Projectteam Zebro onboard navigation system
Literature report

Version 1.0

April 28, 2015

Delft University of Technology

To:
Chris Verhoeven
Maneesh Verma

Abstract
This document contains the first draft
of the literature report for our Bachelor End-Project.
Here we will explain the knowledge obtained from our literature study.

Bart Rijnders, 4103505
Bart van den Bogert, 4215192
Alex Janssen, 4231333
Peter Stijnman, 4215788

Version	Date	Comments
1.0	April 28, 2015	Start of literature report

Contents

1 Introduction

An onboard navigation system is necessary to control a free-roaming robot. This navigation system is connected to the user through an user interface. From here the user is able to control the robot when the user wants to, otherwise the robot should be able to autonomously do tasks on its own.
To be able to accomplish this there are certain subsystems that need to be implemented on the robot. These subsytems are seen in the function tree in figure 1.

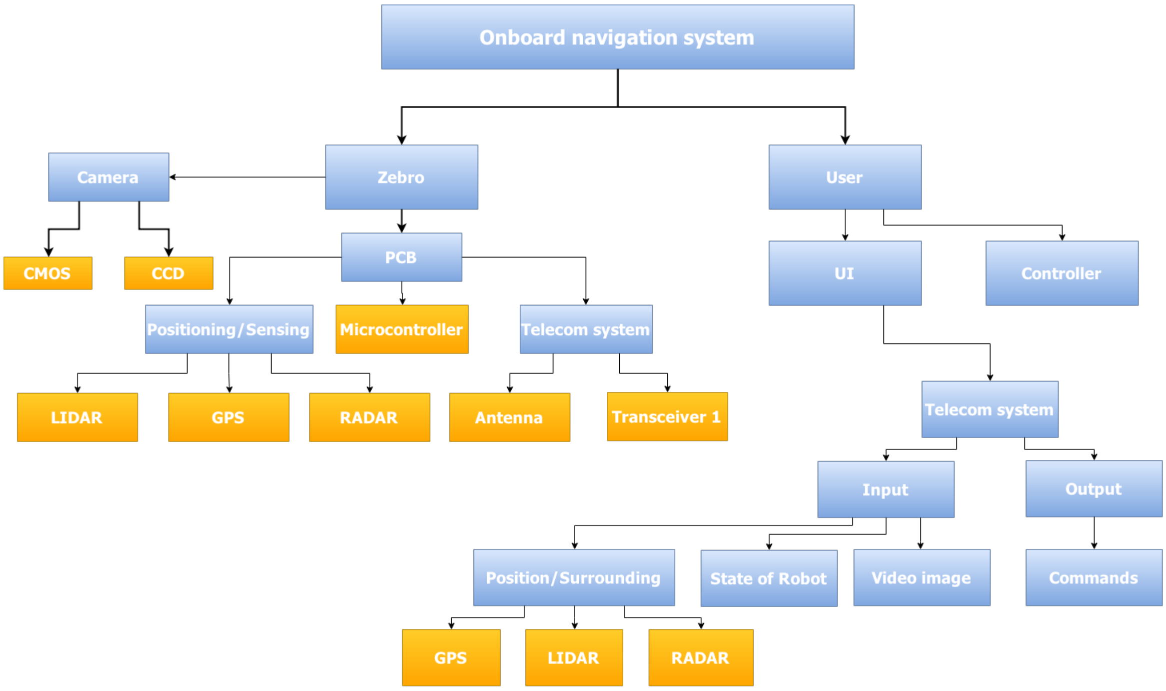

Figure 1: Function tree.

These subsystems are explained throughout this literature report. From this report design choices will be made and these will be implemented on the platform called Zebro. First of the report will start with an explanation of the different methods of sensing the environment. After that the positioning system will be discussed and then the telecommunication link will be explained. In chapter six two different camera types will be analysed and compared. The last part that will contain information on microcontroller specifications.

2 LIDAR

LIDAR is an acronym for Light Detection and Ranging. It is a technique that uses a laser beam to detect or scan objects. It is an active form of remote sensing. So information is obtained by sending a signal from a transmitter which can be reflected by a target. A detector can pick up the reflected signal to determine different types of information about the object. These types are range, chemical properties and velocity. In our case we are only interested in the range of the object, since our goal is to avoid collisions with static objects. The measurement of the range or distance of the object is based on a precise measurement of time. There are two main methods to do this, the timed pulse method and the phase comparison method.

2.1 Timed Pulse Method

As the name indicates, the timed pulse method measures the time required for a very short pulse of laser radiation to travel to the target and back. This method can be used to measure distances of tens of meters to hundreds of kilometers. In figure 2 is the generic type of a laser rangefinder shown, of course the actual design and the specific components that will be used in a particular instrument will depend on the particular application. In our case the laser will be a small low-powered laser instead of a high-powered one.

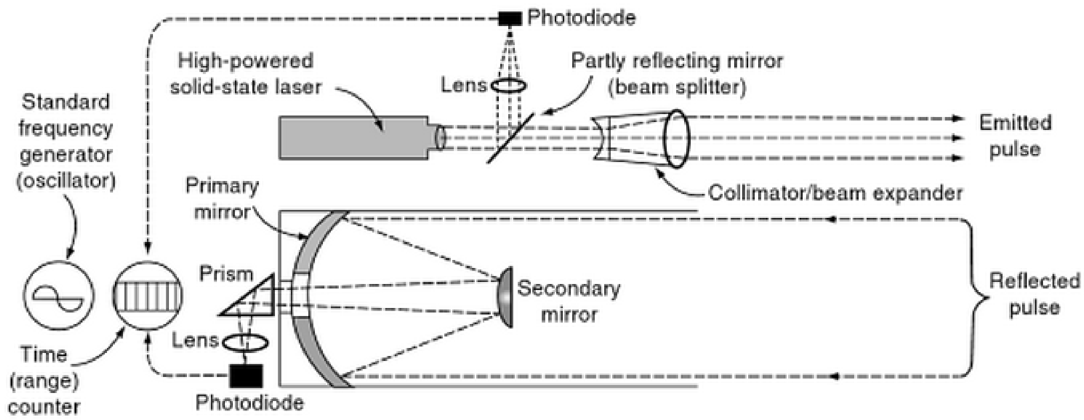

Figure 2: Laser rangefinder.

A lens is placed in front of the laser to first expand and then collimate the laser pulse so the divergence is kept to a minimum. A small part of the energy of an emitted pulse will be diverted by a partly reflecting mirror to a photodiode. This will trigger a timing device with a simple threshold which is usually connected to a very stable oscillator. When the pulse is reflected from the object the returning pulse is picked up by the receiving lens or mirror optics. Because the pulse will collide with the object and tiny particles in the air, the beam will be scattered and very weak in intensity compared to the sent laser pulse. Therefore it is needed to amplify the returned pulse electrically after it has been picked up by a sensitive photodiode. To reduce the noise picked up by this photodiode (e.g. sunlight) narrowband optical filters are placed over it. In simple cases, this picked up reflected pulse will deliver a stop pulse to the timer counter. When this stop pulse reaches a certain threshold the counter will be stopped. When more state-of-the-art systems are used, a more complex triggering mechanism is used of course. Since the time it took for the sent pulse to travel to the object and back is now known, the distance can be calculated using the speed of light.

2.2 Phase Comparison Method

In contrast to the timed pulse method, a continuous laser beam is used in the phase comparison method. A continuous laser beam is often abbreviated to CW laser, which stands for continuous wave laser.Because of the limited power of a CW laser it is almost never used for long range systems. On the other hand, this makes it very suitable for measuring shorter distances (typically <100m). In the phase comparison method, the emitted beam is a basic carrier signal on which a modulation signal has been superimposed. This modulation signal is used to measure the difference in phase of the transmitted and received signal. The frequency of the modulation signal is held at a constant value using a stable oscillator. This means the modulation wave can be used to control the amplitude of the carrier wave, which is known as amplitude modulation. A picture of the modulation process is shown in figure 3.

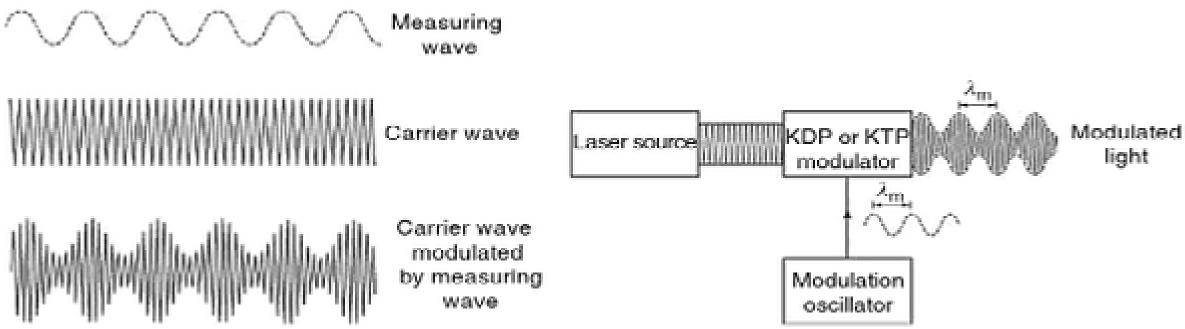

Figure 3: Amplitude modulation.

In the phase comparison method the emitted CW laser is also reflected by an object it collides with. This means a small part of the emitted beam will be reflected and will travel back through the same path as the emitted beam. Just like the timed pulse method the reflected signal is detected using a photodiode and amplified. After amplification the signal will be demodulated, this means the carrier and the modulation signal will be separated. Now the phase of reflected signal is compared to the original emitted signal, from this the phase angle between the two signals can be obtained. The graphs in figure 4 show how the transmitted signal is reflected by an object and picked up by the receiver (a), and how the phase angle is determined between the transmitted and received signal (b).

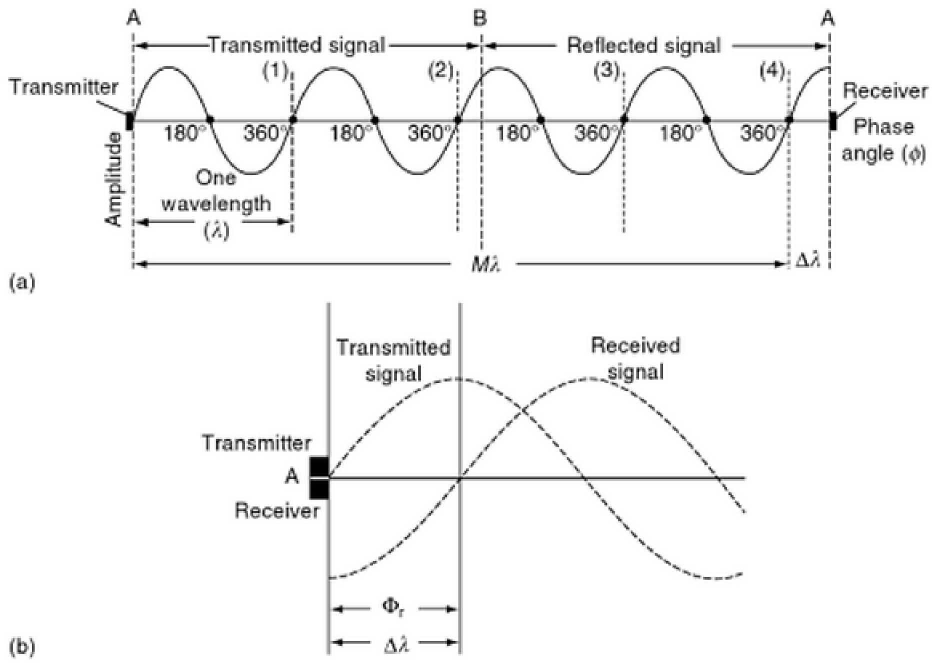

Figure 4: Phase angle measurement method

The phase angle is the fractional part of the total distance.The total distance consists of an integer number multiplied by the wavelength plus the fractional part. This is also shown in graph (a). So the total distance measured can be calculated using the following formula:

$$d = (M\lambda + \delta\lambda) \tag{1}$$

The integer M can be determined by making a number of changes to the wavelength, which in turn changes the frequency of the emitted beam. By doing this very rapidly in succession a system of equations is generated that can be solved for the final distance d.

3 RADAR

Radar stands for radio detection and ranging. It is widely used today in military, telecommunication, medical and many other applications. It utilizes electromagnetic waves to sense the environment. In our case it could be used to detect objects and thus avoid collisions.

3.1 How detect objects with radar

To detect objects, electromagnetic waves have to be emitted first. This is done by radiating electromagnetic energy from an antenna to propagate in space. When these electromagnetic waves collide with an object, the waves will be scattered in many directions. But some of these reflected waves will return to the radar antenna and can be detected. When these reflected waves are picked up by an antenna, the signal will be processed and the location plus possible other information about the object can be obtained. Because the emitted electromagnetic waves are usually sent in all directions, it is possible to detect objects in a certain radius of the radar. So it is not needed to aim the antenna
To determine the range of an object, the time it takes for a radar signal to propagate to the object and back has to be measured. Because electromagnetic waves travel at the speed of light the distance to the object can be calculated.An advantage of radar is that it can measure this distance with great accuracy, even at long range. Because we need to detect objects at a relatively small range, high accuracy becomes even more important.

4 GPS

For the robot to know where it is heading and how it needs to get there it needs to know its own position. For this GPS, otherwise known as global positioning system, can be used to acquire the coordinates. With these coordinates a path can be made for the robot to get to the specified point. GPS uses signals transmitted from satelllites to the GPS receiver that is located on the robot. From this signal the time can be derived it took to get the signal from the GPS receiver to the satellite. With enough of these measurements the correct position can be calculated.

4.1 Measurement of the distance to the satellites

The GPS receiver sends a signal to a satellite. When it reaches the satellite a signal is sent back to the GPS receiver which receives the exact signal back. From this it can look when it sended its own signal an compare the time with the moment the signal from the satellite reached the GPS receiver. From this time the distance can be computed following equation 2.

$$d = \frac{c * t}{2} \tag{2}$$

where d is the distance, c is the speed of light $(3*10^8 m/s)$, t is the total travelling. The divided by 2 comes from the fact that the time the receiver measures is the time it took to go to the satellite and back. This means the required distance is measured twice.

4.2 Calculating the position

To explain how the position is computed an example will be used. When the distance from the satellite is known it results in a geometry problem using spheres. When the distance is known from one satellite the receiver knows it is somewhere located on a sphere with the satellite at its center and the radius equals the distance computed earlier. This is shown in figure 5

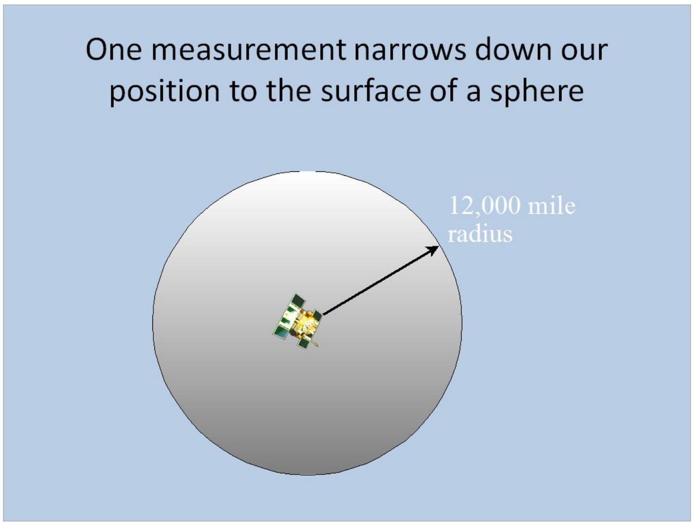

Figure 5: Possible position using 1 satellite.

When a second satellite is added to this cofiguration the receiver knows that is somewhere on the intersection points of those two spheres. These instersection points are located on a circle this is indicated with the black circle shown in figure 6.

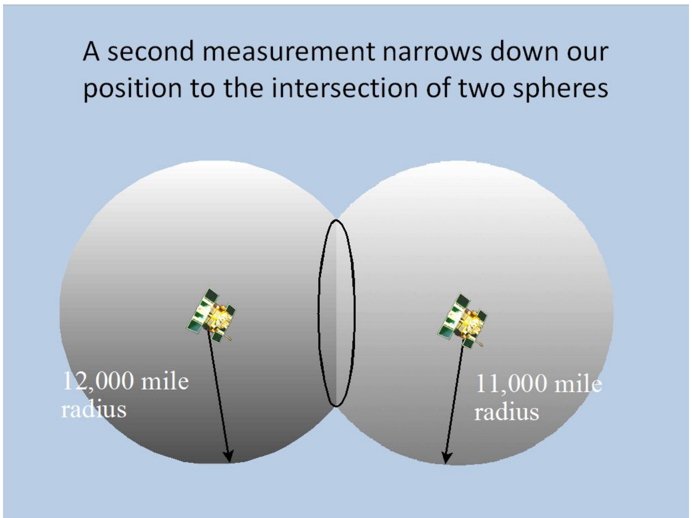

Figure 6: Possible position using 2 satellites.

The next step is to add another satellite. This will result on the receiver knowing it is somewhere on the intersection points of the three spheres around satellite 1,2,3. As shown in figure 7 only two intersection points remain (indicated with the red dots).

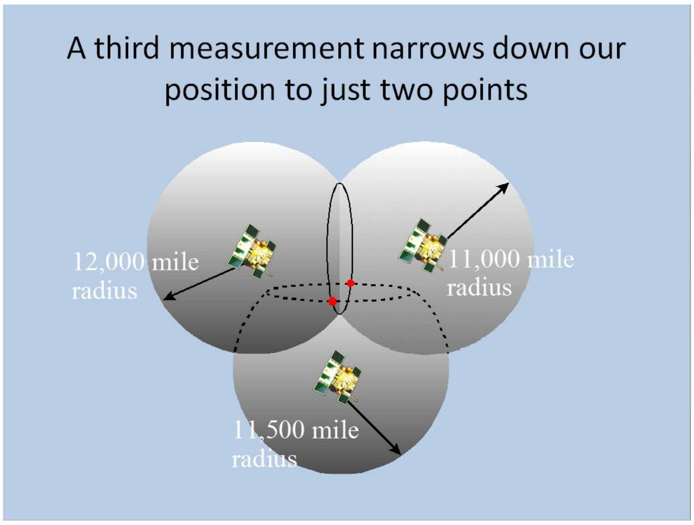

Figure 7: Possible position using 3 satellites.

Most of the time from these two points the receiver can make a choice on its own which point it is located at since the other intersection point has a high chance of being nowhere near earth. If this is not the case the receiver will measure its distance from a fourth satellite and only one intersection point should remain, being the location of the receiver.

5 Telecommunications

We need to communicate with the mars rover over large distances. This will be, of course, scaled down to reasonable distances on earth for the challenge. To make communication possible we will look into the theory first in the form of a link budget analysis and the hardware part which is mainly our two transceivers. One on the mars rover itself, and one on the user side.

5.1 Link budget analysis

For communications from point to point we need to do a Link Budget Analysis.

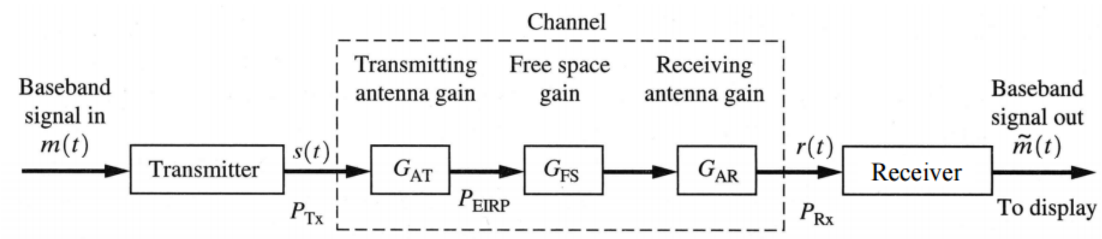

Figure 8: Link Budget Analysis

Here the Transmitter will send a modulated signal, which at the receiver side will be influenced by a couple of factors. The P_{tx} is the output power by the sender, and P_{EIRP} is the total power including the antenna gain transmitted by the sender. After transmitting we will lose power, we call this the free space gain G_{FS}. we can calculate the power at the output of the receiving antenna P_{rx} according to Equation 3. Note that all the variables are in dB in this equation.

$$(P_{Rx})_{dB} = G_{AT} + G_{FS} + G_{AR} + P_{Tx} \tag{3}$$

5.1.1 Antenna gain

The antenna gain is approximated by Equation 4.

$$G_A = \frac{4\pi A_e}{\lambda^2} \tag{4}$$

With this approximation we can view different Antenna options, with the gains in the table below.

Table 1: Antenna gains and effective area's

Antenna Type	G_A	A_e
Isotropic	1	$\lambda^2/4\pi$
Half-wave dipole	1.64	$1.64\lambda^2/4\pi$
Horn, mouth area A	$10A/\lambda^2$	0.81A
Parabola, face area A	$7A/\lambda^2$	0.56A

As we see the Isotropic antenna has a gain of 1, this is because it sends an equal amount of P_{Tx} in all directions thus the isotropic antenna is used as a reference for the antenna gains.

5.1.2 Free space gain

The free space gain is the gain that happens when the wave travels from point to point. We can calulate the free space gain G_{FS} with equation 5.

$$G_{FS} = (\frac{\lambda}{4\pi d})^2 \tag{5}$$

Where $\lambda = c/f$ with c the speed of light which is approximately $3 \cdot 10^8$. Since it is easier to work in dB de formula for the free space loss $(L_{FS})_{dB}$ we can write down the free space loss as following.

$$(L_{FS})_{dB} = 20\log(\frac{4\pi d}{\lambda})dB \tag{6}$$

5.1.3 signal-to-noise ratio

As showed in equation 3 we can see that we will have a P_{Rx} on the other side of the channel. Assuming we know the total noise temperature T_{syst} of the receiver part of the system we can find out how much noise we will have at the output of our total system. This is related by equation 7

$$N = kT_{syst}B \tag{7}$$

If we divide P_{Rx} by N we have our SNR from which we can see if noise will become a problem when receiving data.

5.2 Transceiver

Our transceiver excists out of a load of components. We need an oscillator, a circuit that will handle our chosen modulation and a mixer. We will first handle list some oscillator choices. We can also choose to buy a complete package cointaining all the part or we decide to buy the parts apart, or even to build it ourselves. The signal that we have on the input of the Transceiver will be a digital signal already which is why we won't require any pulse amplitude modulation or sampling.

5.2.1 Oscillator

We can distinguish 3 different kind oscillators. A first order, second order and higher order oscillators. We can base this on the differential equation shown in equation 8.

$$f_1(x)\ddot{x} + f_2(x)\dot{x} + f_3(x) = 0 \tag{8}$$

A first order oscillator would then mean that we only have one dynamic time dependend element. The advantage of first oscillator is that it is it's tunability which might be a very wanted property since there might be other rovers communicating on our channel so we can easily switch the frequency. The downside of these oscillators is their stability. The stability however can be improved by making use of injection-locking of multiple first order oscillators. Second order oscillators have a great stability but their tunability is not that good. The last oscillator group is the third order and higher group which is not very stable since it is prone to chaos and not very tunable. The positive side is that it is relatively easy in design.

5.2.2 Modulation

When modulating we can choose between Amplitude Modulation (AM) and Frequency Modulation (FM). This will our first choice toward a modulation technique. When using an AM modulation technique we change the amplitude in a frequency to be able to transfer where when using FM we use frequency changes in order to transfer bits. The circuit used for each modulation technique differs from eachother which is why we will look into this after making a desicion.

6 CMOS and CCD cameras

For the robot to get from one point to another manually, cameras are needed to see the surroundings and navigate through them. Furthermore, when the the robot needs to use its arm for a task it is convenient if there is a camera in the arm as well. This is because the point of view makes it easier to grab objects or drill in the right spot. The two camera options the team has is a CMOS image sensor or a CCD.

6.1 CCD

When a CCD, short for charge-coupled device, is exposed to light the small photosites on the surface of the CCD capture the light and store it in the form of a charge, as seen in figure 9 .

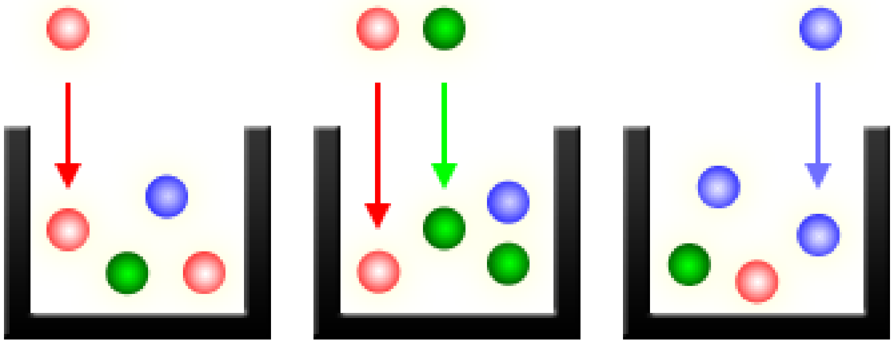

Figure 9: Representation of photosites, charges are collected at the electrodes.

To read out the charge of each photosite a complete row is put into an amplifier one photosite at a time and after that the amplified signal is then passed on to an ADC, analog to digital converter. When the complete row has been read out the row above it moves one down, as do all the rows above, and then that row goes through the amplifier and ADC. This process is done until all rows are read and the image is now digital. This is also where the name comes from , charge-coupled device, because the rows of charges are coupled to the row above them. When one moves down all do. A visual representation is seen in figure 10.

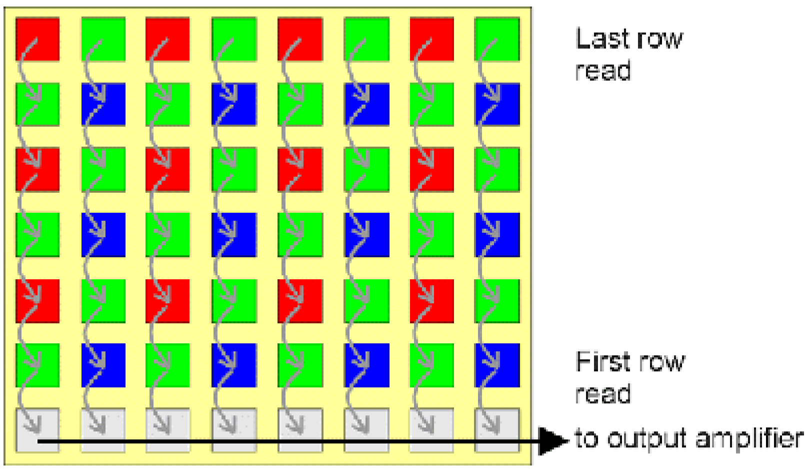

Figure 10: Visual image of how the CCD reads out the array of photosites. This happens one row at a time.

The polysilicon electrodes on the surface of the chip are so small and close together that the charge is kept intact when physically moving from the place where the light was actually cap-

tured to the place where the signal is amplified. To achieve this a clock is needed to move all the charges at the same time, this needs to be done by an off-chip (or secondary chip). This is in order to not interfere with the closely packed polysilicon electrodes on the chip.
Because the CCD needs extra circuitry which needs to be accurate, it can result in a very specified circuit with multiple power sources that need to operate at different, critical values which might not be regular. It is not rare for CCDs to have 5 or 6 of these power sources, this means that a lot of power is needed to operate the chip as intended.

6.2 CMOS image sensor

A newer technology is the CMOS image sensor, it can be easily incorporated with chips and other circuitry made on the same CMOS wafers. There are two basic types of CMOS image sensors, one being passive and the other active. The passive-pixel sensor works along the same principles as the CCD however now the circuitry is on the same chip as the sensors/photosites. This causes noise which can be seen on the image produced. The active-pixel sensors have extra circuitry at each pixel to cancel out the noise, this increases the quality of the image and makes it possible to go for higher resolutions. Here the performance can be equal to that of the CCDs. The negative side here is that this takes up extra space within the area of the photosites/pixels, see figure 11. This results in the photosites not being as close to each other as they are at the CCDs for example (the fillfactor is lower compared to CCDs). Because there is less area capturing the light. This results in a lower charge thus a lower amplitude signal running through the ADC. For the image itself this results in a lower light intensity. This makes the image appear darker than it actually is and therefor makes the CMOS image sensor less qualified to take images of dark environments.

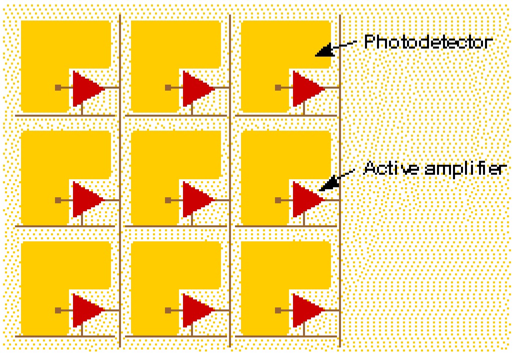

Figure 11: An abstraction of the circuitry around the photosites.

6.3 Comparison

CMOS imagers have a better intergration in circuits compared to the CCD cameras. Plus they use less power and are smaller. The negative here is that CMOS imagers have a lower image quality. Making them they less useful in high end image applications. Another negative for CMOS imagers compared to CCD cameras is that they are less flexible. This due to the circuitry already around the sensor, this is needed for it to work and cannot be changed while a CCD can be implemented in more systems. So the overall trend is that CMOS imagers are used for lower end imaging applications and in mass production while their counterpart the CCD is more suitable for the high end imaging applications.

7 Microcontroller

In order for the a robot to accomplish the different task it will need to decode all the signals that are sent to it. These signals then need to be sent to the right parts (i.e. for the Zebro the signals that are meant for the arm need to be sent to the arm). For this task a microcontroller is perfect.

7.1 Important microcontroller specifications

A microcontroller is a microprocessor that is used to control all different kinds of electronics. A few of the more important specifications are the clock speed, power characteristics, interrupts and operating temperature. For the clock speed, mostly given in ## MHz, and together with the CPI, cycles per instruction, the performance can be given for a microcontroller. This means that it has a certain execution time for a program. This is of course important since if the microcontroller is too slow the robot will get slowed down or signals send back to the user arrive too late in order to make an appropriate decision.
The second important specifications are the power charateristics, this includes the supply voltage and current and the power dissipation. Escpecially for a robot that has a limited power source, a battery, it is not advisable to use a microcontroller that drains the power source quickly. This specification might interfere with the first specification of performance of the mincrocontroller, if there is limited power it might be neccesary to choose a microcontroller that has lower performance. However this is a design option that needs to be made when the time comes.
The third important specification is the way the microcontroller can handle interrupts. Since the microcontroller is used for decoding signals it receives it inevitable that is a signal will be sent from a subsystem to the microcontoller to send back to the user. These signals might be asynchronous and therefore the microcontroller needs to be interrupted and react to the input correctly.
One more specification to take into consideration is the operating temperature. This is straightforward, the minimum and maximum temperature the microcontroller can handle. It will be safe to operate the microcontroller within this range, in other words the microcontroller should correctly in this temperature range. When the temperature is exceeding this range the microcontroller is not guaranteed to operate the way it should. If the temperatures get to extreme it might even get permantly damaged. So for some apllications it is a wise choice to take this specification into consideration.

8 Bibliography

- $http://www.siliconimaging.com/cmos_fundamentals.htm$
- $http://www.cambridgeincolour.com/tutorials/camera-sensors.htm$
- $http://www.globalspec.com/learnmore/semiconductors/microprocessors_microcontrollers/microcontr$ item $Microcontrollersarchitecture, implementation\&programmingauthors: kennethhintzanddanieltab$
- $Skolnik, M.2008, RadarHandbook, McGrawHill$
- $O'Donnel, RobertM.2002, SlidesIntroductiontoRadarSystems, LincolnLaboratoryMassachusettsInstitu$
- $http://www.montana.edu/gps/understd.html$
- $GlobalPositioningSystem: TheoryandPractice, authorsB.Hofmann-Wellenhof, H.Lichtenegger, J.Coll$
- $MITopenlecturehttp://ocw.mit.edu/resources/res-ll-003-build-a-small-radar-system-capable-of-sensing-range-doppler-and-synthetic-aperture-radar-imaging-january-iap-2011/lecture-notes/MITRES_LL_003IAP11_lec03.pdf$
- $DigitalandAnalogCommunicationSystems - LeonW.CouchII$
- $StructuredElectronicDesign, byChrisVerhoevenetal.$