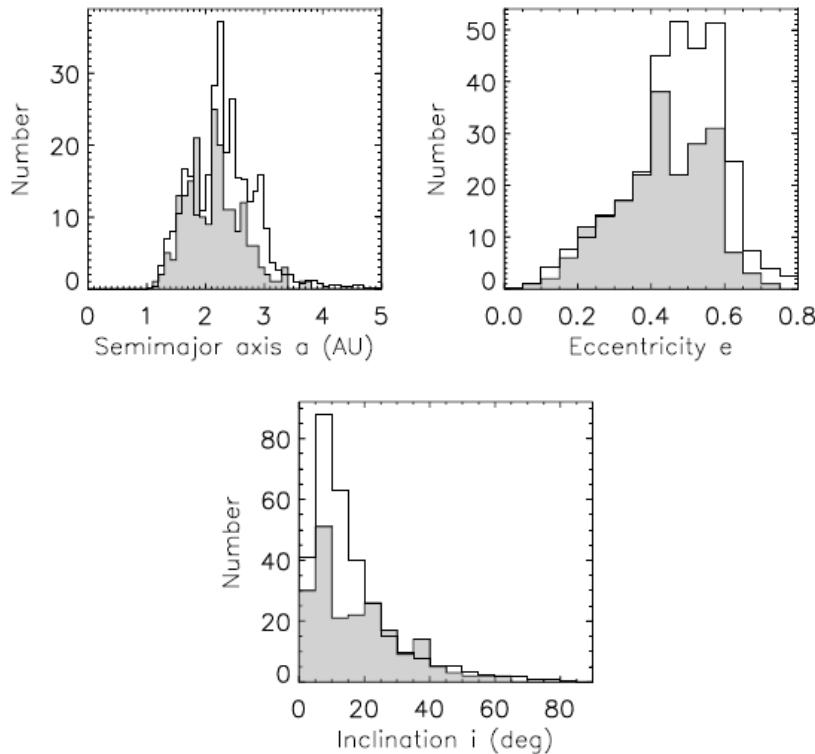


Figure 2.7: Expected (black line) and known as of 2001 (grey histogram) distribution of orbital elements of all Amor asteroids with  $H < 18$ . (**debiased**)

The Amor asteroid group features near-Earth objects conforming to the following criteria:

- The semi-major axis of the objects orbit around the Sun is greater than 1.0 AU.
- The perihelion is greater than Earth's aphelion.

From this follows that the orbital period is longer than one Earth year, and that the objects orbit is not Earth-crossing: the entirety of the orbit lies further away from the Sun than the orbit of Earth. The group is named after the asteroid 1221 Amor. As of March 2021, there are 9270 known Amor asteroids, of which 133 are classified as PHA<sup>6</sup>. The discovery fraction of the Amors is significantly higher than the other asteroid populations. This is most likely because their entire orbit is outside the orbit of Earth, and they are therefore easier to detect **debiased**. Furthermore, according to results from the NEOWISE mission (**subpopulations**), the Amors are proportionally dark, with 35% of the population having  $p_v < 0.01$ . Their size distribution can be fit to Equation 2.2 using  $k = 1.40 \pm 0.18$  for asteroids with  $D < 1.9\text{km}$  and  $k = 5.0 \pm 2.0$  for larger sizes. The distribution of orbital parameters is shown in Figure 2.7.

<sup>6</sup>see: [https://ssd.jpl.nasa.gov/sbdb\\_query.cgi](https://ssd.jpl.nasa.gov/sbdb_query.cgi)

### 2.4.2. Apollo Asteroids

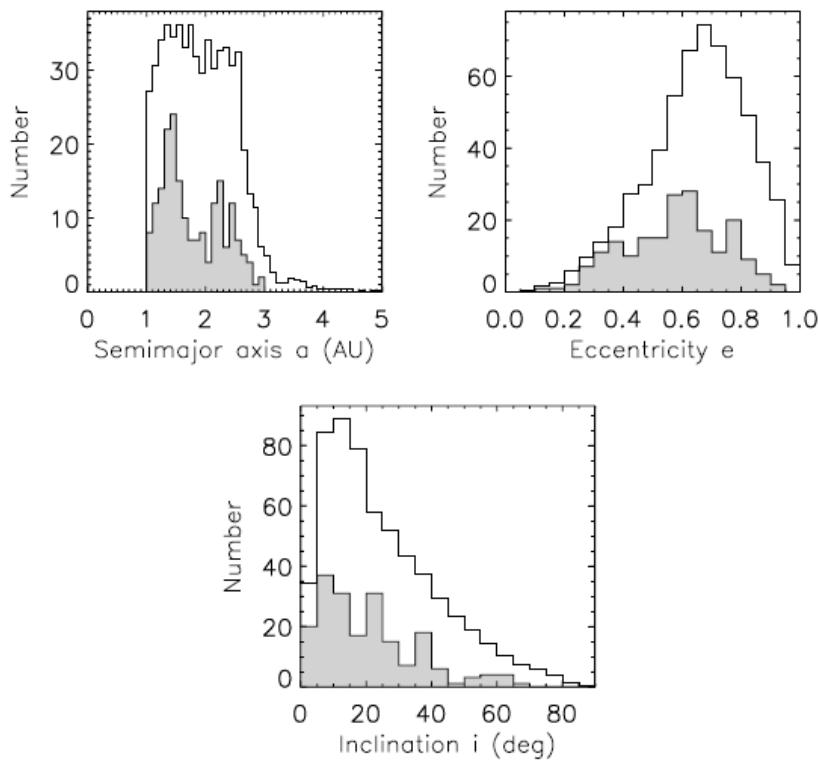


Figure 2.8: Expected (black line) and known as of 2001 (grey histogram) distribution of orbital elements of all Apollo asteroids with  $H < 18$ .(debiased)

The Apollo asteroid group features near-Earth objects conforming to the following criteria:

- The semi-major axis of the objects orbit around the Sun is greater than 1.0 AU.
- The perihelion is smaller than Earth's aphelion.

Like the Amor asteroids, this gives the Apollos an orbital period of greater than one year. Because of their lower perihelion, the Apollos are Earth-crossing. The group is named after 1862 Apollo. As of March 2021, there are 14013 known Apollo asteroids, of which 1836 are classified as PHA<sup>7</sup>. The discovery fraction of Apollos is smaller than the Amors and Atens. This is possibly because their higher eccentricity makes the Apollos have a larger velocity when passing near Earth. This makes them harder to detect. This is supplemented by the fact that especially in the higher eccentricities or inclinations, there are less Apollos detected than expected **debiased**. The Apollos are slightly less dark than the Amors, with 30% of the population having  $p_v < 0.09$ . Furthermore, their size distribution can be modeled with  $k = 1.44 \pm 0.12$  for objects with  $D < 1.6\text{km}$  and  $k = 3.0 \pm 1.5$  for  $D > 1.6\text{km}$ **subpopulations**. The distribution of orbital parameters is shown in Figure 2.8. The asteroid which exploded over Chelyabinsk, Russia in 2013 was an Apollo asteroid.

<sup>7</sup>see: [https://ssd.jpl.nasa.gov/sbdb\\_query.cgi](https://ssd.jpl.nasa.gov/sbdb_query.cgi)

### 2.4.3. Aten Asteroids

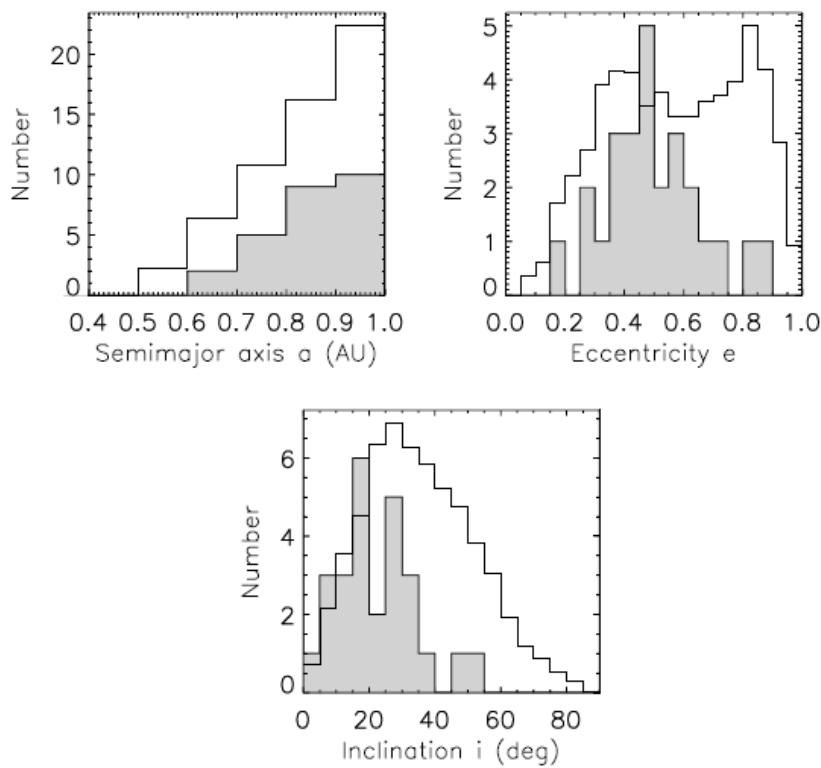


Figure 2.9: Expected (black line) and known as of 2001 (grey histogram) distribution of orbital elements of all Aten asteroids with  $H < 18$ . (**debiased**)

The Aten asteroid group features near-Earth objects conforming to the following criteria:

- The semi-major axis of the objects orbit around the Sun is smaller than 1.0 AU.
- The aphelion is greater than Earth's perihelion.

This gives the Aten asteroids an orbital period less than a year, and makes their orbits Earth-crossing. The group is named after 2062 Aten. As of March 2021, there are 1935 known Aten asteroids, of which 174 are classified as PHA<sup>8</sup>. The discovery fraction of the Atens is quite good, owing to their lower velocity than the Apollos, and the fact that, due to their lower semi-major axis, they can not be far away from Earth at opposition **debiased**. The Atens are a bright group, with only 17% of the population having  $p_v < 0.09$ . Their size distribution can be modelled using Equation 2.2 with  $k = 1.63 \pm 0.30$ . **subpopulations** The distribution of orbital parameters is shown in Figure 2.9. The asteroid 99942 Apophis, which briefly caused a scare in 2004 because of a rumored impact trajectory, is an Aten asteroid.

<sup>8</sup>see: [https://ssd.jpl.nasa.gov/sbdb\\_query.cgi](https://ssd.jpl.nasa.gov/sbdb_query.cgi)

#### 2.4.4. Atira Asteroids

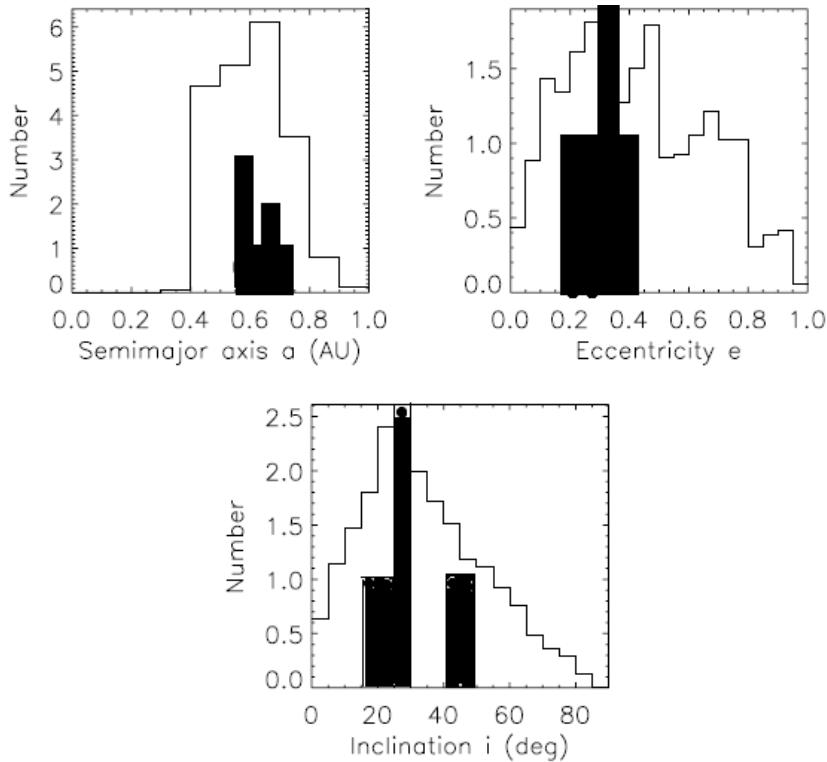


Figure 2.10: Expected (black line) and known as of 2020 (black histogram) distribution of orbital elements of all Atira asteroids with  $H < 18$ .(debiased). The black histogram was not presented in the original paper, but was added by the author of this report. This was done because there were zero known Atira asteroids in 2001. Note that there is still a lot of uncertainty in this population.

The Atira asteroids, sometimes also referred to as Apohele asteroids or interior-Earth objects, are near-Earth objects conforming to the following criteria:

- The semi-major axis of the objects orbit around the Sun is smaller than 1.0 AU.
- The aphelion is smaller than Earth's perihelion.

Similar to the Aten asteroids, the Atira asteroids feature an orbital period shorter than one year. Because of their characteristics, their orbits are entirely contained within Earth's orbit. The group was named after the first confirmed member, 163693 Atira. The Atira's are by far the smallest group, with, as of March 2021, only 23 known objects, of which 6 classified as PHA<sup>9</sup>. Although it should not be impossible to detect Atira asteroids, their observation is more challenging. A simple geometric calculation will show that the Atira asteroids can be viewed at phase angles around 70°. However, as most detections are serendipitous and the population is expected to be small, there are very few known Atira asteroids. **debiased**. Therefore, additional information is scarce. The expected distribution of orbital parameters is shown in Figure 2.10.

<sup>9</sup>see: [https://ssd.jpl.nasa.gov/sbdb\\_query.cgi](https://ssd.jpl.nasa.gov/sbdb_query.cgi)

# 3

## Optical Detection Technology

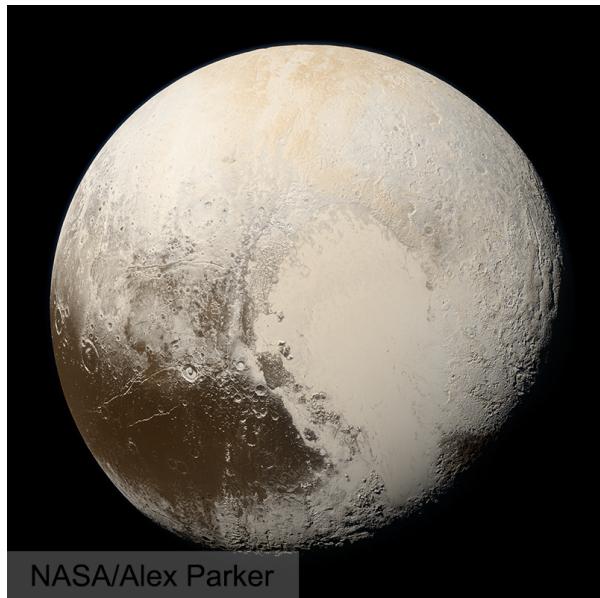


Figure 3.1: True color image of dwarf planet Pluto taken by the New Horizons space probe in 2015.

Optical sensors are some of the most common payloads of modern satellite systems. Ranging from Earth observation, to astronomy, to the cameras used by deep space probes to image their targets, the possibilities of optical sensors to provide detailed and human-interpretable information over large distances have become very mature. Having been used from the start of space exploration (The interested reader is encouraged to read the work of [evolutionofcamera](#)), optical detection and processing, which will be treated in chapter 5, have benefited greatly from the advances in semiconductor manufacturing and processing capabilities of modern space systems. From the Vidicon cathode ray tubes in the Voyager probes, spacecraft optical detectors have advanced greatly into modern day charge-coupled devices (CCD) and complementary metal-oxide-substrate (CMOS) detectors.

In this chapter, firstly the hardware and software used for optical detection will be discussed in section 3.1. Then, a fundamental and mathematical discussion of optics and detection hardware will be given in section 3.2 and lastly the main physical limitation to take into account is laid out in section 3.3.

### 3.1. Optical Systems

Before giving a mathematical treatment of spacecraft optical detection, firstly an overview of developed systems will be given to better understand the underlying process. Firstly, the sensor hardware will be

discussed, followed by a short description of the resulting data, and lastly the ways of magnifying far-away targets are laid out.

The most commonly used sensor in astronomy is the charge-coupled device (CCD) ([CCDimage](#)).



Figure 3.2: CCD sensor from a SONY Alpha commercial camera.

Figure 3.2 shows a CCD sensor. On this sensor, the pixels are light-sensitive metal-oxide substrate capacitors, which upon being hit by a photon produce an electric charge which is stored in the capacitor. CCD's are ideal for high-quality scientific cameras because of their high quantum efficiency, i.e. the conversion from impinging photon to electron in the sensor is almost 100%. This allows for high quality image generation, even in limited light conditions ([CCDimage](#)). However, this quality comes at a cost. Firstly, CCD's are fairly expensive to produce. This is however not a rarity in space systems. On a more practical level, CCD's suffer from an inherent issue known as "Bloom". This might be familiar to the reader where in overexposed images, or images with very bright areas, the brightness *leaks* into neighbouring pixels. This is a result of overflow in the capacitor bins due to overaccumulation of charge. It might lead to very detrimental results when the exposure is bad. On the other hand, this is contrasted by the possibility to reduce the noise from a variety of sources by a lot as described by [CCDsignal](#).

After exposure to the image, the CCD is read out sequentially by transferring the charge to an amplification circuit to obtain a voltage. This voltage is then measured, the value digitized as *data numbers* or *DN* and this result stored in memory ([CCDBook](#)). In practice, this requires selecting an amount of memory to allocate per pixel, the bit-depth (more memory allows a higher resolution in brightness), and storing the subsequent square array. Common values for bit-depth are 8-bit (allowing 256 different values) and 16-bit (allowing 65536 different values).

Optical systems like the CCD commonly require a device to focus incoming light rays onto the focal plane of the sensor. In addition, this allows for capturing photons from a larger area than the area of the sensor itself, therefore increasing the amount of photons coming in, and increasing the image quality. The most common devices are mirrors or lenses. An overview of common arrangements is shown in Figure 3.3. Most of these systems suffer from some aberrations, such as, but not limited to:

- Chromatic abberation: Different wavelengths of light refract at different angles. Therefore, there is a "prism"-like effect in lens telescopes.
- Spherical abberation: Imperfect mirror shapes cause blurring near the edge of the field.

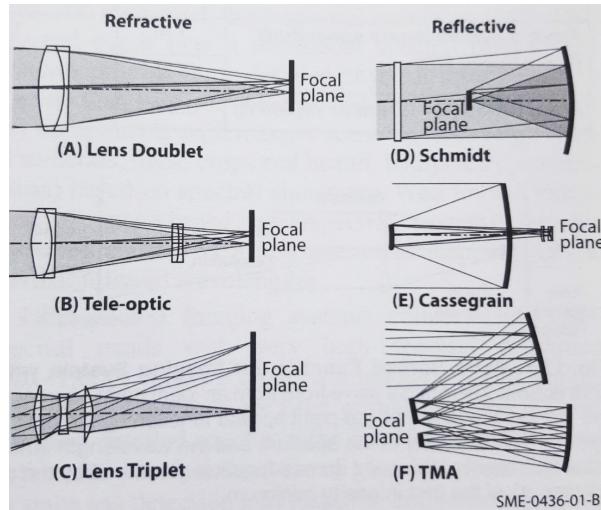


Figure 3.3: Overview of common mirror and lens setups. (**SMAD**)

- Field curvature: Because of the differing distances light has to travel, a flat object can not be brought to focus exactly on a flat detector.
- Coma: Nonuniformity in magnification leads to points near the edge of the field becoming "drawn out".

Although good to keep in mind, these aberrations can all be corrected for using properly and precisely designed optics. Therefore, luckily the mathematical treatment of these is practically consistent, so more detail on these setups is not required (**SMAD**). In section 3.4, an overview will be given of problems that are more troublesome to correct for.

## 3.2. Mathematics of Optics

After discussing the hardware, in this section a mathematical treatment of the optical systems will be presented. These definitions are needed to accurately assess the images in research. The problem is described by **OpNav** as follows:

*"Given the position and velocity of the spacecraft and the target, the camera attitude, and the camera's optical properties, where should the image of the target appear within a picture?"*

### 3.2.1. Target vector

**OpNav** begins his analysis of the problem by determination of the inertial direction to the target. Because the speed of light is finite, the target will appear on the spacecraft's sensors based on the *apparent position* rather than the actual, geometric position. In an inertial reference frame centered on the barycenter of the solar system:

- Calculate the position  $\vec{R}_C(t)$  and velocity  $\dot{\vec{R}}_C(t)$  of the camera at time  $t$ .
- Similarly, compute the position  $\vec{R}_T(t)$  and velocity  $\dot{\vec{R}}_T(t)$  of the target.
- The actual (or geometric) relative position  $\vec{G}(t) = \vec{R}_T(t) - \vec{R}_C(t)$ . Note that this is not the observed position, due to the fact that the target will have moved before the light reaches the camera.
- We can iteratively (because of the magnitude of  $c$  the required accuracy here is not very high, so some quick iterations suffice) calculate the time  $\tau$  which the light takes to reach the sensor from:

$$\tau = |\vec{R}_C(t) - \vec{R}_T(T - \tau)|/c \quad (3.1)$$

- The true position corrected for light time becomes:  $\vec{T} = \vec{R}_C(T) - \vec{R}_T(t - \tau)$ .

- Lastly, the effect of *stellar aberration*, which shifts this direction based on the direction of motion of the spacecraft, has to be taken into account. By performing a vector addition from simple geometry, the following equation for the apparent position  $\vec{A}(T)$  is obtained:

$$\vec{A}(t) = \vec{T}(t) + |\vec{T}(t)|[\dot{\vec{R}}_C(t)/c] \quad (3.2)$$

Equation 3.2 is a formulation of the apparent position of the target as seen from the spacecraft in Newtonian physics. For the proposed research, this is assumed sufficient.

### 3.2.2. Camera Attitude

With the apparent position vector to the target determined, **OpNav** continues by defining the attitude of the camera.

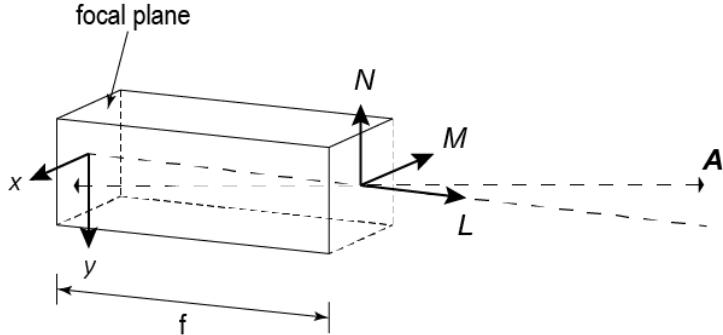


Figure 3.4: Illustration of coordinate frames used in the camera attitude determination

Start by determining a cartesian right-handed coordinate system  $C$ , using axes  $M-N-L$  with the  $L$ -axis oriented along the optical axis of the camera, and  $M$  and  $N$  being arbitrary perpendicular vectors (by convention  $M+$  points to the left and  $N+$  points up, as seen from the focal plane.) Figure 3.4 provides a graphical overview of the coordinate system. The transformation from the of the apparent position in the inertial frame  $\vec{A}^I$  to the apparent position vector in the camera frame  $\vec{A}^C$  is then given by a transformation matrix  $\mathbf{C}$ :

$$\vec{A}^C = \mathbf{C}\vec{A}^I \quad (3.3)$$

The analysis thus continues to obtain matrix  $\mathbf{C}$ . Note that merely finding the direction of the  $L$ -axis is not sufficient; the camera is capable of rotation along this axis, which will alter the resulting matrix. Matrix  $\mathbf{C}$  can be obtained from a sequence of rotations by Euler angles, such as:

$$\mathbf{C} = \mathbf{R}_3(\phi)\mathbf{R}_1(90^\circ - \delta)\mathbf{R}_3(\alpha + 90^\circ) \quad (3.4)$$

Or alternatively:

$$\mathbf{C} = \mathbf{R}_3(\phi + 90^\circ)\mathbf{R}_2(90^\circ - \delta)\mathbf{R}_3(\alpha) \quad (3.5)$$

With the standard rotation matrices:

$$\mathbf{R}_1(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{bmatrix} \quad (3.6)$$

$$\mathbf{R}_2(\theta) = \begin{bmatrix} \cos \theta & 0 & \sin \theta \\ 0 & 1 & 0 \\ -\sin \theta & 0 & \cos \theta \end{bmatrix} \quad (3.7)$$

$$\mathbf{R}_3(\theta) = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (3.8)$$

Here,  $\alpha$  represents the right ascension of the optical axis  $L$  and  $\delta$  represents the declination.  $\phi$  is a thirs twist angle to define the rotation of the camera along the  $L$ -axis. For simplicity, it will be assumed that camera attitude is sufficiently known and controllable that this definition of  $\mathbf{C}$  holds.

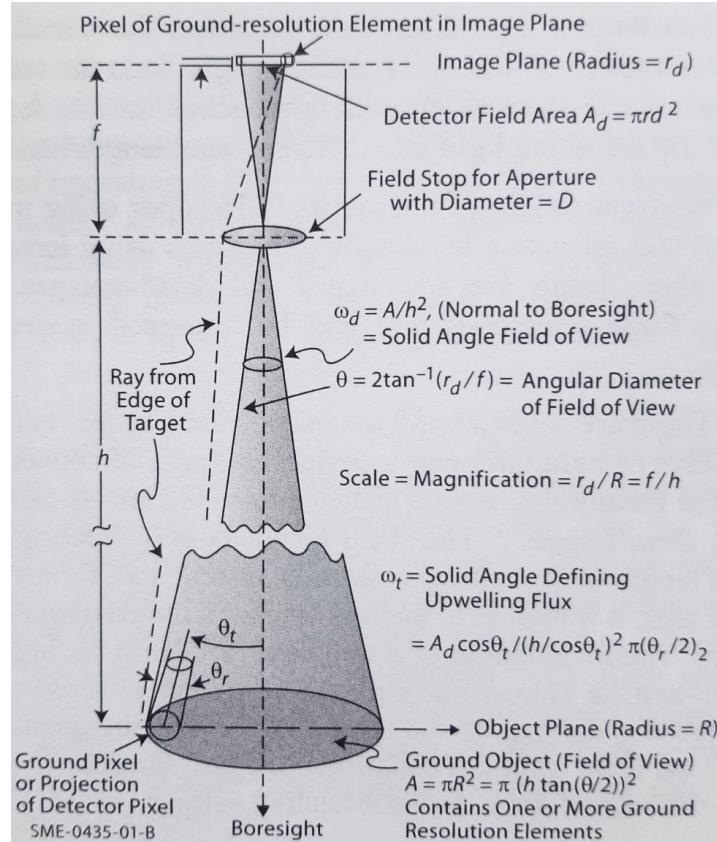


Figure 3.5: Description of elements of an optical system. The presented figure is for a refractive system, although mathematically a reflective system functions similarly. (**SMAD**)

### 3.2.3. Projection

By idealizing the camera's lens to a point, **OpNav** calculates the image's projection as follows. Because of this idealized assumption, Light rays from the target will pass through the point opening in the camera, at the position of the lens (the origin of **C**). As no lens is present but only a infinitely small hole, the light rays will continue in a straight line until reaching the detector in the focal plane, situated at focal length  $f$  behind the camera opening. From simple geometry, Figure 3.4 allows finding the so-called *gnomonic* projection in the coordinates of the detector (note that the L-axis, or optical axis is the third axis in the coordinate system):

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{f}{A_3^C} \begin{pmatrix} A_1^C \\ A_2^C \end{pmatrix} \quad (3.9)$$

The coordinate system shown in Figure 3.4 is "upside down", as the camera will project the vector in an inverted position. Note that because the right-hand side of Equation 3.9 involves a ratio of the components of the apparent position vector, the actual magnitude of this vector is irrelevant, which is confirmed intuitively when considering how a camera normally operates.

### 3.2.4. Lenses and Magnification

In reality, the lens of a camera is not an infinitely small opening in the front of the camera. For this purpose, lenses, mirrors or a combination of these are used to focus the light onto the sensor as described in section 3.1. As the advantages and disadvantages of each have already been described there, the explanation presented here will only focus on lenses, although the mathematics are consistent with mirrors with some trivial adjustment which can be derived from geometry. A perfect lens is a device that will focus incoming parallel light rays into a "focal point", where all the rays will exactly converge. This focal point is located at focal length  $f$  behind the lens. According to **SMAD**, the requirement of focal length is, in design, often dependent on the actual size of the detector and of the required angular resolution or field of view. Let  $M$  be the magnification of the target, defined as the ratio of image

dimensions to object dimensions:

$$M := \frac{r_d}{R} \quad (3.10)$$

With  $R$  the radius of the target and  $r_d$  the radius of the detector. It then follows, that for a perfect lens (see Figure 3.5):

$$M = \frac{f}{h} = \frac{r_d}{R} \quad (3.11)$$

With focal length  $f$  and distance to target  $h$  (**SMAD**). Next to the focal length, a second important property of a lens is how much light it collects. The amount of light is a direct function of the collector area. This is often described as the *infinity F-number* or *F-stop*. It is defined as the ratio of focal length to effective aperture of the main lens or mirror:

$$F\# = \frac{f}{D} \quad (3.12)$$

Assuming a constant photon flux, the number of incoming photons is directly proportional to area. Therefore, image brightness  $\propto F\#^{-2}$ , making this another property to consider. In practice, the lowest value of  $F\#$  used is 0.5 (**SMAD**).

### 3.2.5. From Projection to Pixel Values

Rounding out the description of optics, **OpNav** considers the process of converting from the previously found  $(x, y)^T$  vector to the measured values in the pixels of the detector. transforming the vector into line  $l$  and sample  $s$  is done through a linear transformation:

$$\begin{pmatrix} s \\ l \end{pmatrix} = \begin{bmatrix} K_x & K_{xy} \\ K_{yx} & K_y \end{bmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} s_0 \\ l_0 \end{pmatrix} \quad (3.13)$$

It will be assumed in further analysis that the used detector has perfectly square pixels. In this case  $K_x = K_y$  equalling the side length of a pixel. and  $K_{xy} = K_{yx} = 0$ . The offset  $(s_0, l_0)^T$  corresponds to the offset between the center of the optical axis, where the detector reference frame is located and the origin of the pixel reference frame.

With the reference frame defined, the last steps for determining the digital values are defining the incoming light and converting this to pixel values. Firstly, **OpNav** gives for the signal  $S(s, l)$  in a pixel measured in electrons:

$$S(s, l) = A \iiint I(\alpha, \delta, \nu, t) F(\nu) Q(\nu) d\nu d\Omega dt \quad (3.14)$$

With:

- Camera aperture  $A$
- Solid view angle  $\Omega$  subtended by a pixel.
- Incoming light  $I$  as a function of direction  $\alpha, \delta$ , frequency  $\nu$  and time  $t$  in units of  $\text{m}^{-2}\text{sr}^{-1}\text{Hz}^{-1}\text{s}^{-1}$ .
- A filter on the camera may cause another frequency dependent effect  $F(\nu)$ .
- Because of the usage of the photo-electric effect, there is a *quantum efficiency*  $Q(\nu)$ , the fraction of generated electrons to incident photons, which is also frequency dependent.

The conversion to data numbers is then straightforward from two properties of the conversion process:

$$DN(s, l) = S(s, l)/g + b \quad (3.15)$$

with  $g$  being the gain expressed as  $DN$  per electron, and  $b$  a constant offset term. These values can then be stored by the detector in a numerical array and used for further processing.

### 3.3. Diffraction Limitation

Of relevance to the research is the most extremely angular resolution achievable by the camera. This is important because of the intent to capture images of far away targets of small size. If all optical aberrations are corrected for, which is possible using advanced optics (**SMAD**), and in addition, sufficient noise is removed, which is possibly through a process known as *flat-fielding* (**OpNav**), the limitation to angular resolution is primarily the *Abbe diffraction limit* (**diffractionlimit**). The extent of this limitation is determined by the wavelength of the incoming photons and the aperture of the camera. Figure 3.6 shows a well known example created by shining a laser beam on a tiny puncture in a screen; effectively creating a minuscule aperture. The diffraction of the light will make discerning further detail in the source of the light impossible. With distance to target  $h$ , aperture diameter  $D$ , wavelength  $\lambda$  and target resolution limit  $X'$ , the corresponding relation is:

$$X' = \frac{h\lambda}{D} \quad (3.16)$$

Equation 3.16 can be used to quickly determine the maximum resolution achievable under the diffraction limit. For high-quality optics in space (i.e. no atmospheric disturbance), this is a fair assumption (**SMAD**). In Table 3.1, some examples are given for the diffraction limit for visible light of  $0.5\mu m$ . Even for large apertures<sup>1</sup>, it can be seen that the resolution will almost always be larger than the intended target size, with the exception of very close-by targets.

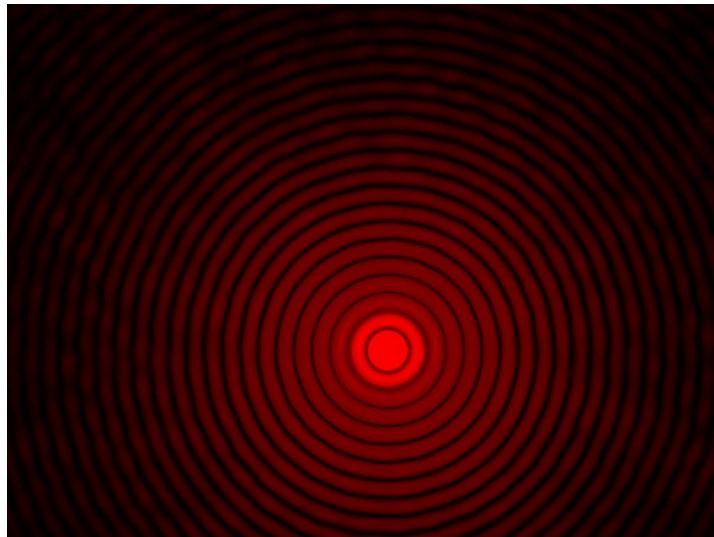


Figure 3.6: An Airy diffraction disk created by shining a laser on a 90-micrometer hole.

Table 3.1: Examples of diffraction target resolution limit for several apertures and distances for  $\lambda = 0.5\mu m$  visible light.

Distance to target	D=0.1m	D=0.5m	D=1m
10.000 km	50 m	10 m	5 m
1 Mm	5 km	1 km	500 m
1 AU	750 km	150 km	75 km

As the population of NEA's above the lower limit for target size (30 m) is approximately  $0.8 \cdot 10^6$ , which can be seen in Figure 2.3, a calculation is possible to judge this distance. Assuming all NEA's are spaced out evenly along Earth's orbit (the best-case scenario), the average distance between two asteroids:

$$\frac{1.5 \cdot 10^8 * 2\pi}{0.8 \cdot 10^6} = 1178 \text{km} \quad (3.17)$$

<sup>1</sup>For reference: the aperture of the Hubble Space Telescope is 2.4m

From this follows that, on average, only roughly 15 50m asteroids will be resolved by the camera of a spacecraft with a 0.1m aperture. Although overly simplified, it is clear that the asteroids should be taken as a point light source in the camera. As will be explained in chapter 5, it is often necessary to obtain more than one pixel of information to carry out detailed processing of the image. Luckily, diffraction offers an important solution in the form of the *Point Spread Function* (PSF). As described by **airyfunction**, the idealized PSF of intensity  $I$  as a function of dimensionless distance from the optical axis  $u$  is given by the airy disk, shown in Figure 3.6:

$$I(u) = \left( \frac{2J_1(u)}{u} \right)^2 \quad (3.18)$$

With:

$$u = \frac{\pi}{\lambda} D \theta \quad (3.19)$$

And  $J_1$  the first Bessel function of the first kind. The presented function is a simplification assuming no obstruction of the aperture (such as would occur in a reflector telescope). These functions are easily generated by e.g. Python, and an example of the result in dimensionless  $u$  quantities is shown in Figure 3.7. This function can be used to model the distribution of incoming light from a target asteroid.

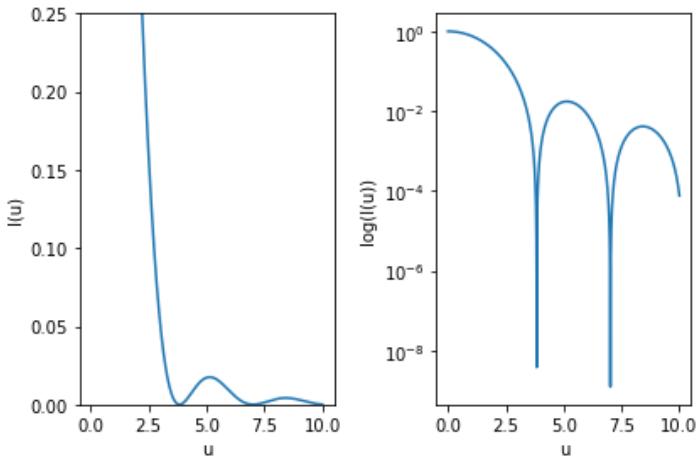


Figure 3.7: Simulated intensity of airy disk in 2D in dimensionless units  $u$

### 3.4. Noise and Bias

In modeling the optical system, there are several factors that need to be considered with regard to the signal to noise ratio of the produced data. There are several sources of noise that can not be compensated for completely, and thus need to be taken into account. In order of decreasing influence, they are (**OpNav; SMAD**):

1. Dark Current  $N_d$ : thermal effects will spontaneously produce electrons in the detector, even in the absence of light. These electrons are indistinguishable from "real" electrons generated by incident light, and therefore deteriorate the signal. Dark current is mainly a function of temperature, increasing with a warmer sensor.
2. Read Noise  $N_r$ : As with any system, interaction will alter the state of the system. Therefore, reading out the amount of charge in the detector will have an influence on the detector, and this is mainly a property of the quality of the detector electronics.
3. Fixed Pattern Noise  $N_f$ : A fixed pattern of dark current, always causing a certain offset in the data. This can have quite a large effect on the image quality, but as the pattern is constant (only being affected slowly by e.g. radiation damaging the sensor), it can be compensated for by calibration.

4. Shot noise  $N_s$ : The photon flux is not a continuous process, but rather a poisson process. Therefore, it has an additional random component in when photons reach the detector. This noise decreases with higher numbers of photons (i.e. brighter targets or longer exposure).
5. Quantization noise  $N_q$ : Because the analog voltage of the detector gets converted to a digital signal, a certain amount of noise is introduced.

The total noise can be determined as:

$$N^2 = N_d^2 + N_r^2 + N_f^2 + N_s^2 + N_q^2 \quad (3.20)$$

In addition, the background light level should be taken into account as an additional noise source. **nightsky** gives an estimate of  $B_e = 3 \cdot 10^{12}$  photons  $\text{sr}^{-1} \text{s}^{-1} \text{m}^{-2}$  in the absence of large bodies nearby (e.g. the Moon). **thesisspacebased** suggests taking the influence of starlight into account. Stars are roughly constant throughout the night sky, with the exception of the galactic plane. **starlight** estimate a mean flux from starlight of between  $B_e = 2 \cdot 10^{12}$  and  $3 \cdot 10^{12}$  photons  $\text{sr}^{-1} \text{s}^{-1} \text{m}^{-2}$  at angular distances over  $30^\circ$  from the galactic plane. When approaching the galactic plane, the intensity increases up to  $6 \cdot 10^{12}$  to  $8 \cdot 10^{12}$  photons  $\text{sr}^{-1} \text{s}^{-1} \text{m}^{-2}$  at the galactic plane.

According to **thesisspacebased**, the sources large enough to consider for the application are the dark current and read noise, as well as the background light. This brings a final estimate of the Signal-to-noise ratio by dividing the signal by the Root-Sum-Square of the noise sources(**OpNav**):

$$SNR = \frac{S_{target}}{\sqrt{N_d^2 + N_r^2 + B_e^2}} \quad (3.21)$$

In practice, SNR's of above three can be used, whereas lower SNR's might yield problems for the detection (**SMAD**). Some examples of typical values from **NASAreport** can be seen in Table 3.2.

Table 3.2: Typical values for spacecraft optical system

Parameter	Typical value
Aperture [m]	0.5
FOV [deg]	10.6 x 5.3
Number of pixels [V x H]	9232 x 9216
Exposure time [s]	24
Dark current [e-/h] at 100 C	2
Readout noise [e-]	4
Quantum efficiency [%]	88



# 4

## Near-Earth Asteroid Surveys

Presently, an not inconsiderable amount of effort is already being spent on cataloguing NEA's (**NEAsurveys**). Therefore, to determine the feasibility of the proposed system, it is important to assess the current capabilities of detection techniques, and to determine where the weaknesses of these systems lie. Firstly, an overview of Earth-based surveys will be given, followed by an overview of space-based surveys. This is by no means an exhaustive list, but these are the most prominent of the surveys.

### 4.1. Earth-based Surveys

By far the majority of current and past NEA surveys have been carried out from Earth. Although there are disadvantages because of Solar glare, the interference of the atmosphere, the fixed orbit of Earth, and effects such as weather and moon phases, logistical reasons often outweigh these.

#### 4.1.1. Catalina Sky Survey

The Catalina Sky Survey of the University of Arizona's Lunar and Planetary lab operates a 1.5m telescope on Mount Lemmon in Arizona. This telescope is covering approximately  $1000\text{deg}^2$  per night with a limiting magnitude of 21.5. In addition, the survey operates two smaller telescopes. The Catalina Sky Survey is based on a limiting size of 140m, per the definition of a potentially hazardous asteroid, and has discovered over 12000 NEA's since the start of operation in 1995 (**catalinaskysurvey**).

#### 4.1.2. Pan-STARRS

Pan-STARRS, the Panoramic Survey Telescope and Rapid Response System features two 1.8m telescopes at the Institute for Astronomy at the University of Hawaii. Their limiting magnitude is slightly better than the Catalina Sky Survey, at a theoretical maximum of 24. It has been operational since 2010 and is capable of imaging approximately  $6000\text{deg}^2$  of sky per night. It is expected to discover a very large number of undiscovered NEA's, among other discoveries. For example, Pan-STARRS was also the telescope which discovered the 2017 interstellar object 'Oumuamua (**Panstarrs**).

#### 4.1.3. ATLAS

ATLAS, the Asteroid Terrestrial-impact Last Alert System designed around the idea of being a last alert system for asteroids approaching Earth. It "will provide one day's warning for a 30-kiloton 'town killer,' a week for a 5-megaton 'city killer,' and 3 weeks for a 100-megaton 'country killer'." (**Atlasite**) The idea here is to provide a last alert for any asteroids missed by other surveys and allow humans time to evacuate. The two 0.5m telescopes of ATLAS are situated on two mountain peaks in Hawaii, and the observatories are run by the University of Hawaii. The impressive aspect of ATLAS is its high cadence: With the two telescopes working together, ATLAS is capable of imaging the night sky as visible from Hawaii four times every two days. This high cadence is required as multiple repeat observations are needed to compute orbits accurately. The team is looking to build more "ATLAS units" worldwide, although issues with weather and solar interference remain (**ATLAS**).

#### 4.1.4. Zwicky Transient Facility

While not a dedicated NEA facility, the Zwicky Transient Facility (sometimes referred to by the name of its predecessor, the Palomar Transient Factory ), yields many observations of NEA's using its 48-inch (approx 1.2 m) telescope. While not particularly powerful in terms of limiting magnitude, the Zwicky Transient Facility shines in its high survey cadence, capable of imaging the night sky every three nights, and the plane of the Milky Way twice per night. The ZTF is the precursor to the upcoming Large Synoptic Survey Telescope, which will feature an 8.4 meter telescope imaging the sky at a high cadence. It is expected to vastly increase the number of mapped objects (by a factor of 10-100), although this is not its primary goal (**LSST**).

### 4.2. Space-based Surveys

Contrary to Earth-based systems, a space-based telescope can be used to image without interference from weather, or the atmosphere in general. It therefore provides superior optical performance, allowing smaller systems to be used effectively.

#### 4.2.1. NEOWISE

NEOWISE was the second mission of the WISE spacecraft. Originally intended to image the entire sky in infrared using cooled optics, the spacecraft was placed in hibernation after the coolant ran out. In 2013, the spacecraft was woken from hibernation and assigned a new mission: to image the solar system for near-Earth objects. From December 2013 through 2017, NEOWISE detected 34000 new objects, among which 262 NEA's, of which 47 classified as a PHA (**NEOWise**).

#### 4.2.2. NEOSSat

Launched in 2013, The Near-Earth Object Surveillance Satellite is a Canadian mission specifically scanning the sky near the Sun for inner-Earth objects, particularly the Aten and Atira classes. Their orbit makes them harder to detect from Earth, making a space-based survey ideal. NEOSSat is an interesting design when contrasted to the other surveys, as it only uses a 0.15m telescope mounted on a microsatellite, and uses its primary telescope also as a star sensor for attitude determination. However, despite these limitations, it is capable of imaging down to apparent magnitude 19 (**Neossat**; **Neossattwo**).

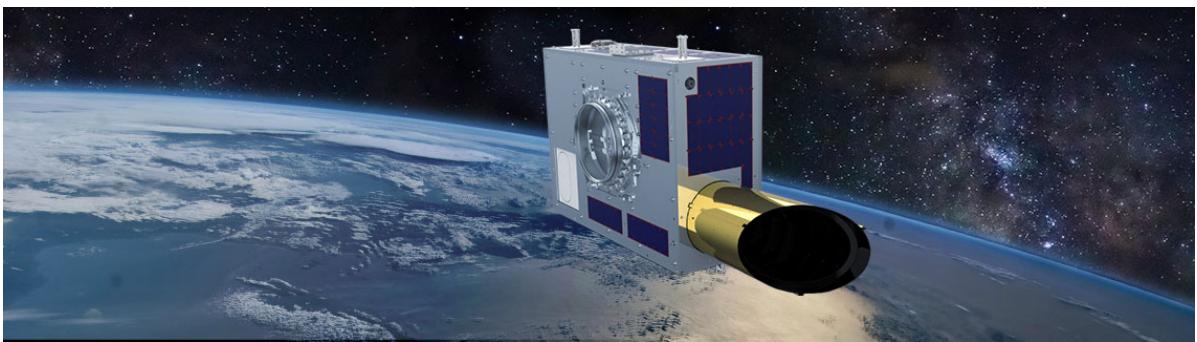


Figure 4.1: Artist rendition of the NEOSSat satellite, showing the aperture of the 0.15m Maksutov-Cassegrain telescope. (Image Courtesy: Canadian Space Agency)

# 5

## Image Processing Techniques

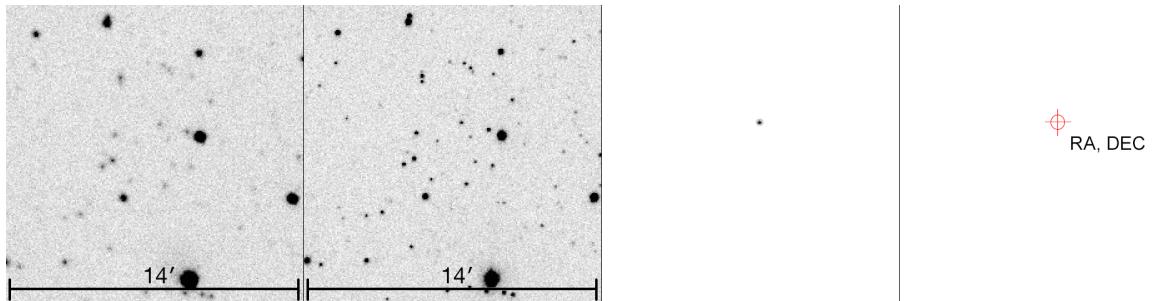


Figure 5.1: Overview of the process of image processing (images from [ATLASBlog](#)). Note that these images were not obtained using the mentioned methods, but merely serve as illustration.

After capturing the image, the process of image processing can begin. This process transforms the data from the sensors to measurements needed for determining the trajectory of the asteroid. This process consists of various steps, of which an overview is provided in Figure 5.1. Firstly, the image is taken by the camera. Then, various steps are undertaken to improve the quality of the image. Next, the background of the image is removed, leaving only the PSF of the target. Lastly, the center of the target is determined. As the first step consists of only a well-described calibration procedure known as *flat-fielding* (e.g **SMAD**, **OpNav**), it will not be discussed in further detail. Therefore in this chapter only the background removal and centerfinding procedures are detailed. In addition, for completeness, some applications of neural networks for image processing will be discussed.

### 5.1. Background Removal

After compensating for known effects such as the fixed pattern noise described in section 3.4, the system obtains an image consisting of three components: The background, the noise and the *transients*, such as asteroids. Thus, the signal in each pixel can be described as a linear combination of signal, background and other noise. The static background signal  $B_e$  comprises primarily background stars and zodiacal light. This background signal can be removed in several ways. The most basic way is to find the value of the background light and subtract this from all signals, such as described by **sextractor**. Various methods for estimating the background value of an image exist, but most methods operate under the assumption that there is always more sky than objects in the image. Therefore, if one is to plot a histogram of pixel values, the mode of this histogram will provide a reasonable assumption for the background signal.

There is, however, a more accurate method available at the cost of more work to remove the background objects: Because of the extremely far distance to other stars, they are essentially unmoving as seen from the space system. Consider a spacecraft in an orbit with a semi-major axis of 1 AU, with an angular resolution of 2 arcseconds per pixel (from the values in Table 3.2). It follows from basic

trigonometry that a target further away than 3.26 lightyears would provide less parallax error than the resolution of the sensor.<sup>1</sup> Of course, this is a shorter distance than even the distance to the nearest other star, Proxima Centauri at more than 4 light years. Therefore, the background can be assumed constant. Thus, it is possible to perform a procedure known as "image differencing", where a reference image is subtracted from the new image, effectively removing the background signal. Thus, the system is left with a combination of signal and the remaining noise components. The reference images can be composed of other images taken by the system itself (e.g. **subtraction**), or the images can be created (e.g. **PalomarPipeline**). This is possible as almost all objects responsible for the background signal are known. These objects are logged in databases such as ESA's GAIA (**GAIA**), which contains positions of over one billion stars. This can be combined with ephemerides from known solar system objects, and thus the difference can be found. Then, through statistical methods targets with a large enough SNR can be found, as is shown in Figure 5.2. This method will of course find more transients than just asteroids; also other phenomena such as, but not limited to, supernovae, variable stars, eclipsing binaries and new quasars will be detected. How to determine the difference between these objects is beyond the scope of this review, although some examples will be given in the final section of this chapter.

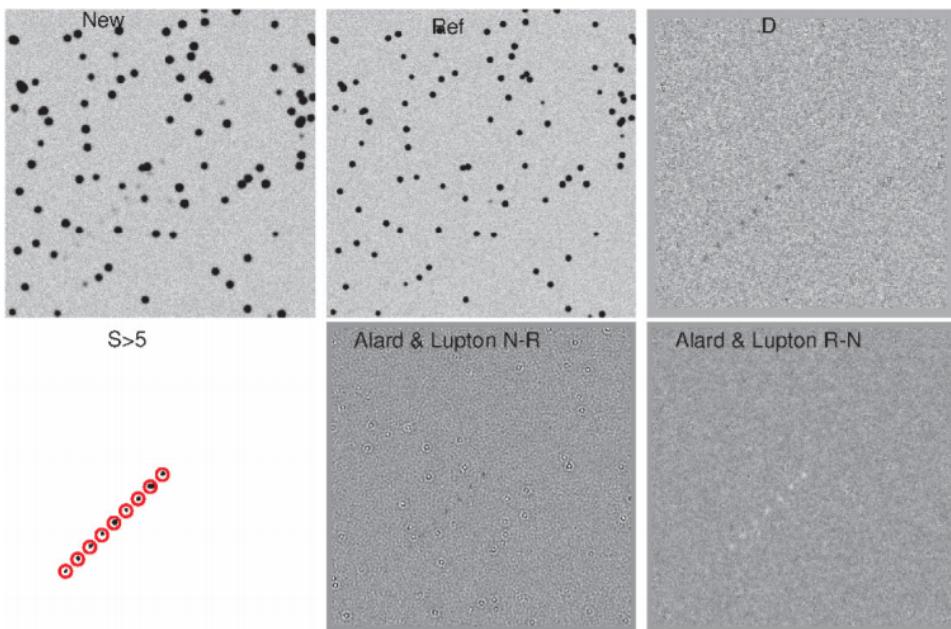


Figure 5.2: Demonstration of the image subtraction process from a series of stacked images. The path of the transient is shown in red (**processingclassic**).

## 5.2. Centerfinding

With the image taken by the camera reduced to a projection of the PSF, the question remains on how to transform this imaged PSF into an angular measurement as accurately as possible. Note that it will only be possible to determine the angular position of the asteroid accurately, not its actual position, as the point spread function is not dependent on the distance to the target, and the target is too small to resolve into multiple pixels. The angular measurements will be combined with knowledge on astrodynamics in the next chapter to solve for the position of the target. However first, this angular position needs to be determined as accurately as possible. **OpNav** provides an overview of the main algorithms suitable for this purpose:

The most simple to implement algorithm is to model a point spread function (resulting in an array **P**) and calculate the correlation across the pixel grid **D** which contains the point spread function. Therefore,

<sup>1</sup>A more convenient unit here would be the *parsec*,  $1pc \approx 3.26ly$ . The parsec is the distance of an object where 1 AU subtends an angle of 1 arcsecond

the challenge is to find  $\mathbf{P}$  in  $\mathbf{D}$ . Let  $\mathbf{P}'$  and  $\mathbf{D}'$  be zero-mean and unit standard deviation normalized versions of the aforementioned arrays. Then, **OpNav** defined the *correlation array*  $\rho(\Delta s, \Delta l)$  as the sum of the products of the arrays, with  $\mathbf{P}'$  shifted by integer amounts  $\Delta s$  and  $\Delta l$  (in such a way that  $\mathbf{P}$  is entirely inside the coordinates of  $\mathbf{D}$ ):

$$\rho(\Delta s, \Delta l) = \sum_i \sum_j \mathbf{P}'_{i,j} \mathbf{D}'_{i+\Delta s, j+\Delta l} \quad (5.1)$$

In practice, this is more convenient by first computing with the unnormalized arrays:

$$m_d = \sum \mathbf{D}_{i+\Delta s, j+\Delta l} \quad (5.2)$$

$$s_d = \sqrt{\sum \mathbf{D}_{i+\Delta s, j+\Delta l}^2} \quad (5.3)$$

$$m_p = \sum \mathbf{P}_{i,j} \quad (5.4)$$

$$s_p = \sqrt{\sum \mathbf{P}_{i,j}^2} \quad (5.5)$$

$$r = \sum \mathbf{D}_{i+\Delta s, j+\Delta l} \mathbf{P}_{i,j} \quad (5.6)$$

Then, with  $N$  the number of pixels in the summation:

$$\rho(\Delta s, \Delta l) = \frac{Nr - m_d m_p}{\sqrt{(Ns_d - m_d^2)(Ns_p - m_p^2)}} \quad (5.7)$$

This will result in a value of  $\rho$  between -1 for perfect anticorrelation, and +1 for perfect correlation. Therefore, the best estimate of  $\Delta s, \Delta l$  is given by the maximum of  $\rho$ , which can be determined to sub-pixel accuracy by e.g. interpolation. For example, consider a parabola fitted to the following points:

$$\rho_0 = \rho(\Delta s, \Delta l) \quad (5.8)$$

$$\rho_- = \rho(\Delta s - 1, \Delta l) \quad (5.9)$$

$$\rho_+ = \rho(\Delta s + 1, \Delta l) \quad (5.10)$$

It follows that the maximum value of  $\rho$  is found at:

$$x = 2 \frac{\rho_+ + \rho_- - 2\rho_0}{\rho_+ - \rho_-} \quad (5.11)$$

Similarly, the second coordinate can be solved for  $y$ . Note that the approach of finding the correlation requires some imprecise interpolation. This might be unwanted, as **veryshortarcs** results indicate a large drop in performance for measurements with a precision worse than 4 or 5 decimals (in degrees). Therefore, a slightly more complicated process which does not require interpolation is to fit some analytical function to the data. **OpNav** mentions that the choice of which function to use is not very important, as long as it is reasonable. Some possibilities include the Lorentzian function (**lorentzian**):

$$B(s, l) = \frac{h}{1 + (r/r_0)^2} + B_e \quad (5.12)$$

Or the two-dimensional Gaussian (**OpNav**):

$$B(s, l) = \frac{h}{2\pi} \exp\left(-\frac{(s - s_c)^2 + (l - l_c)^2}{s\sigma^2}\right) + B_e \quad (5.13)$$

$$\equiv h\mathcal{N}\left(\frac{s - s_c}{\sigma}\right)\mathcal{N}\left(\frac{l - l_c}{\sigma}\right) + B_e \quad (5.14)$$

$$\equiv h\mathcal{N}(\xi(s))\mathcal{N}(\eta(l)) + B_e \quad (5.15)$$

With  $\sigma$  the standard deviation in the pixels,  $h$  the magnitude of the pixel values in *DN* and  $B_e$  the background signal. Where  $\mathcal{N}(z)$  is the zero-mean normal distribution with  $z$  standard deviation. The functions  $\xi(s)$  and  $\eta(l)$  are normalisation functions to zero mean and unit variance. As this constitutes

a density function, it follows that the values of the pixels should be the integral over the area of the pixel. Note that, since background removal is performed prior, it is assumed that  $B_e = 0$ :

$$DN(s, l) = \int_{s-0.5}^{s+0.5} \int_{l-0.5}^{l+0.5} B(x, y) dy dx \quad (5.16)$$

$$= h [erf(\xi(s + 0.5)) - erf(\xi(s - 0.5))] [erf(\eta(l + 0.5)) - erf(\eta(l - 0.5))] \quad (5.17)$$

With  $erf(z)$  the Gaussian error function:

$$erf(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt \quad (5.18)$$

The function fitting concerns the variables  $s_c, l_c, h, \sigma$ . These begin by estimation from the data, from which the expected  $DN$  values can then be computed. Subtraction from the measured values yield an array of residuals, which can be used to iteratively converge to a solution through e.g. gradient descent, or by computing the first-order partial derivatives of  $DN(s, l)$  with respect to  $s_c, l_c, h, \sigma$  and performing a non-linear least-squares fit.

After performing one of these estimations, Equation 3.9 can be used to determine the unit vector to the target in relation to the optical axis of the camera. If the unit vector of the optical axis is known, a solution in e.g. celestial coordinates follows trivially. Note that the vector of the optical axis can be determined very precisely, as the camera can also function as a star tracker, such as demonstrated by the NEOSSat mission exhibited in subsection 4.2.2.

### 5.3. Applications of Neural Networks

Like any field of technical research, astronomy is faced with the challenge and opportunity of exponentially increasing amount of data to process. For this reason, big data techniques such as machine learning and especially artificial neural networks have become widely researched. A further discussion of the workings of artificial neural networks and how to use them for the proposed research is given in section 6.4. In this section, only an overview of the current research regarding neural networks in astronomical image processing is given. It is assumed the reader has a basic familiarity with convolutional neural networks; else they are advised to refer to some of the excellent literature on this topic such as **nnbooktwo** or **nnbookfour**<sup>2</sup>.

Convolutional Neural Networks (CNN's) are the subject of several ongoing research projects. A few interesting results are presented here to give an overview of the current state of research. The first application of CNN's is in the determination of whether or not something is truly a target, or merely noise or an image artifact. CNN's provide an effective way of discriminating between these without human intervention.

Firstly, **AtlasDL** examine the possibility of using a two stage deep learning classifier on data gathered by the ATLAS (subsection 4.1.3) telescopes. In Figure 5.3, an example is given on the types of images the classifier must process. The system is capable of classifying eight different classes: bright comets, asteroid streaks, cosmic rays, and "other astronomical objects", as well as several types of noise and visual artifacts. The network is trained entirely using real-world data. Using a CNN for feature extraction coupled with a dense classifier allows the authors to obtain a 0.4% false negative rate coupled with a 9.1% false positive rate. In practice, this means a reduction of over 90% in workload for humans who would otherwise have to classify the images.

Next to ATLAS, the Zwicky Transient Facility (subsection 4.1.4) is used as a testbed for several processing algorithms to be used on the Large Synoptic Survey Telescope. **supernovacnn** provided the groundwork for application of CNN's at the Zwicky Transient Facility in detecting supernovae, and **processingDOne** expanded greatly on this research. The authors use a deep convolutional neural network to classify images as target or bogus, achieving over 99% accuracy, precision and recall, and outperforming previously used random forest (a type of machine learning) classifiers. The network is

---

<sup>2</sup>An online version of this book is available for free at <https://www.deeplearningbook.org/>.

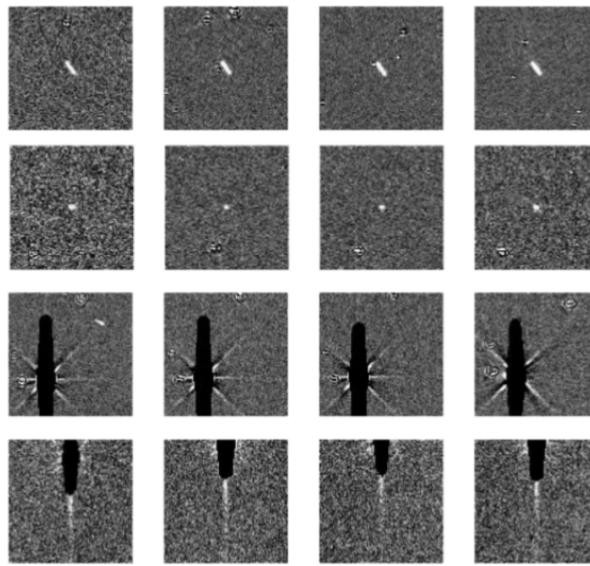


Figure 5.3: Some images from ATLAS data. The first two rows show actual detections of asteroids, whereas the latter two rows show visual artifacts. Note how the white streaks in the bottom images might be mistakes for asteroid tracks in a more simple algorithm (**AtlasDL**.)

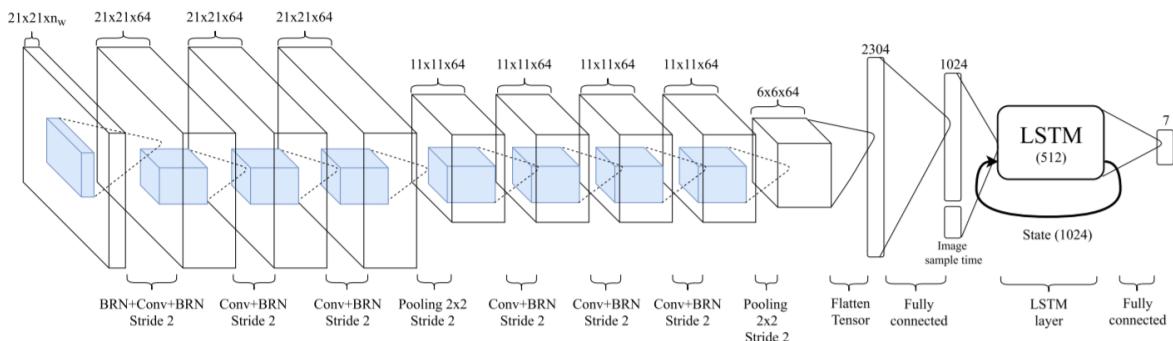


Figure 5.4: Schematic of the network used by **processingDLtwo** for classification into seven classes of observations.

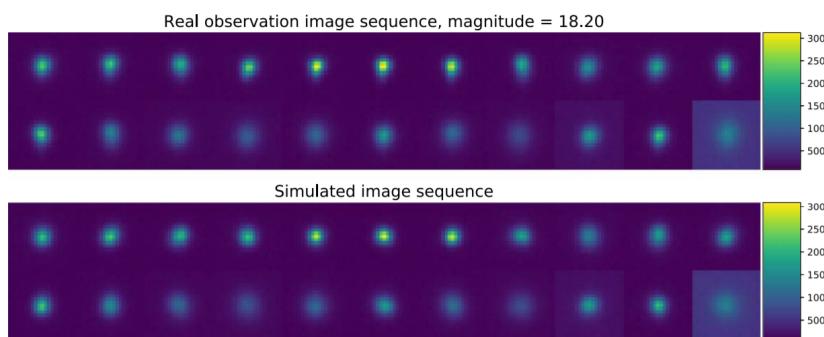


Figure 5.5: Comparison of real and simulated image sequences from **processingDLtwo**.

also trained entirely on real data. An example of using simulated data to effectively train the network was given by **processingDLtwo**: without performing image subtraction, the authors use a convolutional neural network for feature extraction followed by LSTM recurring neural network. A schematic of the network architecture is shown in Figure 5.4. They manage to obtain results similar to classical processing methods without performing the computationally intensive steps associated with those methods. Perhaps more interesting, is that the network was trained on a completely simulated dataset, and later validated on data captured by the telescopes. Figure 5.5 provides an example of a simulated data sequence compared to a real data sequence.

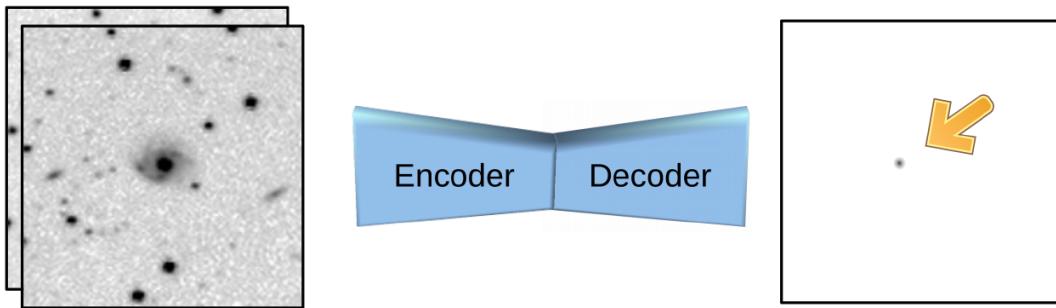


Figure 5.6: The system trained by **autoencoder** and its results.

As a last impressive example of the power of CNN's in image processing, **autoencoder** present another system to be tested at the Zwicky Transient Facility. The author make use of an autoencoder: a type of neural network architecture where the network is trained to replicate its input as accurately as possible. The ingenuity in the autoencoder resides in enforcing a dimensionality reduction on the data (by creating a "bottleneck" in the network), thus forcing the network to learn a more compact representation of the data. It is thus hoped that all patterns in the data are learned and preserved, and any noise will not be encoded and "lost" in the deencoding process. **autoencoder** take a small sidestep from this concept, instead providing the 16-layer fully convolutional autoencoder with both an image taken by the telescope, containing a target and a reference image, as discussed earlier. They then train the system to reproduce the image of just the target. Thus, a method of processing the image into purely the PSF of the target is obtained without the need for any classification or image subtraction. The system is trained on a combination of real data and augmented data from freely available datasets. An example of the system is shown in Figure 5.6, which demonstrates its powerful processing. The authors demonstrate decent performance ( $\approx 80\%$  recall) for relative magnitudes of around -0.1 of the transient compared to the background, and very good (close to 100%) for relative magnitudes of -0.15 and beyond. The system yields vastly cleaner images of the PSF than e.g. **processingclassic**, although the authors admit not being experts in the application of the latter.

In conclusion, neural networks open some very exiting possibilities for processing this kind of data, and can be applied using both real and simulated data. Challenges lie in verifying and validating the data, as well as some common pitfalls encountered in the training of neural networks, which will be discussed in more detail in the next chapter.

# 6

## Trajectory Determination Techniques

As discussed in the previous chapter, after a full sequence of image processing, the system obtains the location of the centroid of the target. As determination of distance is unfeasible, this location is given in terms of exclusively angular measurements. Therefore, to determine the trajectory, the problem reduces to determining the trajectory of the target through purely angular measurements. The reader is encouraged to ponder over this statement for a moment, as its inherent challenges form the basis of this chapter.

Trajectory determination is a wide and often studied field, with applications in for example sports (**trajtwo**), industry (**anntrajectoryone**, **trajone**, **trajthree**), defense (**anntrajectoryone**) and more close to the area of research: in meteor trajectories in Earth's atmosphere (e.g. **trajfour**, **trajfive**, **trajsix**) and orbits of satellites (**phddissertation**). The promising methods have been compiled into this section as a suggestion on the future research methods. Firstly, some methods specific to orbits are discussed, such as Gauss' method. Then, more general techniques which can be applied to this problem are explained. The actual implementation of them, however, is beyond the scope of this report.

### 6.1. Gauss' Method

The first to attempt a solution to this problem was **gaussceres**. Gauss applied his method to solve for the orbit of the dwarf planet Ceres, discovered a few years prior. Its position was known only at each instant  $n$  only in terms of right ascension  $\alpha_n$  and declination  $\delta_n$ , and thus also the unit vector  $\hat{\rho}_n$  from the observation point to the orbiting body could be defined. Gauss' method has been replaced by more accurate methods, but the ideas in it are fundamental enough to warrant inclusion in this report.

Taking three observations, Gauss' method is performed as follows (**curtis**). Firstly, define timesteps:

$$\tau_1 = t_1 - t_2 \quad (6.1)$$

$$\tau_3 = t_3 - t_2 \quad (6.2)$$

$$\tau = t_3 - t_1 \quad (6.3)$$

Then, take the vector products of the unit vectors:

$$\vec{P}_1 = \hat{\rho}_2 \times \hat{\rho}_3 \quad (6.4)$$

$$\vec{P}_2 = \hat{\rho}_1 \times \hat{\rho}_3 \quad (6.5)$$

$$\vec{P}_3 = \hat{\rho}_1 \times \hat{\rho}_2 \quad (6.6)$$

Calculate ten scalar quantities.  $D_0$  the triple product of the unit vectors, and  $D_{11}$  through  $D_{33}$  following from the product of the  $\vec{P}_n$  with the observers position vector at time of observation  $\vec{R}_n$

$$D_{mn} = \vec{R}_m \cdot \vec{p}_n \quad (6.7)$$

From this, calculate the following polynomial coefficients:

$$a = -(A^2 + 2AE + R_2^2) \quad (6.8)$$

$$b = -2\mu B(A + E) \quad (6.9)$$

$$c = \mu^2 B^2 \quad (6.10)$$

with:

$$A = \frac{1}{D_0} \left( -D_{12} \frac{\tau_3}{\tau} + D_{22} + D_{32} \frac{\tau_1}{\tau} \right) \quad (6.11)$$

$$B = \frac{1}{6D_0} \left( D_{12} (\tau_3^2 - \tau^2) \frac{\tau_3}{\tau} + D_{32} (\tau^2 - \tau_1^2) \frac{\tau_1}{\tau} \right) \quad (6.12)$$

$$E = \vec{R}_2 \cdot \hat{\rho}_2 \quad (6.13)$$

$$R_2^2 = \vec{R}_2 \cdot \vec{R}_2 \quad (6.14)$$

Then, the polynomial

$$r_2^8 + a_2^6 + b_2^3 + c = 0 \quad (6.15)$$

Can be solved for the scalar distance to the target at time 2. From here, the slant range and slant position vector can be calculated easily through vector operations. This yields the position vectors  $\vec{r}_n$  of the target at the times of observation. From this, the first two terms of the Lagrange series can be determined:

$$f_1 \approx 1 - \frac{1}{2} \frac{\mu}{r_2^3} \tau_1^2 \quad (6.16)$$

$$f_3 \approx 1 - \frac{1}{2} \frac{\mu}{r_2^3} \tau_3^2 \quad (6.17)$$

$$g_1 \approx \tau_1 - \frac{1}{6} \frac{\mu}{r_2^3} \tau_1^3 \quad (6.18)$$

$$g_3 \approx \tau_3 - \frac{1}{6} \frac{\mu}{r_2^3} \tau_3^3 \quad (6.19)$$

And the velocity of the target follows from:

$$\vec{v}_2 = \frac{1}{f_1 g_3 - f_3 g_1} (-f_3 \vec{r}_1 + f_1 \vec{r}_3) \quad (6.20)$$

The position and velocity vector of the target have been found, and therefore the orbit has been determined. From here, methods exist to improve the accuracy through small corrections. Although this process in itself is not very interesting for the research, it is important to discuss its assumptions and how these make it infeasible. In particular: the estimation of the Lagrange coefficients is not very accurate, the solution assumed a two-body motion and the solution is very sensitive to errors in the initial measurements in the case of a very short arc, as shown by **classicmodernorbits** in Figure 6.1: the initial error in measurement causes a completely erroneous estimate.

Several of the estimation problems can be resolved, e.g. through the methods proposed by **anglesonly**. Der postulates that the efficacy of initial orbit determination might be improved through elimination of numerically sensitive methods such as numerical integration of higher order partial derivatives, inversions of matrices or certain iterative methods. Firstly, the roots of Equation 6.15 might be determined rather than guessed through application of logic (e.g. on what side of the Sun is the orbit going to be). Secondly, instead of determining or guessing a single range vector,  $\vec{r}_2$ , it is possible to instead solve for  $\vec{r}_1$  and  $\vec{r}_3$ . This reduces the problem to Lambert's problem: determine the orbit from two range vectors and time of flight, for which a solution exists (for more information, refer to **curtis**). Lastly, it is possible to solve the Lambert problem for perturbed orbits (e.g. **superlambert** and **superlamberttwo**) using modern numerical algorithms. The accuracy of the resulting calculation is dependent on the measurements, as discussed earlier. For a good set of measurements, Der recommends spacing the measurements by several days to months, but within a single orbit. In addition, the accuracy of the measurement should be in the order of 5 significant digits or better. three to four significant digits might

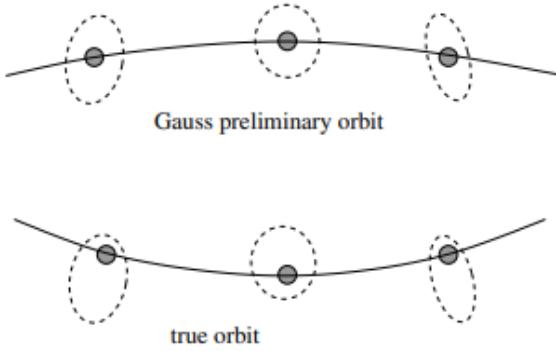


Figure 6.1: Actual orbit and orbit as determined by Gauss' method for a short arc

yield "acceptable" results, anything less will result in unusable results. For the intended research, the angular resolution is not an inherent problem, although the observation window might be. Therefore, methods applicable to shorter arcs should also be considered.

## 6.2. The Method of Very Short Arcs

The theory on trajectory determination for very short arcs was first introduced by **veryshortarcs**, who worked out the solutions in more detail in subsequent papers (e.g. **veryshortarcstwo**). Starting from fundamentals, define a *sequence of observations* as a set of observations belonging to the same object:

$$t_i, \alpha_i, \delta_i, h_i; i = 1, 2, \dots, m; m \geq 2 \quad (6.21)$$

With:

- $t_i$ : time of the observation.
- $\alpha_i, \delta_i$ : right ascension and declination at time  $t_i$ .
- $h_i$ : apparent visual magnitude (optional)

These observations can be said to belong to the same object if they allow construction of a *very short arc*: the observations are close enough together that they can be fit to some smooth curve (**classicmodernorbits**). From here, a vector shall be defined as *attributable* if:

$$\xi = (\alpha, \delta, \dot{\alpha}, \dot{\delta}) \in [-\pi, \pi) \times (-\pi/2, \pi/2) \times \mathbb{R}^2 \quad (6.22)$$

If the apparent magnitude is available, it is also included in the attributable. If the arc described by the observations does not allow for a convergent estimation of the orbit using a classical method as described in section 6.1, the arc is termed a *too short arc*. In this case, it is important to note that the quantities of interest, namely the position and velocity  $r$  and  $\dot{r}$  are completely undetermined. In this case, an attributable can still be used to perform several operations:

- The attributable can be *attributed* to a known object with a known orbit using a least-squares estimation. This essentially identifies the attributable as belonging to that object.
- The attributable can be *linked* to another attributable of (possibly) the same object. This lengthens the arc and might allow orbit computation
- The attributable can be linked to a new attributable from an object in the sky (recovery) or the archives (*precovery*), by the same logic as the previous point.

It might seem like the possibilities end here. However, it is possible to couple the information from the attributable to information on the population of the objects of interest. **veryshortarcs** describe this process for observations of targets in heliocentric orbits from Earth, but the process is easily adapted. Milani et al. define the *admissible region* as the subset of  $(r, \dot{r}) \in \mathbb{R}^2$  for which the following holds:

1. The target is not in an orbit around Earth, i.e.  $\mathcal{E}_{\oplus}(r, \dot{r}) \geq 0$ . This is only useful if:
2. The object is inside the sphere of influence of Earth:  $r \leq 0.01044AU$ .
3. The object is not interstellar:  $\mathcal{E}_{\odot}(r, \dot{r}) \leq 0$ .
4. The object is not very small (and thus very close):  $H \leq H_{max}$  for some reasonable value of  $H_{max}$ .

Logically, it follows that the admissible region is thus defined as  $[(1) \cup (2)^c] \cap (3) \cap 4$ , which is shown qualitatively in Figure 6.2. Importantly, the concept of *admissible region* can be adapted to the pur-

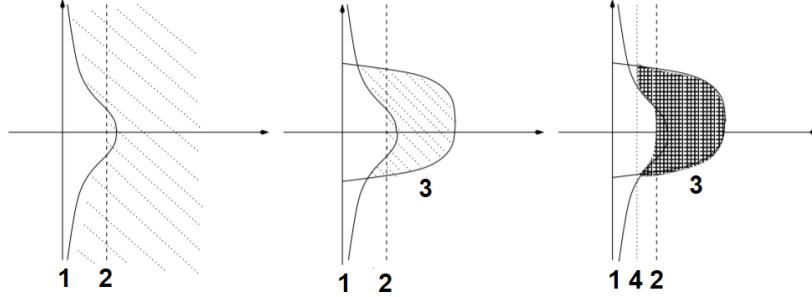


Figure 6.2: Sketch of the construction of the admissible region ([classicmodernorbits](#)).

pose of potentially hazardous near-Earth asteroids quite easily. Most obviously, using Equation 2.1, a reasonable value of  $H_{max}$  can be determined. The computation of absolute from apparent magnitude follows readily from the average apparent magnitude  $h^*$ , allowing a constraint on the radius vector:

$$H = h^* - h \log_{10} r - x(r) \quad (6.23)$$

In addition, condition **1** can be adjusted to fit objects with a certain range of semi-major axes:

$$-\frac{\mu_{\odot}}{2a_{min}} \leq \mathcal{E}_{\odot} \leq -\frac{\mu_{\odot}}{2a_{max}} \quad (6.24)$$

From here, it becomes possible to simulate the entire subspace of *virtual asteroids*. Figure 6.3 shows this region, with a selection of virtual asteroids triangulated on it. Each of these asteroids has a fully determined set of orbital elements, which can be propagated into the past and future for operations such as linkage, or to determine the possibility of the asteroid being on an impact trajectory with Earth. This can be done with e.g. an analogue to the method suggested by **trajsix** or with a least-squares based approach such as suggested by **phddissertation**.

### 6.3. Kalman Filters

Kalman filters (KF) are an essential component of modern control theory and nonlinear state estimation (**kalmangood**). The idea of a Kalman filter was first described by **kalman**. The Kalman filter or linear quadratic estimator (LQE) provides the optimal linear estimation for linear systems with gaussian (white) noise. The KF algorithm is a recursive estimator based around two alternating steps: predict and update.

Firstly, in the predict step, the filter produces an a priori estimate of the evolution of the state based on the current state and any inputs. In addition, the covariance of the state is estimated. Then, in the update step, the predicted state is compared with a measurement. This is known as the innovation residual. Using this innovation residual, a gain is calculated for the correction of the state estimate and covariance. Then, the state is updated and the filter goes back to the predict step.

Firstly, the following are defined:

- $\vec{x}_k$ : the state vector of the system.
- $\vec{u}_k$ : the control-input vector (most likely  $\vec{0}$  for the proposed application)

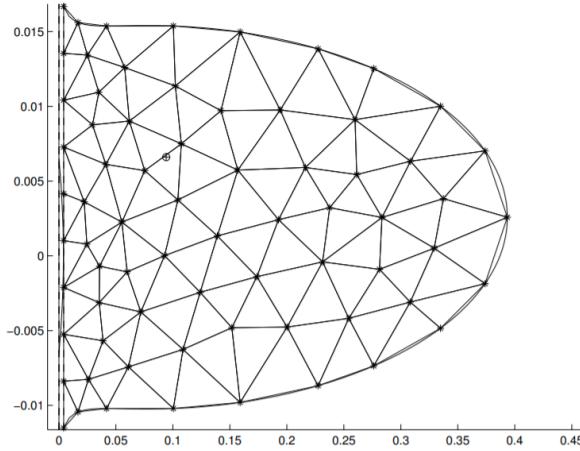


Figure 6.3: An example of an admissible region, for asteroid 2003BH<sub>84</sub> using exclusively data from the day of first observation. The vertical axis is  $r$ [AU/day], the horizontal axis is  $f(r) = 1 - e^{-r^2/(2s^2)}$  where  $s$  is a normalization term. The real position of the asteroid is marked near the top left with a  $\oplus$ . ([classicmodernorbits](#))

- $\vec{w}_k$ : the process noise, which is assumed to follow a normal distribution with covariance  $\mathbf{Q}_k$ :  $\vec{w}_k \sim \mathcal{N}(0, \mathbf{Q}_k)$ . This represents any perturbation in the trajectory not accounted for in the state-transition.
- $\mathbf{F}_k$ : the state-transition matrix.
- $\mathbf{B}_k$ : the control-input model.

Such that the state at time  $k$  is determined from the state at time  $k-1$ :

$$\vec{x}_k = \mathbf{F}_k \vec{x}_{k-1} + \mathbf{B}_k \vec{u}_k + \vec{w}_k \quad (6.25)$$

Then, the system will make an observation at time  $k$  with the following:

- $\vec{z}_k$ : the observation vector.
- $\vec{v}_k$ : the observation noise, assumed to be  $\vec{v}_k \sim \mathcal{N}(0, \mathbf{R}_k)$  with covariance  $\mathbf{R}_k$ .
- $\mathbf{H}_k$ : the observation model, responsible for mapping the state space onto the observation space.

Lastly, there are two important state variables for the filter:

- $\hat{x}_{k|k}$ : the a posteriori state estimate at time  $k$  from the observation at time  $k$  and all preceding observations.
- $\mathbf{P}_{k|k}$ : the a posteriori covariance matrix estimate, which is importantly a measure of the accuracy of the prediction.

Using the notation  $\hat{x}_{n|m}$  being the estimate of  $\vec{x}$  at time  $n$  by all observations preceding and including the observation at time  $m$ , the full filter can be expressed as:

Prediction step:

$$\hat{x}_{k|k-1} = \mathbf{F}_k \hat{x}_{k-1|k-1} + \mathbf{B}_k \vec{u}_k \quad (6.26)$$

$$\mathbf{P}_{k|k-1} = \mathbf{F}_k \mathbf{P}_{k-1|k-1} \mathbf{F}_k^T + \mathbf{Q}_k \quad (6.27)$$

Update step:

$$\bar{y}_k = \vec{z}_k - \mathbf{H}_k \hat{x}_{k|k-1} \quad (6.28)$$

$$\mathbf{S}_k = \mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^T + \mathbf{R}_k \quad (6.29)$$

$$\mathbf{K}_k = \mathbf{P}_{k|k-1} \mathbf{H}_k^T \mathbf{S}_k^{-1} \quad (6.30)$$

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + \mathbf{K}_k \bar{y}_k \quad (6.31)$$

$$\mathbf{P}_{k|k} = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_{k|k-1} \quad (6.32)$$

$$\bar{y}_{k|k} = \vec{z}_k - \mathbf{H}_k \hat{x}_{k|k} \quad (6.33)$$

At this point, it is good to note that the Kalman filter provides optimal estimation for **linear** systems with **zero-mean, independent Gaussian** noise. Of course, in reality these assumptions are not valid, and inaccurate modeling of the system dynamics can lead to disastrous results (**EKF**). Therefore, the Extended Kalman Filter (EKF) might provide a useful solution. In the EKF, the system is linearized about the estimate of the current mean and covariance. It is considered the standard in nonlinear state estimation (**kalmangood**).

The EKF is a generalization of the standard Kalman filter by means of linearization of the state and observation models (**extendedkalman**). At the cost of no longer representing the optimal estimator, the EKF allows for the state-transition and observation matrices to instead be differentiable functions  $f$  and  $h$ , such that:

$$\vec{x}_k = f(\vec{x}_{k-1}, \vec{u}_k) + \vec{w}_k \quad (6.34)$$

$$\vec{z}_k = h(\vec{x}_k) + \vec{v}_k \quad (6.35)$$

The functions  $f$  and  $h$  allow for these transformations to be made. However, they can not be used to recompute the estimate of the covariance. Therefore, the Jacobian of the functions is used instead:

$$\frac{\partial f}{\partial \vec{x}} = \mathbf{J}_f(\vec{x}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \dots & \frac{\partial f_m}{\partial x_n} \end{bmatrix} \quad (6.36)$$

Thus, after substitution of Equation 6.34 and Equation 6.35 into the standard Kalman filter, the remaining state and observation matrices can be approximated by the following Jacobians:

$$\mathbf{F}_k \approx \left. \frac{\partial f}{\partial \vec{x}} \right|_{\hat{x}_{k-1|k-1}, \vec{u}_k} \quad (6.37)$$

$$\mathbf{H}_k \approx \left. \frac{\partial h}{\partial \vec{x}} \right|_{\hat{x}_{k|k-1}} \quad (6.38)$$

Kalman filters are a promising technique for performing the further research. For example, **trajectorythesis**, shows very good results in approximating trajectories using an Extended Kalman Filter, but does note that the filter has problems when the trajectory deviates too much from the used physics model. Therefore, finding a good linearized representation is essential.

## 6.4. Artificial Neural Networks

Artificial neural networks (ANN's) are a class of Machine Learning - or popularly *Artificial Intelligence* - models used to perform complex nonlinear tasks that might be impossible, or at best highly unfeasible, to solve using conventional algorithms. ANN's comprise broad functions which are fit to often very large datasets using *learning* algorithms (**nnbooktwo**<sup>1</sup>). The initial inspiration for ANN's was to very loosely model the neurons in biological brains (see Figure 6.4), but it should not be seen as being an even remotely accurate representation of biological neural circuitry (**nnbookone**).

---

<sup>1</sup>If the reader is interested in learning more about machine/deep learning, this book is highly recommended

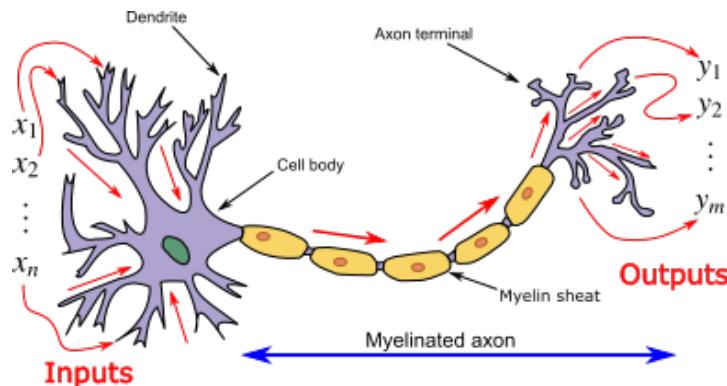


Figure 6.4: Comparison of a biological neuron with an artificial neuron. The Neuron takes inputs, combines those, and forwards them to produce a set of outputs. Note that this is where the analogy ends: biological neurons behave very differently to ANN's. CC BY-SA 4.0 Prof. Vu-Quoc.

In its simplest form, an ANN consists of interconnected neurons. These neurons produce an *activation*  $a$  based on the output of neurons connected to it:

$$a(\vec{o}) = f(\sum w_i o_i + b) \quad (6.39)$$

Here,  $o_i$  and  $w_i$  represent the output and weight of input neurons respectively,  $b$  is a bias term, and  $f()$  is an *activation function*. The purpose of this activation function becomes clear once the structure of an ANN is examined in more detail, such as in Figure 6.5: The input to the network is passed through one or more *hidden layers*, to finally reach an output layer. As any linear combination of linear functions is in itself a linear function, it becomes clear that it is necessary to apply a non-linear activation function to every neuron; else, the entire network would "collapse" into a linear function ([nnbooktwo](#)).

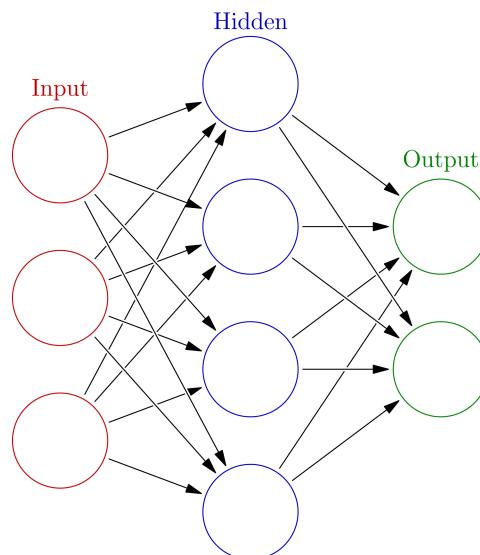


Figure 6.5: Structure of a feedforward neural network. CC BY-SA 3.0 Glosser.ca

In the past, the most commonly used activation function was a sigmoid (an "S"-shape approaching 0 asymptotically as  $x \rightarrow -\infty$  and approaching 1 asymptotically as  $x \rightarrow \infty$ , such as the logistic function ([nnbookfour](#)):

$$S(x) = \frac{1}{1 + e^{-x}} \quad (6.40)$$

However, as the gradient of the sigmoid function approaches zero for large values of  $x$ , this presents problems for training. Therefore, a host of different activation functions are used presently. The most

commonly used presently are the Rectified Linear Unit (ReLU, **Relu**) and Exponential Linear Unit (ELU, **ELU**):

$$\text{ReLU}(x) = \max(0, x) \quad (6.41)$$

$$\text{ELU}(x) = \max(a(e^x - 1), x) \quad (6.42)$$

Both of which are extremely simple and fast to compute. The ELU yields slightly better results, but requires finding a good value of  $a$ .

Fitting an ANN to a dataset generally requires two processes: learning, and tuning the hyperparameters. Because of the scope of the research, only supervised learning, where the network is provided with both input and desired output, is described. Learning is carried out by algorithms based on *backpropagation* (**nnbookthree**) and *gradient descent* (**gradientdescent**). Firstly, the network will process some training data. From the output of the network, the loss function is computed. This are generally well-known metrics such as mean squared error. Then, the backpropagation algorithm iterates backwards through the network, using the chain rule to compute the gradient of the loss function with respect to each network weight. It is important to note that, although not a direct process, the algorithm can compute all gradients in a single pass through the network. Thus, it is computationally very efficient. Then, knowing the gradient, a method like gradient descent can be applied:

$$\vec{w}_{n+1} = \vec{w}_n - \eta \nabla F(\vec{w}_n) \quad (6.43)$$

With *learning rate*  $\eta$  (a hyperparameter). Thus, for a small enough learning rate, the solution will converge to a minimum for the loss function. Note that this process does not guarantee reaching the global minimum; it only does this for a convex function, which deep neural networks are generally not. Additionally, learning may become excessively slow on "plateaus" of the loss function. For these reasons, some other optimization algorithms have been designed, most popular among which the Stochastic Gradient Descent (SGD, **sgd**) or the derived ADAM optimizer (**Adam**, **AdamKeras**).

Next, the hyperparameters, the properties of the network which can not be learned (e.g. number of layers, number of neurons, learning rate) can be tuned. This is generally an iterative process which involves attempting various combinations using a kind of grid search or random search to determine an optimum (**nnbooktwo**). Great care should be taken to not overfit the data in this process, as will be explained in subsection 6.4.2.

#### 6.4.1. Types of Neural Networks to be Used

Having discussed the base theory, it is good to specify which types of ANN's are considered for the research. Because of their properties, and general simplicity, the dense and recurring neural networks have been selected for further investigation. Both of these have already been shown to be usable for trajectory estimation (see e.g. **anntrajectoryone** and **anntrajectorytwo**)

The dense neural network (DNN) can be seen as the simplest of the ANN's. It consists of one or more layers of densely connected simple neurons as described in the previous sections. This has several advantages: it is simple to set up and computationally lightweight, and training is straightforward (**nnbooktwo**). However, dense neural networks are limited in what they can achieve: they are not translationally invariant (such as convolutional neural networks), nor are they capable of processing sequential information coming in in real time: the size of the input and output must be specified beforehand. **anntrajectoryone** achieves very good results using a DNN.

On the contrary, recurring neural networks (RNN's) are capable of holding information that arrives in sequence. They achieve this through feeding the output of a neuron back into itself. Thus, Equation 6.39 changes to:

$$a(\vec{o})_t = f(\sum w_i o_i + b + a_{t-1}) \quad (6.44)$$

Thus, the state of the neuron becomes dependent on the state of the neuron in the previous timestep. This gives the neuron a limited amount of "memory". This is obviously very advantageous for processing sequential data. However, the main drawback becomes apparent when it is considered how RNN's

are trained: as the normal backpropagation algorithm does not account for the recursive information flow, it is necessary to perform a change to the algorithm known as "Backpropagation through time" (**nnbooktwo**). Here, the time-sequence of neurons is "unrolled", and the neurons at different timesteps are treated separately for purposes of calculating the gradient. This leads to the first major problem of RNN's: they are very slow to train. A lot of data is needed (thus limiting the size of training steps, as computers have a limited physical memory), and the *vanishing gradients* problem, where the gradients become so small that gradient descent becomes overly slow, becomes a large issue (**nnbookfour**). Interestingly, **trajectorythesis** achieves a very poor performance on a similar problem using RNN's. However, the graphs of the resulting trajectories are extremely erratic, so it is speculated that this is a result of overfitting the training data. Therefore, it is proposed to start with a simple, robust model and work out from there.

### 6.4.2. Common Problems and Pitfalls

As mentioned in the previous section, there are several pitfalls that can be encountered in fitting and training ANN's. The most common problems and some solutions are listed in this section, based on the work of **nnbooktwo**.

The first, and most common, problem is overfitting the training data. In this case, the network will "learn" the noise in the dataset, rather than learning the useful patterns embedded therein. Because of the way ANN's are trained, this is an oft-encountered problem. Usually, this problem can be identified by an unreasonably effective result on the training data, alongside a very poor generalization to unseen data. The first solution to this problem is to keep a set of validation data completely separate from the training data, and only using this to validate the training. Care should be taken to not use the validation data for other applications, such as tuning the hyperparameters: in this case, the network will repeatedly "see" the validation data indirectly, possibly leading to it fitting to the unseen data as well. Secondly, a number of regularization methods can be applied. Firstly, L1 or Lasso regression, shrinks very small parameters to zero. This means that a parameter which is inconsequential for the result is no longer included, instead of only being included by a very minor amount - which would lead to an overfit. Conversely, L2 or Ridge regularisation, the large values are penalized. This makes sure no unrealistically high values of coefficients can be present. Lastly, there is the idea of "dropout". Here, in every training step, a few neurons are "dropped out" (i.e. set to 0). This forces the remaining neurons to do the work, thereby making sure that all components of the network contribute to a solution.

The second problem is the vanishing or exploding gradients problem. Here, the gradients become excessively small or large during training. Because of how backpropagation works, this is especially an issue for very deep (i.e. many layers) networks. As the gradients become very small or large, training can slow excessively, or the solution can diverge. A part of the solution to this problem is using better activation functions such as ReLU or ELU, as described previously. However, in some cases it might be necessary to include connections to non-adjacent layers, creating a "shortcut" in the network.

Then, there is the problem of *Catastrophic interference*. Here, an ANN spontaneously drops all learned information in favor of new information, essentially overfitting to the last bit of information seen. This manifests as a sudden and dramatic drop in performance during training. This problem is very complex to resolve, and should be accounted for depending on how it manifests. Usually, this is done by placing some restraint on how the model is allowed to be sensitive to new information.

Lastly, a remark should be made with regards to the quality of the data: as neural networks are trained exclusively using the data supplied to them, this means their quality is highly dependent on the quality of the data provided. Obviously, providing very bad data will lead to a very badly trained model. However, in generalizing to a solution useful in practice, it is important to also not feed the network unrealistically good data. In this case, the model will seem to perform far better than what can be expected from it in reality. Thus, an assessment of the input data is necessary to perform a useful evaluation of any ANN-based approach.



# 7

# Conclusion and Thesis Research Questions

After examination of the literature, it has become possible to draw a conclusion on the state-of-the-art of the research field, and therefore the direction and content of the proposed research. In this chapter, firstly an exposition of the identified knowledge gaps in both literature and practical engineering will be given. This will be supported by an overview of methods that can be used for carrying out the work in the research. From the knowledge gap, a main research question with supporting subquestions for the subsequent research will be derived. Lastly, a preliminary work breakdown and Gantt chart of the thesis work will be presented.

## 7.1. Conclusion on Literature and Knowledge Gap

In the preceding chapters, the width and depth of the research field was surveyed. The objective of this survey was to determine interesting avenues for further research, and to find promising methodology to be used in that research.

Firstly, the gaps in the knowledge and current capabilities will be addressed. In chapter 2, it was found that the current completeness of asteroid surveys is highly incomplete; asteroids smaller than the 140m cut-off for PHA-status provide significant threat to human life and possessions, and a large fraction of the population has so far gone unidentified. Relating to this, chapter 4 provided an overview of current efforts to catalogue the asteroid population. Here, it was shown that although very large telescopes are being built and operated from Earth, they suffer from significant drawbacks that prevent them from effectively identifying a large portion of the population of asteroids at any but the last moments before impact. Ideally, plenty of advance warning should be given to defend the Earth against an asteroid, rather than only evacuating the area of impact. Space-based surveys provide a good avenue of research to improve this, and the results from e.g. NEOWISE and NEOSSat are promising. In chapter 5, the issues relating to the processing of the vast amounts of data produced by a sky survey were discussed. The need for automatic processing of large amounts of data has affected all fields of technical research, and space engineering is no exception. Additionally, the capabilities and techniques for automatically processing a sky image into a useful dataset were showcased. It is apparent that this field is at a very advanced level, and provides no applicable research objective. The researched techniques and the knowledge of how the methods perform are still useful, and provide good indicators of how to simulate and handle the data later. Conclusively, as was shown in chapter 6, the methods for actually determining the trajectory of a target from a series of observations are limited. Especially during short or low-quality observations, methods are not suited to processing the enormous amounts of data of an all-sky mission effectively. Therefore, the proposed research is to examine improving autonomous trajectory and impact hazard determination by a space-based telescope.

To carry out this research, methods were found in the literature which can be used. As the general plan, which will be laid out in section 7.3, shows, the idea is to simulate optical observations, process

these, and develop impact hazard determination techniques for these observations. The methods for simulating the images will be based on the information on the asteroid population and properties from chapter 2, and the knowledge of optical systems included in chapter 3. The image processing algorithms are discussed in chapter 5, although these are at a sufficiently advanced level that they are not of particular interest beyond their effect on image quality. Lastly, the techniques proposed for testing are those detailed in chapter 6: A method based on a numerical approach using the method of very short arcs, a control system approach based on the use of an extended Kalman filter, and lastly an application of dense or recurring neural networks. From assessment of literature and comparison with non-space applications of trajectory determination, it was found that these methods provide the best options for examination. The goal of this assessment is to discover if and how these methods can be used in future surveys, and what level of efficacy can be expected.

## 7.2. Thesis Research Questions

With the target knowledge gap identified, and the background knowledge set, the research question will be defined as:

*To what extent can impact-hazardous targets be identified through optical observations by a space system placed strategically in the solar system?*

A breakdown of the components of the question:

- **To what extent can... be identified:** Realistically, it will not be possible to determine the exact impact chance of all (or perhaps any) targets. However, identification of targets for further research or monitoring would be very valuable, as other observational resources can be used to track the targets more accurately.
- **Impact-hazardous targets:** It is desirable to avoid the term “potentially hazardous asteroid”, as it already has a definition. As the problem concerns asteroids impacting Earth, it makes logical sense to see if targets that present an impact hazard can be identified. This hazard is not just a determination of the orbit; it consists of a product of impact energy (i.e. mass/size), impact probability and possibly time to impact.
- **Optical observations by a space system:** It is not necessary to define this further, as it is an opening for this c.q. further research. The important ramification of this specification is that the space system is capable of doing the identification: Mostly, downlinking vast amounts of images to Earth for processing or analysis by humans would make the results of the research less useful: those techniques exist, but they suffer from capacity problems. This links in well with the first point, the identification information obtained by the system should be useful itself.
- **Placed strategically in the solar system:** As measurements are taken from space, the positioning of the satellite can be examined. This would be a later stage of research, moving towards a more complete system. It is not the intention to conduct a full orbit analysis, but broad categories (e.g. Earth-Sun Lagrange points, co-orbital with Earth, co-orbital with Venus, situated in the asteroid belt) should be easy to examine once the simulations are up and running. This is however not seen as crucial in answering the question.

From the main questions, a set of subquestions are derived which are used to set up the work breakdown:

1. How can optical observations of NEA's by a space system be accurately modeled?
  - 1.1. What model should be used for the population of NEA's?
  - 1.2. How can the camera system be modeled?
  - 1.3. How can the simulation be verified and validated?

2. How can the optical observations be processed into angular observations?
  - 2.1. How precise are the methods based on correlation or analytical function fitting?
  - 2.2. Is it possible to obtain a better result using a CNN?
  - 2.3. What other information related to the impact hazard can be derived from the optical observation?
3. To what extent can impact hazard be determined from these observations?
  - 3.1. How can a system based on estimating and numerically searching the parameter space be implemented?
  - 3.2. How can a system based on a Kalman filter be implemented?
  - 3.3. How can a system based on an artificial neural network be implemented?
4. What is the effect of strategic placement of the system on its performance?

## 7.3. Project Plan

Before commencing the research, a planning is made to facilitate finishing the work in a timely manner. The planning is based on the TU Delft Thesis Procedure ([thesisprocedure](#)). Firstly, an overview of all planned periods of the work is given, supplemented by national and planned holidays and contingency time. This is used to estimate the full timeframe of the project. From here, dates for the milestones are defined. Then, a work breakdown is given, and these are combined into a Gantt chart.

### 7.3.1. Time planning

Table 7.1 shows the preliminary time estimate for the thesis work, culminating in a defence around the last week of January 2022. Note that this planning was made on the assumption that restrictions due to Covid-19 will be lifted around the summer of 2021. In case of unforeseen developments in this area, no claims will be made about the validity of this planning, but a new one will be developed based on best insight at the time.

Table 7.1: Thesis timeframe planning.

Topic	Date	Weeks	Milestone
Start to Kick-off	19/04 - 30/04	2	29/04: Kick-off Meeting
Kick-off to Mid-term	03/05 - 15/10	16	14/10: Mid-term Review
Mid-term to Draft	18/10 - 10/12	6	10/12: Draft Thesis
Draft to Green Light	13/12 - 24/12	2	23/12: Green Light Meeting
Green Light to Final	27/12 - 14/01	2	14/01: Final Thesis
Final to Defence	17/01 - 28/01	2	28/01: Defence
Summer Holiday	15/07 - 28/07	2	-
Study Tour	16/08 - 05/09	3	-
Holidays Q4	-	1	-
Christmas Break	25/12 - 09/01	2	-
Contingency	-	3	-
Total	19/04 - 28/01	41	-

### 7.3.2. Work Breakdown

Following is a preliminary work breakdown based on the previously established subquestions. Included are some thoughts or elements to keep in mind when starting the components.

#### 1. Determination of mathematical foundations:

- Before working out methods or writing code, the mathematical basis for the various methods will need to be defined.
- This is especially important for the ANN-based solutions, as the shape of the inputs can be elemental in determining their structure.

#### 2. Establishing trajectory determination methods:

- Before creating the simulation software, the methods should be defined in more detail.
- This order allows more time to think about the important aspects (i.e. the methods) over the simulation software.
- Care should be taken not to make the simulation fit the methods.

#### 3. Creation of observation simulation:

- From the chapters in this review, it is known how to construct a simulation of how a satellite would observe a target.
- The software should be capable of producing a large amount of data, potentially for training the ANN, as well as evaluating it.
- Initially, only sets of angular measurements and accompanying uncertainties will be used; the further processing of optical measurements can be added later. This allows more time to resolve problems.
- Brightness of the target can be included as an additional parameter.

#### 4. Creation of threat assessment simulation:

- The working of this simulation will provide answers to most of the questions. Therefore it is the integral part of the research work.
- Note that the problem is not to determine the orbit of the target as accurately as possible; the problem is to assess the risk the target poses to Earth.
- This step might require iteration with previous steps, depending on outcome.
- It should be judged early whether the code needs to be parallelizable through e.g. CUDA, should a large number of simulations be desirable.
- Experimentation with regards to the position of the satellite can be carried out.

#### 5. Verification and validation

- Arguably the toughest topic, and the most important one. So far a solid conclusion has not been found as to the best method for verifying and validating the solution.
- Some ideas include tests on real-world data or using other simulation software that is freely available.

#### 6. Evaluate and conclude

- Results of the simulations should be placed in context, and conclusions drawn on what would be the best suited method for this problem.
- To assess the feasibility of the suggested methods, they should be compared to existing surveys and background risk.
- Other factors such as computational complexity or time needed to image a target should be considered as well, thereby placing the research in context.

### 7.3.3. Gantt Chart

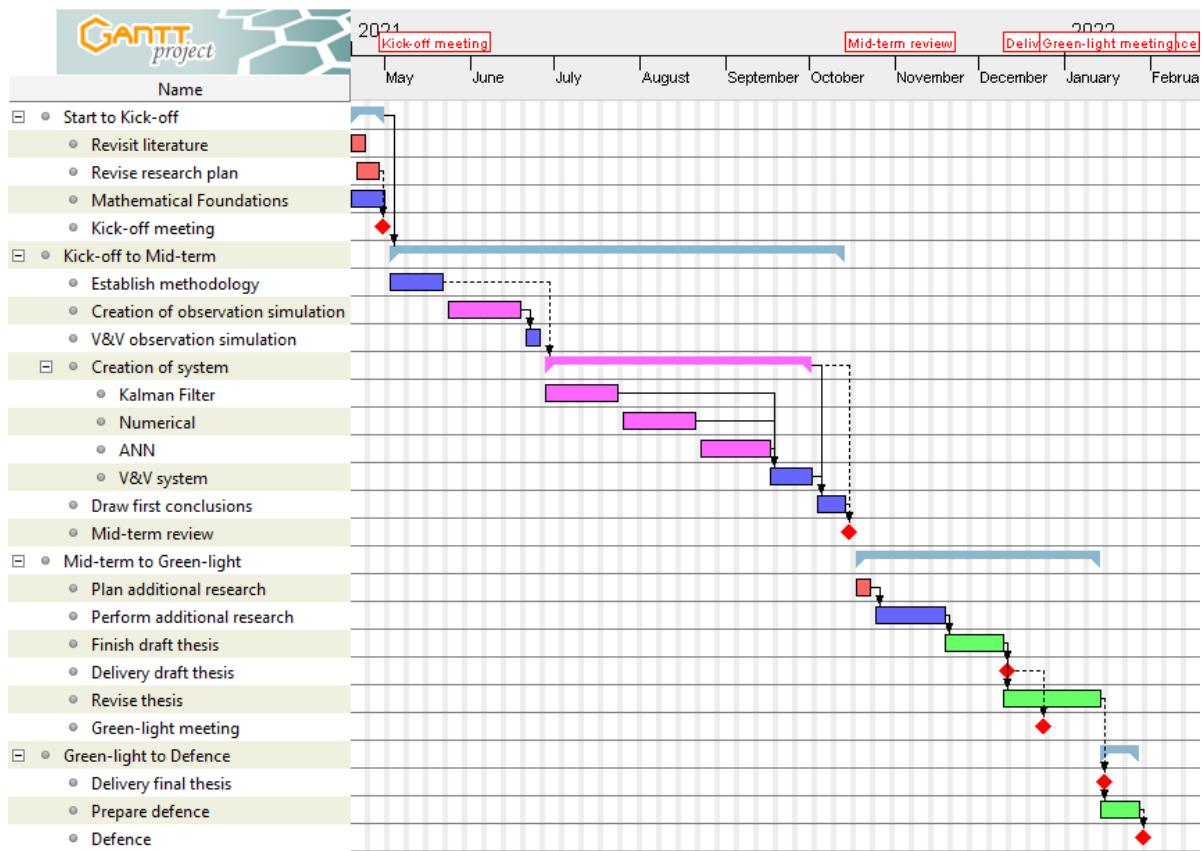


Figure 7.1: Gantt chart for the thesis work. Shown in red is literature research work, blue is theoretical research work and pink the programming work. Milestones are indicated as red diamonds. Each rectangle represents a week.