Radar target classification using Support Vector Machines and Mel Frequency Cepstral Coefficients

SEBASTIAN EDMAN
Radar target classification using Support Vector Machines and Mel Frequency Cepstral Coefficients

EDMAN SEBASTIAN
Abstract

In radar applications, there are often times when one does not only want to know that there is a target that reflecting the out sent signals but also what kind of target that reflecting these signals. This project investigates the possibilities to from raw radar data transform reflected signals and take use of human perception, in particular our hearing, and by a machine learning approach where patterns and characteristics in data are used to answer the earlier mentioned question. More specific the investigation treats two kinds of targets that are fairly comparable namely smaller Unmanned Aerial Vehicles (UAV) and Birds. By extracting complex valued radar video so called I/Q data generated by these targets using signal processing techniques and transform this data to a real signals and after this transform the signals to audible signals. A feature set commonly used in speech recognition namely Mel Frequency Cepstral Coefficients are used two describe these signals together with two Support Vector Machine classification models. The two models where tested with an independent test set and the linear model achieved a overall prediction accuracy 93.33 %. Individually the prediction resulted in 93.33 % correct classification on the UAV and 93.33 % on the birds. Secondly a radial basis model with a overall prediction accuracy of 98.33 % where achieved. Individually the prediction resulted in 100% correct classification on the UAV and 96.76 % on the birds. The project is partly done in collaboration with J. Clemedson [2] where the focus is, as mentioned earlier, to transform the signals to audible signals.
Sammanfattning

Klassificering utav radarmål genom Support Vector Machines och Mel Frequency Cepstral Coefficients

I radarapplikationer räcker det ibland inte med att veta att systemet observerat ett mål när en reflekterad signal deketeras, det är ofta också utav stort intresse att veta vilket typ av föremål som signalen reflekterades mot. Detta projekt undersöker möjligheterna att utifrån rå radardata transformera de reflekterade signalerna och använda sina mänskliga sinnen, mer specifikt våran hörsel, för att skilja på olika mål och också genom en maskinöknings approach där med hjälp av mönster och karaktärsdrag för dessa signaler används för att besvara frågeställningen. Mer ingående avgränsas denna undersökning till två typer av mål, mindre obemannade flygande farkoster (UAV) och fåglar. Genom att extrahera komplexvärd radar video även känt som I/Q data från tidigare nämnda typer av mål via signalsbehandlingsmetoder transformera denna data till reella signaler, därefter transformeras dessa signaler till hörbara signaler. För att klassificera dessa typer av signaler används typiska särdrag som också används inom taligenkänning, nämligen, Mel Frequency Cepstral Coefficients tillsammans med två modeller av en Support Vector Machine klassificerings metod. Med den linjära modellen uppnåddes en prediktions noggrannhet på 93.33%. Individuellt var noggrannheten 93.33 % korrekt klassificering utav UAV:n och 93.33 % på fåglar. Med radial bas modellen uppnåddes en prediktions noggrannhet på 98.33%. Individuellt var noggrannheten 100% korrekt klassificering utav UAV:n och 96.76% på fåglar. Projektet är delvis utfört med J. Clemedson [2] vars fokus är att, som tidigare nämnt, transformera dessa signaler till hörbara signaler.
Acknowledgement

I would like to thank my supervisor Johan Karlsson, associate professor at the department of mathematics at KTH for supervising this thesis and by discussing ideas and giving feedback resulting to a thesis at the best of my abilities.

I wound also like to thank my two supervisors as SAAB, Stefan Eriksson and Niklas Broman. When staring this thesis my knowledge for radar system where small or even close to zero. By always taking time to answer my questions and give feedback on my reflections I’ve broaden my knowledge further to a direction who was not the main area in my field of study.

Thanks to friends and family for supporting me not only during this thesis but also under my five years at KTH. Lastly a sincere thanks to Johan Clemedson for his collaboration in this project.

Stockholm, August 2017

Sebastian Edman
Contents

1 Introduction and Problem Description 3
   1.1 Outline of Thesis . . . . . . . . . . . . . . . . . . . . . . . . . . . 4

2 Background on RADAR technology 5
   2.1 Pulse radar sets . . . . . . . . . . . . . . . . . . . . . . . . . . . 5
   2.2 Antenna . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 6
   2.3 Transmitter . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 8
   2.4 Transmission of signal . . . . . . . . . . . . . . . . . . . . . . . . 8
   2.5 Radar equation . . . . . . . . . . . . . . . . . . . . . . . . . . . . 9
   2.6 Range and Bearing . . . . . . . . . . . . . . . . . . . . . . . . . . 10
   2.7 Radar resolution . . . . . . . . . . . . . . . . . . . . . . . . . . . 11
   2.8 Doppler effect . . . . . . . . . . . . . . . . . . . . . . . . . . . . 12
      2.8.1 In-phase/Quadrature demodulation . . . . . . . . . . . . . . 14
      2.8.2 µ-Doppler effect . . . . . . . . . . . . . . . . . . . . . . . . 16

3 Signal Processing Background 18
   3.1 Filters . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 18
      3.1.1 FIR Filters . . . . . . . . . . . . . . . . . . . . . . . . . . . 18
      3.1.2 Matched Filter . . . . . . . . . . . . . . . . . . . . . . . . . 19
   3.2 Spectral Transforms . . . . . . . . . . . . . . . . . . . . . . . . . . 22
      3.2.1 Discrete Fourier Transform . . . . . . . . . . . . . . . . . . . 22
   3.3 Features . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 23
      3.3.1 Mel Frequency Cepstral Coefficients . . . . . . . . . . . . . 23

4 Classification Method 25
   4.1 Support Vector Machines (SVM) . . . . . . . . . . . . . . . . . . . 25
   4.2 Dual Formulation . . . . . . . . . . . . . . . . . . . . . . . . . . . 27
   4.3 Kernel Trick . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 28

5 Radar Signal Processing 30
   5.1 MTI-filter . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 31
   5.2 Pulse Compression . . . . . . . . . . . . . . . . . . . . . . . . . . . 32
   5.3 Signal Extraction . . . . . . . . . . . . . . . . . . . . . . . . . . . . 33
   5.4 Spectral Modifications . . . . . . . . . . . . . . . . . . . . . . . . . 34

6 Data Sets and Implementation 37
   6.1 Data Set . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 37
   6.2 Implementation . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 37
   6.3 Training and Parameter Evaluation . . . . . . . . . . . . . . . . . . . 39

7 Results 41

8 Discussion and Future Work 45
### Appendix A

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resampling</td>
<td>47</td>
</tr>
<tr>
<td>Up-Sampling by integer factor</td>
<td>47</td>
</tr>
<tr>
<td>Interpolation Filter</td>
<td>50</td>
</tr>
<tr>
<td>Polyphase Interpolation filter</td>
<td>51</td>
</tr>
<tr>
<td>Down-sampling (or Decimation) by an Integer Factor</td>
<td>52</td>
</tr>
<tr>
<td>Resampled with non-integer factor</td>
<td>52</td>
</tr>
<tr>
<td>Kaiser window</td>
<td>52</td>
</tr>
</tbody>
</table>
Nomenclature

\((\cdot)^H\) Hermitian transpose  
\(\alpha\) Lagrange multipliers  
\(*\) Convolution  
\(\beta_0\) bias  
\(\beta\) Weight vector  
\(\lambda\) Wave Length  
\(\sigma_t\) Radar Cross Section  
\(\tau\) Pulse Width  
\(c\) Speed of light  
\(f_0\) Carrier frequency  
\(f_D\) Doppler frequency  
\(G\) Antenna Gain  
\(h[\cdot]\) Impulse response  
\(L_s\) Loss factor  
\(P\) Signal power  
\(S_A\) Angular resolution  
\(S_r\) Range resolution
## Acronyms

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CC</td>
<td>Cepstral Coefficients.</td>
</tr>
<tr>
<td>CW</td>
<td>Continuous Wave.</td>
</tr>
<tr>
<td>DFT</td>
<td>Discrete Fourier Transform.</td>
</tr>
<tr>
<td>DTW</td>
<td>Dynamic Time Wrapping.</td>
</tr>
<tr>
<td>FIR</td>
<td>Finite Impulse Response.</td>
</tr>
<tr>
<td>HMM</td>
<td>Hidden Markov Models.</td>
</tr>
<tr>
<td>IDFT</td>
<td>Inverse Discrete Fourier Transform</td>
</tr>
<tr>
<td>LOS</td>
<td>Line Of Sight.</td>
</tr>
<tr>
<td>LSS</td>
<td>Low, Slow flying and Small.</td>
</tr>
<tr>
<td>MDS</td>
<td>Minimum Discernible Signal.</td>
</tr>
<tr>
<td>MFCC</td>
<td>Mel Frequency Cepstral Coefficients</td>
</tr>
<tr>
<td>MTI</td>
<td>Moving Target Indicator.</td>
</tr>
<tr>
<td>PRF</td>
<td>Pulse Repetition Frequency.</td>
</tr>
<tr>
<td>PRT</td>
<td>Pulse Repetition Time.</td>
</tr>
<tr>
<td>RF</td>
<td>Radio Frequency.</td>
</tr>
<tr>
<td>SNR</td>
<td>Signal to Noise Ratio.</td>
</tr>
<tr>
<td>SVM</td>
<td>Support Vector Machine.</td>
</tr>
<tr>
<td>TX</td>
<td>Transmission.</td>
</tr>
<tr>
<td>UAV</td>
<td>Unmanned Aerial Vehicles.</td>
</tr>
</tbody>
</table>
1 Introduction and Problem Description

The radar has for more than half a century been the go to technology for surveillance and target recognition. This is due to the many advantages of radar systems, for example a radar system can detect and track moving objects in all weathers during day and night. A radar transmits and receives electromagnetic waves, that interacts with targets and the surrounding environment. The velocity of moving targets can be calculated by measuring the Doppler frequency shift of the received echo signal. Detecting the target is often not enough, one also wants to distinguish between targets. In today’s radar systems this can be done in various ways, for example discriminate targets by their velocity or specific motion patterns which are unique for a target. Such methods are almost as old as the radar itself and in later years more sophisticated methods have been developed by the use of machine learning algorithms, for example image classification in Synthetic Aperture Radar (SAR). But classification is not necessary performed in an automatic approach. Some radar system, possess an audio output. For example the MSTAR Battlefield Surveillance Radar, which is a man portable lightweight Doppler radar used by an operator who can listen to the audio output. The audio signal is a representation of the echo from the illuminated target which contains the Doppler frequencies. The operator makes the classification by recognize specific patterns in the audio signal, which is similar to the techniques used in SONAR applications. The operators ability to perform auditory classification is based on speech phonemics principles. A phoneme is a specific sound pattern that the human brain is able to recognize. For the operator to be able do distinguish between different targets (and also specific actions by targets) extensive training is needed. The human factor may also increase the error rate, especially on the battlefield where external factors can have an impact on the human senses. Today it exist speech recognition methods like Dynamic Time Wrapping (DTW) and Hidden Markov Models (HMM). But those methods have been optimized for speech signals and the aural Doppler signal is not a conventional speech signal. Another challenge may be to classify targets during the scan of a surveillance radar which results in rather short time frames of the signal (milliseconds) and contains relatively long discontinuities between the frames. This can lead to that existing methods may be possible to use in a mode when the radar stares at a target but does not imply that it would be possible to use in scanning mode.

The overall aim of this project is to investigate the possibilities to transform radar data to an aurial output and make use of feature driven classification algorithms based on aurial signals to discriminate between Low, Slow flying and Small (LSS) targets such as smaller Unmanned Aerial Vehicles (UAV) and birds that is fairly comparable (size, velocity, flight pattern) and thus hard to distinguishing between in a radar system application. The project has three main parts, and is partly done in collaboration with J. Clemedson. The included parts is stated as follows.

i) From raw radar data, with help of signal processing techniques extract sig-
nals from UAV:s and birds

ii) Transform those signals so it would be possible for an operator to listen to an aurial signal

iii) Based on these aurial signals, develop an automatic classification method on the same theme to serve as supplement to the operator

Part i) is done in collaboration with J. Clemedson and included in this thesis, part ii) can be found in [2] and part iii) are alone included in this thesis.

1.1 Outline of Thesis

The thesis contains 8 Sections. Section 2 treats basic radar technology including transmission of signals, range measurements and the Doppler effect. In Section 3 signal processing techniques are introduced that are used in the thesis. Section 4 treats theory on the Support Vector Machine classification method and subjects like duality and kernel trick. Section 5 describes the process used for extracting radar data from complex raw radar video, so called I/Q data. In section 6 the datasets for classification are described and the implementation of the classification method training phase is outlined. In Section 7 the result from the classification is presented and finally in Section 8 some own thoughts on the project and future work are discussed. Section 2, 3 (except 3.3) and 5 are done in collaboration with J. Clemedson.
2 Background on RADAR technology

A radar (RADio Detecting And Ranging) transmits electromagnetic energy in the radio-frequency (RF) interval. When the transmitted RF energy hits a target the energy is reflected. The radar receives a small quantity of the reflected energy, called the echo, which is used to determine the distance and direction of the object. This section aims to introduce some basic radar theory needed understand the terminology used in this thesis. For starters general signal transmission is introduced with theory gathered from [18] and the included components in a radar system from [19] and [21] then general range measurements in a radar system from [20].

Radar systems can be divided in to two different types of radar systems, primary and secondary radar systems. Primary radar systems transmits a signal and receives the reflected echo. While secondary radar systems receives an coded reply signal from an transmitter on the illuminated targets. Hence secondary radar systems are not used to detect unknown targets but instead tracking and identifying friendly targets. Primary radar systems are divided into Continuous Wave (CW) and pulsed radar systems. A CW radar set transmits a high frequency signal continuously, the received echo signal is also processed continuously. CW radars can normally only measure the targets speed and not the distance to targets, since there are no pulse sets to time. This problem can be solved by constantly shifting the frequency in the transmitted signal. These frequencies can be extracted from the echo and by knowing when in the past that particularly frequency was sent out, one can do a range calculation. This types of CW radar sets are called frequency modulated CW radars, hence CW radar sets can be divided in modulated and unmodulated CW radar systems. The typical use of unmodulated CW radar sets, which only can measure speed, are speed gauges for the police.

Pulse radar transmits a high frequency impulse signal of high power, after the impulse a longer break follows in which the echoes are received before then next impulse is sent out. Properties such as direction, range and speed can be determined by using a pulse radar. The following theory will only consider pulse radar theory.

2.1 Pulse radar sets

A powerful transmitter generates the radar signal, which is transmitted from antenna through a duplexer. The function of the duplexer is to switch the antenna between the transmitter and receiver, which means that only one antenna is needed. Switching between transmitting and receiving signals is also necessary, since the high power pulses produced by the transmitter would destroy the receiver if energy were allowed to enter the receiver during transmission. The antenna illuminates the target with the radio frequency pulse, which is reflected at the target. The backscattered echo signal is picked up by the receiver in the antenna. The received radar pules is amplified and demodulated in the receiver, which produces video signals that can be displayed. The operating principles of
2.2 Antenna

One of the most important parts of the radar system is the antenna. The antenna performs the following crucial functions [19]:

- The antenna transfers the energy to electromagnetic waves in space with required distribution and efficiency. The antenna also transforms the electromagnetic waves in the echo signal to electric signals.

- The required signal pattern in space is ensured by antenna. The signal pattern in angle has to be sufficiently narrow to provide the required angular resolution.

- A scanning antenna has to provide the required frequency of target position updates, i.e., the revolution rate.

- The antenna has to measure the pointing direction with a high degree of accuracy.

One important antenna characteristic is the antenna gain (Directivity, Direction gain), which describes how well the antenna can focus the outgoing energy in a certain direction. The antenna gain is defined as the ratio between the amounts of energy propagated in the radar direction compared to energy propagating in other directions. The antenna will also have the same gain for receiving signals, if a transmitting antenna is used as a receiving antenna.

Antennas usually emit a stronger radiation in one direction than in other directions. Such radars are called anisotropic. A radiation pattern is formed by the energy radiated from the radar. The shape of the radiation pattern depends on the type of the antenna. When the transmitted energy is measured in various angles at a constant distance from the antenna an illustration of the radiation
pattern can be plotted, usually in polar coordinates. In such plots three key features can be observed:

- In the region that is within 3 dB of the maximum radiation, is called the main lobe or main beam.

- Smaller beams in other directions than the main lobe are called sidelobes, which are usually radiation in undesired directions. These sidlobes can never be completely eliminated.

- The portion of the radiation pattern that are directed in the opposite direction of the main lobe is called the backlobe.

One important characteristic of the radiation pattern is the beam width. The beam width is defined as the angular range of the radiation pattern in which at least half of the maximum power is still emitted. Hence the bordering points of the lobe are therefore the points at which the field strength fallen 3 dB compared to the maximum field strength. The notation of the beam width (or half power angle) is $\Theta$. The beam width can be determined in both the horizontal $\Theta_{AZ}$ and vertical plane $\Theta_{EL}$.

The ratio of power gain between the main lobe and the backlobe is called the front to back ratio. One desires a high front to back ratio, since its means that a minimum amount of energy is radiated in the undesired direction.

Air-surveillance radar system usually uses a cosecant squared pattern, to achieve a more uniform signal strength from a target that moves with a constant elevation. The height information from a detection can be calculated by knowing the elevation of the returned echo. This is done with the agile multiple beam concept, where the height information is divided into multiple parts (beams), see Figure 2. The cosecant squared pattern can be achieved by stacking beams according to the figure.

![Cosecant Pattern Illustration](image-url)

Figure 2: Illustration of cosecant pattern [Image]
2.3 Transmitter

The task of the transmitter is to produce high power Radio Frequency (RF) pulses of energy that are radiated into space by the antenna. There are different kinds of transmitter with different properties. Depending on what type of transmitter used in the radar set, the radar set can be classified as coherent or non-coherent [21]. The radar system is said to be coherent if every transmitted pulse starts with the same phase and non-coherent if the phase of each successive pulse is random. The reason to use a coherent system is to keeping track of the phase change of the reflected pulses generated by a moving target and hence the Doppler frequencies. Radars that emit coherent pulses to measure the Doppler shift are known as pulse Doppler radars. The difference between coherent and non-coherent pulses can be seen in Figure 3.

![Figure 3: Illustration of coherent and non-coherent radar transmitters.](image)

2.4 Transmission of signal

Each transmitting pulse is radiated from the radar, during the transmit time (or pulse width \(\tau\)). The radar is waiting for return echo during the listening time, after each transmitted pulses. There is a short rest time between the listening time and the next pulse. The time between two pulses are called the Pulse Repetition Time (PRT). The number of pulses transmitted per second is called Pulse Repetition Frequency (PRF), the relationship between the pulse repetition time and pulse repetitive frequency is given by \(PRT = PRF^{-1}\). An illustration of a transmission of a pulse with previously mentioned variables can be seen in Figure 4.
2.5 Radar equation

The relation between the transmitted power $P_{tx}$ and the power in the echo signal $P_{rx}$ is given by the radar equation [20],

$$P_{rx} = P_{tx} \frac{G^2 \cdot \lambda^2 \cdot \sigma_t}{(4\pi)^3 \cdot R^4 \cdot L_s}.$$  \hspace{1cm} (2.1)

Where $G$ is the antenna gain, which is a measure of the antenna’s ability to focus outgoing energy into the direction of the beam. The antenna gain is given by the maximum radiation intensity divided by the average radiation intensity. How well an antenna can pick up power from an incoming electromagnetic wave are described by the antenna aperture, $G \cdot \lambda^2/(4\pi)$. Where $\lambda$ is the wave length of the electromagnetic wave. The parameter $\sigma_t$ is the radar cross section (RCS), which is the size and the ability of a target to reflect radar energy summarized in one term. The factor $1/(4\pi \cdot R^2)^2$ describes the free space path loss, where $R$ is the range to the target. The free space path loss is the loss in signal strength of an electromagnetic wave propagating in the line of sight path through free space. All the internal losses of the radar are summarized in the loss factor $L_s$.

The Minimum Discernible Signal (MDS) is the smallest signal that the radar can detect. If the power is smaller than this $P_{MDS}$, the signal will not be usable since it will be lost in the background noise. By rewriting the radar equation (2.1) and setting the power of the echo signal equal to the $P_{MDS}$ one obtains,

$$R_{\text{max}} = \sqrt[4]{\frac{P_{tx} \cdot G^2 \cdot \lambda^2 \cdot \sigma_t}{(4\pi)^3 \cdot P_{MDS} \cdot L_s}}.$$ \hspace{1cm} (2.2)

This gives the relation between the maximum range $R_{\text{max}}$ and the transmitted power for a radar system. Due to the fourth root one must increase the trans-
mitted power 16 times to double the maximum range, if the other parameters are constant.

2.6 Range and Bearing

The slant range $R$ is defined as the line of site distance between the target and the radar antenna. It is possible to calculate the slant range from the time delay $t_{\text{delay}}$ between the transmitted and the reflected pulse, with the following equation

$$ R = \frac{ct_{\text{delay}}}{2}, \quad (2.3) $$

where $c$ is the speed of light. It is required to know the target’s elevation to calculate the horizontal distance between the target and radar (ground range).

Since a pulse radar usually transmits a sequence of pulses and measures the time between the last transmitted pulses and the echo pulse. It is possible that the received echo is from a long range target, so that the received signal arrives at the radar after or during the transmission of the next transmitting pulse. This means that the radar is measuring the wrong time interval and therefore the wrong range. The radar assumes that the pulse is the reflection of the second transmitted pulse and declares a reduced range for the target. This occurs when strong targets are located outside the range that corresponds to the pulse repetition time, and are called range ambiguity. Hence a maximum unambiguity range is defined by the pulse repetition time. The relationship between the pulse repetition time and the unambiguity range $R_{\text{uamb}}$ is given by

$$ R_{\text{uamb}} = \frac{(PRT - \tau)c}{2}. \quad (2.4) $$

There are two types of echo signals that arrive after the reception time:

- Echo signals that arrive during the transmission time. These signals will not be registered since the receiver is turned off during transmission.

- Echo signals that arrive during the following reception time. These signals will give range measuring failures (ambiguous returns) illustrated in Figure 5.

Still, it is possible to discriminate the true range of targets by using different Transmission (TX) modes. The different transmission modes have various pulse repetition frequencies (PRF) explained in Section 2.4. Hence targets at ambiguous ranges will appear at different ranges for each TX-mode allowing the radar system to compute and solve the ambiguity and extract the true range.
There is also a minimum detectable range (or blind distance) to consider. This is due to that the echo signal will not be registered, if the echo from the beginning of the pulse falls inside the transmitting pulse. Since the receiver is turned off during the transmitting time. The blind distance can be calculated with the following equation,

\[ R_{\text{min}} = \frac{(\tau + t_{\text{rest}})c}{2}, \]

where \( t_{\text{rest}} \) is the resting time. Both the horizontal and elevation angle between the antenna and target, can be determined by measuring the direction in which the antenna is pointing when the echo is received. The accuracy of the radars angular measurements is determined by the antennas directional gain. The angle measured in the horizontal plane is refereed to as bearing, which can be measured in true or relative bearing. True bearing is the angle between the true north and a line pointing directly at the target, measured in a clockwise direction. Relative bearing is the angle between the centerline of the own ship ore aircraft and a target, measured in a clockwise direction.

### 2.7 Radar resolution

A radar is not able to distinguish between targets that are very close in bearing or range. The ability to distinguish between close target are given by the target resolution of the radar, which is divided in angular resolution and range resolution. The minimum angular separation of two equal targets at the same range, are called angular resolution. The angular resolution of an radar is determined by the half power beam width \( \Theta \) of the radar. Which is the angle between the half power (-3 dB) points of the main lobe. Which means that two targets can be resolved in angle if they are separated more than one beam width. Hence the smaller the beam width, the higher directivity of the antenna, the better angular resolution. The distance between two targets corresponding to the angular resolution is a function of the slant range and given by,

\[ S_{A} = 2R\sin \frac{\Theta}{2}. \]

Where \( S_{A} \) is the angular resolution given as a distance between two targets.
The ability to distinguish between two or more targets at different ranges but on the same bearing, is called range resolution. The range resolution is a factor since the echo’s from close targets in range gets mixed up as illustrated in Figure 6.

![Figure 6: Illustration of range resolution](image)

Hence the primary factor in range resolution is the pulse width of the transmitted pulse. But the range resolution also depends on types and sizes of the target, and the efficiency of the receiver and indicator. Targets separated by half the pulse width, can be separate distinguished by an well designed radar with all other factors at maximum efficiency. Hence, the theoretical range resolution of a radar system is given by

$$S_r = \frac{c\tau}{2}.$$  \hspace{1cm} (2.7)

Where $S_r$ is the range resolution given as distance between two targets. A method to improve the range resolution is by using a pulse compression system, using pulse compression allows high range resolution with long pulses, but with a higher average power.

### 2.8 Doppler effect

The Doppler effect was discovered by Christian Doppler 1842 and has been used in electromagnetic since 1930 and the first Doppler radar was produced in 1950. The use of the Doppler radar is mainly for targets in motion e.g in military use targets of interest can be hostile air, sea and land targets such as airplanes, ships and tanks but also smaller targets like rockets, artillery and mortars. The principle behind the Doppler shift is that a target with a motion relative to the radar will induce a frequency shift in the reflected signal i.e the Doppler shift which depends on the wavelength of the transmitted signal and the radial velocity (the velocity in the Line Of Sight (LOS) of the radar), of the illuminated target. Suppose that the transmitted signal is
\( s_i(t) = u(t) \cos(2\pi f_0 t) \) \hspace{1cm} (2.8)

where \( u(t) \) is the signal envelope, \( f_0 \) is the carrier frequency, and \( t \) is the time. The received backscattered echo signal from a target can then be expressed as

\[ s_r(t) = \sigma s_i(t - t_r) = \sigma u(t - t_r) \cos(2\pi f_0(t - t_r)) \] \hspace{1cm} (2.9)

where \( \sigma \) is the target reflection coefficient and \( t_r \) is the time delay of the echo w.r.t the transmitted signal. If the target would be stationary relative to the radar the time delay would be constant and also the phase of the echo signal. If the target would be in motion relative to the radar with a velocity \( v_r \) in the line of sight of the radar the echo signal received at time \( t \) transmitted at time \( t - t_r \) and the the time that the target is illuminated is \( t_i = t - t_r/2 \) and hence the distance between the origin of the radar and the illuminated target is

\[ R(t_i) = R_0 - v_r t_i \] \hspace{1cm} (2.10)

where \( R_0 \) is the distance between the origin of the radar and the illuminated target at time \( t = 0 \). The propagation time of the signal for traveling the distance from the radar to the target and back is the time delay of the echo \( t_r \) and thus

\[ t_r = \frac{2R(t_i)}{c} \] \hspace{1cm} (2.11)

where \( c \) is the speed of light in the case of electromagnetic waves. By combining (2.10) with (2.11) and substitution into the echo signal (2.9) yields

\[ s_r(t) = \sigma u \left( \frac{t}{c+v_r} - \frac{2R_0}{c-v_r} \right) \cos \left[ 2\pi f_0 \left( \frac{t}{c-v_r} - \frac{2R_0}{c-v_r} \right) \right] \] \hspace{1cm} (2.12)

The echo signal (2.12) possesses two important properties \[24\]

i) From the phase term of (2.12) one can see that the frequency is shifted from \( f_0 \) to \( f_0 \frac{c+v_r}{c-v_r} \)

ii) There is a scaling change for the echo signal envelope in terms of time.

The envelope change of the echo signal described in [3] can in most radar applications be ignored since the process of the phase term will not heavily depend on the scaling change of the echo signal envelope \[24\].

Usually the radial velocity \( v_r \) of the illuminated target is significantly smaller than the propagation speed of an electromagnetic wave \( c \) thus one can argue to approximate the frequency shift

\[ f_0 \frac{c+v_r}{c-v_r} - f_0 = f_0 \frac{2v_r}{c-v_r} \approx f_0 \frac{2v_r}{c} = \frac{2v_r}{\lambda_0} \] \hspace{1cm} (2.13)

Where \( \lambda_0 \) is the carrier wavelength, the right hand side in (2.13) is the defined as the Doppler frequency
If \((2.14)\) combined with \((2.13)\) is inserted in \((2.12)\) and assume that the envelope of the signal \(u(s) = 1\) we arrive at

\[
s_r(t) = \sigma \cos \left[ 2\pi (f_0 + f_D) t - \frac{2R_0}{c - \nu_r} \right] = \sigma \cos \left[ 2\pi (f_0 + f_D) t - \theta \right]
\]

\[
(2.15)
\]

In order to distinguish between positive and negative Doppler frequencies, an I/Q representation of the signal is used and introduced in the next section.

### 2.8.1 In-phase/Quadrature demodulation

In Doppler radar applications it is of great importance to be able to distinguish between positive and negative frequencies. Since the sign of the Doppler frequency represents in which direction the targets is moving. But a signal representation just using a series of samples of the momentary amplitude of the signal (see Figure 7a) will not be able to distinguish between positive and negative frequencies (for example \(\cos(x) = \cos(-x)\)). Hence the In-phase/Quadrature (I/Q) representation of the signal is introduced. Where the signal is compared to a reference signal which gives an in-phase (I) component and a quadrature (Q) component of the signal [9]. The I/Q-signal can be illustrated as spiral in three dimensions, see Figure 7b.

![Figure 7](image_url)

(a) Projection of I/Q representation  
(b) I/Q representation

Figure 7: Signal representation

One can see that the projection of the spiral on to the vertical plane is the "real" signal (Figure 7a), which is the I component. The projection of the spiral on to
the vertical plane gives the Q component of the signal. The direction of rotation of the spiral in Figure 7b determines the sign of the frequency.

To extract the I and Q component of a signal and hence the Doppler frequency shift a quadrature detector is used. The quadrature detector produces a signal with an In-phase (I) component and a Quadrature (Q) component from the input signal. The extraction of the Doppler frequency shifts with the quadrature detector are illustrated in Figure 8.

\[ s_r(t) = \sigma \cos(2\pi f_0 t + \phi(t) - \theta) \]

\[ s_t(t) = \cos(2\pi f_0 t) \]

The quadrature detector consists of two mixers, called synchronous detectors. In the synchronous detectors the received signal is mixed with a reference signal, the transmitted signal in the first synchronous detector and a 90° shift of the transmitted signal in the second mixer. After each synchronous detector a low pass filter is applied to filter out the carrier frequency \( f_0 \) of the transmitted signal

\[ s_r(t) = \sigma \cos(2\pi(f_0 + f_D)t - \theta). \quad (2.16) \]

In the first synchronous detector the received signal is mixed with the transmitted signal

\[ s_t(t) = \cos(2\pi f_0 t), \quad (2.17) \]

which gives the output

\[ s_r(t)s_t(t) = \frac{\sigma}{2} \cos(4\pi f_0 t + \phi(t) - \theta) + \frac{\sigma}{2} \cos(\phi(t) - \theta). \quad (2.18) \]

Applying the low pass filter to the signal gives the I-channel output,
\[ I(t) = \frac{\sigma}{2} \cos(\phi(t) - \theta). \] (2.19)

In the other synchronous detector, the 90° phase shifted transmitted signal,
\[ s_{t90}^0(t) = \sin(2\pi f_0 t) \] (2.20)

which gives the output
\[ s_r(t)s_{t90}^0(t) = \frac{\sigma}{2} \sin(4\pi f_0 t + \phi(t) - \theta) - \frac{\sigma}{2} \sin(\phi(t) - \theta). \] (2.21)

Applying the low pass filter to the signal gives the Q-channel output,
\[ Q(t) = -\frac{\sigma}{2} \sin(\phi(t) - \theta). \] (2.22)

By combining the I and Q part the following signals is obtained
\[ s_D(t) = I(t) + iQ(t) = \frac{\sigma}{2} e^{-i\phi(t)} e^{-i(2\pi f_D t - \theta)} = \frac{\sigma}{2} e^{-i(2\pi f_D t - \theta)}. \] (2.23)

From the complex Doppler signal in Equation (2.23) it is possible in to extract positive and negative frequencies, i.e. positive and negative velocities.

### 2.8.2 µ–Doppler effect

By extracting information about the illuminated target via the Doppler shift in terms of radial velocity and range useful information is gained, but in real radar applications targets with single motion pattern is quite rare. For example man made aerial targets like helicopters and UAV:s consist of more complex motions than just the bulk motion, like engine vibrations and rotations of propellers. And biological targets such as personnel or birds generates complementary motions like swinging arms and flapping wings. These so called micro-motions can be useful when trying to distinguish between many more classes of targets mention above and even different types of the same kind of target due to unique characteristics. Pursuant to Doppler theory beyond the bulk motion of a target micro-motions from parts of the target or the target itself can cause frequency modulation on the echo signal from a radar system. Which is in fact a Doppler sideband besides the main Doppler frequency induced by the bulk motion of the target [24]. These frequency modulations generated by micro-motions are in literature and research called micro-Doppler effect and a target is said to have a specific micro-Doppler signature.

The micro-Doppler effect has its origin in coherent laser detecting and ranging systems (LADAR) [1] who transmit electromagnetic waves at optical frequencies and by the backscattered wave from an object one can measure properties such as range, velocity similar to a radar system by preserve phase information. Since the phase of a backscattered signal in a coherent system is sensitive to the variation in range a half wavelength change in range can cause 2π change in phase. In LADAR systems where the wavelengths is typically short e.g 2 µm and thus a
change in radial distance by 1 $\mu$m can generate a shift in phase by $2\pi$ leading to extremely high sensitivity in LADAR systems where for example tiny vibrations can be observed rather easily. The micro-Doppler frequency is a time varying property and can be extracted from the output from a quadrature detector used in standard Doppler processing [1].

The author of [1] [24] validated the $\mu$-Doppler effect in radar systems by using an X-band radar to detect trigonometric scatter target with a vibration amplitude of 1 mm and vibration frequency of 10 Hz and successfully extracted micro-Doppler frequency shift through time-frequency analysis technique, later on the concerned also put forth results of micro-Doppler analysis results of a pedestrian with X-band radar. Since then many papers related to the research field have been published, not only in the research of micro-Doppler but also with the previous combined with various classification methods. The authors of [13] uses a speech recognitions techniques, Dynamic Time Wrapping (DTW) and a $k$-NN classifier to classify baseband audio output signal from a radar with help of micro-Doppler signatures. Speech processing algorithms exploit the time variance in speech patterns to classify signals and identify words and the intend was to exploit the time variance in the micro-Doppler signature in a similar manner. Here the classification set consisted of three classes namely wheeled vehicles, tracked vehicles and personnel. The correct classification rate where 80%, 70% and 100% for the incoherent DTW classifier and 86%, 68% and 94% for the coherent DTW classifier for respective class. The DTW classifiers outperformed the $k$-NN classifier by far. Worth mention is that that the data used consisted of 80,000 samples of complex data when the velocity of the three classes where comparable and moving radially towards the radar. The duration of the data was also far longer than a typical radar dwell in scanning mode and data was divided into frames of reasonable times to increase the realism. The random nature of the initial phase of the micro-motions and the angle of LOS is also discussed in terms of the challenges it entails. However the aim of this thesis is not to make direct use of the micro-Doppler analysis as above but instead this section serves as motivation for the possibilities to distinguish targets by the "unique" time varying nature of the aural output generated by the micro-motions. In [7] the authors analyses the Doppler sound and uses cepstrum features and a Hidden Markov Model (HMM) together with a track based classifier to distinguish between personnel, land and air based vehicles. However a standalone analysis of the Doppler sound classifier showed good result (around 90 to 95 % correct classification of respective class).
3 Signal Processing Background

Linear and time invariant (LTI) systems are a tool used in signal processing, for example filters are almost always LTI systems. A system model is said to be linear if the model can be described as an linear mapping $w: \mathcal{U} \rightarrow \mathcal{Y}$, where $\mathcal{U}$ is an input space and $\mathcal{Y}$ an output space [4]. Causal and time invariant systems can be represented as

$$y[n] = \sum_{k=0}^{\infty} h[k] x[n-\ell] = \sum_{k=0}^{\infty} h[k] z^{-k} x[n] = h(z) x[n],$$

(3.1)

where $x[n] \in \mathcal{U}$ for $n \in \{0, 1, 2, \ldots\}$ is the a sequence of inputs and $y[n] \in \mathcal{Y}$ for $n \in \{0, 1, 2, \ldots\}$ is a sequence of outputs. The $z$ in Equation 3.1 denotes the forward shift operator. The values in the sequence $h[n]$ are called impulse response of the system and $h(z)$ the transfer function of the system.

3.1 Filters

Filters has the purpose to changes a signals frequency content, a filter can be ether analog or digital. In this thesis only digital filters are considered. A digital filter is a linear time invariant discrete system with the purpose of letting frequencies in a specific range pass and stop frequencies outside this range [5]. The filter can be described by

$$y[n] = \sum_{i=0}^{N} b_i x[n-i] - \sum_{j=1}^{M} a_j y[n-j],$$

(3.2)

where $x$ is the input and $y$ the output of the system. The parameters $a_j$ and $b_i$ are filter specific parameters, which characterize the filter. By taking the $z$-transform of (3.2) one can obtain the filter transfer function

$$H(z) = \frac{Y(z)}{X(z)} = \frac{\sum_{i=0}^{N} b_i z^{-i}}{1 + \sum_{j=1}^{M} a_j z^{-j}}$$

(3.3)

The filter is designed by choosing the coefficients $a$ and $b$ so that the desired filter characteristics are fulfilled.

3.1.1 FIR Filters

One type of digital filters is the Finite Impulse Response (FIR) filter, which is a non-recursive filter, hence the filter only depends on previous values of the input signal [5]. One characteristic property of the FIR filter is that the impulse response is equal to the filter coefficients and are zero outside a bounded interval. A causal FIR filter of order $N$ are described by the following convolution sum,

$$y[n] = \sum_{i=0}^{N} b_i x[n-i],$$

(3.4)

where $y$ is the output signal, $x$ is the input signal and $b_i$ is the value of the impulse response at the $i$:th instant. Since the FIR filter is causal and finite it
also holds that \( b_i = 0 \) if \( i < 0 \) or \( i > N \). A system illustration of an FIR filter is presented in Figure 9.

![Figure 9: System overview of a FIR filter](image)

Where each unit delay is a \( z^{-1} \) operator in z-transform notation. The transfer function of a FIR filter are calculated with the z-transform and given by

\[
H(z) = \frac{Y(z)}{X(z)} = \sum_{i=0}^{N} b_i z^{-i}. \quad (3.5)
\]

Equation 3.5 can be rewritten as

\[
H(z) = \frac{\sum_{i=0}^{N} b_i z^{-i}}{1} = \frac{\sum_{i=0}^{N} b_i z^{N-i}}{z^N}. \quad (3.6)
\]

Hence a FIR filter has equally many poles and zeros, but all the poles are located at the origin. Causal discrete system are stable if all poles lie in the open unit disk. Hence all causal FIR filters are stable, since all the poles for a causal FIR filter are located at the origin.

### 3.1.2 Matched Filter

In signal processing, and in particular radar systems, a technique called pulse compression which is an example a matched filter is often used. A matched filter is obtained by correlating the sent out signal with a received echo to detect the signal in the presence of noise, which is equivalent to convolve the echo with a conjugated and time reversed form of the signal. Moreover Matched filtering is used to maximize the Signal to Noise Ratio (SNR) with linear time invariant filters and the characteristics of the filter can be designed by either frequency response or impulse response [3].

The matched filter can be expressed as the following convolution sum,

\[
y[n] = \sum_{k=-\infty}^{\infty} h[n - k] x[k]. \quad (3.7)
\]

Where \( h \) is the filter impulse responds, \( x \) is the input and \( y \) the output. The idea of the matched filter is to suppress the noise and amplify the signal at some time sample \( n_0 \), as can be seen in Figure 10.
The desired matched filter is a complex-valued $N$-point FIR filter, $g$, which maximizes the signal to noise ratio. The output of the filter is the conjugate inner product of the filter and the $N$-point observed signal $x$. The observed signal $x$ consist of a deterministic signal $s$ and stochastic noise $w$. Which means that the signal can be expressed as

$$x[n] = s[n] + w[n], \text{ for } n \in \{0, 1, \ldots, N - 1\}.$$  \hfill (3.8)

For convenience the time index is dropped. The following definition is used in the derivation of the matched filter.

**Definition 1.** A matrix $A$ is said to be Hermitian symmetric if $A = A^H$, where $A^H$ is the conjugate transpose of the matrix $A$.

If the noise mean value is assumed to be zero, the covariance matrix is given by

$$R_w = E \{ww^H\}. \hfill (3.9)$$

Note that the covariance matrix is Hermitian symmetric. The output of the filter $y$ is given by the convolution of the filter $g$ and the observed signal $x$,

$$y = \sum_{n=0}^{N-1} \bar{g}[n]x[n] = g^H x = g^H s + g^H w = y_s + y_w, \hfill (3.10)$$

where $\bar{g}$ is the conjugate of $g$. The output can be split into $y_s$ and $y_w$, generated by the signal and the noise respectively. The Signal to Noise Ratio (SNR) is given by the ratio of the power of the desired signal and the power of the noise, which can be expressed as

$$\text{SNR} = \frac{|y_s|^2}{E \{|y_w|^2\}} = \frac{|g^H s|^2}{E \{|g^H w|^2\}}. \hfill (3.11)$$

The denominator can be in the following way

$$E \{|g^H w|^2\} = E \{(g^H w)(g^H w)^H\} = g^H E \{ww^H\} g = g^H R_w g, \hfill (3.12)$$

which gives the SNR expression

$$\text{SNR} = \frac{|g^H s|^2}{g^H R_w g}. \hfill (3.13)$$
Since the objective of the matched filter is to maximize the SNR, the problem is to solve the following optimization problem,

$$\max_g \text{SNR}$$ \hspace{1cm} (3.14)

or

$$\max_g \frac{|g^H s|^2}{g^H R_w g}.$$ \hspace{1cm} (3.15)

By using the property of Hermitian symmetry Equation 3.13 can be rewritten as

$$\text{SNR} = \frac{|g^H (R_w^{-\frac{1}{2}} R_w^{-\frac{1}{2}} s)^2|}{g^H (R_w^{-\frac{1}{2}} g)} = \frac{|(g R_w^{-\frac{1}{2}})^H (R_w^{-\frac{1}{2}} s)|^2}{(g^H R_w^{-\frac{1}{2}})^H (R_w^{-\frac{1}{2}} g)}.$$ \hspace{1cm} (3.16)

The Cauchy-Schwarz inequality is used to find an upper bound for the objective function. For an complex inner product the Cauchy-Schwarz inequality is given by

$$|u^H v|^2 \leq \|u\|^2 \cdot \|v\|^2 = (u^H u) \cdot (v^H v),$$ \hspace{1cm} (3.17)

the equality holds if and only if the vectors $u$ and $v$ are linear dependent. Hence an upper bound for the objective function is given by

$$\text{SNR} \leq s^H R_w^{-1} s.$$ \hspace{1cm} (3.19)

The upper bound is achieved if and only if

$$R_w^{-\frac{1}{2}} g = \alpha R_w^{-\frac{1}{2}} s,$$ \hspace{1cm} (3.20)

for an arbitrary scalar $\alpha$. The optimal filter coefficients to the filter in Equation (3.10) can be expressed as

$$g = \alpha R_w^{-1} s.$$ \hspace{1cm} (3.21)

If the noise is assumed to be zero mean withe noise $(E y_w = 0)$, the expected value of the power can be expressed as the standard deviation of the noise $\sigma_w$, since

$$\sigma_w = E \{|y_w|^2\} - (E \{y_w\})^2 = E \{|y_w|^2\}.$$ \hspace{1cm} (3.22)

The expectation value of the power of the noise can also be expressed as

$$E \{|y_w|^2\} = E \{|g^H w|^2\} = (\alpha R_w^{-1} s)^H R_w (\alpha R_w^{-1} s) = \alpha^2 s^H R_w^{-1} s = \sigma_w^2$$ \hspace{1cm} (3.23)

giving

$$\alpha = \frac{\sigma_w}{\sqrt{s^H R_w^{-1} s}}.$$ \hspace{1cm} (3.24)

Which gives the normalized filter coefficients

$$g = \frac{\sigma_w}{\sqrt{s^H R_w^{-1} s}} R_w^{-1} s.$$ \hspace{1cm} (3.25)

The matched filters impulse responds $h$ is given by the complex conjugate time reversal of $g$. 

21
3.2 Spectral Transforms

In order to be able to analyze frequency content of time signals the Discrete Fourier Transform (DFT) is introduced. This provides a method to analyze discrete signals since in real world applications continuous signals are rarely the case hence a method for managing discrete signals is needed.

3.2.1 Discrete Fourier Transform

If we start by the definition of complex Fourier series for a function \( x(t), \ t \in \{\mathbb{R}, 0 \leq t \leq T \} \)

\[
x(t) = \sum_{k=-\infty}^{\infty} c_k e^{i2\pi nt/T}
\]

\[
c_k = \frac{1}{T} \int_{0}^{T} x(t) e^{-ik2\pi t/T} dt.
\] (3.26)

Now instead of the continuous function \( x(t) \) consider discrete samples of \( x \) taken at time intervals \( \delta t \). Introducing \( x[n] \) for the \( n \):th sample of \( x(t) \) i.e. \( x(n\Delta t) \) of length \( N \) where \( N\Delta t = T \). If now we calculate the Fourier coefficients \( c_k \) in equation (3.26) as the sum over the samples \( x[n] \) instead of the integral

\[
c_k = \frac{1}{N} \sum_{n=0}^{N-1} x[n] e^{-ik2\pi n \frac{1}{N} \Delta t}
\] (3.27)

Where the right hand side of equation (3.27) is the definition of the Discrete Fourier Transform (DFT) explicitly given by

\[
X[k] = \frac{1}{N} \sum_{n=0}^{N-1} x[n] e^{-ik2\pi n /N}.
\] (3.28)

Corresponding to the DFT taking us from time domain to frequency domain the Inverse Discrete Fourier Transform (IDFT) taking us the other way around is defined by

\[
x[n] = \sum_{k=0}^{N-1} X[k] e^{ik2\pi n /N}.
\] (3.29)

Note the scaling by \( 1/N \) in equation (3.28) and 1 in equation (3.29). It is quite common in literature to use the scaling the other way around, but as proposed here the scaling is consistent with the Fourier series [5].

A useful property of the DFT is the frequency shift property defined as

\[
X[k - l] = \frac{1}{N} \sum_{n=0}^{N-1} x[n] e^{-i2\pi nk /N} e^{i2\pi nl /N} = \frac{1}{N} \sum_{n=0}^{N-1} x[n] e^{-i2\pi n(k-l) /N}
\] (3.30)
that is in words by multiplying the discrete signal $x[n]$ by the complex exponential $e^{\frac{i2\pi l}{N}}$ will generate a frequency shift in the spectrum with $l$.

### 3.3 Features

To be able to distinguish between various sound signals a set of latent variables is needed, i.e. variables that are characteristic for the signal that can be used for comparison without comparing the signal itself. The idea is as previously mentioned a operator is listening to the sound signature of a target and as a compliment to the human classification a feature driven classifier can act as support in situations when the operator is not certain on what kind of target is being listened to. In digital speech processing a commonly used feature set are the **Mel Frequency Cepstral Coefficients (MFCC)** [22]. However the signal in this case are not a speech signal and one can argue that it would not be a valid feature set for this kind of classification. To motivate the choice of MFCC it have been shown that using MFCC as a feature set for sound signal classification not being speech signals. For example in [12] MFCC is used as features when classifying respiratory sounds signals and delivered satisfactory classification results. And in [8] MFCC are used and also compared to ordinary Cepstral Coefficients (CC) as a feature set for an classification of heart sound and in particular heart diseases, the result showed that MFCC was superior to CC in a classification task. The concept of MFCC is to capture the human perceptual scale of frequencies, and incorporating this scale makes the features match more closely to what humans actually hear. To capture this a filter bank containing triangular filters evenly spaced in the Mel scale is applied to the spectrum of the signal.

#### 3.3.1 Mel Frequency Cepstral Coefficients

The MFCC feature extraction process used in this thesis is described below.

Let $x[n]$ be the discrete signal of length $N$, take the DFT of the signal

$$X[k] = \sum_{n=0}^{N-1} x[n] e^{-i2\pi n k/N}, \quad 0 \leq k < N. \quad (3.31)$$

Construct a filter bank consisting of $M$ triangular filters ranging from frequencies $f_{\text{start}}$ to $f_{\text{max}}$ given by

$$H_m[k] = \begin{cases} 0 & k < f(m-1) \\ \frac{k-f[m-1]}{f[m]-f[m-1]} & f[m-1] \leq k \leq [m] \\ \frac{f[m+1]-k}{f[m+1]-f[m]} & f[m] \leq k \leq f[m+1] \\ 0 & k > f[m+1] \end{cases} \quad (3.32)$$

the boundary points for the $m$:th filter is

$$f[m] = \frac{N}{F_s} B^{-1} \left( B(f_{\text{start}}) + m \frac{B(f_{\text{max}}) - B(f_{\text{start}})}{M+1} \right). \quad (3.33)$$
Here $F_s$ is the sampling frequency and $B(f)$ is the mapping from physical frequency $f$ to Mel scale

$$B(f) = 2595 \log \left(1 + \frac{f}{700}\right)$$

and thus

$$B(b)^{-1}(b) = 700(e^{b/2595} - 1).$$

An example of such a filter bank is illustrated in Figure 11. The filter bank is uniformly spaced in the Mel scale and consequently the bandwidth of each filter will increase logarithmically in the normal scale.

The energy in each band ($m$:th filter) is computed as

$$S[m] = \log \left( \sum_{k=0}^{N-1} |X[k]|^2 H_m[k] \right) \quad 0 < m \leq M.$$  

The resulting $S[m]$, is a representation of the energies that is sensitive in a way human hearing works in terms of frequencies.

The final step to extract the MFCC is to apply the Discrete Cosine Transform to the filter bank energies

$$c[n] = \sum_{m=0}^{M-1} S[m] \cos \left(\frac{\pi n (m - 0.5)}{M}\right) \quad 0 < n \leq M.$$
Where $c[n]$ are the MFCC, typically 8-13 coefficients are used when used as a feature set for classification \cite{15} with a filter bank consisting of 26-40 filters.

## 4 Classification Method

A method for classification commonly used in binary classification problems is a method named Support Vector Machines, reviewed in literature such as \cite{16}, \cite{23} and \cite{11}, which have its origin in linear discriminant analysis where the idea is to construct linear decision boundaries that separates data into their respectively classes as well as possible with a separating hyperplane. Considering a binary classification problem and a dataset in $\mathbb{R}^n: S = \{(x_i, y_i); i = 1, 2, ..., N\}$ where $x_i$ is the $i$:th data vector (feature vector) and $y_i$ the corresponding response or label (typically \{1,-1\}). The hyperplane is defined by

$$f(x) = x^T \beta + \beta_0 = 0. \tag{4.1}$$

By choosing $\beta$ and $\beta_0$ to lets say $\hat{\beta}$ and $\hat{\beta_0}$ we achieve a hyperplane, defined by, 

$$\hat{f}(x) = x^T \hat{\beta} + \hat{\beta_0} = 0$$

that separates the classes. For a new set of data points we are then able to make a prediction on the corresponding class by evaluating the sign of $\hat{f}(x_{new})$,

$$\hat{G}(x_{new}) = \text{sign}(\hat{f}(x_{new})) = \text{sign}(x_{new}^T \hat{\beta} + \hat{\beta_0}) \tag{4.2}$$

to clarify if $\hat{G}(x_{new}) < 0$ then $x_{new}$ have a class corresponding to -1 and respectively if $\hat{G}(x_{new}) > 0$ the class corresponding to 1. Classifiers such this, that returns a sign, are typically called perceptions in engineering literature \cite{16}. The problem is to choose $\beta$ and $\beta_0$ such that the best hyperplane for the given data can be found, which is the idea of the Support Vector Machine classifier.

### 4.1 Support Vector Machines (SVM)

The standard way of expressing the support vector machine classifier is as the hard margin SVM and by formulating the following optimization problem \cite{23}

$$\min_{\beta, \beta_0} \frac{1}{2} ||\beta||^2 \text{ s.t } y_i(x_i^T \beta + \beta_0) \geq 1 \text{ for } i = 1, 2, ..., N \tag{4.3}$$

which is a convex optimization problem (quadratic criterion with linear constraints). The constrains define an empty space or margin around a linear decision boundary. To illustrate this, think of an example dataset in $\mathbb{R}^2$ and consider Figure \cite{12} hence the objective is to choose $\beta$ and $\beta_0$ to maximize the margin.
This formulation is fine as long as the data for the corresponding classes are linearly separable, which is rarely the case in real world applications. If we now suppose that the classes overlap in the feature space one way to deal with this and still be able to maximize the margin is to introduce slack variables \( \varepsilon = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_N) \), \( \varepsilon_i \geq 0 \), \( \sum_{i=1}^{N} \varepsilon_i \leq K \), where \( K \) is some constant. In this way we modifying the constraint in Equation (4.3) and thus allow for miss classification for some reasonable amount bounded by the constant. We can include miss classification in the objective function and express Equation (4.3) as

\[
\min_{\beta, \beta_0} \frac{1}{2} ||\beta||^2 + C \sum_{i=1}^{N} \varepsilon_i \\
\text{s.t} \quad y_i(x_i^T \beta + \beta_0) \geq 1 - \varepsilon_i \quad \text{for} \quad i = 1, 2...N \\
\varepsilon_i \geq 0 \quad \text{for} \quad i = 1, 2...N,
\]

(4.4)

where the cost parameter \( C \) replaces the bounded constant \( K \).
Figure 13: Illustration of the introduction of slack variables $\varepsilon_i$ which allow for miss classification

Figure 13 illustrates this concept, points on the margin will have $\varepsilon_i = 0$ and points inside (or even pass the separating hyperplane) will have $\varepsilon > 0$

### 4.2 Dual Formulation

By formalizing Equation (4.4) as its dual formulation will result in a simpler convex quadratic optimization problem and can be solved with standard techniques [16], but the main benefit is that it will be possible to express the problem in terms of dot products of vectors in the feature space $x_i^T x_j$, further more this will be the key concepts when, explained further down, specify a non linear transformation of the feature vectors allowing for non linear decision boundaries and take use of the Kernel Trick. A note on Lagrangian duality in correspondence to the SVM can be found in [11]. The Lagrangian primal function is

$$L_p = \frac{1}{2} ||\beta||^2 + C \sum_{i=1}^{N} \varepsilon_i - \sum_{i=1}^{N} \alpha_i \left[ y_i (x_i^T \beta + \beta_0) - (1 - \varepsilon_i) \right] - \sum_{i=1}^{N} \mu_i \varepsilon_i \quad (4.5)$$

where $\alpha_i \geq 0$, $\mu_i \geq 0$ are the Lagrange multipliers and $L_p$ is to be minimized with respect to $\beta, \beta_0$ and $\varepsilon_i$. Setting the derivatives to zero we achieve

$$\beta = \sum_{i=0}^{N} \alpha_i y_i x_i \quad (4.6)$$

$$\sum_{i=1}^{N} \alpha_i y_i = 0$$

$$\alpha_i = C - \mu_i \quad \forall i.$$
Inserting the Equations in (4.6) in Equation (4.5) and simplifying we obtain the following dual form

$$\max_{\alpha} \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j y_i y_j x_i^T x_j$$

s.t. \[ 0 \leq \alpha_i \leq C \quad \forall i \]

$$\sum_{i=1}^{N} \alpha_i y_i = 0.$$  \hspace{1cm} (4.7)

In addition to the Equations in (4.6), the Karush-Kuhn-Tucker conditions include the constraints

$$\alpha_i (y_i (x_i^T \beta + \beta_0) - (1 - \varepsilon_i)) = 0 \hspace{1cm} (4.8)$$

$$\mu_i \varepsilon_i = 0$$

$$(y_i (x_i^T \beta + \beta_0) - (1 - \varepsilon_i)) \geq 0.$$  \hspace{1cm} (4.8)

After solving the dual problem we need to obtain the primal solution to classify new points \textit{i.e.} we need $\beta$ and $\beta_0$, inspecting the first Equation in (4.6) we notice that the solution for $\beta$ is on the form

$$\hat{\beta} = \sum_{i=1}^{N} \hat{\alpha}_i y_i x_i$$  \hspace{1cm} (4.9)

which includes the Lagrange multipliers $\hat{\alpha}_i$ with the corresponding $i$:th observation. These observations are, in literature, denoted as the support vectors or support points. As can be seen in Figure 13 points on the boundary will have $\hat{\varepsilon}_i = 0$ which corresponds to $0 < \hat{\alpha}_i < C$ which follows by the second Equation in (4.8) and the last Equation in (4.6). Hence the first Equation in (4.8) can be used to solve $\hat{\beta}_0$.

The decision function for classifying new inputs is now

$$\hat{f}(x_{\text{new}}) = x_{\text{new}}^T \hat{\beta} + \hat{\beta}_0 = \sum_{i=1}^{N} \hat{\alpha}_i y_i x_{\text{new}}^T x_i + \hat{\beta}_0$$  \hspace{1cm} (4.10)

and by investigating the sign of the above the class can be determined as explained in Equation (4.2).

### 4.3 Kernel Trick

SVM:s are, as mentioned, a linear classifier and to improve classification one would like to relax this condition in a way not only being capable to handle overlapping data sets with the method of introduce slack described above. Rather than using our original features $x$ we may want to learn some features $\phi(x)$ [11].

This process consist of two steps, first the input vectors are transformed, or mapped, to some high dimensional feature vectors by the mapping $\phi$, then the SVM finds the hyperplane that maximizes the margin in the new feature space.
The separating hyperplane will be a linear function in the transformed but a non linear in the original feature space.

Consider the optimization problem in (4.7) and an optional transformation \( \phi(\cdot) \) of the input data

\[
\max_{\alpha} \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j y_i y_j \phi(x_i)^T \phi(x_j)
\]

s.t \( \sum_{i=1}^{N} \alpha_i y_i = 0 \)

and the solution function, Equation (4.10) yields

\[
\hat{f}(x_{\text{new}}) = \sum_{i=1}^{N} \hat{\alpha}_i y_i \phi(x_{\text{new}})^T \phi(x_i) + \hat{\beta}_0
\]

In both Equation (4.11) and (4.12) we see that the objective function and the solution function includes the transformation \( \phi(\cdot) \) in terms of dot products of the transformed feature vectors. Instead of explicitly define the transformation \( \phi(\cdot) \), which can lead to very high dimensional feature vectors in the transformed feature space, one can instead make use of the Kernel trick which returns the dot product directly computed in the original feature space.

Now given the feature mapping \( \phi(\cdot) \) we define the corresponding Kernel to be

\[
K(x_i, x_j) = \phi(x_i)^T \phi(x_j).
\]

Instead of explicitly find or represent \( \phi(x_i) \) and \( \phi(x_j) \) we can directly define the Kernel in terms of \( x_i \) and \( x_j \) by some Kernel function.

Popular Kernels to use are in SVM literature [16],

- \( K(x_i, x_j) = (1 + x_i^T x_j)^p \), \( p \)-Degree polynomial Kernel,
- \( K(x_i, x_j) = e^{-\frac{|x_i - x_j|^2}{2\sigma^2}} \), Radial basis or Gaussian Kernel,
- \( K(x_i, x_j) = \tanh(\kappa_1 x_i^T x_j + \kappa_2) \), Sigmoid Kernel.

with corresponding parameters \( p, \sigma, \kappa_1 \) and \( \kappa_2 \) to choose for the model. In general the Radial basis kernel is a reasonable first choice since it has fewer numerical difficulties and the Sigmoid Kernel is not valid under some parameters \( \kappa_1 \) and \( \kappa_2 \) [6].
5 Radar Signal Processing

When the radar system send out electromagnetic signals $s_t(t)$ in the free space and receive echoes $s_r(t)$ from the surrounding environment. The received echoes are transformed with a quadrature detector to an I/Q signal $s_D$ which is represented in the free space as $x[n,m]$ where $n$ represents the shift in bearing as the radar revolves and $m$ represents a range. Much of the returned signals is from objects of non interest such as buildings, trees and even the ground itself. Echoes from these kinds of objects are often called ground clutter or clutter. Due to the presence of clutter, further signal processing is needed on the raw radar video before extracting the radar echoes (I/Q data) from targets, since clutter echoes can be many orders of magnitude larger than the target itself. This can be seen in Figure 14 where the instantaneous power of $x[n,m]$ is plotted in decibel

$$P_{\text{inst}} = 20 \log_{10}|x[n,m]|.$$ (5.1)

The plot it hardly dominated by clutter and it is impossible to resolve echoes from targets. To be able to resolve targets and extract I/Q data from a target two signals processing methods are used, MTI-filtering and pulse compression. Echoes from targets can then be extracted as complex time series, by taking the Fourier Transform of these signals the Doppler spectrum are achieved.

![Figure 14: Raw radar video](image-url)
5.1 MTI-filter

The Moving Target Indicator (MTI) filter is used to suppress echoes from clutter, a property of clutter is that it is stationary or close to stationary, i.e., the Doppler frequencies induced by echoes from clutter is zero or close to zero. The MTI filter is a high pass filter that filters out the low Doppler frequencies, i.e., the objects with low velocities. The MTI filters used are FIR filters of order $N$ (number of filter coefficients) explained in section 3.1.1 and for a set of impulse response coefficients $b_i$ and the signal $x[n, m]$ the filtered signal $x_{mti}[n, m]$ can be expressed as the following convolution sum

$$x_{mti}[n, m] = \sum_{i=0}^{N} b_i x[n - i, m], \text{ for all } m,$$  \hspace{1cm} (5.2)

A graphical representation of the instantaneous power of the MTI filtered signal earlier seen in Figure 14 as its raw form can be seen in Figure 15.

![Figure 15: MTI filtered radar video](image)

As can be seen in Figure 15 most of the clutter is removed and it is possible to distinguish echoes returned from targets.
5.2 Pulse Compression

As can be seen in Figure 15 a target is clearly visible at range 3000 meters and bearing 214°. However the target is smeared out over a large range interval, i.e., the range resolution is poor.

The range resolution is improved by shorten the pulse width, Equation (2.7), and the maximum range detection is improved by increasing the energy of the pulse, Equation (2.2). Where make the pulse longer is the easiest way to increase the energy of the pulse. Hence in pulsed radar systems, the characteristics of the transmitted pulse is compromise between good range resolution and the maximum range detection [3]. To handle this trade off a method called pulse compression is used. In pulse compression techniques the received signal is processed using a so called matched filter (see section 3.1.2). Hence pulse compression in radar systems is a practical implementation of a matched-filter system [10]. The output $x_{pc}[n, m]$ of the matched filter is given by the convolution between the MTI filtered signal $x_{mti}[n, m]$ and the impulse response $h$ of the matched filter

$$x_{pc}[n, m] = \sum_{k=-\infty}^{\infty} h[m - k]x_{mti}[n, k], \text{ for all } n. \quad (5.3)$$

Where the impulse response of the matched filter is derived from the transmitted signal. The result of the pulse compression is presented in Figure 16, where the instantaneous power of the signals $x_{pc}$ is plotted. The range resolution is significantly improved compared too the MTI filtered signal $x_{mti}$, also targets that where not clearly visible in Figure 15 are visible after the pulse compression.

![Figure 16: MTI filtered and pulse compressed radar video](image-url)
5.3 Signal Extraction

To be able to point out targets in the I/Q data a approximatively synchronization was made, between the I/Q data and a data set containing the targets spatial positions at certain times. A signal from the target \( x_{I/Q} \) can be extracted by taking an interval in bearing around the targets position. Hence if the target is in range bin \( m_p \) and bearing bin \( n_p \), are the extracted signal \( x_{I/Q} \) given by

\[
x_{I/Q}[n] = x_{pc}[m_p, n], \text{ for } n_p - I \leq n \leq n_p + I,
\]

(5.4)

where \( 2I \) is the bearing interval, in this case \( 2I = 3^\circ \). This method is illustrated in Figure 17 for an UAV. In the top of Figure 17 are the position of the target marked with a cross and the end points of the extracted interval is marked with the circles. The signal extracted from this interval are presented in the bottom of Figure 17.

![Figure 17: Top: Radar video and the position of the UAV. Bottom: Extracted I/Q data](image-url)
5.4 Spectral Modifications

Due to the fact that different TX-modes (different PRF) are used to neglect range unambiguity explained in Section 2.6, the signals $x_{IQ}$ is up-sampled to a new sampling frequency $F_s = 20000$ Hz, leading the new signals $x_{20k}$. This ensures that all signals, $x_{20k}$, is sampled with the same sample frequency. The up-sampled version of the signal $x_{IQ}$ form Figure 17 is illustrated in Figure 18.

![Figure 18: Up sampled I and Q with their corresponding originals](image)

The Doppler spectra of the signal is obtained by taking the Fourier transform of the up-sampled signal $x_{20k}$,

$$X_{20k} = \mathcal{F}\{x_{20k}\}. \quad (5.5)$$

Where the Doppler frequencies explained in section 2.8.1 are related to the targets velocity. The Doppler frequency related to the targets bulk motion is filtered out, since the investigation is mainly focused on the $\mu$-Doppler effect discussed in section 2.8.2. A function in the radar system, not explicitly explained here, called a tracker function contains information about the velocity of the object.
The targets velocity can be represented as a Doppler frequency, by Equation (2.14).

The filtering is made by shifting the spectra using the frequency shift property of the DFT described in Section 3.2.1. The Doppler frequency that corresponds to the current velocity of the object is shifted to DC (frequency = 0), by

\[ X_{\text{shift}}[k] = X_{20k}[k - l], \quad (5.6) \]

where \( l \) is the frequency shift. The spectral content at the DC frequency in the shifted spectra is filtered out with a FIR filter, described in Section 3.1.1. Giving the spectra of the shifted and filtered signal \( X_{\text{shift,FIR}}[k] \). The spectra is shifted back with the same frequency shift property of the DFT, creating the filtered signal,

\[ X_{\text{FIR}}[k] = X_{\text{shift,FIR}}[k + l], \quad (5.7) \]

where \( l \) is the same frequency shift. The process is graphically illustrated in Figure 19 below.

Figure 19: Top: Original spectrum and velocity corresponding to the target. Middle: Shifted spectrum and filtered spectrum with a FIR filter. Bottom: Spectrums shifted to the original position.
The signal that generates the Doppler spectrum is complex and to be able to extract a real signal from the Doppler spectrum the spectrum need to be symmetric. In this thesis the method to achieve a real signal that have some correlation in a meaningful way is to use shift the whole spectrum by the frequency shift property of the DFT and then mirror the spectrum, denoted $X_{mirror}$, illustrated in Figure 20.

\[ x_{real} = F^{-1}\{X_{mirror}\} \]  

(5.8)

It should also be noted that depending on the sign on the velocity (moving from or towards the radar) i.e., the sign on the Doppler frequency corresponding to the velocity the spectrum is shifted to right or left to not let the direction of movement induce characteristics on the two objects.
6 Data Sets and Implementation

The signals, $x_{\text{real}}$, resulting from section 5 are transformed to signals, $x_{\text{audio}}$, that are audible and could be listened to by an operator. In complement to the human perception a machine driven data classification is needed to support the operator when not being able to distinguish these signals for various reasons. To be able to do this, characteristics for the signals is needed commonly known as features. In this thesis a feature set commonly used in speech processing is used, namely Mel Frequency Cepstral Coefficients (see Section 3.3). This feature set will be used as input for a feature driven classification method who then can be used to predict and classify new input data. The classification method used is called Support Vector Machine (SVM) who have its origin in Linear Discriminant Analysis (see Section 4). The idea of the SVM method formulating the discrimination problem as a optimization problem where the aim is to find the best separating discriminator in the feature space. Moreover, a dual formulation of the optimization problem will provide more flexibility of the SVM method since it allow us to use the Kernel Trick and thus achieve the non linear SVM method. The classification method training phase includes a set of parameters depending on how you model with respect to the data, these parameters are chosen with a method called cross validation in the implementation, finally the classifier is tested on the ability to classify input data.

6.1 Data Set

The dataset originated from the signal processed radar echoes of UAV:s and birds as the results described in section 5 is transformed to signals that would be audible, and thus possible for an radar operator to listen to, explicitly described in the thesis of Johan Clemedson [2], with the idea that by human perception a radar operator would be able to classify these sounds. In short the signals $x_{\text{real}}$ are extrapolated with an AR-filter with Burg’s method. This is done, without going in on details, by estimating so called AR parameters from the $x_{\text{real}}$ signals. For each signal an AR-filter of order 40 is estimated with Burg’s method. Each signal $x_{\text{real}}$ is then extrapolated using the AR-filter and presented as an audio output $x_{\text{audio}}$. The data set consist of 181 audio signals related to radar echoes from a UAV and 232 audio signals related to birds. This data set consisting of 413 audio signals in total will will serve as basis for the development of a feature driven classifier.

6.2 Implementation

MFCC Features from the signals, $x_{\text{audio}}$, where extracted as described in section 3.3.1 to represent the characteristics of each class, with 40 filters and the first 13 MFCC each signal where represented as a 13-dimensional vector $c_i$ for the $i$:th signal. The implementation of a SVM classifier model, defined by the optimization problem
\[
\max_{\alpha} \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j y_i y_j \phi(c_i)^T \phi(c_j)
\]
\[
s.t \quad 0 \leq \alpha_i \leq C \quad \forall i
\]
\[
\sum_{i=1}^{N} \alpha_i y_i = 0,
\]
is made in MATLAB and in particular with the Statistics and Machine Learning toolbox which allow modeling with the linear kernel (no transformation \(\phi\) of the feature vector) and the radial basis kernel which corresponds to a non linear transformation \(\phi\) in the sense
\[
\phi(c_i)^T \phi(c_j) = e^{-\frac{|c_i - c_j|^2}{2\sigma^2}}.
\]  
(6.2)

The analysis is performed on a so called linear model (linear kernel), and a transformation related to the radial basis kernel with the corresponding parameters \(C\) and \((C,\sigma)\) respectively to be tuned for both models. To train, tune and test the models the following approach is used.

- **Split data into a training and testing set**
  - The total set of available signals containing of 413 signals, hence 413 MFCC feature vectors. 30 signals of each class is used for validation of the models and the remaining 353 to build the model.

- **Consider the linear and radial basis model**
  - The linear model consist of the parameter \(C\) to be determined. The radial basis model have, beyond \(C\), a parameter from the kernel function, hence \(\sigma\) also needs to be determined.

- **Use Cross-validation to find best parameters for the models**
  - By splitting the training data in folds (subsets) to measure the prediction error and sequentially swap these folds in terms of what folds are used for training and testing and then take the average of these sequentially measured prediction error a more generalized parameter selection are achieved that takes the whole data set in account *i.e.* not just one subset are used for testing.

- **Use the best parameters to train the models**
  - The best parameters (who gives lowest prediction error) are used to train the models.

- **Test the models**
  - The models are tested with the 60 signals that are not used in the training phase and hence the 60 signals can be seen as new observations and serves as a final validation on how good the models are to predict new data.
6.3 Training and Parameter Evaluation

Two classification models were trained, the SVM without any transformation of the features (linear) and the SVM with the radial basis kernel. The first models only consist of the cost parameter for miss classification $C$ and the radial basis model also have the scale parameter $\sigma$.

The performance of the of the models can be represented by a confusion matrix which contains information of how well the model performed on data you feed to the model, known as the test data. An illustration of such a representation can be seen in Table 1.

<table>
<thead>
<tr>
<th>Input</th>
<th>Predicted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class 1</td>
<td>$P_{11}$ $P_{12}$</td>
</tr>
<tr>
<td>Class 2</td>
<td>$P_{21}$ $P_{22}$</td>
</tr>
</tbody>
</table>

- $P_{ii}$: Number of occurrences of Class $i$ that is predicted as Class $i$
- $P_{ij}$: Number of occurrences of Class $i$ that is predicted as Class $j$, $i \neq j$

Table 1: Illustration of a confusion matrix for the two class problem

By examine the confusion matrix one will get an idea of how good the model predicted on the test data. The diagonal values in the confusion matrix contains "overall" information of how good the model is to predict new data, but useful information can also be gathered to examine off diagonal values, for example to find out if a model predicting a specific class better than other classes.

To adjust model parameters the prediction error of the model is used, which simply are the number of wrongly classified observations in the test set divided by the total number of observations in the test set. The approach is simply to test parameters $C$ and $(C, \sigma)$ for the respective models and then use those parameters when evaluating future performance of the classification models. In relation to Table 1 the prediction error is defined as

$$\text{prediction error} = \frac{P_{12} + P_{21}}{N},$$

where $N$ is the size of the test set.

To improve parameter selection a cross validation approach is used. This means that you instead of test and evaluate the prediction error your model with a single test set the test data is partitioned in $M$ subsets or folds. Moreover the training part now consist of $M - 1$ of these folds and 1 fold used for evaluation of the prediction error, now one iteration or trail has been completed. The next step is to use $M - 1$ new folds for the training part and 1 new fold for evaluation. After completed $M$ trails in this process the average of the prediction error is used to chose model parameters in the sense that parameters that generate a low average prediction error will be result in the model that best fits the data.
with the "whole" data set in mind. An illustration of the cross validation scheme where \( M = 5 \) can be seen in Figure 21 below.

<table>
<thead>
<tr>
<th>data set</th>
<th>class 1 (UAV)</th>
<th>class 2 (Bird)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>fold 1</td>
<td>fold 2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>trail</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Average</th>
<th>Overall error</th>
</tr>
</thead>
<tbody>
<tr>
<td>trail 1</td>
<td>error 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>trail 2</td>
<td>error 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>trail 3</td>
<td>error 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>trail 4</td>
<td>error 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>trail 5</td>
<td>error 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 21: Illustration of cross validation scheme
7 Results

The linear model where tested for the parameter $C$ in the interval $[1, 100]$ consisting of 10 logarithmically equal spaced values for $C$. The cross validated models overall test error, explained in Section 6.3 where the training data set where split in 5 folds and the prediction error was measured by Equation (6.3) and hence the overall test error is the average of these 5 prediction errors, was evaluated for different values of $C$. The resulting test error is illustrated in Figure 22.

As can be seen the model is not particular sensitive for the parameter $C$, since the test error do not change significantly while varying the parameter $C$ which is more clear when inspecting the lower plot on Figure 22. However the best parameter value, who gives the lowest test error $C$ corresponds to $\log(C) = 0.4444$.

The same approach where implemented for the radial basis (RB) model with the parameter pair $(C, \sigma)$, also here the test values for the parameter $\sigma$ in the interval $[1, 100]$ consisting of 10 logarithmically equal spaced test parameters, illustrated in Figure 23.
As can be seen the test error does not change significantly for a large range of parameter values. To further compare the two models, for a set of $\sigma$ values the test error depending on $C$ is plotted for the radial basis model together with the linear model in Figure 24.
Figure 24: comparison of the linear model and the RB model for a set of values of the parameter $\sigma$

For some values on $\sigma$ the radial basis model perform better than the linear model. The final parameters for the two models is presented in Table 2 with corresponding test error.

<table>
<thead>
<tr>
<th></th>
<th>$\log(C)$</th>
<th>$\log(\sigma)$</th>
<th>error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>0.4444</td>
<td>−</td>
<td>0.0567</td>
</tr>
<tr>
<td>Radial Basis</td>
<td>0</td>
<td>0.4444</td>
<td>0.0255</td>
</tr>
</tbody>
</table>

Table 2: Choice of parameters for the two models

To further validate the two models with corresponding parameters are tested with a data set that have not been included in the training phase consisting of 30 observations of a UAV and 30 observations of birds and compared in terms of confusion matrices. The results from the test can be seen in Table 3 and 4 below.
Table 3: Test performance linear model

<table>
<thead>
<tr>
<th>Input</th>
<th>Predicted</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>UAV</td>
<td>28</td>
<td>2</td>
</tr>
<tr>
<td>BIRD</td>
<td>2</td>
<td>28</td>
</tr>
</tbody>
</table>

Table 4: Test performance RB model

<table>
<thead>
<tr>
<th>Input</th>
<th>Predicted</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>UAV</td>
<td>30</td>
<td>0</td>
</tr>
<tr>
<td>BIRD</td>
<td>1</td>
<td>29</td>
</tr>
</tbody>
</table>

In general the radial basis model perform better than the linear model, both in terms of test error in the training phase where the radial basis model achieved a test error of 2.55% compared to the linear model with a test error of 5.67%. Also in when the models were compared with an independent test set the radial basis performed better with 100.00% correct prediction on the UAV and 96.67% on birds compared to the linear model with 93.33% correct prediction on the UAV and 93.33% on birds. Overall the final models predicted correct on 98.33% and 93.33% with the radial basis model and the linear model respectively.


8 Discussion and Future Work

In this project a method for extracting radar returns from raw radar data have been proposed in combination with an aurial output, explicitly explained in the thesis of Johan Clemedson [2], a feature driven classification method with MFCC as a descriptor for these signals and a Support Vector Machine as classification method have also been implemented. There have been many moving pars through out this project, data processing, signal processing, machine learning and by this, a small change in the beginning of the process can change a lot further down the road which have made this project to an iterative and changing process.

The project was heavily dependent on the data available and since most of the data was poor and even bad as illustrated in in Figure 25 and 26. A reason for this may be that the field test when gartering data, many radar systems where used at the same time and hence this could be interferences due to this cause.

Figure 25: Illustration 1 of interference in the available data

Figure 26: Illustration 2 of interference in the available data
It was hard to find good methods to come up with a purposed solution and many ideas have been purposed but later discarded. Indeed, the available data set was small but methods that works for this data set was developed. In future work more data have to be gathered to validate this methods before a real time implementation can be done.

The data from the UAV consist of only one type of UAV, more data is needed to investigate how these methods will perform when there are different types of UAV:s to discriminate between. One possibility is that this kind of UAV examined in the project implies such discrimination in the radar returns that it is possible to use these methods to distinguish between this UAV and birds but will fail when include others kinds of UAV:s in the analysis.

The feature set used in this thesis has its origin in speech processing, different ways of implementing this can be used but the purposed method in this thesis seems to work good for this kind of problem setup. In further analysis one should investigate other feature set that have the possibility to serve as input to a classification method and provide perhaps even better results. Further more on this theme, one may want to use another kind of classification method more adapted to aural classification for example Hidden Markov Models that are used in speech processing.

The grid search method for finding good parameters for respective SVM model are time consuming, but in this case when having a small amount of data it is affordable. However if more data is gathered it is not sure that the parameters used in this thesis will be the best for the new data set and it is recommended to re do the grid search or use another method to determine the parameters. One could also use a finer grid, or even a completely different spacing, than the grid used to see if it is possible to achieve even lower classification error, however this is not done in this thesis due to the fact that the results are satisfactory.

To summarize, this project can be seen to serve as a pilot for future work in a new way in the area of radar classification, ideas, methods and thoughts on how to attack a problem like this.
A Appendix A

A.1 Resampling

The signals are sampled with the PRF for the specific illumination. Since the radar transmits different PRF’s for different revolutions, the signals is sampled with different sampling frequencies. For further analysis the signals have to be sampled with the same sample frequency and of the same length. Hence the signals have to be resampled with some factor. It is not guaranteed that the resampling factor is an integer. Hence the resample method has to be able to resample the signals with an non-integer factor. This is done by first up-sample the signal with an integer $K$ and then down-sampled with an integer $L$, resulting in an resample factor of $K/L$.

The methods presented in this section is based on the method described in [13]. Up-sampling by integer factors are described in the first part of this section. In part two and three are a type of filter described, which is used in the up-sampling method. The Down-sample part is presented in the fourth part of this section. The last part of this section summarize the method for resample signals with an non-integer factor.

A.1.1 Up-Sampling by integer factor

The up-sampling algorithm used is a time domain method that may be implemented in real-time applications. For a function $f(t)$ sampled at intervals $\Delta T$, will generate a sequence of samples ${f[n]} = \{f(n\Delta T)\}$. By inserting $K - 1$ zeros between each sample in ${f[n]}$ a new sequence ${\hat{f}[k]}$ is generated. The extended sequence is sampled at intervals $\Delta T/K$, hence the new sequence $\{\hat{f}[k]\}$ is $K$ times longer than $\{f[n]\}$. The method of inserting zeros between each sample, is illustrated with an example in Figure 27 below. The example signal in Figure 27 (a) is of length $N = 16$ and is up-sampled with a factor $K = 3$, hence in Figure 27 (b) two zeros are inserted between every sample of the original signal.
Figure 27: (a) A signal with length $N = 16$, and (b) the same signal with two zero samples interpolated between each of the original samples.

To examine the effect of inserting $K - 1$ samples of zeros, the DFT of the sequences are analyzed. The DFT for the original data sequence is given by

$$F[m] = \frac{1}{N} \sum_{n=0}^{N-1} f[n] e^{-i\frac{2\pi mn}{N}}, \quad m = 0, \ldots, N - 1. \quad (A.1)$$

For the sequence $\{\hat{f}[k]\}$ the DFT is given by

$$\hat{F}[m] = \frac{1}{KN} \sum_{n=0}^{KN-1} \hat{f}[n] e^{-i\frac{2\pi mn}{KN}}, \quad m = 0, \ldots, KN - 1. \quad (A.2)$$

But since $\{\hat{F}[m]\}$ only has $N$ non zero elements and $\hat{f}[Kk] = f[k]$, the DFT can be rewritten as

$$\hat{F}[m] = \frac{1}{KN} \sum_{k=0}^{N-1} \hat{f}[Kk] e^{-i\frac{2\pi nk}{N}}, \quad m = 0, \ldots, KN - 1. \quad (A.3)$$

One can note that $\{F[m]\}$ is periodic with period $N$ and $\{\hat{F}[m]\}$ is periodic with period $KN$. Which gives that $\hat{F}[m] = F[m]/K$, for $m = 0; \cdots, N - 1$ and $\{\hat{F}[m]\}$ will contain $K$ repetitions of $\{F[m]\}$. A illustration of the DFT of the example signal from Figure 27 (a) is shown in Figure 28 (a). The DFT of the interpolated signal is scaled with the factor $K$ and plotted in Figure 28 (b) (Note in this example $K = 3$). It can be seen that the DFT of the original signal is repeated three ($K$) times in the DFT of the interpolated signal.
By inserting $K - 1$ zeros has resulted in a signal with sampling interval $\Delta T/K$ and a Nyquist frequency of $K/(2\Delta T)$. The frequency resolution is unchanged and the original DFT is replicated $K$ times within the frequency span of $K/\Delta T$. The interpolated signal can be reconstructed by elimination of the replications of the original spectral components. This is done by applying a low-pass filter to the sequence $\{\hat{F}[m]\}$ to preserve only the base band part of the spectrum. A representation of the two sided spectrum of $\hat{f}[m]$ are presented in Figure 29. Which illustrates the replications of $F[m]$ and the low-pass filtering process.

Figure 28: DFT of the signal shown in 27 (a), which is periodic with period 16. DFT of the interpolated signal shown in 27 (b), which is periodic with period 48.

Figure 29: Illustration of low-pass filtering used to eliminate spectral replications introduced by the interpolation of the signal.
The algorithm used to upsample a signal with an integer factor $K$ can be summarized in the block diagram, displayed in Figure 30.

Figure 30: Block diagram of the algorithm for upsample a sequence with an integer factor $K$.

The interpolation filter used as the low-pass filter in the upsample algorithm is described in the next section.

### A.1.2 Interpolation Filter

The ideal filter for the interpolated sequence, is a low pass filter with cut off frequency at the Nyquist frequency of the original signal $1/(2\Delta f)$. The impulse response of such a filter is given by

$$ h[n] = \frac{\sin(\pi n/K)}{\pi n/K} $$

But this kind of filters are acausal, hence a [FIR](https://en.wikipedia.org/wiki/Finite_impulse_response) approximation is made. This is done by truncating $h[n]$ to a finite length. If one sets the finite length to $N = 2KM + 1$ and delaying it with $KN$ samples the FIR approximation is obtained as

$$ h[n] = \frac{\sin(\pi (n - KM)/K)}{\pi (n - KM)/K}, \quad n = 0,\ldots,N - 1. $$

The resampled signal is also windowed by a Kaiser window in the filtering process. The Kaiser window is explained in more detail in Section A.1.6.

The FIR interpolation filter for a up-sample factor $K$ is illustrated in Figure 31 below.

Figure 31: Illustration FIR interpolation filter.

One can note that the values of $\hat{f}[n]$ is zero for a lot of the filter coefficients. This fact can be used to implement the filtering in an efficient way. The more ef-
ficient implementation of the interpolation filter is called polyphase interpolation filter and are described in the next section.

### A.1.3 Polyphase Interpolation filter

To improve the computational speed of the algorithm a polyphase filter structure is used. In the filter computation above one can see that many of the up-sampled data values are zeros at each filter step. It is clear that the zero values does not contribute to the filter output. Therefore only a subset of the filter coefficients are required at each of the computational steps. This leads to the polyphase filter structure, where the filter is split into sub sets of filter which are applied to the signal. The polyphase filter structure can be summarized in the following steps:

1. Based on the up-sampling factor $K$, design a $M-1$ order FIR interpolation filter.

2. From the filter coefficients $b_m$, $m = 0, \ldots, M-1$, form $K$ sub filters, in the following way

   Sub filter 0: $b_0 \ b_K \ b_{2K} \ b_{3K} \ \cdots$
   Sub filter 1: $b_1 \ b_{K+1} \ b_{2K+1} \ b_{3K+1} \ \cdots$
   Sub filter 2: $b_2 \ b_{K+2} \ b_{2K+2} \ b_{3K+2} \ \cdots$
   \vdots
   Sub filter $K-1$: $b_{K-1} \ b_{2K-1} \ b_{3K-1} \ b_{4K-1} \ \cdots$

   Each sub filter is a FIR filter with the coefficients given above, the coefficients $b_i$ that is not included are equal to zero.

3. At each major time step the interpolation data stream may be created by passing the data sequence $\{f[n]\}$ through each of the sub filters, so that the interpolated value $\hat{f}[Kn+i]$ is the output of sub filter $i$.

The polyphase interpolation filter can be represented with Figure 32.

![Figure 32: Illustration poly interpolation filter.](image-url)
When using the polyphase structure it is not necessary to preform the up-sampling step of inserting zeros in to the data sequence.

A.1.4 Down-sampling (or Decimation) by an Integer Factor

By down-sample a signal one want to increase the sampling interval $\Delta T$ by a factor $L$ (or decreasing the samplings frequency with $L$). This is done by retain every $L$ sample of the original sequence $\{f[n]\}$ by letting $\hat{f}[n] = f[Ln]$, where $\{\hat{f}[n]\}$ is the down-sampled sequence. However caution must be taken to prevent aliasing effects in the down-sampled sequence. One can not directly down-sample a sequence unless it is known that the spectrum of the data sequence is equal to zero for all frequencies at and above the new Nyquist frequency defined by the new sampling frequency. To prevent aliasing effects in the resampled data, are the original sequence passed through a digital low-pass filter before down-sampled.

A.1.5 Resampled with non-integer factor

Assume that one want to resample a sequence to non-integer factor $P$ and assume that the factor $P$ can be expressed as a rational factor

$$P = \frac{K}{L},$$  \hspace{1cm} (A.6)

where $K$ and $L$ are integers. Then the resampling can be achieved by up-sample the sequence by a factor $K$ followed by a down-sample of the sequence by a factor $L$. Since the low pass filtering parts of the up- and down sapling are cascaded, see Figure 33a. The two filters may be replaced by a single filter, see Figure 33b.

Figure 33: Illustration of filtering process for the resample algorithm.

For resampling factors $P$ that are irrational one first has to approximate the factor $P$ with a rational number.

A.1.6 Kaiser window

A Kaiser window is given by the following equation

$$w[n] = \frac{I_0 \left( \beta \sqrt{1 - \left(\frac{n-N/2}{N/2}\right)^2} \right)}{I_0}, \hspace{1cm} n = 0; \ldots, N-1.$$  \hspace{1cm} (A.7)
Where $\beta$ is the Kaiser window parameter and $I_0$ is the zeroth-order modified Bessel function of the first kind, given by

$$I_0(z) = \sum_{k=0}^{\infty} \frac{\left(\frac{z^2}{4}\right)^k}{k!\Gamma(k + 1)}.$$  \hspace{1cm} (A.8)

Where $\Gamma$ is the Gamma function, $\Gamma(n) = (n - 1)!$. 

53
References


