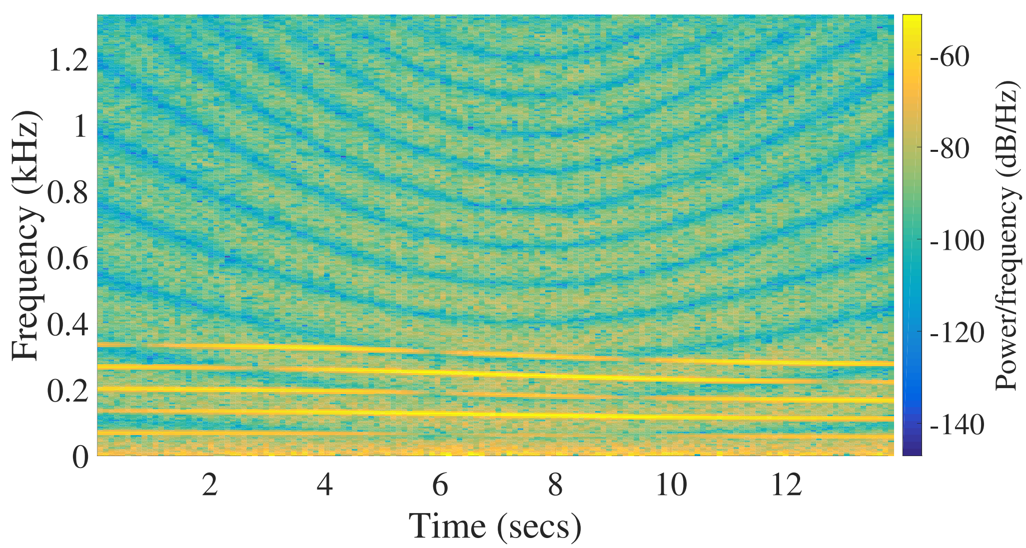
Optimizing Features for the Classification of Aircraft Noise

May 2016

Anders Nikolai Christensen

Master's Thesis



Aalborg University
Sound & Music Computing
Rendsburggade 14
DK-9000 Aalborg

Architecture Design & Media Technology
Rendsburggade 14
DK-9000 Aalborg
http://smc.aau.dk

AALBORG UNIVERSITY
STUDENT REPORT

Title:
Optimizing Features for the Classification of Aircraft Noise

Theme:
Master's Thesis

Project Period:
Spring Semester 2016

Project Group:
SMC161031

Authors:

Anders Christensen

Supervisors:
Mads Græsbøll Christensen (AAU)
Karim Haddad (B&K)
Woo-Keun Song (B&K)

Number of Prints:
3

Number of Pages:
83

Date for Submission:
May 25, 2016

Abstract:

This project deals with methods for extracting features in signals, which then can be used in a machine learning approach for recognizing an aircraft in a signal. Noise is often unwanted and needed to be hold at a minimum, especially in areas close to where people are living and staying, for example around an airport. In that case, a maximum accepted sound pressure level will be set and the issue is now to automatically recognize when an aircraft sound is present in a recorded signal. A signal having an aircraft present will often have temporal variations, which we believe, when utilized in the analysis, improves the detection of an aircraft in a signal. We have investigated different feature extraction methods which utilizes temporal variation. We have used the Cyclic Analysis for obtaining information about (hidden) periodicities, in 1 s segments, which returns a Cyclic Spectral Coherence. We have also used a Harmonic Chirp Model (HCM) chirp pitch estimator for obtaining the instantaneous frequencies from an *event* of an aircraft in a signal and used it in a flight parameter estimator for estimating flight parameters from a recording when an aircraft is passing by, by using parameters from its trajectory. The Cyclic Analysis obtained same classification error rates with a parametric classifier as the Mel-Frequency Cepstral Coefficients, used as baseline, did with a non-parametric classifier. We have shown the advantages of using the HCM chirp pitch estimator, we have tested the flight parameter estimator on synthetic signal and we have obtained great results of detecting aircraft within an event when an aircraft is passing by. All in all, we believe we have shown the benefits of using temporal variations in feature extraction from a signal when the purpose is to detect an aircraft within a signal.

Preface

This master thesis is written by Anders Nikolai Christensen, student at the master program Sound and Music Computing at Aalborg University (AAU) within Architecture Design & Media Technology at AAU. The work has been carried out in corporation with Brüel & Kjær (B&K) Denmark. B&K has a project running on classification of noise from different sources. This work focus on classifying aircraft noise within a noisy environment. B&K has provided a database on which the classification is wanted to be performed. The central aspect of this work was to investigate optimal features for classification of aircraft noise, where B&K also has provided a baseline for being able to draw conclusions.

I would like to thank my two supervisors Karim Haddad and Woo-Keun Song from B&K for providing me with support, guidance and feedback, which indeed has been needed throughout this work. At my working days at B&K in Nærum both the supervisors always had the time to discuss my work and I had a great feeling of being respected and treated in a professional manner and always with a smile, which for me is crucial for making a great working environment where I can perform maximum.

I would also like to thank my supervisor professor Mads Græsbøll Christensen, who has been giving me helpful guidance and feedback throughout my thesis work. It has been very interesting to take the step from studying the research professor Mads Grøsbøll Christensen has provided me as one of his students at the master program Sound & Music Computing, and use these state of the art methods for solving engineering problems.

Reading Guide

The report is divided into following chapters:

- Introduction
- Physics
- Theory
- Classification Experiments
- Conclusion

In the Introduction chapter a small introduction to the problem is given followed by a section containing related work, from which the proposed improvements are introduced, which lead to the problem formulation. In the following Physics chapter the signal of interest is introduced, from which physics can be derived. A trajectory model forms the basis of an later on proposed estimator, which uses the Doppler model derived from same trajectory model. The Theory chapter describes all the estimation and analysis theory used for answering the problem formulation. Small tests are performed where it is meaningful and lastly the used classification theory is introduced. In the Classification Experiments chapter, the theory from the previous chapter is used for classifying aircraft within a specific database. The classification is done on 1 seconds blocks for the baseline and proposal for improvements, but also classification on events has been tested since this gives new possibilities for estimating unique features in a signal where an aircraft is present. A directly comparison and discussion ends this chapter. Lastly, the Conclusion chapter provides the conclusions which can be tracked throughout the report and lastly, proposal for future work is given.

All implementation and tests in the thesis has been carried out in MATLAB. Since this work uses a database and some implementations which are covered in a Non-Disclosure Agreement (NDA) with B&K, it has not been possible to share all work carried out in MATLAB.

Appendices

Appendices are found after the main report and on the attached CD. The appendices contains the A/V product, report and source code from MATLAB not enclosed by the NDA. A complete list of the CD content can be found here Appendix D.

All figures, tables and equations are referred to by the number of the chapter they are used in, followed by a number indicating the number of figure, table or equation in the specific chapter. Hence, each figure has a unique number, which is also printed at the bottom of the figure along with a caption. An example is Figure 2.1, which means the first figure in chapter 2. The same applies to tables and equations, the latter of which have no captions. Appendices are referred to by capital letters instead of chapter numbers.

Bibliography

At the end of the main report, a Bibliography is listed which contains all sources of information used in the report. In the Bibliography books are indicated with author, title, publisher and year. Web pages are indicated with author, title and link. Articles are indicated with author, title, publisher and year. All information sources are referred to by the number which they feature in the list. This will look like this: [number].

Acronyms

AAU Aalborg University.
aNLS approximately Non-linear Least Squares.

B&K Brüel & Kjær.

CPA Closest Point of Approach.
CSD Cyclic Spectral Density.

DFT Discrete Fourier Transform.

FFT Fast Fourier Transform.

HCM Harmonic Chirp Model.
HM Harmonic Model.
HOG Histogram Of Gradients.

IF Instantaneous Frequency.
IP Instantaneous Phase.

KNNC K Nearest Neighbour Classifier.

MAP Maximum A Posteriori.
MFFC Mel Frequency Cepstral Coefficients.
ML Maximum Likelihood.
MMSE Minimum Mean Square Error.

NFFT N length of Fast Fourier Transform.
NLS Non-linear Least Squares.

PCA Principal Component Analysis.
PSD Power Spectral Density.

QDC Quadratic Distance Classifier.

RMSE Root Mean Square Error.

SCD Correlation Density.
SNR Signal-to-Noise-Ratio.

Nomenclature

$(\cdot)^H$	Hermitian transposition
$*$	Convolution
α	Cyclic Frequency
$\boldsymbol{\theta}$	Parameters of the source
$\circ$	Hadamard Product
$\det(\cdot)$	Determinant
$\hat{(\cdot)}$	Estimate
$\ln$	Natural logarithm
$\mathbf{X}$	Matrix
$\mathbf{x}$	Vector
$\mathbf{Z}$	Vandermonde Matrix of the source
ω_0	Fundamental radian frequency of the source
ϕ_l	Phase of lth harmonic of the source
ψ_l	Frequency of the lth harmonic of the source
τ	Emission time
τ_c	Emission time for when aircraft is at Closest Point of Approach (CPA)
τ_d	Emission time for the direct path
τ_r	Emission time for the reflected path
θ	Phase
A_l	Amplitude of lth harmonic of the source
a_l	Complex amplitude of lth harmonic of the source
b	Chirp coefficient
c	Speed of sound

d_c	The ground distance (horizontal range) when aircraft is at CPA.
$e(n)$	Broadband Radiated Noise signal
$E[\cdot]$	Expectation operator
f_0	Fundamental frequency of the source
h_m	Height of the microphone ($\cdot_m$ for microphone)
h_t	Height of the aircraft, ($\cdot_t$ for transmitter)
L	Number of harmonics in source
M	Number of samples in segment
N	Number of samples
n	Time index
$p(\cdot;\cdot)$	Log-likelihood function
R	The path from source to origin. Also called slant range.
R_c	The slant range when aircraft is at CPA.
R_d	The direct distance from aircraft to microphone
R_r	The ground-reflected distance from aircraft (via ground) to microphone. Or modelled from image source to microphone.
$s(n)$	Sinusoidal part of observed signal
t	Reception time
v	Velocity
$v(n)$	White Noise signal
x	Scalar
$X(\omega)$	Fourier transform of $x(n)$
$x(n)$	Observed signal
$\mathbf{I}$	Identity Matrix
$\mathbf{Q}$	Covariance matrix of noise $e(n)$
$\mathbf{R}$	Covariance matrix $x(n)$
Var($\cdot$)	Variance operator

Contents

1 Introduction **1**
- 1.1 Motivation . . . 1
- 1.2 Related Work . . . 1
- 1.3 Proposed Improvements . . . 3
- 1.4 Problem Formulation . . . 3

2 Physics **5**
- 2.1 Source Signal . . . 5
- 2.2 Trajectory Model . . . 6
- 2.3 Doppler Model . . . 10

3 Theory **13**
- 3.1 Mel-Frequency Cepstral Coefficients . . . 13
- 3.2 Cyclic Spectral Analysis . . . 14
- 3.3 Chirp Signal Model . . . 19
- 3.4 Pitch Estimator . . . 21
- 3.5 Flight Parameter Estimation . . . 29
- 3.6 Classification . . . 34

4 Classification Experiments **39**
- 4.1 General in Experiments . . . 39
- 4.2 Classification by Baseline . . . 42
- 4.3 Classification by Cyclic Spectral Coherence . . . 44
- 4.4 Classification by Cyclic Spectral Coherence Mean Envelope . . . 47
- 4.5 Classification on Flight Parameters by IF Estimates . . . 49
- 4.6 Comparison and Discussion on Classification Experiments . . . 50

5 Conclusion **53**

Bibliography 57

Appendix 59

A PRTools in MATLAB 61

B Image Processing Approach - HOG 63

C Classification by Pitch and Related Features 67

D CD Content 71

Chapter 1

Introduction

In this chapter a motivation for the project is given, followed by a section containing related work. We believe improvements can be done on existing solutions, which we will introduce and lastly the problem formulation is presented.

1.1 Motivation

Noise is often unwanted and needed to be hold at a minimum, especially in areas close to where people are living for example around an airport. For keeping down the noise, authorities can make rules for the maximum accepted sound pressure level from a noisy source. In this work, the focus is the maximum accepted sound pressure level from aircraft around an airport. If the maximum is exceeded, the owner and responsible of the aircraft should have a fine. In that case, it becomes interesting to make a system for detecting aircraft, from which the sound pressure level can be measured in the same signal and compared to a maximum allowed level.

In the context of noise impact around airports, noise terminals are distributed in order to estimate the noise from airports. However all kind of acoustic events are possibly recorded together with the target aircraft noise, and they may lead to a misleading estimate aircraft noise impact. One way to overcome this issue is to listen to the recordings and select intervals with aircraft noise, but it is a cumbersome and time-consuming task. Another way consists in setting up a process that automatically recognize when an aircraft sound is present. This is the purpose of this project, to investigate features for such a system, which can be used in a machine learning approach for recognizing an aircraft in signal.

1.2 Related Work

Since this report deals with several approaches for detecting an aircraft, the following will be divided into a number of subsections. First a feature part, which deals with different features possible to extract from an acoustical signal having a sinusoidal part. The next subsection deals with feature extraction and classification by the acoustic signature and lastly a subsection describing other approached for detecting an aircraft for example by using images and RADAR.

1.2.1 Features from Acoustical Signal

Pitch estimation is a well-known and long time study. One method for doing pitch estimation is the autocorrelation which is a well-known approach for doing pitch estimation, which easily goes back to the sixties and seventies as the paper by M. R. Schroeder [30] and by Georgije Lukatela [24]. In the book by Mads Græsbøll Christensen et al. [7] the authors presents several different approaches for doing pitch estimation, where the statistical one forms the basis of the pitch estimation in this report. In the article [17] the Mel Frequency Cepstral Coefficients (MFFC) is used in an audio search engine, which classifies based on similarity. The authors emphasizes their method's simplicity, due to no perceptual measures, like pitch and brightness. Another advantage of their method, is the low computational cost and therefore believed to be appropriate in case of big audio database. In each block of samples they calculate 12 MFFC plus one energy term, on which a supervised, known labels, tree quantizer is used for doing the classification.

1.2.2 Aircraft Feature Extraction and Classification by Acoustic Signature

In the article [25] the authors propose a 1/24 octave analysis and MFFC analysis for doing the feature extraction on an aircraft signal. The total number of features were 136, 96 for the octave analysis and 40 for the MFFC analysis. The approach is based on take-off noise from aircraft, where the classification was done using neural network. The number of classes, and thereby number of different aircraft, was 13, on which the authors obtained a total efficiency of $83\,\%$. For improving the results the authors centred the signal, which they did by following the standard in the International Civil Aviation Organization where it is stated that only the part of the signal, which is within $10\,\mathrm{dB}$ of the maximum sound pressure, should be used.

In the article [32] the authors used a Linear Predictive Coding method on a signal, which had been segmented into four parts by using an energy and zero-crossing method for detecting the center of the signal containing the aircraft taking off. The number of features was 140 for each segment and the number of aircraft classes was 13, on which the authors used a neural network approach for doing the classification and obtained a minimum efficiency of $85\,\%$

In the article [5] the authors propose what they call an *aircraft noise likeness detector*, which is based on the similarity between the observed signal and a generic aircraft sound giving a value between zero and one. The features that are extracted for doing so are MFFC, where the authors used 13 coefficients and a bandwidth with the lower limit at $0\,\mathrm{Hz}$ due to components in aircraft noise being low frequency. The authors used a statistical classifier, which uses the Bayes decision rule for decide the class. The authors ran into problems making generic model for non-aircraft sound, and decided to use a one-class classifier approach. In the case of Signal-to-Noise-Ratio (SNR) above $8\,\mathrm{dB}$ the correct aircraft sounds detected was $93\,\%$.

One of the approaches we will propose in this report to use along with our proposed pitch estimator is based on the paper by Brian G. Ferguson et al. [13], where a flight parameter estimator is given, which uses the Instantaneous Frequency (IF) estimates. The flight parameter estimation is based on a Non-linear Least Squares (NLS) approach on which the authors of [13] obtains acceptable Root Mean Square Error (RMSE) values on the flight parameters. The author has written several other papers like [14], [15] and [12] on the topic of estimating flight parameters from the Doppler Frequency and the destructive interference frequencies. Observe, we also use the article [15] for deriving a trajectory model, on which the flight parameters are presented and thereby a synthetic signal can be produced used for testing the flight parameter estimator.

In the article [20] an approach of event detection is introduced. This is based on the GABOR filter-bank, which perform well in noisy environment. The method is based on spectral a temporal modulation frequency in the signal. The database used for testing contained breaking glass, explosion, gunshot, scream and tube station noise. They compared their results with MFFC, and showed GABOR can outperform MFFC in noisy environment.

The authors of the article [6] proposed to use Histogram Of Gradients (HOG) on a spectrogram image together with Sub-band Power Distribution in acoustic scene classification. The classification was done with a support vector machine. They described how HOG will capture different significant parts in a spectrogram image because of differences of aircraft sounds, but when used together with the Sub-band Power Distribution they obtained an accuracy of $93.4\,\%$.

1.2.3 Other Methods for Classification of Aircraft

Instead of basing the classification or detection of aircraft on the acoustic signature other approaches are possible. For example in the article [18] where a Very High Resolution image provides the measurement data. A circle frequency filter is then used to extract candidates of location in the entire image, which goal is to decrease the complexity of the image. The features for classification are extracted with Robust Hue Descriptor and HOG, used to find color and shape information in image. It is also possible to use infra-red signature from an aircraft for detect and classify an aircraft. In the article [23] the authors use a low resolution infra-red sensor for constructing an image for doing aircraft classification and lastly different RADAR approach are used for aircraft classification like [28] and [33] where both uses the Doppler effect for doing the detection and classification of aircraft.

1.3 Proposed Improvements

When an aircraft is wanted to be detected or classified within a signal, several approaches exist, where some has been introduced in Section 1.2. We believe much of the earlier work based on acoustical signature, can be improved by using the temporal variation that will be present in the signal when an aircraft is passing by. The temporal variation can both be seen as the time-varying pitch received at the observer point, but is can also be seen as time-varying amplitude in the observed signal, which in most cases will be a hidden periodicity. By investigate these two different views of temporal variation we believe the success rate of detection an aircraft can be improved. We will use MFFC as a baseline where we will extract MFFC features for training and testing a classifier. What we will contribute with, is an investigation of features on signal blocks of $1\,\mathrm{s}$, where we will perform a classification for each block. We will also investigate the use of the acoustical signature in a model for extracting flight parameters. This cannot be done with $1\,\mathrm{s}$ blocks of signal, so in this case we use the *event* of an aircraft, which is an information provided in the database we have used, and the basic meaning of an event is simply an interval where an aircraft is present. In this case, the signal blocks are minimum several seconds.

1.4 Problem Formulation

We have derived the following problem formulation:
Will the utilization of the temporal variation in the feature extracting based on the acoustic signature in a signal, improve the detection of an aircraft and how to use these features in a optimum manner for doing detection of an aircraft in signal?

Chapter 2

Physics

The goal in this chapter is to introduce the physics behind the problem to be solved. First a real signal is visualized in form of a spectrogram, to visualize the components in the signal. A few assumptions are introduced on which a geometric model for an aircraft's trajectory is presented using flight parameters. Since these flight parameters are believed to be significant and unique, a synthetic signal based on these flight parameters are produced, which make it possible to later on test if the flight parameters can be estimated. For making the synthetic more realistic, broadband noise which creates destructive interferences frequencies will be added to the synthetic signal.

2.1 Source Signal

A true signal can be seen in the spectrogram in Figure 2.1. This signal contains an aircraft passing by an observer position. In Figure 2.1 two major things can be observed.

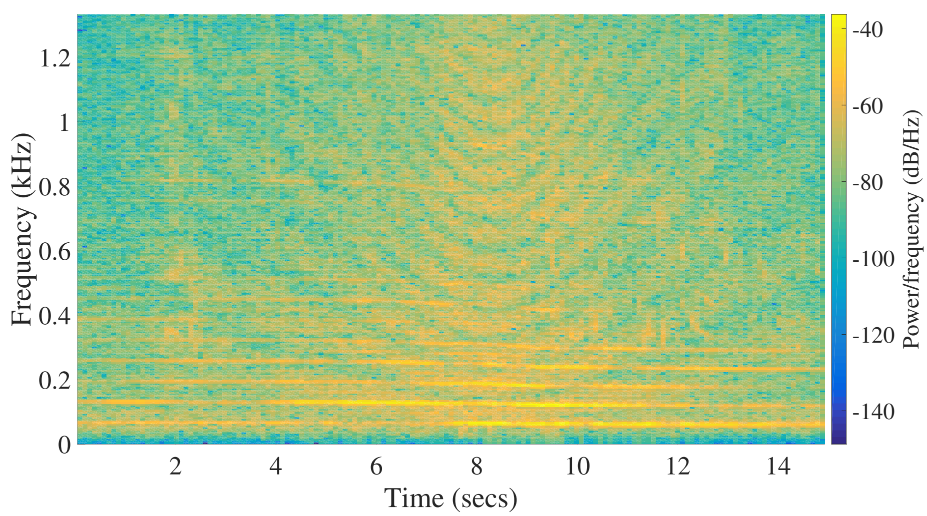


Figure 2.1: Spectrogram of an aircraft passing by.

- Firstly, a pitch, sum of harmonics, is present in the low frequency area. The received pitch is actually changing with time according to the Doppler Effect.

- Next to be observed, is the parabolic patterns which is due to destructive inference of broadband noise radiated from the aircraft. The destructive inference happens due to the ground-reflected signal, which will interfere with the direct signal from source to microphone.

Also, the sound pressure, amplitude of signal, is changing over time due to changing distance between observer and aircraft; starts low, increases, reaching maximum when closest to microphone and decreases again. This tells, the waves needs to be modelled as spherical waves.

The goal is now to describe something similar, which means the observed signal $x(n)$ will be

$$x(n) = s(n) + e(n) + v(n), \tag{2.1}$$

where $s(n)$ is the sinusoidal part of the observed signal containing the pitch, $e(n)$ is the signal containing the colored (pink) noise giving the interference frequency curves and lastly $v(n)$ which is additive white Gaussian noise.

We want to make a parametric model of the signal and from that, an estimator on those parameters should be possible to construct. Thereby, a number of features can be found in the signal of interest and hopefully these features are useful for doing classification, where the goal is to distinguish between non-aircraft sound and aircraft sound.

2.2 Trajectory Model

The basic behind the following from the paper by Brian G. Ferguson et al. [15] is that knowing the height of the microphone, height of object, speed of object, distance from object to microphone at the Closest Point of Approach (CPA) and the emission time for it, a model can be made for the received signal. The model has the following assumptions

- **Stationary height.**
- **Stationary speed.**
- **Straight trajectory.**

The trajectory geometry model seen in Figure 2.2 is not using an image source as usual, but rather an additional image receiver, which means the model is made by having one emission time and two reception times. The model is then build from that the reception time at origin in Figure 2.2 is known. From this, it is possible to derive expressions for the two ranges, $R_d(\tau)$ and $R_r(\tau)$, from which their emission times can be found and used for producing a synthetic signal.

2.2.1 Flight Parameter Equations from Trajectory Model

Before deriving the equations for different flight parameters seen in Figure 2.2, equations for the IF and Instantaneous Phase (IP) are presented here as the authors of the paper [14] have stated. IF can be written as a function of the reception time t as

$$f(t) = f_0 \frac{\partial \tau(t)}{\partial t}, \tag{2.2}$$

and the IP can be expressed as

$$\theta(t) = 2\pi f_0 \tau(t) + \phi, \tag{2.3}$$

where ϕ is the initial phase. $\tau(t)$ is the time-varying emission time, and is expressed as

$$\tau(t) = t - R(t)/c, \tag{2.4}$$

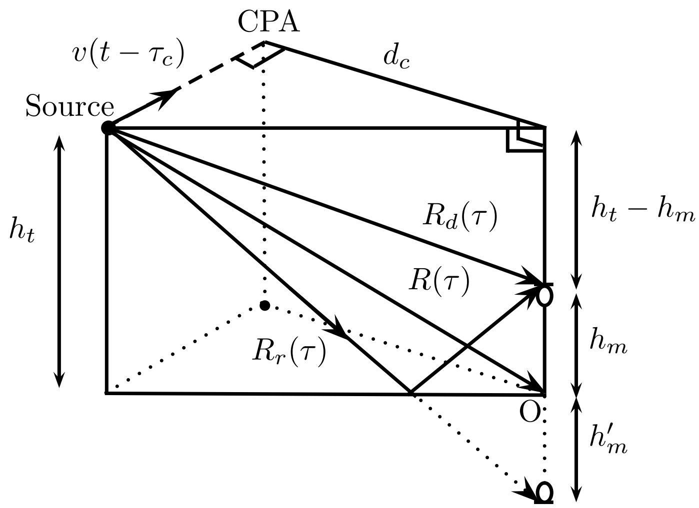


Figure 2.2: Source Geometry used for deriving path equations and expression for emission time

where $R(t)$ is the time varying distance from object to origin (origin can be though as the position of the microphone, but with a height equal zero, see Figure 2.2). That means the time of emission, $\tau(t)$, is the time of reception, t, minus the time it takes the sound for travelling $R(t)$. The reception time t is known, so it is the emission time $\tau(t)$ that are of interest to be found, which can then be used to model two sources (one true and one image), from which a synthetic signal be produced.

Starting with the path $R(\tau)$ (slant range) which goes from the source to origin. This path is described as

$$R(\tau) = \sqrt{v^2(\tau - \tau_c)^2 + d_c^2 + h_t^2}, \tag{2.5}$$

where τ_c is the emission time for CPA. It is clear τ, the emission time, is needed be found or the expression for R needs to be redefined. By using the relation from eq. 2.4 and substitute eq. 2.5 into that one, the results is

$$\tau(t) = \tau_c + \frac{c^2(t - \tau_c) - \psi(t)}{c^2 - v^2}, \tag{2.6}$$

where it gets clear why τ is a function of time t. The function $\psi(t)$ is derived as:

$$\psi(t) = \sqrt{(h_t^2 + d_c^2)(c^2 - v^2) + v^2c^2(t - \tau_c)^2}. \tag{2.7}$$

Now the path range R can be described as a function of reception time t by using eq. 2.4 and eq. 2.6

$$R(t) = \frac{c}{c^2 - v^2}[\psi(t) - v^2(t - \tau_c)]. \tag{2.8}$$

When the $R(t)$ is found, the emission time $\tau(t)$ can be found, with eq.2.4, and from that the two paths $R_d(\tau)$ and $R_r(\tau)$ can be found, which are the two time-varying paths, direct and reflected. Direct path range R_d can be expressed as:

$$R_d(\tau) = \sqrt{v^2(\tau - \tau_c)^2 + d_c^2 + (h_t - h_m)^2}, \tag{2.9}$$

and reflected is

$$R_r(\tau) = \sqrt{v^2(\tau - \tau_c)^2 + d_c^2 + (h_t + h_m)^2}, \tag{2.10}$$

where the derivation of these two is clear from the drawing in Figure 2.2.

The Two Emission Times

When having the $\mathbf{R}_d$ and $\mathbf{R}_r$ vectors containing all the path distances, one can change to have **two** sources, the true and image one, which means having **two** emission times

$$\boldsymbol{\tau}_d = \mathbf{t} - \mathbf{R}_d/c, \tag{2.11}$$

and

$$\boldsymbol{\tau}_r = \mathbf{t} - \mathbf{R}_r/c. \tag{2.12}$$

2.2.2 Signal Model with Direct and Reflected Path

The sinusoidal Doppler part of the signal $s(n)$ with L harmonics, contains at any given time index n, the direct and reflected signal

$$s(n) = s_d(n) + s_r(n), \tag{2.13}$$

where s_d is the direct signal and s_r is the ground-reflected signal. The IF for the two paths are

$$\theta_{dl}(n) = \omega_0 l \tau_d(n) + \phi_{dl}, \tag{2.14}$$

and

$$\theta_{rl}(n) = \omega_0 l \tau_r(n) + \phi_{rl}, \tag{2.15}$$

where the two emission times follows

$$\tau_d(n) = n - R_d(n)/c, \tag{2.16}$$

and

$$\tau_r(n) = n - R_r(n)/c, \tag{2.17}$$

where $R_d(n)$ is the direct path distance at the reception time t indexed by n and $R_r(n)$ is the reflected path distance at the reception time t indexed by n. The signal $s(n)$ becomes

$$s(n) = \sum_{l=1}^{L} a_{dl}(n) e^{j(\omega_0 l \tau_d(n))} + a_{rl}(n) e^{jl(\omega_0 l \tau_r(n))}, \tag{2.18}$$

where $a_{dl}(n) = \frac{A}{r_d(n)} e^{j\phi_{dl}}$, having the real part A which is divided by the distance between source and receiver $r_d(n)$, which thereby models the spherical wave. The reflected complex amplitude $a_{rl}(n) = \frac{A}{r_r(n)} e^{j\phi_{rl}}$.

2.2.3 Alternative Expression by Reception Time

Instead of using the emission times $\tau_d(n)$ and $\tau_r(n)$, the signal model in eq. 2.18 can be described with the time of reception t indexes by n as

$$s(n) = \sum_{l=1}^{L} a_{dl}(n) e^{jl(\omega_0 (n - \frac{R_d(n)}{c}))} + a_{rl}(n) e^{jl(\omega_0 (n - \frac{R_r(n)}{c}))}, \tag{2.19}$$

where $a_{dl}(n) = \frac{A}{r_d(n)} e^{j\phi_{dl}}$ and $a_{rl}(n) = \frac{A}{r_r(n)} e^{j\phi_{rl}}$. The two path ranges, $R_d(n)$ and $R_r(n)$ has been described earlier in respectively eq. 2.9 and eq. 2.10, both are quadratic equations. It is clear the $\frac{R_x}{c}$, distance divided by speed of sound, gives a delay in time, which can be seen in Figure 2.3. In Figure 2.3, it can be seen that the delay can indeed be seen as quadratic equation, which also

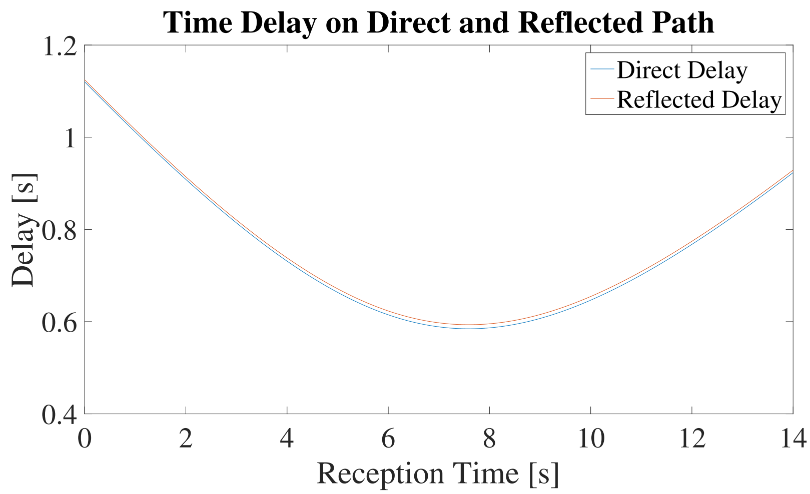


Figure 2.3: Time delay as a function of reception time. Both Direct and Reflected path delay.

gets clear if observing the equation for the delay of the direct path

$$\frac{R_d(n)}{c} = \frac{\sqrt{v^2(\tau(n)-\tau_c)^2 + d_c^2 + (h_t - h_m)^2}}{c}. \tag{2.20}$$

Since $\frac{R_d(n)}{c}$ gives a time, it might be denoted as

$$T_d(n) = \left(\frac{R_d(n)}{c}\right), \tag{2.21}$$

as the delay for the direct path. Applying eq. 2.4 to replace $\tau(n)$ in eq. 2.20 and us the new delay T gives

$$T_d(n) = \sqrt{v^2\left(n - \frac{R(n)}{c}\right)^2 - 2v^2\tau_c\left(n - \frac{R(n)}{c}\right) + d_c^2 + (h_t - h_m)^2 + v^2\tau_c^2}, \tag{2.22}$$

which is on the form of a 2-order polynomial as can be observed in Figure 2.3. For the reflected path, the time delay can be expressed:

$$T_r(n) = \sqrt{v^2\left(n - \frac{R(n)}{c}\right)^2 - 2v^2\tau_c\left(n - \frac{R(n)}{c}\right) + d_c^2 + (h_t + h_m)^2 + v^2\tau_c^2}. \tag{2.23}$$

Observe, an expression for $R(n)$ has earlier been derived in eq. 2.8. Now, the left side of eq. 2.22 and eq. 2.23 can be inserted into eq. 2.19, which becomes

$$s(n) = \sum_{l=1}^{L} a_{dl}(n)e^{jl(\omega_0(n-T_d(n)))} + a_{rl}(n)e^{jl(\omega_0(n-T_r(n)))}, \tag{2.24}$$

and thereby the signal model is described by the reception time and delay on both direct and reflected paths.

2.3 Doppler Model

The derivation of the following Doppler model is found in paper [16]. The Doppler effect can be observed when an object is moving relative to the observer. The distance between the two and the relative speed will give the change of frequency. The Doppler frequencies can be described with flight parameters from the aircraft's trajectory seen in Figure 2.2 in Section 2.2.

If observing from the receiver point of view, then the reception time t is given by the emission time τ plus the time it takes for the sound to reach the observer due to the distance R

$$t = \tau + R/c, \tag{2.25}$$

where c is speed of sound and thereby a delay is added to the emission for obtaining the reception time. Observe that eq 2.25 is a rewritten and simplified version of eq. 2.4 in Section 2.2, where we here neglected t in both τ and R. The distance R can be found with Pythagoras by using the Trajectory Model given in Figure 2.2 in Section 2.2 from which the parameters in focus are the slant range at CPA R_c and the horizontal range r, which is the same as $v^2(\tau - \tau_c)^2$, where τ_c is the emission time for the CPA. Observe, R_c is found by eq. 2.8 at $t = \tau_c$, and thereby the R_c is found with the parameters h_t, d_c and v.

$$R_c = \frac{c}{c^2 - v^2}\sqrt{[(h_t^2 + d_c^2)(c^2 - v^2)]}. \tag{2.26}$$

It is now possible to derive R

$$R = (R_c^2 + r^2)^{1/2} = (R_c^2 + v^2(t - \tau_c)^2)^{1/2}. \tag{2.27}$$

By substituting eq. 2.27 into eq. 2.25 an expression for τ can be found

$$\tau = \frac{c^2 t - v^2\tau_c - (R_c^2(c^2 - v^2) + v^2c^2(t - \tau_c)^2)^{1/2}}{c^2 - v^2}. \tag{2.28}$$

The IF at the reception time t is given by

$$f(t) = f_0\frac{d\tau}{dt} = \frac{f_0 c^2}{c^2 - v^2}\left(1 - \frac{v^2(t - \tau_c)}{(R_c^2(c^2 - v^2) + v^2c^2(t - \tau_c)^2)^{1/2}}\right) \tag{2.29}$$

The IF in eq. 2.29 can be expressed in a different way following the paper [14]. At the same time we also change time t by its index n.

$$f(n) = \gamma + \beta p(n; \tau_c, s), \tag{2.30}$$

where

$$\gamma = f_0 c^2/(c^2 - v^2) \tag{2.31a}$$

$$\beta = -f_0 cv/(c^2 - v^2) \tag{2.31b}$$

$$s = \frac{R_c(c^2 - v^2)^{1/2}}{vc} \tag{2.31c}$$

$$p(n; \tau_c, s) = \frac{n - \tau_c}{[s^2 + (n - \tau_c)^2]^{1/2}}, \tag{2.31d}$$

where v is velocity of object, c speed of sound, n is time index for reception time, R_c is the slant range at CPA and τ_c emission time for the CPA.

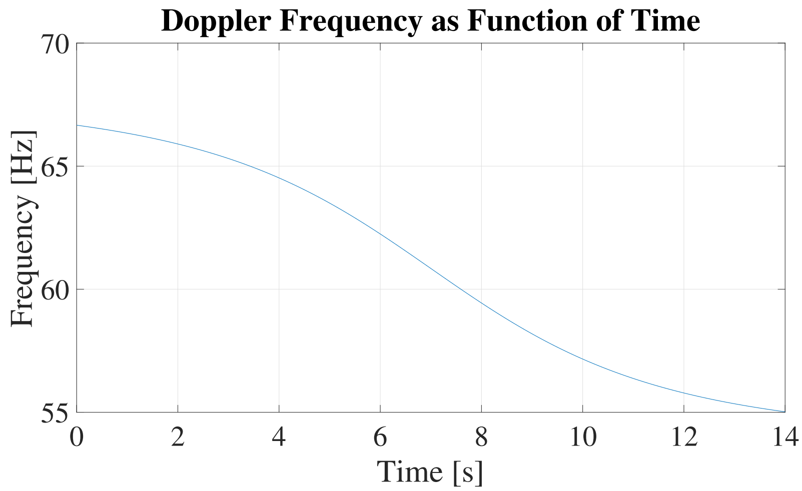


Figure 2.4: Example of a Doppler Frequency curve with true $f_0 = 60\,\mathrm{Hz}$.

In Figure 2.4 a Doppler frequency curve following eq.2.30 is seen. The parameter γ given in eq. 2.31a is the center frequency, which in Figure 2.4 is at $7\,\mathrm{s}$ and is slightly above $60\,\mathrm{Hz}$ as expected. The parameter β given in eq. 2.31b is the deviation between the center frequency γ and the maximum Doppler frequency at infinite time. The maximum Doppler frequency is never possible to observe, so the maximum observed Doppler frequency is given by scaling β by $p(n; \tau_c, s)$, given in eq. 2.31d. The $p(n; \tau_c, s)$ is a time varying variable with maximum range -1 to 1 in case of infinite time. It is zero at the center Doppler frequency time and is symmetric around the center time. That means $p(n; \tau_c, s)$ is a scaled and mirrored form of the Doppler frequency curve. One thing to observe is, eq. 2.31c will increase as the *slant range* R_c increases. As R_c gets close to infinite, the parameter s will be close to infinite which results in a linear or even flat Doppler frequency curve, since $p(n; \tau_c, s)$ will be very close to zero and thereby the scaling β is almost not added to γ. Also if the velocity v is low, the Doppler frequency curve will tend to be linear. That means, when the distance between microphone and aircraft is large or the aircraft has a low ground speed, the Doppler Frequency curve will tend to be approximately linear or even flat.

It should be stressed that assuming the Doppler frequency curve to be linear is most likely no good, and the following two approaches are believed to be a better solution in the case of pitch estimation. The first approach is to describe the curve with a higher order polynomial, which will make it possible to describe full Doppler directly. However, we will propose to work on the full signal in smaller segments, where linearity can be assumed. By that, the Harmonic Chirp Model (HCM) can be applied, on which a pitch estimator is proposed, which is based on the paper [8]. All this is to come in the following chapter.

Chapter 3

Theory

The baseline in this work is the MFFC, which will be introduced in this chapter. In Section 1.3 we introduced two approaches for taking the temporal variation into account. Those two approaches will be described in details in the following. For the (hidden periodic) time-varying amplitude, cyclic spectral analysis is used. For the time-varying pitch, a chirp pitch estimator, which can be provide the IF, is proposed. The proposed chirp pitch estimator uses a parametric HCM, which will be described before the chirp pitch estimator. The proposed chirp pitch estimator is based firstly on a NLS Estimator working on a plain Harmonic Model (HM), followed by the chirp pitch estimator working on the HCM. It is well-known in pitch estimation that the order is often the key to solve the pitch estimation problem, so one approach for estimating the model order is also described and in addition this approach also serves as a pitch detection. One great advantage of the proposed chirp pitch estimator is the directly derivation of the IF, which can be used to estimate flight parameters, based on the Trajectory and Doppler model introduced in Chapter 2. The goal of this report is to investigate features for detecting aircraft. Therefore, to be able to directly compare the different features which can be extracted from a signal when an aircraft is present, we perform classification based on these features. For that, a short description about the used classification theory is presented as the last part in this chapter.

3.1 Mel-Frequency Cepstral Coefficients

Before looking into the implementation of MFFC, we want to start with a short introduction of the unit *mel*. The author of the book [22] describes the pitch having the unit *mel*, as a perceived sound produced by frequency, intensity, duration and spectral content. The reference sine wave, is a $1\,\mathrm{kHz}$, which is defined as the pitch of $1000\,\mathrm{mel}$. The pitch is then measured subjective by comparing it to this given sine wave. The relation between frequency and pitch is often described as in the paper [31]:

$$mel(f) = 2595 \log_{10}\left(1 + \frac{f}{700}\right), \tag{3.1}$$

where f is the frequency. Observe in Figure 3.1 a plot is made for frequency range $0\,\mathrm{Hz}$ to $10\,\mathrm{kHz}$, which visualises the non-linear pitch curve. As explained in the article [17] the steps for doing the MFFC follows:

1. Calculate power spectrum by Discrete Fourier Transform (DFT).

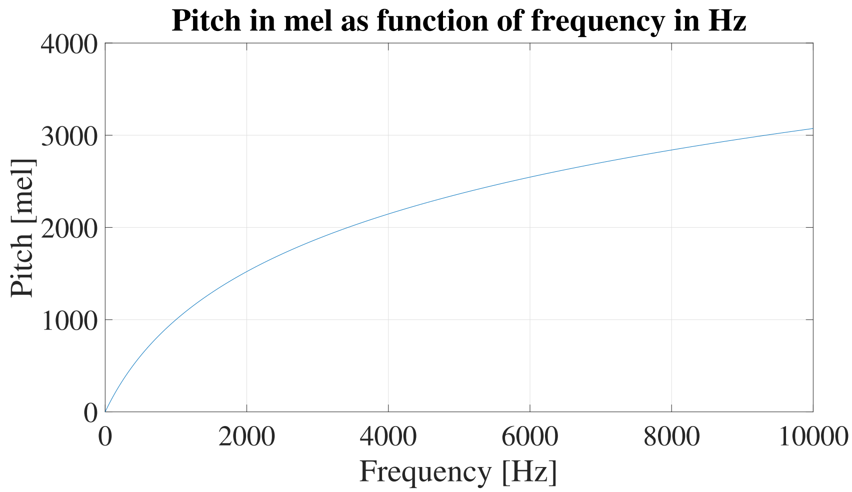


Figure 3.1: Typical pitch curve as a function of frequency. Curve is produced with eq. 3.1. Until $1\,\mathrm{kHz}$ one assume linearity and above that logarithm scale is assumed.

2. Divide data into the weighted mel-scale filter banks, which have equally bandwidth sizes below $1\,\mathrm{kHz}$ and logarithm increasing bandwidths above. This is due to the perceptual mel-frequency curve seen in Figure 3.1 or eq. 3.1
3. Apply logarithm to the signal.
4. Lastly, a discrete cosine transform is added, to obtain the cepstral coefficient. The discrete cosine transform is here used instead of the inverse DFT, since using the power spectrum which has absolute (real) values.

It is now clear that the idea behind the MFFC is similar to the (real valued) Cepstrum analysis which can be found in e.g. the book by Alan V. Oppenheim [27] and Applications Notes by B&K[29]. The Cepstrum equation can be written as

$$C = |\boldsymbol{\mathcal{F}}(\log(|\boldsymbol{\mathcal{F}}\{f(t)\}|^2)|^2, \tag{3.2}$$

from which similarity between MFFC and Cepstrum analysis can be observed.

3.2 Cyclic Spectral Analysis

The Cyclic Spectral Analysis is a tool for taking the temporal variation into account. In the following an introduction to the Cyclic Spectral Analysis is given with the article [19] as source.

3.2.1 Cyclo-Stationary Signals

Often a stationary signal is assumed in signal processing, which the authors of [2] states, that the assumption is more of a convenient matter then it is the truth. Often the signal is non-stationary and random in the waveform, but observing the energy and flow might show hidden periodicity. This periodicity is what is referred to as a cycle, as introduced in [1]. A typically example of a cyclo-stationary signal is a mechanically signal from a rotating machine, e.g. from a propeller, from which is gets clear why the cyclic analysis is of big interest when detecting aircraft. In time domain, the cyclic analysis provides information about instantaneous auto-correlation function, instantaneous power and envelope. These together gives the temporal variations. In the frequency domain, the spacing between spectral components will give the cyclic frequency. The modelling of a cyclo-stationary signal, is explained in the article [19] and follows

$$x(t) = A\cos(2\pi\alpha t + \theta), \quad \alpha \neq 0, \tag{3.3}$$

where α is the frequency of the finite strength additive sine wave. Having the Fourier

$$M_x^\alpha = \{x(t)e^{-i2\pi\alpha t}\}, \tag{3.4}$$

which, when exist, gives $M_x^\alpha = \frac{1}{2}Ae^{i\theta}$. In addition, $\{\bullet\}$ in eq. 3.4 is a time averaging operation following

$$\{\bullet\} \triangleq \lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{T/2} (\bullet)dt. \tag{3.5}$$

Now, if $f = \alpha$ and $f = -\alpha$ are present in the Power Spectral Density (PSD) of $x(t)$ as spectral lines, that means the PSD contains

$$|M_x^\alpha|^2[\delta(f-\alpha) + \delta(f+\alpha)], \tag{3.6}$$

where δ is the Dirac impulse function. A signal with the terms from eq. 3.6 in its PSD, is referred to as a first-order periodicity signal having frequency α. Now, the signal $x(t)$ is given as

$$x(t) = A\cos(2\pi\alpha t + \theta) + n(t), \tag{3.7}$$

where $n(t)$ is the remaining part, noise, which is uncorrelated with α. If $n(t)$ is most significant in the signal $x(t)$, a hidden periodicity is present in the signal since it cannot be observed by visual inspection of the signal. Still due to spectral lines in the PSD, it is possible to use spectral analysis for estimating the hidden periodicity. When the hidden periodicity is more complex, a quadratic transformation can be applied to signal, e.g. squaring the signal, which will convert the hidden periodicity in the signal to a first-order periodicity, where spectral lines can be found. The squaring approach will not always work, so a better solution might be to use a quadratic transformation including delays. The example given in the article [19] is having pulse-amplitude modulated signal

$$x(t) = \sum_{n=-\infty}^{\infty} a(nT_0)p(t - nT_0), \tag{3.8}$$

where $a(nT_0)$ is the pulse amplitudes, T_0 is the pulse repetition interval and $p(t - nT_0)$ is pre-shaped pulse. Now let all amplitudes be 1, which means having a constant for all t when the squaring transformation is done. By that, a spectral line is present at the DC, but not at any harmonics. Therefore, for finding the hidden periodicity the quadratic transform should be done as

$$y(t) = x(t)x(t - \Delta t), \tag{3.9}$$

where Δt is the delay and spectral lines will be present at m/T_0

$$M_y^\alpha = \{y(t)e^{-i2\pi\alpha t}\} \quad \neq 0, \tag{3.10}$$

where α is m/T_0 for integer values m. A general time-invariant quadratic transform can be written as

$$y(t) = \int h(\Delta t)x(t)x(t - \Delta t)d\Delta t, \tag{3.11}$$

which is a linear combination of delay products from eq. 3.9 and weighted by $h(\Delta t)$, which is like the impulse response function for a linear transformation. The definition of a second-order periodicity follows that, if the PSD of the signal in eq. 3.9 with some delays Δt has spectral lines at none-zero α frequencies, meaning eq. 3.10 is satisfied, then the signal $x(t)$ contains second-order periodicity. For convenience the symmetric delay products is introduced as

$$y_{\Delta t}(t) = x(t + \Delta t/2)x^*(t - \Delta t/2), \tag{3.12}$$

where * is the complex conjugated for making the model fit to complex signals. It should be observed that the author of the used MATLAB implementation [3] of cyclic analysis, recommend to use complex (analytic) signal, which can be obtained with the Hilbert transform [21]. The next is that eq. 3.10 can be rewritten to

$$R_x^\alpha \triangleq \{x(t + \Delta t/2)x^*(t - \Delta t/2)e^{-i2\pi\alpha t}\}, \tag{3.13}$$

which is the cyclic auto-correlation function, related to the general auto correlation function, but with the cyclic weighting factor included, $e^{-i2\pi\alpha t}$. If $\alpha = 0\,\mathrm{Hz}$ eq. 3.13 becomes the general auto correlation function. For making the direct link to the general auto correlation function, the eq. 3.13 can be rewritten into

$$R_x^\alpha \triangleq \{[x(t + \Delta t/2)e^{-i\pi\alpha(t+\Delta t/2)}][x(t - \Delta t/2)e^{+i\pi\alpha(t-\Delta t/2)}]^*\}, \tag{3.14}$$

which can be written as the conventional cross correlation function

$$R_{uv}(\Delta t) \triangleq \{u(t) + \Delta t/2)v^*(t - \Delta t/2)\} = R_x^\alpha, \tag{3.15}$$

with $u(t) = x(t)e^{-i\pi\alpha t}$ and $v(t) = x(t)e^{+i\pi\alpha t}$. It can be expressed, that $x(t)$ has second-order periodicity if $v(t)$ and $u(t)$ are correlated. A signal having this form is referred to as a cyclo-stationary signal and α is the cyclic frequency, see example of cyclic frequency estimate in Subsection 3.2.2

The Correlation Density (SCD) function in cyclic analysis is of great interest in this work, since it leads to the (cyclic) Spectral Coherence, on which several approaches can be uses for doing classification. Still, the article [19] forms the basis of the following. For obtaining the SCD, the PSD for $x(t)$ is needed, which can be written for any f as

$$S_x(f) \triangleq \lim_{B\to 0}\frac{1}{B}\{|h_B^f(t) * x(t)|^2\}, \tag{3.16}$$

where $*$ is convolution, $h_B^f(t)$ the impulse response of a one-sided bandpass filter, having the center frequency f and lastly B is the bandwidth. Letting $u(t)$ and $v(t)$ pass through the same sets of bandpass filters the result is the SCD

$$S_x^\alpha(f) = \lim_{B\to 0}\frac{1}{B}\{|h_B^f(t) * u(t)||h_B^f(t) * v(t)|^*\}. \tag{3.17}$$

The expression in eq. 3.16 can be obtained as the Fourier transform of the cyclic auto-correlation written as

$$S_x^\alpha(f) = \int_{-\infty}^{\infty} R_x^\alpha(\Delta t)e^{-2\pi f\Delta t}d\Delta t. \tag{3.18}$$

The relation seen in eq. 3.18 also leads to that SCD is expressed as the Cyclic Spectral Density (CSD) function. In addition, the SCD is the Fourier transform of the cross-correlation function of the two parts $v(t)$ and $u(t)$. Due to that, SCD is equal to the cross-spectral density function

$$S_x^\alpha(f) = S_{uv}(f), \tag{3.19}$$

where right hand side of eq. 3.19 is equal to the right hand side of eq. 3.17 and by that the spectral correlation between $v(t)$ and $u(t)$. The next part is now to do a normalization, which will change the correlation of spectral components to be equal to the covariance since mean of spectral components are zero, when no spectral lines at frequency f are present in the PSD of $v(t)$ and $u(t)$, equal to no spectral lines at $f \pm \alpha/2$. One normalization is the geometric mean of the variances, S_u and S_v, by which the covariance ends up being the correlation coefficient

$$\frac{S_{uv}(f)}{\sqrt{S_u(f)S_v(f)}} = \frac{S_x^{\alpha}(f)}{\sqrt{S_x(f+\alpha/2)S_x(f-\alpha/2)}} \triangleq \rho_x^{\alpha}(f), \tag{3.20}$$

on which absolute value is normally calculated giving a value between 0 and 1. That means eq. 3.20 gives an measure of the spectral redundancy. The author of the article [2] also calls eq. 3.20 for the (Cyclic) Spectral Coherence, where the author points that this relation is only meaningful when having second-order cyclo-stationary signal in which case a strong linear dependence can be observed.

3.2.2 Cyclic Spectral Analysis Test

The used MATLAB implementation of Cyclic Spectral analysis [3] is wanted to be tested for obtaining the Cyclic Spectral Coherence introduced in eq. 3.20 on which the result can be seen in Figure 3.2. By observing the Figure 3.2 we can conclude that the signal used in this test has a hidden periodicity having cyclic frequency $\alpha = 87\,\mathrm{Hz}$. The Cyclic Spectral Coherence visualized

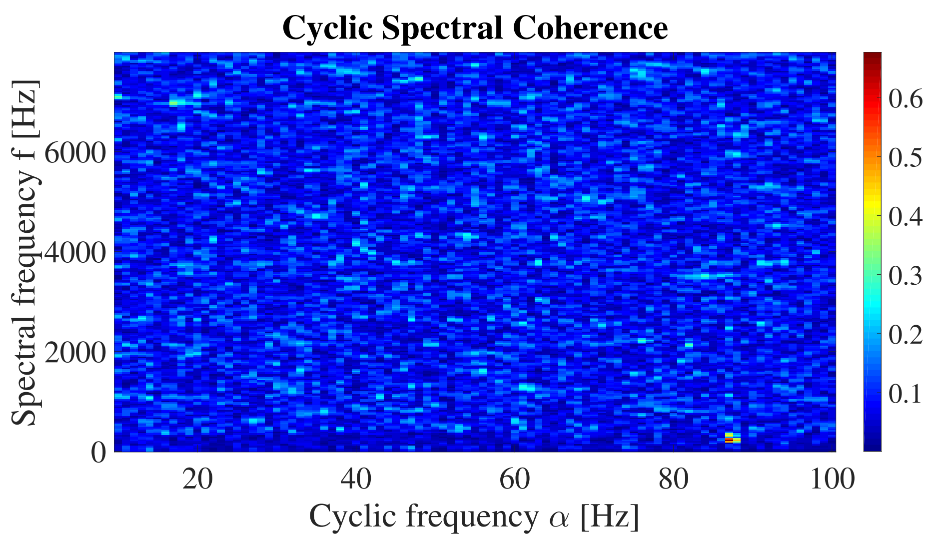


Figure 3.2: Cyclic Analysis by Cyclic Spectral Coherence given a Cyclic Frequency $\alpha = 87\,\mathrm{Hz}$, the red dot to the right lower corner.

in Figure 3.2 can be used directly as a feature, which will give a high number of features, on which Principal Component Analysis (PCA) is needed to be applied, see Subsection 3.6.4. Another solution for reducing the number of features is to mean the Cyclic Spectral Coherence for constructing an additional feature, mean envelope. For an easily observed periodicity in a signal, one such mean envelope will look like the one seen in Figure 3.3, which is not produced with the same signal as in Figure 3.2. However, if applying the mean to Figure 3.2, the result ends up as can be seen in Figure 3.4, which is very different in the envelope. It is clear from Figure 3.4 that the mean levels for the different cyclic frequency are not significant and should be rejected if comparing it to the

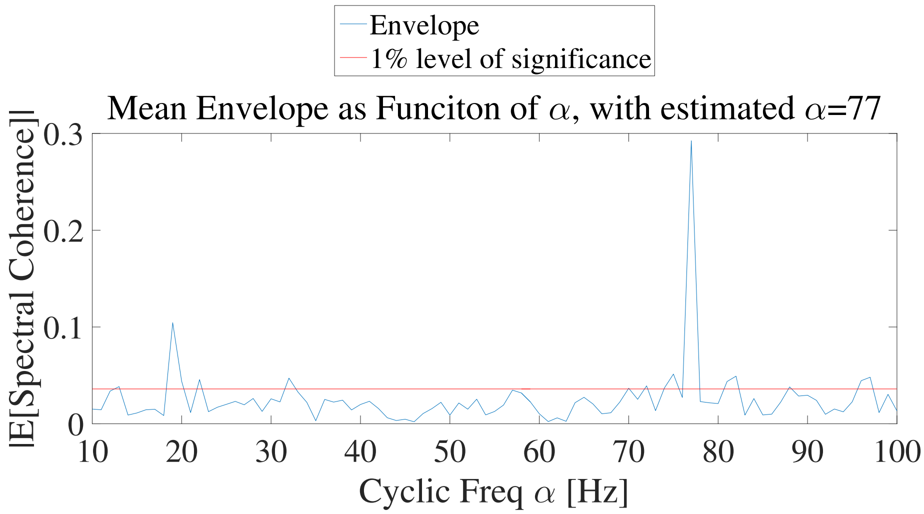


Figure 3.3: Mean of the Cyclic Spectral Coherence when an aircraft is present in the signal. Also, the 1 % level of significance is shown.

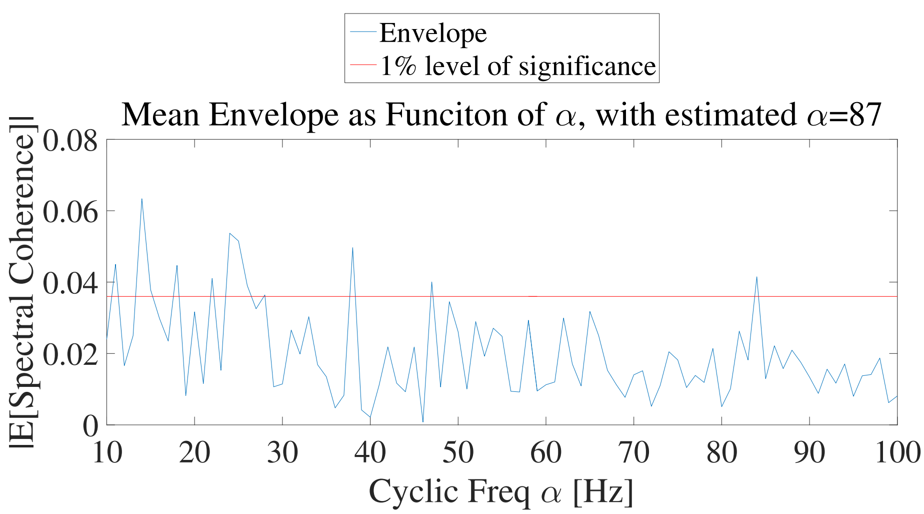


Figure 3.4: Mean of the Cyclic Spectral Coherence when an aircraft is present in the signal. Also, the 1 % level of significance is shown.

1 % level of significance line; the red line in Figure 3.4. When no aircraft is present a similar mean envelope can be observed, see Figure 3.5 It was already clear from the differences of Figure 3.3 and Figure 3.4 and the similarity of Figure 3.5 and Figure 3.4 that a separation between aircraft and non-aircraft based on the mean envelope alone will be a hard challenge, but still a test was carried out which can be seen in Section 4.4.

Lastly, the Cyclic Spectral Coherence seen in for example Figure 3.2 can be transformed to an image, on which image processing tools can be used for extracting features. We did a small test on such a method, which can be found in Appendix B.

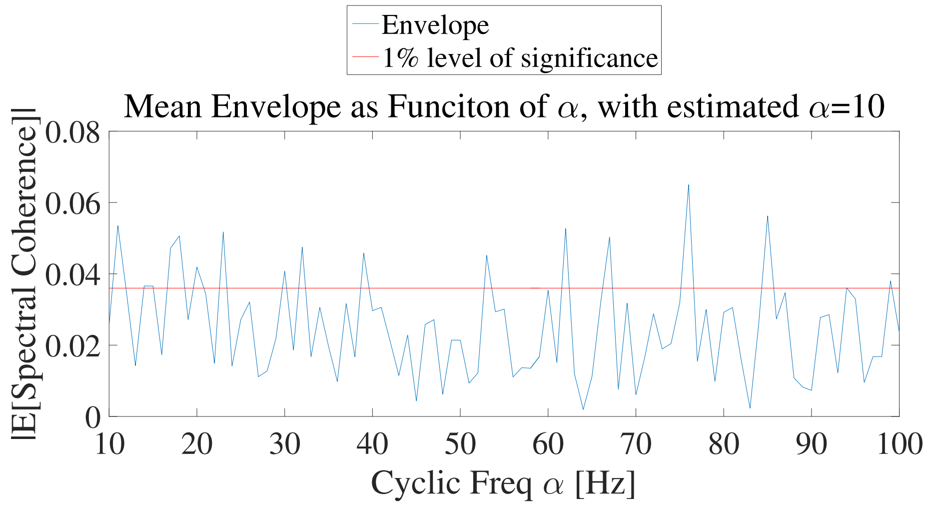


Figure 3.5: Mean of the Cyclic Spectral Coherence when no aircraft is present in the signal. Also, the 1 % level of significance is shown.

Cyclic Frequency Estimation on Signal where Aircraft is Passing By

The cyclic frequency estimate can be used as a feature for doing the classification, but is believed to not be significant by itself. It might however be an alternative to the proposed pitch estimator in Subsection 3.4.4, even tough it lags the IF information. An example of a cyclic analysis frequency estimate on a signal with an aircraft present, can be seen in Figure 3.6.

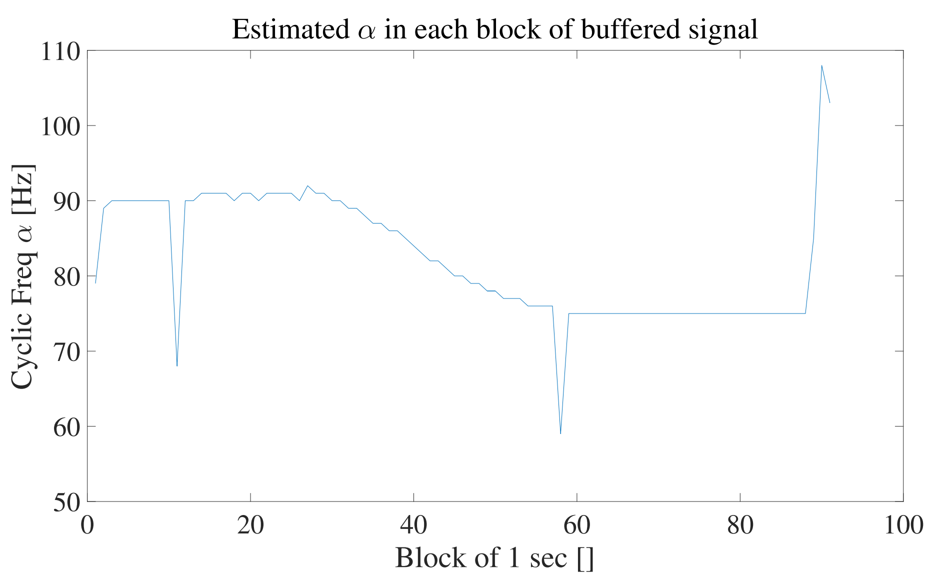


Figure 3.6: Cyclic Analysis estimating Cyclic Frequency α on a signal having an aircraft present in the signal, divided into 1 s blocks including 67 % overlap.

3.3 Chirp Signal Model

Before introducing the proposed chirp pitch estimator a signal model is needed to be described. Starting with the Harmonic Chirp Model (HCM) for Instantaneous Frequency (IF), found in the paper [26], where lth fundamental frequency is

$$\omega_l(n) = l(\omega_0 + bn), \tag{3.21}$$

where ω_0 is the fundamental frequency and b is the chirp coefficient. Eq. 3.21 can be seen as a first-order Taylor approximation [8]. The precision of the model will increase as segments gets shorter, since the linearity assumption fits better. If looking at the instantaneous Doppler frequency in eq. 2.30, it can be seen that ω_0 should be described as the first observed Doppler frequency, which can be expressed as

$$\omega_{Dopp,0} = \gamma + \beta p(n; \tau_c, s), \quad n = 1. \tag{3.22}$$

Now the problem of the chirp b. It has already been mentioned that when aircraft is far away from microphone, nearly infinite distance, one can assume a linear received Doppler frequency curve, which means the chirp model in eq. 3.21 can be used even for the whole signal. However, if the distance is not great, this assumption is not valid for bigger segments and another approach is needed. One is to use a higher order for the Taylor approximation for large segments. If using large segment, b would be a time varying factor, which is defined as

$$b(n) \triangleq \beta p(n; \tau_c, s), \tag{3.23}$$

according to flight parameters. It is clear in eq. 3.23 that the product of β and $p(n; \tau_c, s)$ must be a time varying scaling factor of the initial $\omega_{Dopp,0}$ frequency. The IF model is given as

$$\omega_l(n) = l(\omega_{Dopp,0} + b(n)). \tag{3.24}$$

where $b(n)$ needs to be approximated with a higher order Taylor Approximation, e.g. 3rd or 4th order.

Instead of modelling with a 3rd or 4th order Taylor Approximation, we propose to work on smaller segments, which is believed to be a better solution due to the knowledge about the Doppler frequency curve, which will change character, Taylor order, due to distance from aircraft and microphone and therefore no stationary order is optimal. Therefore we propose to divide data into smaller segments and apply the 1-order chirp Taylor approximation seen in eq. 3.21. The segments will have the length $\mathbf{M}$, which typically is $20\,\mathrm{ms}$ to $30\,\mathrm{ms}$ for speech, but in this case a larger segment can be used and it is also needed due to a limitation of an approximation done in the pitch estimation, introduced in Subsection 3.4.2. Observe, we will later on make a test with different segments length, see Subsection 3.4.6.

Inside this block with $n = n_0, ...M$, one can assume linearity of the fundamental frequency and thereby the IF, becomes

$$\omega_l(n) = l(\omega_{Dopp,0} + bn). \tag{3.25}$$

Observe, that the linear chirp coefficient b will be negative due to Doppler Effect. Now, the IP, is needed before having a model. That is the integral of IF.

$$\theta_l(n) = l\left(\frac{1}{2}bn^2 + \omega_{Dopp,0}n\right) + \phi_l, \tag{3.26}$$

from which the signal model for a signal length M can be made

$$s(n) = \sum_{l=1}^{L} a_l(n) e^{jl(\omega_{Dopp,0}n + b/2n^2)}, \tag{3.27}$$

where $a_l(n) = \frac{A_l}{r(n)}(n)e^{j\phi_l}$. Observe the real part of the complex amplitude, $\frac{A_l}{r(n)}$ is a function of time, which is due to the changing distance $r(n)$ between aircraft and microphone and therefore modelled as a spherical sound wave amplitude.

The signal model in eq. 3.27 can be used to produce a synthetic signal with random $\omega_{Dopp,0}$, b and ϕ_l values in a decided range, which will be used for testing the estimator proposed in Section 3.4.4. This will prove the performance of the estimator on a 1-order chirp signal model, which can be assumed on small blocks of the true signal of interest.

3.4 Pitch Estimator

The observed signal $x(n)$ introduced in section 2.1 contains the sinusoidal Doppler frequency part $s(n)$, a colored noise part $e(n)$ from which the destructive interference frequencies will be made and a Gaussian noise $v(n)$ part with covariance matrix $\mathbf{Q}$. However, since there is no sinusoidal part in $e(n)$ we will here treat the observed signal $x(n)$ as a sinusoidal part $s(n)$ and a noise part $v(n)$ which will be modelled as white Gaussian noise.

$$\mathbf{x}(n) = \mathbf{s}(n) + \mathbf{v}(n) \tag{3.28}$$

In the following we will assume complex signal, which gives simpler equations. If having a real valued signal, the complex signal can be obtained by using the Hilbert Transform [21]. The problem is at the form of single pitch estimation, which we solve as proposed in the book [7] for the harmonic case. However, due to the time-varying Doppler frequency, a chirp coefficient is added to the model, on which an estimator in proposed in the paper [8].

We have already claimed the best solution to describe the observed signal is to divide it into sub-vectors having length M.

$$\mathbf{x} = [x(n_0)\ x(n_0+1)\ ...\ x(n_0+M-1)]^T \tag{3.29}$$

where T is the transpose. Observe, since the full signal is divided into segments, the first time index in the sub-vector, should be denoted n_0 and for obtaining minimum estimation error one should choose $n_0 = -(M-1)/2$ and M is odd [10]. The lth amplitude will normally be written as

$$\alpha_l = A_l e^{j\phi_l}, \tag{3.30}$$

where A is the real amplitude, ϕ is the phase. Observe, we here neglect the spherical wave theory when deriving the estimator. Next, a frequency matrix representing the signal can be written in Vandermonde form. The Vandermonde frequency matrix is described by the radian frequency ω_0 and the chirp coefficient b

$$\mathbf{Z} = [\mathbf{z}(\omega_0, b)\ \mathbf{z}(2\omega_0, 2b)\ ...\ \mathbf{z}(L\omega_0, Lb)] \tag{3.31}$$

where $\mathbf{z}$ follows

$$\mathbf{z}(l\omega_0, lb) = \begin{bmatrix} e^{j(\frac{1}{2}bln_0^2+\omega_0 ln_0)} \\ e^{j(\frac{1}{2}bl(n_0+1)^2+\omega_0 l(n_0+1)} \\ \vdots \\ e^{j(\frac{1}{2}bl(n_0+M-1)^2+\omega_0 l(n_0+M-1)} \end{bmatrix} \tag{3.32}$$

The unknown parameters can be written in vector-form as

$$\boldsymbol{\theta} = [\omega_0\ \alpha_l\ \phi_l\ b\ ...\ \alpha_L\ \phi_L]^T. \tag{3.33}$$

3.4.1 Maximum Likelihood Estimator based on Harmonic Model

It is clear that the Doppler frequencies are not stationary, and thereby a plain Harmonic Model (HM) will not be sufficient, as described in Section 3.3. However, one suitable solution for estimating a chirp pitch, is actually to do an initial pitch estimate based on the harmonic model and afterwards using an iterative approach for estimating the true pitch and chirp [8]. Therefore the following is based on the HM statistical pitch estimate approach presented in the book [7]. Observe the HM is equal to HCM when chirp $b = 0$.

Since focus is the estimation of the time-varying Doppler frequency, the signal in eq. 3.28 can be written as

$$\mathbf{x}(n) = \mathbf{s}(n) + \mathbf{v}(n) = \mathbf{Z} \circ \boldsymbol{\alpha} + \mathbf{v}(n), \tag{3.34}$$

where $\boldsymbol{\alpha}$ is a matrix with the L columns, each being $[A_l(n_0)e^{j\phi_l} \; A_l(n_0+1)e^{j\phi_l} \; ... \; A_l(n_0+M-1)e^{j\phi_l}]^T$ and the product $\mathbf{Z} \circ \boldsymbol{\alpha}$ is a element-wise multiplication (Hadamard product). The noise part $\mathbf{v}(n)$ is Gaussian noise having the covariance matrix $\mathbf{Q}$. However, in this work the pitch is the main feature to be estimated and we will not focus on time-varying amplitude and thereby the Hadamard product is neglected and eq. 3.34 becomes

$$\mathbf{x}(n) = \mathbf{s}(n) + \mathbf{v}(n) = \mathbf{Z}\boldsymbol{\alpha} + \mathbf{v}(n), \tag{3.35}$$

where $\boldsymbol{\alpha}$ is a vector with length L due to the number of harmonics.

The estimator proposed here is a Maximum Likelihood (ML) estimator which work on a signal sub-vector, as introduced in eq. 3.29. Now the vector is given at time index n instead of n_0, since we here do not focus on which time indexes to use.

$$\mathbf{x}(n) = [x(n) \cdots x(n+M-1)]^T \tag{3.36}$$

The likelihood function is expressed as a function of the unknown parameters $\boldsymbol{\theta}$ introduced in eq. 3.33

$$p(\mathbf{x}(n); \boldsymbol{\theta}) = \frac{1}{\pi^M \det(\mathbf{Q})} e^{-\mathbf{v}^H(n)\mathbf{Q}^{-1}\mathbf{v}(n)}. \tag{3.37}$$

Having sub-vectors with the length M, and a full signal length N, then $G = N - M + 1$ sub-vectors are needed to describe the full signal $\{\mathbf{x}(n)\}_{n=0}^{G-1}$. Now, an assumption is introduced, which is only valid for the HM and not the HCM. If $\mathbf{s}(n)$ is stationary and $\mathbf{v}(n)$ is independent and equally distributed over n, then the likelihood for all sub-vectors can be written

$$\begin{aligned} p(\{\mathbf{x}(n)\}; \boldsymbol{\theta}) &= \prod_{n=0}^{G-1} p(\mathbf{x}(n); \boldsymbol{\theta}) \\ &= \frac{1}{\pi^{MG} \det(\mathbf{Q})^G} e^{-\sum_{n=0}^{G-1} \mathbf{v}^H(n)\mathbf{Q}^{-1}\mathbf{v}(n)}, \end{aligned} \tag{3.38}$$

which can be simplified by apply the logarithm of the likelihood.

$$\boldsymbol{\mathcal{L}}(\boldsymbol{\theta}) = -GM \ln \pi - G \ln \det(\mathbf{Q}) - \sum_{n=0}^{G-1} \mathbf{v}^H(n)\mathbf{Q}^{-1}\mathbf{v}(n). \tag{3.39}$$

The unknown, but deterministic parameters can now be estimated by maximizing the log-likelihood as

$$\hat{\boldsymbol{\theta}} = \arg\max \boldsymbol{\mathcal{L}}(\boldsymbol{\theta}) \tag{3.40}$$

which becomes the ML estimator.

It might be needed to estimate the noise $\hat{\mathbf{v}}(n)$ which is constructed from the estimated parameters

$$\hat{\mathbf{v}}(n) = \mathbf{x}(n) - \hat{\mathbf{s}}(n). \tag{3.41}$$

How to estimate the noise covariance matrix is shown in the book [7], but for now a white signal model is believed to be sufficient.

3.4.2 Estimator in White Noise

For white noise, the noise covariance matrix is known, and it is reduced to the scaled diagonal matrix $\mathbf{Q} = \sigma^2\mathbf{I}$. Furthermore, there is no need for sub-vectors, which means $M = N$ and also thereby $G = 1$, observe again this is not true for HCM. As a result of that the log-likelihood is now on a simpler form

$$\boldsymbol{\mathcal{L}}(\boldsymbol{\theta}) = -N\ln(\pi) - N\ln(\sigma^2) - \frac{1}{\sigma^2}||\hat{\mathbf{v}}||_2^2. \tag{3.42}$$

This method shown in eq. 3.42 where the goal is to minimize 2-norm is referred to as being the NLS Method. It should be noted that the estimate depends on the order L. For at given L the noise is a variance, estimated as

$$\hat{\sigma}^2 = \frac{1}{N}||\mathbf{x} - \mathbf{\Pi}_Z\mathbf{x}||_2^2, \tag{3.43}$$

where $\mathbf{\Pi}_Z$ is defined as

$$\mathbf{\Pi}_Z = \mathbf{Z}(\mathbf{Z}^H\mathbf{Z})^{-1}\mathbf{Z}^H. \tag{3.44}$$

which means $\mathbf{\Pi}_Z\mathbf{x}$ is the projection of $\mathbf{x}$ onto the range of $\mathbf{Z}$. In fact, the $\mathbf{\Pi}_Z$ can be approximated due to the fact that complex sinusoids are asymptotically orthogonal. The approximation gives

$$\lim_{N\to\infty} N\mathbf{\Pi}_Z = \lim_{N\to\infty} N\mathbf{Z}(\mathbf{Z}^H\mathbf{Z})^{-1}\mathbf{Z}^H \quad = \mathbf{Z}\mathbf{Z}^H. \tag{3.45}$$

Observe for a given N, the approximation gets worse as ω_0 gets smaller, and might end up being a poor approximation when having very low ω_0.

By using the approximation in eq. 3.45, the variance estimate is now an approximately Nonlinear Least Squares (aNLS) and is written as

$$\hat{\sigma}^2 = \frac{1}{N}||\mathbf{x} - \frac{1}{N}\mathbf{Z}\mathbf{Z}^H\mathbf{x}||_2^2. \tag{3.46}$$

The aNLS in eq. 3.46 also simplifies the log-likelihood function in eq. 3.42, since it only depends on ω_0 now, to

$$\boldsymbol{\mathcal{L}}(\omega_0) \approx -N\ln(\pi) - N\ln(\hat{\sigma}^2) - N. \tag{3.47}$$

Finally, the pitch estimator for $\hat{\omega}$ is now an aNLS and described by

$$\begin{aligned}\hat{\omega}_0 &= \arg\max_{\omega_0} \mathbf{x}^H\mathbf{\Pi}_Z\mathbf{x} \\ &\approx \arg\max_{\omega_0} \mathbf{x}^H\mathbf{Z}\mathbf{Z}^H\mathbf{x} \\ &= \arg\max_{\omega_0} ||\mathbf{Z}^H\mathbf{x}||_2^2.\end{aligned} \tag{3.48}$$

3.4.3 Implementing the aNLS as Harmonic Summation

Now the aNLS pitch estimator given in eq. 3.48 is wanted to be implemented. For doing so, the $\mathbf{Z}$ frequency matrix is convenient since it can be implemented in a computational efficient matter, by only using one Fast Fourier Transform (FFT).

$$\begin{aligned} ||\mathbf{Z}^H\mathbf{x}||_2^2 &= \sum_{l=1}^{L} |\sum_{n=0}^{N-1} x(n)e^{-j\omega_0 ln}|^2 \\ &= \sum_{l=1}^{L} |X(\omega_0 l)|^2. \end{aligned} \tag{3.49}$$

The pitch estimate $\hat{\omega}_0$ is then found by

$$\hat{\omega}_0 = \arg\max_{\omega_0} \sum_{l=1}^{L} |X(\omega_0 l)|^2, \tag{3.50}$$

where L is a fixed number of harmonics and X is the FFT the $\mathbf{x}$. Observe, the implementation will zero-pad the data to obtaining a wanted number of FFT points, N length of Fast Fourier Transform (NFFT).

3.4.4 Chirp Pitch Estimator

The aNLS estimator is only valid for a stationary pitch, which is almost never the case and indeed not when observing a Doppler changing pitch as in this study. One part of the solution is to take non-stationary pitch into account, by using the HCM in eq. 3.27 seen in Section 3.3. The proposed estimator, working on a such signal, follows the one in the paper [8].

HCM Pitch Estimator

Again a ML estimator is wanted in the assumption of white noise. The estimation of the parameters thereby follow a NLS estimator

$$\{\hat{\mathbf{a}}, \hat{b}, \hat{\omega}_0\} = \arg\min_{\mathbf{a},b,\omega_0} ||\mathbf{x} - \mathbf{Z}\mathbf{a}||^2. \tag{3.51}$$

If neglecting the time-varying amplitudes and say that the amplitudes are of no interest, these can be substituted by a least square fit, which changes the estimator in eq. 3.51 to

$$\{\hat{b}, \hat{\omega}_0\} = \arg\min_{b,\omega_0} ||\mathbf{x} - \mathbf{Z}(\mathbf{Z}^H\mathbf{Z})^{-1}\mathbf{Z}^H\mathbf{x}||^2, \tag{3.52}$$

which results in a 2-D optimization problem for the non-linear parameters, $\hat{b}, \hat{\omega}_0$. One way to overcome the computational problem in the estimator in eq. 3.52 is to first introduce the orthogonal projection matrix

$$\mathbf{\Pi}(\omega_0, b) = \mathbf{Z}(\mathbf{Z}^H\mathbf{Z})^{-1}\mathbf{Z}^H, \tag{3.53}$$

and the orthogonal complement one

$$\mathbf{\Pi}^{\perp}(\omega_0, b) = \mathbf{I} - \mathbf{\Pi}(\omega_0, b), \tag{3.54}$$

from which a simple iterative method for doing the 2-D optimization problem can be introduced

$$\hat{b}^{(i)} = \arg\min_{b} = \mathbf{x}^H\mathbf{\Pi}^{\perp}(\hat{\omega}_0^{(i-1)}, b)\mathbf{x} \tag{3.55a}$$

$$\hat{\omega}_0^{(i)} = \arg\min_{\omega_0} = \mathbf{x}^H\mathbf{\Pi}^{\perp}(\omega_0, \hat{b}^{(i)})\mathbf{x} \tag{3.55b}$$

The iterations in eq. 3.55a and eq. 3.55b are repeated until convergence, which in practice will be a value defined from the cost function. Observe, that in eq. 3.55a an initial pitch estimate is needed, which is where the aNLS pitch estimator will be used. By testing on synthetic signals with known parameters, the improvement of the proposed HCM pitch estimator can easily be shown, since both estimates are done in the process here. The chirp factor b can be assumed to be zero as a start value. The grid area for eq. 3.55a and eq. 3.55b should start with a rough precision. Around the previous estimate a search area is used, in which the minimum is wanted to be found. Afterwards a dichotomous search is used, which has the constrain that it can only work in the convex region around the minimum of the cost function.

Dichotomous Search

The dichotomous search method is described in the book [4]. The general problem to solve is

$$\min F = f(x), \tag{3.56}$$

which means having a function of one variable; the goal is to find the variable that minimizes the function. For finding one such minimum, the function needs to be *unimodal*, which means only one minimum is present in a given area. Having a search area expressed as an interval, $[x_{Lower}, x_{Upper}]$ or in short $[x_L, x_U]$ in which a minimizer x^* is present. The search method approach for finding a minimum is to reduce this search area, until having a very narrow search area, in which the minimum is assumed to be located in the middle of the minimized search area. One such method for solving a minimization problem, is the dichotomous search approach. The interval $[x_L, x_U]$ is the range of uncertainty and the minimizer x^* can be found by reducing the range of uncertainty until a small range is left. The approach uses a number of values of the function $f(x)$. If the function value is known at one point, x_a, inside the range $[x_L, x_U]$ then the minimizer x^* will most likely be in the area x_L to x_a or in the area x_a to x_U. It is clear this is not sufficient information for doing any reduction of the range. For doing any reduction, two values made from two points, x_a and x_b in the function, are needed. This can be done due to the following

$$f(x_a) < f(x_b) \tag{3.57a}$$
$$f(x_a) > f(x_b) \tag{3.57b}$$
$$f(x_a) = f(x_b). \tag{3.57c}$$

If having the case seen in eq. 3.57a, the minimizer x^* might be located in range $x_L < x^* < x_a$ or $x_a < x^* < x_b$, which means $x_L < x^* < x_b$. Since the function only has one minimum, the range $x_b < x^* < x_U$ can be neglected. For the case in eq. 3.57b the opposite of the just described can be used. If having the case seen in eq. 3.57c, $x_a < x^* < x_b$, which means both $x_L < x^* < x_b$ and $x_a < x^* < x_U$ needs to be satisfied. For reducing the range, the dichotomous search evaluates a function a two points $x_a = x_1 - \epsilon/2$ and $x_b = x_1 + \epsilon/2$, where ϵ is a small (positive) value, for example 0.01. If having the case seen in eq. 3.57a the range x_L to $x_1 + \epsilon/2$ is selected or if having the case seen in eq. 3.57b, the range $x1 - \epsilon/2$ to x_U is selected. If having the case seen in eq. 3.57c any of the range can be used. One example of the the reduction this approach can perform is that if $x_1 - x_L = x_U - x_1$. Then this means $x_1 = (x_L + x_U)/2$ and the range is already half the starting range. At the next iteration the function will be evaluated at $x_2 - \epsilon/2$ and $x_2 + \epsilon/2$, where again x_2 is located at the center of the range. For each iteration the range of uncertainty is reduced by half, which reduces the uncertainty to

$$I_k = (\frac{1}{2})^2 I_0, \tag{3.58}$$

where I_0 is $x_U - x_L$. By using the dichotomous search along with the iteration steps in eq. 3.55a and eq. 3.55b, we have our proposed HCM pitch estimator.

3.4.5 Maximum A Posteriori (MAP)

Order selection is often seen to be the key of solving estimation problems. One method for overcome this issue, is to apply the Maximum A Posteriori (MAP), which is here introduced as it is in the book [7]. The following is based on a single pitch source.

Starting with the model candidate vector $\mathbb{Z}_q = 0, 1, ..q-1$, indexed by $\boldsymbol{\mathcal{M}}_m$. Now the goal is to maximize the a posteriori probability, when having the observed signal $\mathbf{x}$:

$$\widehat{\boldsymbol{\mathcal{M}}} = \arg \max_{\boldsymbol{\mathcal{M}}_m, m \in \mathbb{Z}_q} p(\boldsymbol{\mathcal{M}}_m|\mathbf{x}), \tag{3.59}$$

which can be rewritten with Bayes rule

$$\widehat{\boldsymbol{\mathcal{M}}} = \arg \max_{\boldsymbol{\mathcal{M}}_m, m \in \mathbb{Z}_q} \frac{p(\mathbf{x}|\boldsymbol{\mathcal{M}}_m)p(\boldsymbol{\mathcal{M}}_m)}{p(\mathbf{x})}. \tag{3.60}$$

Assume an uniform prior for all models, meaning $p(\mathbf{x}_k)$ becomes a constant when $\mathbf{x}$ is observed. By that, the Maximum A Posteriori (MAP) model reduces to the following likelihood function seen as a function of $\boldsymbol{\mathcal{M}}_m$

$$\widehat{\boldsymbol{\mathcal{M}}} = \arg \max_{\boldsymbol{\mathcal{M}}_m, m \in \mathbb{Z}_q} p(\mathbf{x}|\boldsymbol{\mathcal{M}}_m). \tag{3.61}$$

The next step is to integrate out the unknown parameters $\boldsymbol{\theta}$, which in fact is depending on $\boldsymbol{\mathcal{M}}_m$ but fore simplification here neglected. The integration becomes

$$p(\mathbf{x}|\boldsymbol{\mathcal{M}}_m) = \int_{\Theta} p(\mathbf{x}|\boldsymbol{\theta}, \boldsymbol{\mathcal{M}}_m)p(\boldsymbol{\theta}|\boldsymbol{\mathcal{M}}_m)d\boldsymbol{\theta} \tag{3.62}$$

No analytic solution can be given for eq. 3.62, which is why in [7] a Laplace integration is proposed to be used with the assumption that it will give the most significant peak in the likelihood function around the maximum likelihood estimated $\hat{\boldsymbol{\theta}}$ when N is high. By that, eq. 3.62 can now be rewritten to be equal to

$$(2\pi)^{D/2} \det\left(\widehat{\mathbf{H}}\right)^{-1/2} p(\mathbf{x}|\hat{\boldsymbol{\theta}}, \boldsymbol{\mathcal{M}}_m)p(\hat{\boldsymbol{\theta}}|\boldsymbol{\mathcal{M}}_m), \tag{3.63}$$

where D is the number of parameters and $\widehat{H}$ is the Hessian of the log-likelihood function which is evaluated at the $\hat{\boldsymbol{\theta}}$. The $\widehat{H}$ is given by

$$\widehat{\mathrm{H}} = -\frac{\partial^2 \ln(p(\mathbf{x}|\boldsymbol{\theta}, \boldsymbol{\mathcal{M}}_m)}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^T}\Bigg|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} \tag{3.64}$$

Next, the logarithm is taken of eq. 3.63. Also, the two terms $\boldsymbol{\mathcal{O}}(1)$ and $\frac{D}{2}\ln(2\pi)$ are ignored when N is large. All that gives

$$\widehat{\boldsymbol{\mathcal{M}}} = \arg \max_{\boldsymbol{\mathcal{M}}_m, m \in \mathbb{Z}_q} -\ln p(\mathbf{x}|\hat{\boldsymbol{\theta}}, \boldsymbol{\mathcal{M}}_m) + \frac{1}{2}\ln \det(\widehat{\mathrm{H}}) \tag{3.65}$$

where the first term in eq. 3.65 is the log-likelihood, on which a penalty term is needed to be added, which is the last part. By using eq. 3.65 it is possible to select between various models and orders.

What is now needed is a criterion for selecting the model order for the signal introduced in eq. 3.27 in Section 3.3. Furthermore, a pitch detection criterion is shown. Starting by observing eq. 3.64, which is related to the Fisher matrix

$$\widehat{\mathrm{H}} \approx -\mathrm{E}\left\{\frac{\partial^2 \ln(p(\mathbf{x}|\boldsymbol{\theta}))}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^T}\right\}\Bigg|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} \tag{3.66}$$

where the diagonal terms can be found by using the normalized matrix:

$$\mathbf{K}_N = \begin{bmatrix} N^{-3/2} & \mathbf{0} \\ \mathbf{O} & N^{-1/2}\mathbf{I} \end{bmatrix} \tag{3.67}$$

where $\mathbf{I}$ is a $2L \times 2L$ identity matrix. The information given in the matrix in eq. 3.67 is on the first diagonal entry due to the fundamental frequency and the remaining is due to L amplitudes and phases. Now the determinant of the Hessian in eq. 3.65 is calculated as

$$\det(\widehat{\mathrm{H}}) = \det(\mathbf{K}_N^{-2})\ \det(\mathbf{K}_N\widehat{\mathrm{H}}\mathbf{K}_N). \tag{3.68}$$

A simplification can be done, since the last term in eq. 3.68 is equal to $\boldsymbol{\mathcal{O}}(1)$. Now taking the logarithm, the result is

$$\begin{align} \ln\det(\widehat{\mathrm{H}}) &= \ln\det(\mathbf{K}_N^{-2})\ \ln\det(\mathbf{K}_N\widehat{\mathrm{H}}\mathbf{K}_N) \tag{3.69}\\ &= \ln\det(\mathbf{K}_N^{-2}) + \boldsymbol{\mathcal{O}}(1) \tag{3.70}\\ &= 3\ln N + 2\,L\,\ln N + \boldsymbol{\mathcal{O}}(1) \tag{3.71} \end{align}$$

When having additive white complex Gaussian noise, the log-likelihood in eq. 3.65 is $N\ \ln\ \sigma_k^2$. Next, σ^2 needs to be replaced by the estimate for each candidate order L, which is denoted as $\hat{\sigma}^2(L)$. By substitute eq. 3.71 into eq. 3.65 the selecting order criteria is

$$\hat{L} = \arg\max_L N\ln\hat{\sigma}^2(L) + \frac{3}{2}\ln N + \hat{L}\ln N \tag{3.72}$$

where the first term is the log-likelihood and the last two terms are the penalty terms.

As a pitch detection control, which means a check for if a harmonic is present in the signal, eq. 3.72 should be compared to log-likelihood of a zero order model

$$N\ln\hat{\sigma}^2(0) < N\ln\hat{\sigma}^2(L) + \frac{3}{2}\ln N + \hat{L}\ln N \tag{3.73}$$

where $\hat{\sigma}^2(0)$ is the variance of the observed signal.

3.4.6 Chirp Pitch Estimator Test

In the following we will perform a two tests for proving the improvements the HCM chirp pitch estimator adds compared to the aNLS, when estimating pitch in a chirp signal. The results for each test will presented as a RMSE for each estimator; the aNLS and the HCM. In both tests the signal used for testing is produced with eq. 3.27. The tests are performed using a Monte Carlo approach, which means a test is repeated many times and each time randomly parameter values are used. In the following test these random parameter values are f_0, b and ϕ_l within a decided range of $f_0\ \epsilon\ [50, 100]$ Hz, $b\ \epsilon\ [-200, 200]$ Hz^2 and $\phi_l\ \epsilon\ [0, 2\pi]$.

Test 1 - RMSE as Function of SNR

The purpose of this test is to obtain the RMSE for each estimator, where RMSE will be a function of SNR. Synthetic test signal is produced with eq. 3.27, on which additive white Gaussian noise is added. The following parameters gives the entire signal used in the test:

- Sample frequency = $8\,\mathrm{kHz}$.
- L = 10.
- NFFT = 2^{20} for aNLS including zero-padding.
- Duration = $50\,\mathrm{ms}$.
- Number of Monte Carlo iterations = 100.
- SNR in the range: $[-10 : 100]$ dB, with steps of 10 dB.

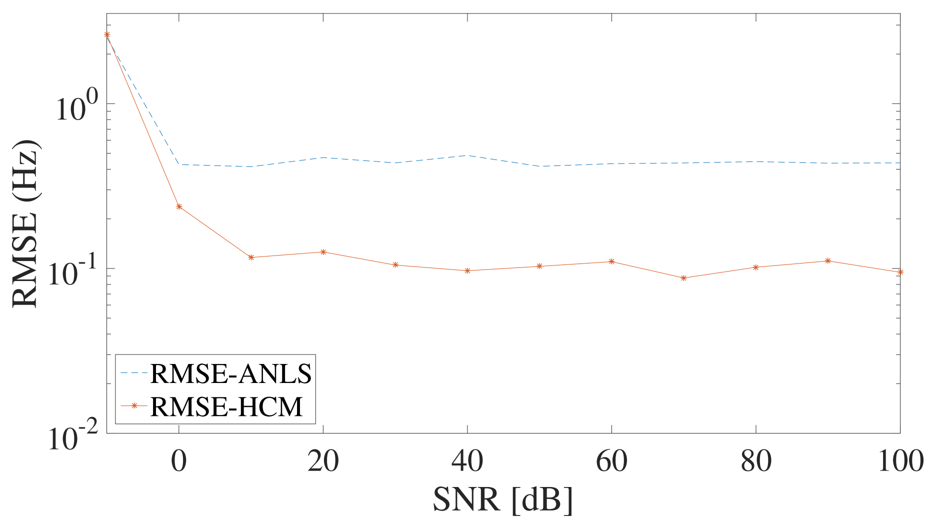


Figure 3.7: RMSE as function of SNR in the range $[-10 : 100]$ dB on a chirp signal.

It is clear from Figure 3.7, that the HCM chirp pitch estimator outperforms the aNLS pitch estimator based on HM, when it comes to RMSE value on a chirp signal.

Test 2 - RMSE as Function of N

The purpose of this test is to obtain the RMSE for each estimator, where RMSE will be a function of N. Synthetic test signal is produced with eq. 3.27, on which additive white Gaussian noise is added. The following parameters gives the entire signal used in the test:

- Sample frequency = $8\,\mathrm{kHz}$.
- L = 10.
- NFFT = 2^{20} for aNLS including zero-padding.
- Number of samples N in the range $[100 : 500]$.
- Number of Monte Carlo iterations = 100.
- SNR =$100\,\mathrm{dB}$.

Again, it is clear from Figure 3.8, that the HCM chirp pitch estimator outperforms the aNLS pitch estimator based on HM when it comes to RMSE value on a chirp signal.

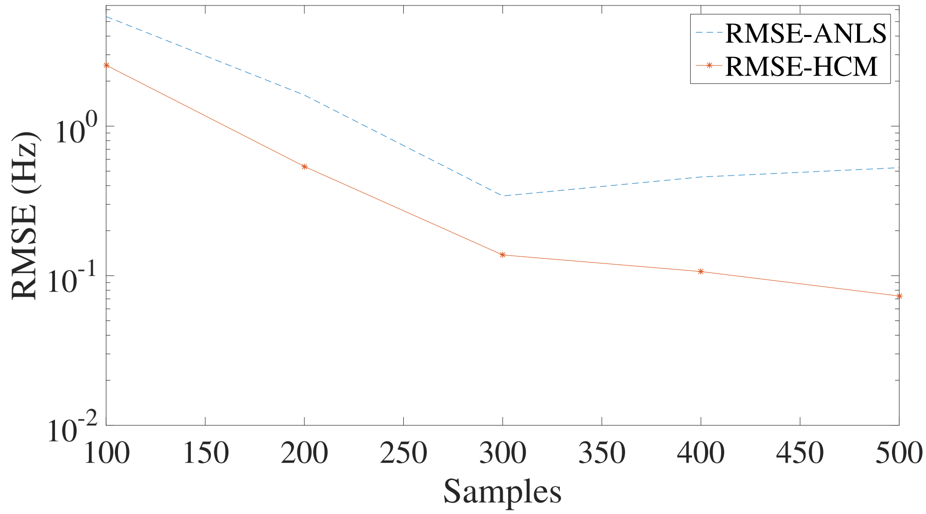


Figure 3.8: RMSE as function of N in the range $[100:500]$ on a chirp signal.

3.5 Flight Parameter Estimation

If having an estimator which can estimate the IFs in a signal on the model in eq. 3.27, the next is to use all the IFs to fit the model in eq. 2.30, from which the flight parameter can be estimated. These parameters are: The true f_0 from the source, velocity v, emission time of CPA, τ_c, and slant range for that time, R_c. The three first parameters in $[f_0, v, R_c, \tau_c,]$ are also equal to γ,β and s introduced in Section 2.3. Brian G. Ferguson in [13] proposes a NLS between the model eq. 2.30 in Section 2.3 and the IFs, $\hat{f}$, which we will estimate with the proposed HCM chirp pitch estimator in Subsection 3.4.4.

3.5.1 Flight Parameter Estimator

The proposed NLS estimator minimizes the sum of the squared errors between the IF estimates and the predicted values according to the model in eq. 2.30. The estimator ends up the following 4 dimensional minimization problem

$$\{\hat{\gamma}, \hat{\beta}, \hat{\tau}_c, \hat{s}\} = \arg \min_{\gamma', \beta', \tau_c', s'} \sum_{k=1}^{K} \left(\gamma' + \beta' p(n_k; \tau_c', s') - \hat{f}(n_k)\right)^2, \tag{3.74}$$

where K is the number of IF estimates, denoted as $\hat{f}(n_k)$ and n_k is the indexing of IF estimates. This 4 dimensional maximization problem in eq. 3.74 can be reduced to a 2 dimensional maximization problem

$$\{\hat{\tau}_c, \hat{s}\} = \arg \min_{\tau_c', s'} \frac{|\sum_{k=1}^{K} \left(\hat{f}(n_k) - \overline{\hat{f}}\right) p(n_k)|^2}{\sum_{k=1}^{K} \left(p(n_k) - \overline{p}\right)^2} \tag{3.75a}$$

$$\hat{\beta} = \frac{|\sum_{k=1}^{K} \left(\hat{f}(n_k) - \overline{\hat{f}}\right) \hat{p}(n_k)|^2}{\sum_{k=1}^{K} \left(\hat{p}(n_k) - \overline{\hat{p}}\right)^2} \tag{3.75b}$$

$$\hat{\gamma} = \overline{\hat{f}} - \hat{\beta}\overline{\hat{p}}, \tag{3.75c}$$

where $\overline{\hat{f}}$ is the mean of IF $\hat{f}$ estimates, $p(n_k)$ is $p(n_k; \tau_c', s')$ given in eq. 2.31d in Section 2.3, $\overline{p}$ is the mean of all $p(n_k)$, $\hat{p}(n_k)$ is $p(n_k; \hat{\tau}_c, \hat{s})$ and $\overline{\hat{p}}$ is the mean of $\hat{p}(n_k)$. When the four estimates in eqs. (3.75a) to (3.75c) are found, the estimated flight parameters can be found with

$$\hat{v} = -(\hat{\beta}/\hat{\gamma})c \tag{3.76a}$$

$$\hat{f}_0 = \hat{\gamma}(1 - \hat{v}^2/c^2) \tag{3.76b}$$

$$\hat{R}_c = \hat{s}\hat{v}c/(c^2 - \hat{v}^2)^{1/2}, \tag{3.76c}$$

which are all derived from the Doppler eqs. (2.31a) to (2.31d) in Section 2.3. Next, we want to test this flight parameter estimator on a synthetic produced signal with known parameter values.

3.5.2 Synthetic Signal based on Flight Parameters

Using eq. 2.24 or eq. 2.18 a synthetic signal can be made of an aircraft passing by. A spectrogram of one such signal can be seen in Figure 3.9. Observe the Doppler Effect is easily seen. In Figure 3.9, the following parameter values where used

- $v = 40\,\frac{\text{m}}{\text{s}}$
- $c = 340\,\frac{\text{m}}{\text{s}}$
- $h_t = 200\,\text{m}$
- $h_m = 1.5\,\text{m}$
- Starting horizontal distance to CPA $= 280\,\text{m}$
- Stop horizontal distance to CPA $= 280\,\text{m}$
- $f_s = 16\,\text{kHz}$
- Reception time is: $\mathbf{t} = 0 : \frac{1}{f_s} : \frac{(280+280)}{v}$
- $d_c = 10\,\text{m}$
- $\tau_c = \frac{t(end)}{2}$, where $t(end)$ is basically the time from start to stop horizontal position.
- $L = 5$

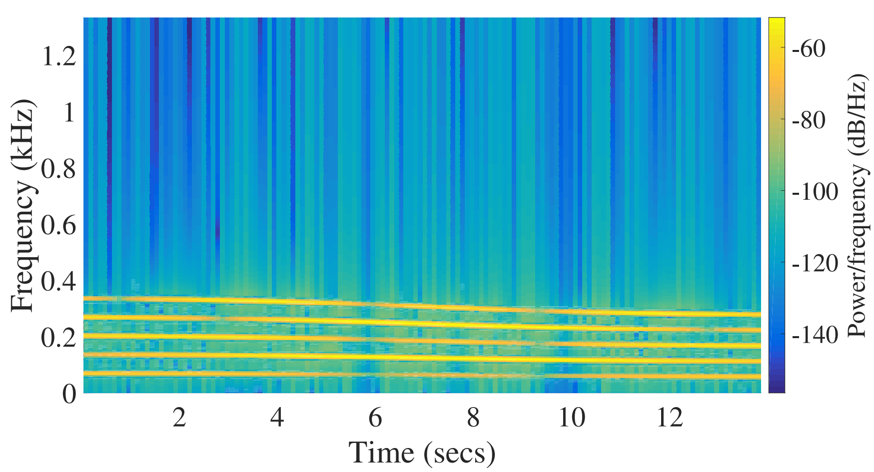


Figure 3.9: Spectrogram based on the signal model in eq. 2.18. The Doppler Effect is easily seen in the Spectrogram after a down sampling with a factor of 6.

3.5.3 Model The Destructive Interference Frequencies in the Noise Radiated from Aircraft

The signal model seen in eq. 2.18 clearly only contains the $s(n)$ part of the observed signal $x(n)$. In the visualization of the signal in Figure 3.9, it becomes clear that Figure 2.1 and Figure 3.9 are not fully similar, yet. The next step is therefore to model the destructive inference due to ground-reflection. The good news, is that no new parameters are needed to be found, since all that is needed here is the time difference between emission time for the direct path and reflected path, which gives the time-varying delay $D(n)$.

$$D(n) = \tau_d(n) - \tau_r(n) \tag{3.77}$$

Observe, the emission time of the direct path will be larger value than emission time of the reflected path. The delay seen in eq. 3.77 will end up being a fractional delay, which gives the following issues:

1. Making the signal $e(n)$ online is not as straight forward as for the source signal $s(n)$.
2. The delay in time ends up being a fractional sample delay due to sample frequency, which means one need to do time domain fractional delay e.g. interpolation or going to the frequency domain. In the frequency domain one could make a fractional delay by multiplying a response equal to a linear phase filter onto the signal. The linear phase needs a gradient equal to the wanted delay.

Instead of trying to making the signal online, we propose to do a noisy signal, $e_{color}(n)$, offline and then afterwards apply the fractional delay in the frequency domain. Based on visible observations of spectrograms, we will use pink noise.

We also propose a simple approach for doing the fractional delay in frequency domain. If having fractional sample 45.3, then take sample 45 and add a phase change, equal to 0.3 samples. Implementation in MATLAB will be to do a FFT on a small part of the signal, calculate absolute value and angle of the FFT and then add the phase change to the angle of the FFT signal. See Algorithm 1 for the implementation.

Before starting on the fractional delay, we delay the wanted observed noise signal $e(n)$, equal to the amount of the delay for the first received wave. This is simply done by ceiling the time delay to an index idx_0, and start to construct the wanted signal, $e(n)$, with idx_0 of zeros. Now the rest of the signal is needed to be done as described in the Algorithm 1

3.5.4 Destructive Interference Frequency Signal Model

The total noise signal will be

$$e(n) = e(n)/r_d(n) + e(n - D(n))/r_r(n), \tag{3.78}$$

where $D(n)$ is added as described in Algorithm 1.

3.5.5 Realistic Synthetic Signal

In Section 2.1 we introduced the observed signal as

$$x(n) = s(n) + e(n) + v(n), \tag{3.79}$$

where the part $s(n)$ has been derived with the HCM described in segments of length M in Section 3.3.

Algorithm 1 Implement $D(n)$ as a fractional delay in the frequency domain

1. Floor the delay-time of the index n, e.g. index $45.3 = 45$, call it idx.
2. A small range of samples are needed, so at index n, idx is subtracted giving the lower index. Add 2 to that index and the range is now 3 samples. The fractional delay that is wanted is in between index 1 and 2.
3. $Y = FFT(e_{color}(range))$
4. Apply the fractional delay as a phase. $|Y|e^{\phi+\phi_{delay}}$
5. $y = iFFT(Y)$
6. Construct the signal $e(n)$ at n, by using first sample in y, the n sample of noise signal e_{color}, and divide respectively with $r_r(n)$ and $r_d(n)$.
7. Continue until all delays have been applied

The total observed signal $x(n)$ is now possible to construct, and if using the same parameter values as earlier on and do at down-sample with a factor of 6, the resulting spectrogram will look like:

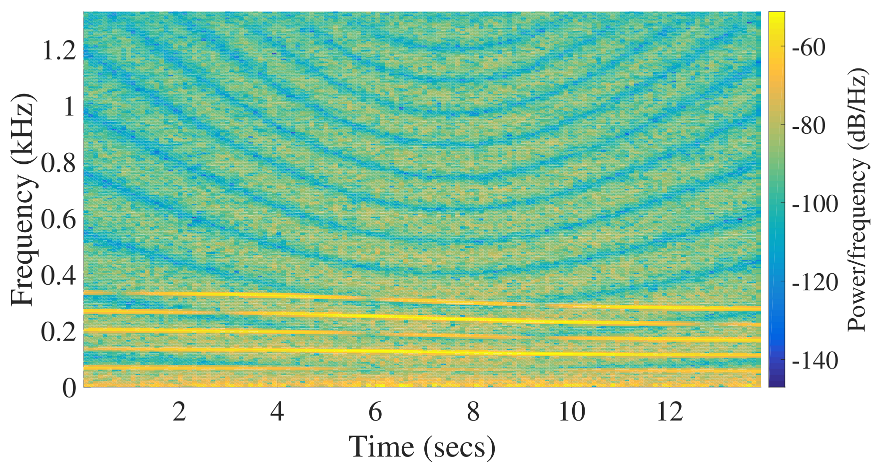


Figure 3.10: The synthetic observed signal $x(n)$ including additive white Gaussian noise with $SNR = 100$ has been added.

3.5.6 Flight Parameters Estimation in Synthetic Signal

The produced signal seen in Figure 3.10 has known parameter values and we now use the flight parameters estimator described in eqs. (3.75a) to (3.75c) and eqs. (3.76a) to (3.76c). Observe, the key in the flight parameter estimation is to obtain the IF estimates, which is given directly from the model in 3.21. We want to stress this great advantage of the used HCM chirp pitch estimator. See Figure 3.11 for a comparison of true and estimated IF. Observe, in the HCM chirp pitch estimation the signal has been down sampled with a Q factor at 6, the signal was buffered into blocks of $50\,\mathrm{ms}$, the order was not found with MAP but directly taken from the producing of the synthetic signal. In Table 3.1 the true and estimated flight parameters are shown Table 3.1, where the first 4 parameters are the flight parameters and the last three are parameters from the Doppler model in Section 2.3, which are used for estimate the 4 flight parameters. Observe τ_c and s has been estimated with

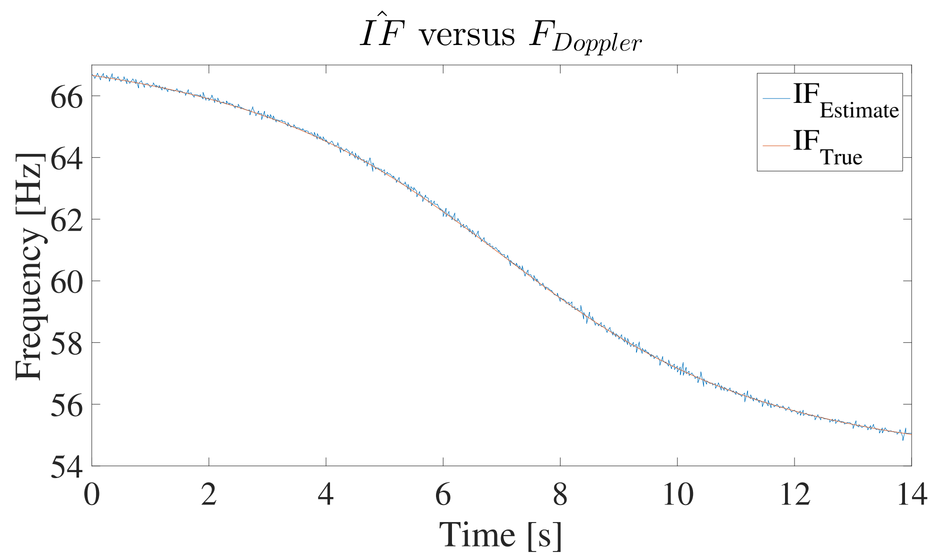


Figure 3.11: IF estimates compared to frequencies from Doppler Model in eq. 2.30.

a rough search, therefore no decimals. It is clear from Table 3.1, that the proposed HCM chirp

	True Value	**Estimated Value**
f_0	$60\,\text{Hz}$	$60.0136\,\text{Hz}$
v	$40\,\text{m/s}$	$40.0459\,\text{m/s}$
R_c	$201.6505\,\text{m}$	$201.6329\,\text{m}$
τ_C	$7\,\text{s}$	$7\,\text{s}$
γ	$60.8421\,\text{Hz}$	$60.8578\,\text{Hz}$
β	$-7.1579\,\text{Hz}$	$-7.1680\,\text{Hz}$
s	$5.0063\,\text{s}$	$5\,\text{s}$

Table 3.1: True flight parameters versus estimated parameters.

pitch estimator together with the flight parameter estimator seems to be able to estimate useful parameters, but a better test is to observe the RMSE, which now follows.

3.5.7 RMSE on Flight Parameters

We have made a small Monte Carlo test for obtaining RMSE values on the estimated flight parameters. The following parameters and intervals are used for testing

- Sample frequency = $16\,\text{kHz}$.
- L = 5.
- Down sampled signal with factor 6.
- Buffered Signal into $50\,\text{ms}$ segments with $67\,\%$ overlap.
- Number of Monte Carlo iterations = 50.
- SNR $200\,\text{dB}$.
- Random aircraft speed in the range : $40\,\frac{\text{m}}{\text{s}}$ to $60\,\frac{\text{m}}{\text{s}}$.
- Random f_0 in the range : $55\,\text{Hz}$ to $65\,\text{Hz}$.

- Random aircraft height in the range : $100\,\mathrm{m}$ to $300\,\mathrm{m}$.
- Other constant parameters: height of microphone $h_m = 1.5\,\mathrm{m}$, horizontal range $400\,\mathrm{m}$, depth distance $d_c = 10\,\mathrm{m}$,
- HCM chirp pitch estimation has maximum 20 iterations.
- s has a rough search, 1 to 10 with step size 0.1.
- τ_c search from 1 to half the total signal length, plus one.

Results

In Table 3.2 the result of the test can be seen. It is clear from Table 3.2, except for $\hat{R}_c$ which needs

	RMSE
f_0	0.2149
v	0.9142
R_c	9.4100
τ_C	0.1345
γ	0.2028
β	0.1619
s	0.0989

Table 3.2: RMSE on flight parameters.

a very precise s, that all RMSE values are acceptable low, and we can conclude our HCM chirp pitch estimator together with the flight parameter estimator performs in an acceptable manner on estimating the flight parameters and we believe it can be used for a classification problem on real signal.

3.6 Classification

The machine learning methods used in this report is described in the book [11], where the authors have made a MATLAB toolbox, PRTools, which is directly used in this report. In Appendix A a small introduction on how to use PRTools in MATLAB is given. The used classifiers are

1. Quadratic Distance Classifier (QDC), Bayes Normal 2
2. K Nearest Neighbour Classifier (KNNC)

where the first one can be referred to as a parametric learning approach and the second one a non-parametric learning approach.

The QDC is based on the basic Bayes decision function with a uniform cost function. It is a MAP probability classifier but in fully terms it is described by the conditional probability densities and prior probabilities

$$\hat{\omega}_{MAP}(\mathbf{z}) = \underset{\omega \in \Omega}{\arg\max}\{p(\mathbf{z}|\omega)P(\omega)\}. \tag{3.80}$$

The Bayes normal quadratic classifier is based on having measurement vectors with class ω_k which is normally distributed with the expectation vector $\boldsymbol{\mu}_k$ and a $\boldsymbol{C}_k$ covariance matrix, which gives the following conditional probability densities

$$p(\mathbf{z}|\omega_k) = \frac{1}{\sqrt{(2\pi)^N|\boldsymbol{C}_k|}} \exp\left(\frac{-(\boldsymbol{z}-\boldsymbol{\mu}_k)^T\boldsymbol{C}_k^{-1}(\boldsymbol{z}-\boldsymbol{\mu}_k)}{2}\right). \tag{3.81}$$

3.6.1 Quadratic Distance Classifier

The quadratic classifier, can also be thought as the full Bayes normal classifier, since no assumptions are made and thereby substitute eq. 3.81 into eq. 3.80, the result is the following minimum error rate classification:

$$i = \text{argmax}\left\{\frac{1}{\sqrt{(2\pi)^N|\boldsymbol{C}_k|}} \exp\left(\frac{-(\boldsymbol{z}-\boldsymbol{\mu}_k)^T\boldsymbol{C}_k^{-1}(\boldsymbol{z}-\boldsymbol{\mu}_k)}{2}\right) P(\omega_k)\right\}. \tag{3.82}$$

Now the terms containing k are irrelevant and can be neglected. Also the logarithm is applied

$$i = \text{argmax}\left\{-\log|\boldsymbol{C}_k| + 2\log P(\omega_k) - \boldsymbol{\mu}_k^T\boldsymbol{C}_k^{-1}\boldsymbol{\mu}_k + 2\boldsymbol{z}^T\boldsymbol{C}_k^{-1}\boldsymbol{\mu}_k - \boldsymbol{z}^T\boldsymbol{C}_k^{-1}\boldsymbol{z}\right\} \tag{3.83}$$

$$i = \text{argmax}\left\{\omega_k + \boldsymbol{z}^T\boldsymbol{w}_k + \boldsymbol{z}^T\boldsymbol{W}_k\boldsymbol{z}\right\}, \tag{3.84}$$

where the last term reveals why the classifier is a quadratic one.

3.6.2 K Nearest Neighbour Classifier (KNNC)

A Nearest Neighbour Classifier is a method for having high resolution in the regions of dense training set, and otherwise low resolution. Having a hypersphere $R(\mathbf{z}) \subset \mathbb{R}^N$ with volume V, centred around $\mathbf{z}$ and a training set T_k with N_k number of samples, then it is a binomial distribution describing the probability of having n samples within $R(\mathbf{z})$. The expectation is

$$E[n] = N_k \int_{y\in R(\mathbf{z})} p(\mathbf{y}|\omega_k)d\mathbf{y} \approx N_k V p(\mathbf{z}|\omega_k). \tag{3.85}$$

The size of the sphere, still around $\mathbf{z}$, is selected so having exactly K samples inside the sphere. The position of $\mathbf{z}$ will be directly related to the volume, and thereby volume is written $V(\mathbf{z})$, on which the density is given

$$\hat{p}(\mathbf{z}|\omega_k) = \frac{K}{N_k V(\mathbf{z})}, \tag{3.86}$$

which tells that when $p(\mathbf{z}|\omega_k)$ is large, the volume is most likely small. That means in some cases the sphere needs to grow, which is controlled with K, which again controls balance between bias and variance. Another option is to use the KNNC, where the volume of the sphere is optimized to have exactly K number of samples inside. Letting K_k denote number of neighbours found for a class ω_k, then an estimate of the conditional density can be done

$$\hat{p}(\mathbf{z}|\omega_k) \approx \frac{K_k}{N_k V(\mathbf{z})}, \tag{3.87}$$

on which the sub-optimal classifier is made again as Bayes Classification

$$\hat{\omega}(\mathbf{z}) = \omega_k \tag{3.88a}$$

$$k = \arg \max_{i=1,\dots,K} \left\{ \hat{p}(\mathbf{z}|\omega_i)\hat{P}(\omega_i) \right\} \tag{3.88b}$$

$$k = \arg \max_{i=1,\dots,K} \left\{ \frac{k_i}{N_i V(\mathbf{z})} \frac{N_i}{N_s} \right\} \tag{3.88c}$$

$$= \arg \max_{i=1,\dots,K} \{k_i\}, \tag{3.88d}$$

where the class of $\mathbf{z}$ is simply decided based on which class has the maximum number of votes from K samples best fitted to $\mathbf{z}$. One problem with the KNNC is that, due to it is non-parametric it is computational heavy compared to the QDC.

3.6.3 Detection

Since there is only two classes, aircraft and non-aircraft, then the classification problem is a detection problem. Having two classes, ω_1 and ω_2 and each having a prior probabilities $P(\omega_1)$ and $P(\omega_2)$ the simplified Bayes decision rule becomes:

$$p(\boldsymbol{z}|\omega_1)P(\omega_1) > p(\boldsymbol{z}|\omega_2)P(\omega_2) \tag{3.89}$$

If test is passed, it is decided for ω_1, otherwise ω_2. When having a detection problems, two different errors can be made:

- Type I: Estimated non-Aircraft – Aircraft present
- Type II: Estimated Aircraft – NO aircraft present

3.6.4 Principal Component Analysis

If having a large number of features, it might be useful to do feature reduction, both because of the computational complexity, but also to try to highlight specific characteristic within a data set. One such method is the linear method, Principal Component Analysis (PCA). Having a high dimensional measurement vector $\mathbf{z}$, the goal is now to construct a lower dimensional feature vector $\mathbf{y}$. That can be done with transform matrix $\mathbf{W}_D$, so that

$$\mathbf{y} = \mathbf{W}_D\mathbf{z}. \tag{3.90}$$

The transformation matrix $\mathbf{W}_D$ has the dimensions $D \times N$, so that it transform from N-dimension space to a D-dimension space. The optimal transformation, will make $\mathbf{y}$ a perfect representation of $\mathbf{z}$. The selection of $\mathbf{W}_D$ can be described with a linear Minimum Mean Square Error (MMSE) estimator

$$\mathbf{W}_D = \arg \min_{\mathbf{W}} \left\{ \mathbf{E}\left[||\hat{\mathbf{z}}_{lMMSE}(\mathbf{y}) - \mathbf{z}||^2 \right] \right\}, \tag{3.91}$$

where $\mathbf{y} = \mathbf{W}\mathbf{z}$. Observe that eq. 3.91 does not give a unique solution for $\mathbf{W}_D$ since any invertible matrix $\mathbf{A}$ multiplied onto $\mathbf{W}_D$ will give the same minimum. If a unique solution is requested a few requirements are needed. First requirements is an *adding* requirement, which states that elements of $\mathbf{y}$ needs to add up in a way, that if $\mathbf{y}$ has the dimension D, then the dimension D-1 is obtained by simply neglecting the least informative element. If $\mathbf{y}$ is ordered so this least informative element is placed at the last element, then $\mathbf{W}_{D-1}$ can be obtained by removing the last element of $\mathbf{W}_D$.

Observe, columns of $\mathbf{y}$ needs to be uncorrelated for this to be possible. That means the covariance matrix $\mathbf{C}_y$ of $\mathbf{y}$ needs to be a diagonal matrix Λ, which gives

$$\mathbf{C}_y = \mathbf{W}_D \mathbf{C}_z \mathbf{W}_D^T = \Lambda_D. \tag{3.92}$$

When $D = N$, then $\mathbf{C}_z \mathbf{W}_N^T = \mathbf{W}_N^T \Lambda_N$, since $\mathbf{W}_N$ is an invertible orthogonal matrix and $\mathbf{W}_N^T \mathbf{W}_N$ is a diagonal matrix. Since Λ_N is diagonal, then the columns in $\mathbf{W}_N^T$ needs to be eigenvectors of $\mathbf{C}_z$ and thereby the diagonal elements of Λ_N are the eigenvalues. A second requirement is needed for obtaining a unique solution, which is that columns of $\mathbf{W}_N^T$ has unit length. The eigenvectors are orthogonal and unit length is obtained when $\mathbf{W}_N^T \mathbf{W}_N = \mathbf{I}$, where $\mathbf{I}$ is a $N \times N$ unit matrix. The rows, eigenvectors, in $\mathbf{W}_N$ needs to be sorted so eigenvalues forms a non-ascending sequence. Thereby, $\mathbf{W}_D$ can be constructed by deleting $N - D$ rows from $\mathbf{W}_N$. In words, $\mathbf{W}_N$ performs a rotation on $\mathbf{z}$, which will make the orthonormal basis align with the principal axes of the ellipsoid which is connected to the covariance matrix of $\mathbf{z}$. The coefficients of the representation of $\mathbf{z}$ are referred to as principal components. The MMSE approximation of $\mathbf{z}$ is obtained by

$$\hat{\mathbf{z}}_{IMMSE}(\mathbf{y}) = \mathbf{W}_D^T \mathbf{y} = \mathbf{W}_D^T \mathbf{W}_D \mathbf{z}. \tag{3.93}$$

Chapter 4

Classification Experiments

In the following the different features proposed in Chapter 3 are used for classification of aircraft versus non-aircraft, which means it is essentially a detection problem. First a section describing the used signal, how the testing was done and how the result is evaluated. Then the baseline results are presented, since these will be important in the discussion of the performance of the proposed features in this report. The baseline is important for providing information about the hypothesis given in this report, that taking temporal variation into account when extraction features, will outperform the more well-known approaches for doing the detection and classification of aircraft sound. In the following sections all classifications results will be presented, followed up by a directly comparison on which a discussion follows. We have used blocks of $1\,\mathrm{s}$ *in which we want to label the class, aircraft or non-aircraft.*

The IF in a signal has also been of great interest, since it can be used for estimating flight parameters on an event. An event is a signal, where an aircraft is passing by the observer, microphone. We have proposed an estimator for obtaining the flight parameter and in this chapter, the parameters obtained from the estimator is used for classification on events. The flight parameter estimator uses the HCM chirp pitch estimator, but the HCM chirp pitch estimator also gives other features, which might be useful for classification, and for that reason, we have made classification test with these features, which can be found in Appendix C.

4.1 General in Experiments

The goal is to test the different approaches for extracting features introduced in Chapter 3. We have decided on a specific database to train and test a classifier, which will be introduced in the following.

4.1.1 Classification by PRTools

As described in Section 3.6 we used PRTools for doing classification and we wanted to test both a parametric classifier and a non-parametric classifier. When training a classifier we used $80\,\%$ percent of the total training data for training and $20\,\%$ for directly testing the classifier. Observe, testing the classifier in that respect, is not the same as our experiment test since this is on a different test data set. For the cyclic analysis we needed to do feature reduction, where we used PCA. The measurement experiment test data is restricted to never be used for any training of a classifier including cross validation. Therefore, we have a separate data set referred to as testing data, which is only used for testing and obtaining classification results. PRTools provides a brilliant function for present the results of a detection problem, which is a confusion matrix. Please see Appendix A for a basic introduction on how to use the toolbox.

4.1.2 Training and Testing Data

The training and testing data has been provided by B&K. All data are recorded with a sample frequency of 16 kHz. The database is huge, so earlier on has B&K decided on specific test data. Based on earlier work B&K has performed and preliminary investigation of the problem, we found that three different sets of training was of great interest. The three training sets are

- Training data with high similarity of the testing data.
- Training data based on high SNR between noise and aircraft sound within a signal. Observe, SNR has not been measured, and the used signals are based on a basic listening test.
- Training data constructed as a mix of the two above.

We decided to have these three different training data, since it was of interest to observe if clean aircraft signal having high SNR would outperform the intuitive selection of training data based on high similarity of the testing data, and lastly a mixture of the two was of interest since both mentioned argument for training data could be valid and in that case the mixed training data was believed to be usable.

The original data are recordings done over several days. Based on audible and maybe visible detection, the data contains an information called *event*. An event means an aircraft is present in the signal and an interval is given for when the aircraft is present in the signal. In some models for extracting features and classification, for example the used flight parameter estimation model, it is an advantage to use these given intervals within a signal where an aircraft is present and extract flight parameters based on several seconds of recording. This give the problem of doing the event detection automatically and that classification cannot be done second for second. If this, classifying second by second, is the goal the investigated signal should be divided into blocks containing 1 s, in which some features can be extracted, for example MFFC. Since MFFC is the baseline, and we were more interested in the features than the problem of classifying, we decided to base the direct comparison between features extracting methods on the latter approach, where we buffered the signal of interest into 1 s blocks having 67 % overlap and did classification for each block. However, for the flight parameter estimation case, we used events with known larger intervals for testing and training a classifier and for doing the experiment.

Aircraft Training and Test Set

The mentioned events have in earlier work carried out by B&K lead to a number of signals on which the goal is to classify. It is based on 200 events. We aggregated the data and did some cleaning, meaning we removed non-aircraft in aircraft signal, of the signal and ended up having one large signal containing 1244 seconds, which we buffered into 1 second intervals having 67 % overlap, which gave 3770 blocks of 1 second. As earlier stated we also wanted to train on similar data to the test data in focus. The size of the training data, was decided to be 946 blocks of 1 second, since this was the maximum blocks to get from the clean and high SNR data and we wanted to have the same amount of blocks for the three training sets. Therefore a randomly selecting of 946 blocks of the 3770 blocks in test data were done, which then became the training set, called *noisyAirCraftSignal_for_training*. The remaining became the test data called *noisyAirCraftSig-*

nal_for_testing having 2824 blocks. Please observe, no same intervals are present in training and test data, since this would bias the results.

Earlier work carried out by B&K has lead to 20 intervals with clean and significant aircraft within a signal. These 20 intervals was aggregated into one signal, containing 312 seconds, which gave 946 blocks of 1 s, when the buffering with 67 % overlap was perform. As already mentioned, this amount of blocks, decided the amount of training data in the two other training set cases. These 946 blocks of clean aircraft sound, will from now be denoted as *cleanAirCraftSignal_for_training*.

The last training set, the mixed one, was constructed by randomly selecting 473 blocks in *noisyAirCraftSignal_for_training* and *cleanAirCraftSignal_for_training*. The results was 946 blocks, and this training set will be referred to as *mixedAirCraftSignal_for_training*

Non-Aircraft Training and Test Set

The last data set to be constructed is the non-aircraft data and then divide this into a training part and a testing part. A large aggregated signal containing background noise, containing 1677 seconds, gave a total of 5082 blocks when the 67 % overlap was used. The signal contains speech, trucks passing by, wind, birds etc. Next, the 5082 blocks were needed to be divided into a training and testing part, having the two names *noiseSignal_for_training* and *noiseSignal_for_testing*. For the training 946 blocks were randomly selected, and these blocks were at the same time deleted from the testing part, to ensure no biasing. Furthermore for having same amount of blocks in both test data set, the non-aircraft testing set was shorten to 2824 blocks in a randomly process.

Overview of used Data Sets

To give a short overview of the different signal, and number of 1 seconds intervals in each signal please see Table 4.1.

Names	Number of 1 s Intervals
cleanAirCraftSignal_for_training	946
noisyAirCraftSignal_for_training	946
mixedAirCraftSignal_for_training	946
noiseSignal_for_training	946
noisyAirCraftSignal_for_testing	2824
noiseSignal_for_testing	2824

Table 4.1: Overview of testing and training data, and the number of 1 s intervals in each.

4.1.3 Event Training and Test Data for Flight Parameters

As earlier described, the flight parameters cannot be extracted from a signal of 1 s, and therefore a larger signal length is needed. We used the same database of signals for finding training and test data. The aircraft test data are original produced from 100 events, only used half the events here, where observers have observed an aircraft and written that into the database as an event. As earlier mentioned we did a manually cleaning of the data, to get free from too much non-aircraft sound in aircraft label sound. We used 40 events for training a classifier, including calculating cross validation and preserved 10 events for the experiment test. Observe, it is of no interest to test different training set for doing flight parameter estimation, since we assume a signal with an aircraft passing by, which is crucial for the estimator to work.

Overview of used Data Sets For Flight Parameter Estimation

To give a short overview of the different events see Table 4.2.

Names	Number of Events
noisyAirCraftSignal_for_training	40
noiseSignal_for_training	40
noisyAirCraftSignal_for_testing	10
noiseSignal_for_testing	10

Table 4.2: Overview of testing and training data, and the number of events in each.

4.2 Classification by Baseline

The baseline is Mel Frequency Cepstral Coefficients (MFFC), which is used as features for the recognition of aircraft. The baseline is essentially an earlier project carried out by B&K, who provided an implementation of MFFC in MATLAB. We then simply applied that to our training and test data.

4.2.1 MFCC Results with Clean Aircraft Signal for Training

The 5-fold, 100 iterations cross validation errors for the two trained classifiers are

$$\text{QDC Cross Validation Error} = 10.65\,\%$$
$$\text{KNNC Cross Validation Error} = 7.88\,\%$$

When applying the two trained classifier on the test data, the QDC returns the confusion matrix seen in Table 4.3.

True Labels	Estimated Labels		
	Aircraft	non-Aircraft	Totals
Aircraft	1806	1018	2824
non-Aircraft	347	2477	2824
Totals	2153	3495	5648

Table 4.3: Confusion matrix on MFFC by QDC trained on Clean Aircraft Signal, having error rate $24.17\,\%$

And the KNNC returns the confusion matrix seen in Table 4.4.

True Labels	Estimated Labels		
	Aircraft	non-Aircraft	Totals
Aircraft	1459	1365	2824
non-Aircraft	315	2509	2824
Totals	1774	3874	5648

Table 4.4: Confusion matrix on MFFC by KNNC trained on Clean Aircraft Signal, having error rate $29.75\,\%$

MFCC Results with Noisy Aircraft Signal for Training

The 5-fold, 100 iterations cross validation errors for the two trained classifiers are

$$\text{QDC Cross Validation Error} = 18.28\,\%$$
$$\text{KNNC Cross Validation Error} = 17.75\,\%$$

When applying the two trained classifier on the test data, the QDC returns the confusion matrix seen in Table 4.5.

True Labels	**Estimated Labels**		
	Aircraft	non-Aircraft	Totals
Aircraft	2233	591	2824
non-Aircraft	423	2401	2824
Totals	2656	2992	5648

Table 4.5: Confusion matrix on MFFC by QDC trained on Noisy Aircraft Signal, having error rate $17.95\,\%$

And the KNNC returns the confusion matrix seen in Table 4.6.

True Labels	**Estimated Labels**		
	Aircraft	non-Aircraft	Totals
Aircraft	2356	468	2824
non-Aircraft	467	2357	2824
Totals	2823	2825	5648

Table 4.6: Confusion matrix on MFFC by KNNC trained on Noisy Aircraft Signal, having error rate $16.55\,\%$

MFCC Results with Mixed Aircraft Signal for Training

The 5-fold, 100 iterations cross validation errors for the two trained classifiers are

$$\text{QDC Cross Validation Error} = 15.13\,\%$$
$$\text{KNNC Cross Validation Error} = 15.64\,\%$$

When applying the two trained classifier on the test data, the QDC returns the confusion matrix seen in Table 4.7.

True Labels	**Estimated Labels**		
	Aircraft	non-Aircraft	Totals
Aircraft	2206	618	2824
non-Aircraft	488	2336	2824
Totals	2694	2954	5648

Table 4.7: Confusion matrix on MFFC by QDC trained on Mixed Aircraft Signal, having error rate $19.58\,\%$

And the KNNC returns the confusion matrix seen in Table 4.8.

True Labels	**Estimated Labels**		
	Aircraft	non-Aircraft	Totals
Aircraft	2084	740	2824
non-Aircraft	392	2432	2824
Totals	2476	3172	5648

Table 4.8: Confusion matrix on MFFC by KNNC trained on Mixed Aircraft Signal, having error rate $20.04\,\%$

4.2.2 Conclusion of MFCC Classification

The best performing training data when it comes to the cross validation error rate, was the clean aircraft data, where the cross validation error rate was $10.65\,\%$ for the QDC and $7.88\,\%$ for the KNNC. However, when the trained classifier was applied to the test data, the error rate became much higher, respectively $24.17\,\%$ for QDC and $29.75\,\%$ for KNNC. Next, we trained a classifier on a more similar data set in relation to the test data. The trained classifier had a higher cross validation error rate compared to the trained classifier on clean aircraft sound, but the difference in this case was that the error rate, when applied the trained classifier to the test data, did not change significant from the cross validation error rate. The error rates results in this case were respectively $17.95\,\%$ for QDC and $16.55\,\%$ for the KNNC. This result was not improved when the mixed training data, was used for training classifier and thereby the conclusion, at least when extracting MFFC, is that one should use training data with the highest similarity to the test data and we saw the lowest error rate was $16.55\,\%$ for KNNC, but the faster parametric QDC performed almost as good, with error rate of $17.95\,\%$.

4.3 Classification by Cyclic Spectral Coherence

The author of the article [1] has developed an estimator for doing Cyclic Spectral Analysis on the basis of the same article. The MATLAB implementation is available at MATLAB's file exchange web-page [3]. We use the estimator *as-it-is* and only changed parameters for doing the Cyclic Analysis on the provided sound files from B&K. We only searched for cyclic frequency in the range of $50\,\mathrm{Hz}$ to $110\,\mathrm{Hz}$ in which range the cyclic spectral coherence is calculated, giving a matrix $N \times M$ where N is spectral frequency and M cyclic frequency. This matrix was rearranged to a vector with size $MN = G$ and this vector became the measurement vector with very high number of features. For that reason, it was needed to reduce the number of features by doing feature reduction with PCA.

4.3.1 Cyclic Analysis Results with Clean Aircraft Signal for Training

In the following the PCA has reduced the features matrix dimension to 500. The 5-fold, 100 iterations cross validation errors for the two trained classifiers are

$$\text{QDC Cross Validation Error} = 9.85\,\%$$
$$\text{KNNC Cross Validation Error} = 2.79\,\%$$

When applying the two trained classifier on the test data, the QDC returns the confusion matrix seen in Table 4.9.

And the KNNC returns the confusion matrix seen in Table 4.10.

True Labels	**Estimated Labels**		
	Aircraft	non-Aircraft	Totals
Aircraft	1439	1385	2824
non-Aircraft	194	2630	2824
Totals	1633	4015	5648

Table 4.9: Confusion matrix on Cyclic Analysis Features by QDC trained on Clean Aircraft Signal, having error rate $27.96\,\%$

True Labels	**Estimated Labels**		
	Aircraft	non-Aircraft	Totals
Aircraft	385	2439	2824
non-Aircraft	52	2772	2824
Totals	437	5211	5648

Table 4.10: Confusion matrix on Cyclic Analysis Features by KNNC trained on Clean Aircraft Signal, having error rate $44.10\,\%$

Cyclic Analysis Results with Noisy Aircraft Signal for Training

In the following the PCA has reduced the features matrix dimension to 500. The 5-fold, 100 iterations cross validation errors for the two trained classifiers are

$$\text{QDC Cross Validation Error} = 27.48\,\%$$
$$\text{KNNC Cross Validation Error} = 19.32\,\%$$

When applying the two trained classifier on the test data, the QDC returns the confusion matrix seen in Table 4.11.

True Labels	**Estimated Labels**		
	Aircraft	non-Aircraft	Totals
Aircraft	2429	395	2824
non-Aircraft	540	2284	2824
Totals	2969	2679	5648

Table 4.11: Confusion matrix on Cyclic Analysis Features by QDC trained on Noisy Aircraft Signal, having error rate $16.55\,\%$

And the KNNC returns the confusion matrix seen in Table 4.12.

True Labels	**Estimated Labels**		
	Aircraft	non-Aircraft	Totals
Aircraft	1813	1011	2824
non-Aircraft	180	2644	2824
Totals	1993	3655	5648

Table 4.12: Confusion matrix on Cyclic Analysis Features by KNNC trained on Noisy Aircraft Signal, having error rate $21.09\,\%$

Cyclic Analysis Results with Mixed Aircraft Signal for Training

In the following the PCA has reduced the features matrix dimension to 500. The 5-fold, 100 iterations cross validation errors for the two trained classifiers are

$$\text{QDC Cross Validation Error} = 26.85\,\%$$
$$\text{KNNC Cross Validation Error} = 16.91\,\%$$

When applying the two trained classifier on the test data, the QDC returns the confusion matrix seen in Table 4.13.

True Labels	**Estimated Labels**		
	Aircraft	non-Aircraft	Totals
Aircraft	2161	663	2824
non-Aircraft	519	2305	2824
Totals	2680	2968	5648

Table 4.13: Confusion matrix on Cyclic Analysis Features by QDC trained on Mixed Aircraft Signal, having error rate $20.93\,\%$

And the KNNC returns the confusion matrix seen in Table 4.14.

True Labels	**Estimated Labels**		
	Aircraft	non-Aircraft	Totals
Aircraft	1470	1354	2824
non-Aircraft	195	2629	2824
Totals	1665	3983	5648

Table 4.14: Confusion matrix on Cyclic Analysis Features by KNNC trained on Mixed Aircraft Signal, having error rate $27.43\,\%$

4.3.2 Conclusion of Cyclic Spectral Coherence Classification

In this classification experiment the cross validation error rate was minimum when using clean aircraft data, as it was observed in Section 4.2, but in this case the cross validation was even lower. It was respectively $9.85\,\%$ for the QDC and $2.79\,\%$ for KNNC. However, as it has been seen so far, the error rate on the test data was much higher and it ended up being $27.96\,\%$ for the QDC and $44.10\,\%$ for KNNC. For the next training data, noisy aircraft, the cross validation errors were respectively $27.48\,\%$ for the QDC and $19.32\,\%$ for the KNNC, but when the classifier was applied to data the error rates became $16.55\,\%$ for the QDC and $21.09\,\%$ for the KNNC. In the mixed training data, the cross validation became a bit lower again, but the error rate on test data, was not as good as the noisy aircraft training data. The conclusion in this case, is that the training data should be the data with highest similarity of the test data, and in that case the best result was an error rate of $16.55\,\%$ obtained with the QDC.

4.4 Classification by Cyclic Spectral Coherence Mean Envelope

We use the same estimator as in Section 4.3. We searched for cyclic frequency in the range of $50\,\mathrm{Hz}$ to $110\,\mathrm{Hz}$ in which range the cyclic spectral coherence is calculated and afterwards we calculated the mean across the cyclic frequency range. This gave a vector, which became the measurement feature vector.

4.4.1 Mean Envelope Results with Clean Aircraft Signal for Training

The 5-fold, 100 iterations cross validation errors for the two trained classifiers are

$$\text{QDC Cross Validation Error} = 16.16\,\%$$
$$\text{KNNC Cross Validation Error} = 23.84\,\%$$

When applying the two trained classifier on the test data, the QDC returns the confusion matrix seen in Table 4.15.

True Labels	**Estimated Labels**		
	Aircraft	non-Aircraft	Totals
Aircraft	1682	1142	2824
non-Aircraft	458	2366	2824
Totals	2140	3508	5648

Table 4.15: Confusion matrix on Mean Envelope from Cyclic Analysis by QDC trained on Clean Aircraft Signal, having error rate $28.33\,\%$

And the KNNC returns the confusion matrix seen in Table 4.16.

True Labels	**Estimated Labels**		
	Aircraft	non-Aircraft	Totals
Aircraft	1442	1382	2824
non-Aircraft	351	2473	2824
Totals	1793	3855	5648

Table 4.16: Confusion matrix on Mean Envelope from Cyclic Analysis by KNNC trained on Clean Aircraft Signal, having error rate $30.68\,\%$

4.4.2 Mean Envelope Results with Noisy Aircraft Signal for Training

The 5-fold, 100 iterations cross validation errors for the two trained classifiers are

$$\text{QDC Cross Validation Error} = 26.16\,\%$$
$$\text{KNNC Cross Validation Error} = 29.42\,\%$$

When applying the two trained classifier on the test data, the QDC returns the confusion matrix seen in Table 4.17.

True Labels	**Estimated Labels**		
	Aircraft	non-Aircraft	Totals
Aircraft	1745	1079	2824
non-Aircraft	519	2305	2824
Totals	2264	3384	5648

Table 4.17: Confusion matrix on Mean Envelope from Cyclic Analysis by QDC trained on Noisy Aircraft Signal, having error rate $28.29\,\%$

And the KNNC returns the confusion matrix seen in Table 4.18.

True Labels	**Estimated Labels**		
	Aircraft	non-Aircraft	Totals
Aircraft	1531	1293	2824
non-Aircraft	442	2382	2824
Totals	1973	3675	5648

Table 4.18: Confusion matrix on Mean Envelope from Cyclic Analysis by KNNC trained on Noisy Aircraft Signal, having error rate $30.72\,\%$

4.4.3 Mean Envelope Results with Mixed Aircraft Signal for Training

The 5-fold, 100 iterations cross validation errors for the two trained classifiers are

$$\text{QDC Cross Validation Error} = 20.85\,\%$$
$$\text{KNNC Cross Validation Error} = 26.67\,\%$$

When applying the two trained classifier on the test data, the QDC returns the confusion matrix seen in Table 4.19.

True Labels	**Estimated Labels**		
	Aircraft	non-Aircraft	Totals
Aircraft	1712	112	2824
non-Aircraft	491	2333	2824
Totals	2203	3445	5648

Table 4.19: Confusion matrix on Mean Envelope from Cyclic Analysis by QDC trained on Mixed Aircraft Signal, having error rate $28.38\,\%$

And the KNNC returns the confusion matrix seen in Table 4.20.

True Labels	**Estimated Labels**		
	Aircraft	non-Aircraft	Totals
Aircraft	1677	1147	2824
non-Aircraft	505	2319	2824
Totals	2182	3466	5648

Table 4.20: Confusion matrix on Mean Envelope from Cyclic Analysis by KNNC trained on Mixed Aircraft Signal, having error rate $29.25\,\%$

4.4.4 Conclusion of Cyclic Spectral Coherence Mean Envelope Classification

Again in this classification experiment the cross validation error rate was minimum when using clean aircraft data, as it was observed in Section 4.2. The cross validation error rate was $16.16\,\%$ for the QDC and $23.84\,\%$ for KNNC. However, when the trained classifier was applied to the test data, the error rates were respectively $28.33\,\%$ for QDC and $30.68\,\%$ for KNNC. The next training data was the noisy aircraft, the one with higher similarity of the test data, and in this case the cross validation error rate on training data and error rates on test data were approximately equal. In this case the error rates on the test data were respectively $28.29\,\%$ for QDC and $30.72\,\%$. Thereby the QDC performed slightly better when trained on noisy aircraft data, but with almost lowest possible margin, so the improvement is not valid. Lastly, the mixed data set was used for training a classifier, where the cross validation error rates were respectively $20.85\,\%$ for QDC and $26.67\,\%$ and the error rates when classifiers applied to test data, respectively $28.38\,\%$ for QDC and $29.25\,\%$ for KNNC. The conclusion in this experiments is that no training data perform better than another training data, and the minimum error rate, $28.29\,\%$, was not as good as in Section 4.2 and in Section 4.3.

4.5 Classification on Flight Parameters by IF Estimates

Instead of classifying segment by segment with a specific length, we here proposed to work on events, where segment length can vary. One such event should be an aircraft passing by, which will give a time varying frequency. We use the flight parameter estimator as in Section 3.5 for extracting 7 features, which are all used for the classification problem. Observe, the flight parameter estimator uses the IF estimated from the proposed HCM chirp pitch estimator, which is a crucial key for estimating the flight parameter.

Flight Parameters Results with Noisy Aircraft Signal for Training

The 5-fold, 100 iterations cross validation errors for the two trained classifiers are

$$\text{QDC Cross Validation Error} = 10.51\,\%$$
$$\text{KNNC Cross Validation Error} = 19.00\,\%$$

When applying the two trained classifier on the test data, the QDC returns the confusion matrix seen in Table 4.21 And the KNNC returns the confusion matrix seen in Table 4.22

4.5.1 Conclusion of Flight Parameters Classification

The cross validation error rate is small both for QDC and KNNC, respectively $10.51\,\%$ for QDC and $19.00\,\%$ for KNNC. When the trained classifier is applied to the test data, the error rates becomes $5\,\%$ for QDC and $5.00\,\%$ where the QDC is closest to its cross validation. We have earlier shown, how the flight parameter estimator can gives useful estimates, which we tested on synthetic

True Labels	**Estimated Labels**		
	Aircraft	non-Aircraft	Totals
Aircraft	9	1	10
non-Aircraft	0	10	10
Totals	9	11	20

Table 4.21: Confusion matrix on Flight Parameters by QDC trained on Noisy Aircraft Signal, having error rate $5\,\%$

True Labels	**Estimated Labels**		
	Aircraft	non-Aircraft	Totals
Aircraft	10	0	10
non-Aircraft	1	9	10
Totals	11	9	10

Table 4.22: Confusion matrix on Flight Parameters Features by KNNC trained on Noisy Aircraft Signal, having error rate $5.00\,\%$

signal. When applied to a few real events, we cannot say if the parameters are estimated well, but we can conclude the estimates can be used for classification indicating this approach is very useful and that the QDC is the preferred classifier in this case due to the lower computational complexity compared to the KNNC.

4.6 Comparison and Discussion on Classification Experiments

Firstly, the classification experiments based on $1\,\mathrm{s}$ segments will be compared. The best classification results from Section 4.2, Section 4.3 and Section 4.4 are shown in Table 4.23 It is clear from

	Best Cross Validation	**Best Error Rate**
Baseline	$7.88\,\%$	$16.55\,\%$
Cyclic Coherence	$2.79\,\%$	$16.55\,\%$
Mean Cyclic Coherence	$16.16\,\%$	$28.29\,\%$

Table 4.23: Comparison of Classification Results on 1 seconds blocks.

Table 4.23 that the Cyclic Coherence by Cyclic Analysis did not performed better than MFFC on test data, but gave exactly the same error rate. However, the Cyclic Coherence obtained the result with QDC where the MFFC obtained the results with KNNC, where we believe the QDC is preferable due to lower computational time. We also claim that Cyclic Coherence contains other possibilities for outperforming the MFFC, which we argue by the lower cross validation error rate compared to MFFC.

All best cross validation error rates were observed when training a classifier on the clean aircraft data set, but all the best experiment test error rates were observed when classifier was trained on highly similar training data compared to test data. We also want to point that possibilities of optimizing in the Cyclic Analysis is possible, like segments length, filtering, cyclic frequency range, but also how to use the features from the analysis, where we have shortly looked into an image processing approach, see Appendix B.

In Section 4.5, the goal was not to classify $1\,\mathrm{s}$ segments, but rather classifying events. This, we have done as a prove of concept, and for showing the advantages the proposed HCM pitch estimator gives, since this estimator directly gives the IF. In earlier work a Short Time Fourier Transform and interpolation was used for obtaining the IF [13], where we believe our method is superior to that. We estimated 4 flight parameters, plus 3 Doppler parameters. We used 50 events in total, 40 for training and test classifier and 10 for experiment. The cross validation error rate was as best $10.51\,\%$ and the error rates on experiment test data was $5\,\%$ both for QDC and for KNNC, which is very promising. We have assumed in this study, that events and segments are known, but it real world applications it will be needed to investigated how to detect events. It might be possible to simply detect by maximum sound pressure level and use the part of the signal, which obey the International Civil Aviation Organization standard which states that the part of the signal, which is within $10\,\mathrm{dB}$ of the maximum sound pressure should be used [25]. Other methods to consider for finding the center are GABOR [20] or zero-crossing method [32]. Lastly, it might also be possible to use some of the proposed theory in this report by simply investigate a larger block, for example 6 seconds, do pitch detection, model order estimation and flight parameter estimation on which a threshold for the NLS of the flight parameter is used to detect an event of an aircraft passing by.

Chapter 5

Conclusion

We have by studying the physics of an aircraft passing by, obtained an understanding of the problem to solve. It was clear that temporal variations was present in one such signal, which we believed could improve aircraft detection compared to the baseline feature in this work, Mel Frequency Cepstral Coefficients (MFFC). We have shown by using Cyclic Analysis, how a cyclo-stationary signal can be analysed with Cyclic Spectral Coherence, on which a classifier can be trained. This approach showed to have equally good classification results as the baseline, on signal-segments of $1\,\mathrm{s}$ but the results was obtained by using a parametric classifier, where for the baseline a non-parametric classifier gave the best result. If limiting the time consuming is an issue, the cyclic analysis therefore outperforms the baseline, MFFC. Furthermore, the Cyclic Analysis is capable of finding a hidden periodicity, which makes it very usable in noisy environment.

A cyclic frequency was also possible to estimate where a Doppler frequency was clearly present. However, we proposed a Harmonic Chirp Model (HCM) chirp pitch estimator for estimating the Instantaneous Frequency (IF), from which flight parameters could be estimated on the basis of the physics when an aircraft is passing by, its trajectory. The HCM chirp pitch estimator was tested and compared to a normal Harmonic Model (HM) approximately Non-linear Least Squares (aNLS) pitch estimator, which lag the incorporation of time varying pitch and thereby gets higher RMSE values on a chirp signal, and it does not directly gives the IF. We have shown the IF estimate from the HCM chirp pitch estimator can be used in a proposed flight parameter estimator, from which unique and significant features could be found with acceptable precession in a synthetic produced signal having two time-varying phenomena, Doppler frequencies and destructive interference frequencies, where only the first one has been of interest in the study. We also tested the flight parameter estimation on real signals, where we obtained promising results for one such method.

Even though we have utilized and shown some of the advantages the temporal variation gives when incorporated into models and estimators, we still believe a lot of work and optimization can be done. The Cyclic Analysis provides different measurements, where we mainly focused on the Cyclic Spectral Coherence but others are available, like the cyclic spectral density. It could also be beneficial to do an investigation for an optimal method on how to use the cyclic frequency. We also did a short investigation on image processing from Cyclic Spectral Coherence image, but also here we believe a dedicated investigation is needed for being able to answer if this approach is useful or not. We do believe the HCM chirp pitch estimator together with the flight parameter estimator is very useful. In our experiment we did not focus on how to find the event, intervals of signal to work on, but we believe that using the proposed methods in this report like pitch detection, model order estimation and comparing the NLS from the flight parameter estimator to a specific threshold, it will be possible to do the event detection and thereby having a signal on which the parameters should be used in a machine learning approach. In addition, for getting more features we also believe the destructive interference frequencies is useful and can be used for extracting more flight parameters.

All in all, we believe we have shown the benefits of investigating temporal variation in a signal when the purpose is to detect an aircraft within a signal. We believe we have proposed useful methods for doing so, both for smaller segments and on larger segments, which we argue by our experiments, where both the Cyclic Spectral Coherence and the flight parameter estimation showed promising results.

Bibliography

[1] Jérôme Antoni. *Cyclic spectral analysis in practice*. Article from Mechanical Systems and Signal Processing 21 (2007) 597–630, 2007.

[2] Jérôme Antoni. *Cyclostationarity by examples*. Article from Mechanical Systems and Signal Processing 23 (2009) 987–1036, 2009.

[3] Jérôme Antoni. *Cyclic Spectral Analysis*, 2014. MATLAB File Exchange - http://www.mathworks.com/matlabcentral/fileexchange/48909-cyclic-spectral-analysis.

[4] A. Antoniou and W.-S. Lu. *Practical Optimization: Algorithms and Engineering Applications*. Springer Verlag, 2007.

[5] C. Asensio and M. Recuero M. Ruiz. *Real-time aircraft noise likeness detector*. Article from Applied Acoustics 71 (2010) 539–545, 2010.

[6] Victor Bisot, Slim Essid, and Gaël Richard. *HOG AND SUBBAND POWER DISTRIBUTION IMAGE FEATURES FOR ACOUSTIC SCENE CLASSIFICATION*. 23rd European Signal Processing Conference (EUSIPCO), 2015.

[7] Mads G. Christensen and Andreas Jakobsson. *Multi-Pitch Estimation*. Book from Morgan & Claypool, 2009.

[8] Mads Græsbøll Christensen and Jesper Rindom Jensen. *Pitch Estimation for Non-Stationary Speech*. Article from Asilomar Conference on Signals, Systems and Computers. Conference Record. 1400-1404. [TP5a-3]. 10.1109/ACSSC.2014.7094691, 2014.

[9] Navneet Dalal and Bill Triggs. *Histograms of Oriented Gradients for Human Detection*. Article from the Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), 2005.

[10] PETAR M. DJURIC and STEVEN M. KAY. *Parameter Estimation of Chirp Signals*. Article from IEEE TRANSACTIONS ON ACOUSTICS, SPEECH. AND SIGNAL PROCESSING. VOL. 38 . NO. 17. DECEMBER, 1990.

[11] D. de Ridder F. van der Duin, R.P.W Duin and D.M.J. Tax. *Classification, Parameter Estimation and State Estimation - An Engineering Approach using MATLAB*. Book Published by WILEY, 2004.

[12] Brian G. Ferguson and Kam W. Lo. *Passive Estimation Of Aircraft Motion Parameters using Destructive Interference Between Direct And Ground-Reflected Sound Waves*. Article from Information, Decision and Control, 1999. IDC 99. Proceedings., 1999.

[13] Brian G. Ferguson and Kam W. Lo. *Turboprop and rotary-wing aircraft flight parameter estimation using both narrow-band and broadband passive acoustic signal-processing methods*. Journal Acoustic Society of America 108 (4), October, 2000.

[14] Brian G. Ferguson and Kam W. Lo. *Flight Path Estimation Using Frequency Measurements from a Wide Aperture Acoustic Array*. Article from IEEE TRANSACTIONS ON AEROSPACE AND ELECTRONIC SYSTEMS VOL. 37, NO. 2 APRIL, 2001.

[15] Brian G. Ferguson and Kam W. Lo. *Aircraft Flight Parameter Estimation Using Acoustical Lloyd's Mirror Effect*. Article from IEEE TRANSACTIONS ON AEROSPACE AND ELECTRONIC SYSTEMS VOL. 38, NO. 1 JANUARY, 2002.

[16] Brian G. Ferguson and Barry G. Quinn. *Application of the short-time Fourier transform and the Wigner-Ville distribution to the acoustic localization of aircraft*. Article from . Acoust. Soc. Am. 96 (2), Pt. 1, 1994.

[17] Jonathan T. Foote. *Content-Based Retrieval of Music and Audio*. Article from SPIE Vol. 3229, 1997.

[18] Feng Gao, Qizhi Xu, and Bo Li. *Aircraft Detection from VHR Images Based on Circle-Frequency Filter and Multilevel Features*. Hindawi Publishing Corporation The Scientific World Journal Volume 2013, Article ID 917928, 7 pages, 2013.

[19] William A. Gardner. *Cyclostationarity by examples*. Article from IEEE Signal Processing Magazine, 1991.

[20] Jürgen T. Geiger and Karim Helwani. *IMPROVING EVENT DETECTION FOR AUDIO SURVEILLANCE USING GABOR FILTERBANK FEATURES*. 23rd European Signal Processing Conference (EUSIPCO), 2015.

[21] Julius O. Smith III. *Analytic Signals and Hilbert Transform Filters*. Book from Stanford University, 2000. https://ccrma.stanford.edu/~jos/mdft/Analytic_Signals_Hilbert_Transform.html.

[22] Mendel Kleiner. *Acoustic and Audio Technology - Third Edition*. Book Published by J.Ross Publishing, 2012.

[23] Sidonie Lefebvre, Stéphanie Allassonnière, Jérémie Jakubowicz, Thomas Lasne, and Eric Moulines. *IGINAL PAPER Aircraft classification with a low resolution infrared sensor*. Article from Machine Vision and Applications (2013) 24:175–186, Springer-Verlag, 2013.

[24] Georgije Lukatela. *Pitch Determination by Adaptive Autocorrelation Method*. Article from Haskins Laboratories Status Report on Speech Research, 1972.

[25] Miguel Márquez-Molina, Luis Pastor Sánchez-Fernández, Sergio Suárez-Guerra, and Luis Alejandro Sánchez-Pérez. *Aircraft take-off noises classification based on human auditory's matched features extraction*. Article from the Journal Applied Acoustics 84 (2014) 83–90, 2013.

[26] Sidsel Marie Nørholm, Jesper Rindom Jensen, and Mads Græsbøll Christensen. *Enhancement of Non-Stationary Speech using Harmonic Chirp Filters*. Article from INTERSPEECH, 2015.

[27] Alan V. Opperheim and Ronald W. Schafer. *Discrete-Time Signal Processing*. Book from Pearson Education, 2010.

[28] PIOTR PTAK, JUHA HARTIKKA, MAUNO RITOLA, and TUOMO KAURANNE. *Aircraft Classification Based onRadar Cross Section of Long-Range Trajectories*. Article from IEEE TRANSACTIONS ON AEROSPACE AND ELECTRONIC SYSTEMS VOL. 51, NO. 4 OCTOBER, 2015.

[29] R.B. Randall. *Application Notes - Cepstrum Analysis and Gearbox Fault Diagnosis*. Application Notes by Brüel & Kjær, 1980.

[30] Manfred R.Schroeder. *Generalized Short-Time Power Spectra and Autocorrelation Functions*. Article from the Journal of the Acoustical Society of America, 34, 1962.

[31] Ben J. Shannon and Kuldip K. Paliwal. *A Comparative Study of Filter Bank Spacing for Speech Recognition*. Article from the Conference MICROELECTRONIC ENGINEERING RESEARCH CONFERENCE, 2003.

[32] Luis Alejandro Sánchez-Pérez, Luis Pastor Sánchez-Fernández, Sergio Suárez-Guerra, and José Juan Carbajal-Hernández. *Aircraft class identification based on take-off noise signal segmentation in time*. Article from Expert Systems with Applications 40 (2013) 5148–5159, 2013.

[33] Shu zhu Shi, Zheng yu Zhao, Yan Liu, Gang Chen, Ting Li, Jing nan Liu, and Ming Yao. *Experimental Demonstration for Ionospheric Sensing and Aircraft Detection With a HF Skywave Multistatic Radar*. Article from IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, VOL. 11, NO. 7 July, 2015.

Appendix

Appendix A

PRTools in MATLAB

The MATLAB Toolbox PRTools is a simple toolbox to use and the authors have also written a book about the function and the use of these [11]. Here we will give a short introduction the basic use of the toolbox in a detection problem.

Train Data, Error and Confusion Tables

By using one of the two classifiers, from now referred to as Quadratic Distance Classifier (QDC) and K Nearest Neighbour Classifier (KNNC), the features can be used for doing classification of aircraft. Firstly, the estimated features are needed to be divided into training and test part. We decided to split data into $80\,\%$ training and $20\,\%$ test data. The procedure for dividing data, set labels, set prior probability etc., all follows the procedure given in [11].

- Full data set : z
- Training data set : x
- Test data set : y

Now, the a classifier can be trained, tested and the result of each classification can be presented in a confusion matrix as seen in Listing A.1

```
trainedClassifier = qdc(x);
errorRateQDC = testc(y*trainedClassifier);
labelTraining = getlabels(y);
labelClassification = y*trainedClassifier*labeld;
confmat(labelTraining,labelClassification);
```

Listing A.1: Training and testing QDC-Classifier and present result in a confusion matrix. Observe the highlighted words are function in PRTools Toolbox

Cross Validation

When testing a classifier a cross validation is an optimal test, since it will show the expected error rate when training a classifier on same data. The cross validation provided in PRTools is a *k-fold cross validation* where the full data z is randomly partitioned into k number of equally sized blocks. The blocks are then divided into training and test. The training number of blocks will be $k-1$, which gives one block left for testing. Observe, this gives $80\,\%$ and $20\,\%$. When blocks are divided into training and test, basically the classifier is trained and thereafter tested. This is repeated i number of times, e.g. 100 or 1000. See Listing A.2 for the MATLAB code for doing cross validation with PRTools:

```
i=100; kFolding=5;
classifier=qdc([]);
crossvalidationerror=prcrossval(z,classifier,kFolding,i);
```

Listing A.2: MATLAB code for doing cross validation using PRTools Toolbox

Principal Component Analysis

When having a very large number of features, a principal component analysis can be used for doing feature reduction. In PRTools this is done as

```
dimensions=500; %The wanted dimension after PCA map
pca_map = pcam(z,dimensions);% the PCA map, which data will be mapped to
new_z = z*pca_map; %new data, which has been mapped to pca_map
```

Listing A.3: MATLAB code for doing PCA using PRTools Toolbox

Appendix B

Image Processing Approach - HOG

In the article [9] the authors propose the use of Histogram Of Gradients (HOG) as features for doing human detection in an images. The contour of the human will end up being the boundaries, and we wanted to investigate if significant contours can be found in the Cyclic Spectral Coherence, when this is transformed to an image. This of course assume great similarity in Cyclic Spectral Coherence across all signals when any aircraft is present is the signal.

The HOG method is made from the idea that local objects and shapes is often possible to describe by the distribution of local intensity gradients. The image is divided into a number of regions referred to as cells. This region size should be set so one have the wanted amount of details, which means if the aim is small details, the region size should also be small. For each region a HOG directions is made. The features can be extracted in MATLAB, by using the function *extractHOGFeatures*, which is part of the vision toolbox. In this case, the images processed are the Cyclic Spectral Coherence image introduced in Section 3.2

One drawback in this approach is the large number of features, which meant it was needed to do features reduction on data when training and testing a classifier, which we did as described in Section 3.6. In Figure B.1 an images of the Cyclic Spectral Coherence, when an aircraft is strongly represented in a signal, can be seen. Observe, we did a visible inspection of a few images, and even when having a significant clean aircraft signal, the images does not have this nice clear vertical line at cyclic frequency $\alpha = 77\,\mathrm{Hz}$, since frequency will be changing, both due to Doppler both also aircraft type. Thereby, the assumption about high similarity of Cyclic Spectral Coherence for any aircraft signal is not obeyed, but since we only could inspect a few images, because it is very time demanding to observe several thousand Cyclic Spectral Coherence images, we did not knew if the HOG features would still be able to outperform the MFFC approach for classifying aircraft in signal. In Figure B.2 an image of the cyclic spectral coherence, when only noise is present, can be seen. Please observe, that when extracting HOG features, all labels, axis values etc. are neglected in the producing of the image on which the HOG feature extraction is applied.

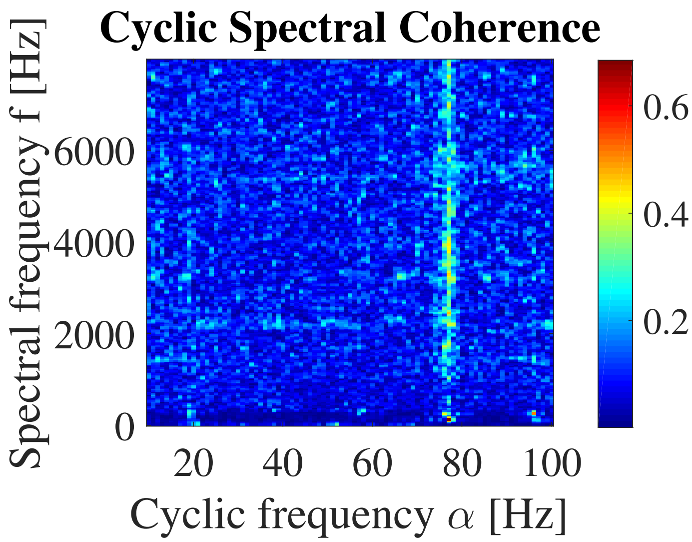


Figure B.1: Example of an images produced by Cyclic Spectral Coherence on Clean Aircraft signal. Observe, y-axis is spectral frequency and x-axis, is cyclic frequency, but these information are not wanted in the image for the HOG extracting. Observe, this images represent closely to a perfect image, when there is an aircraft present. Many images extracted, even from the clean aircraft signal, does not have this nice clear vertical line.

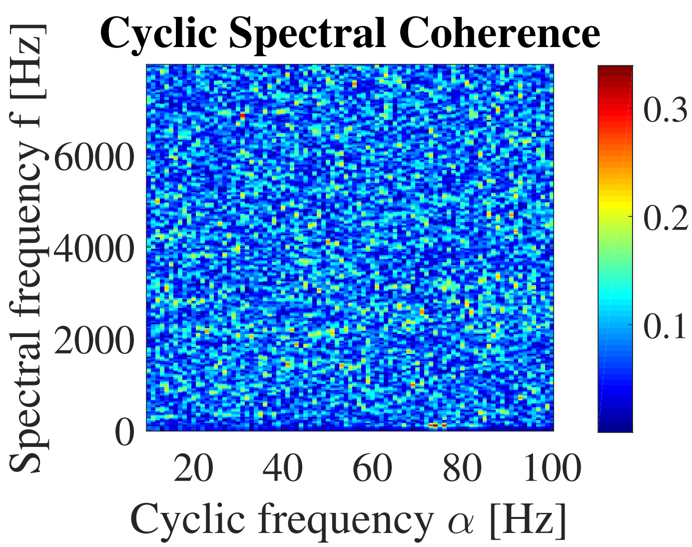


Figure B.2: Example of an images produced by Cyclic Spectral Coherence on noise signal. Observe, y-axis is spectral frequency and x-axis, is cyclic frequency, but these information are not wanted in the image for the HOG extracting. No significant cyclic frequencies are present, since the image represent a noise.

Classification by HOG

We wanted to investigate if the HOG can be used on the Cyclic Spectral Coherence, when transformed to an image, so for blocks in training and test data, we extracted the cyclic spectral coherence transformed all does figures to images, so the HOG features could be extracted.

HOG Results with Clean Aircraft Signal for Training

In the following the PCA has reduced the features matrix dimension to 50. The 5-fold, 100 iterations cross validation errors for the two trained classifiers are

$$\text{QDC Cross Validation Error} = 17.50\,\%$$
$$\text{KNNC Cross Validation Error} = 21.59\,\%$$

When applying the two trained classifier on the test data, the QDC returns the confusion matrix seen in Table B.1 And the KNNC returns the confusion matrix seen in Table B.2

True Labels	**Estimated Labels**		
	Aircraft	non-Aircraft	Totals
Aircraft	955	1869	2824
non-Aircraft	61	2763	2824
Totals	1016	4632	5648

Table B.1: Confusion matrix on HOG features by Quadratic Distance Classifier trained on Clean Aircraft Signal, having error rate $34.17\,\%$

True Labels	**Estimated Labels**		
	Aircraft	non-Aircraft	Totals
Aircraft	950	1874	2824
non-Aircraft	333	2491	2824
Totals	1283	4365	5648

Table B.2: Confusion matrix on HOG features by K nearest neighbour classifier trained on Clean Aircraft Signal, having error rate $39.08\,\%$

HOG Results with Noisy Aircraft Signal for Training

In the following the PCA has reduced the features matrix dimension to 50. The 5-fold, 100 iterations cross validation errors for the two trained classifiers are

$$\text{QDC Cross Validation Error} = 29.46\,\%$$
$$\text{KNNC Cross Validation Error} = 35.98\,\%$$

When applying the two trained classifier on the test data, the QDC returns the confusion matrix seen in Table B.3 And the KNNC returns the confusion matrix seen in Table B.4

True Labels	**Estimated Labels**		
	Aircraft	non-Aircraft	Totals
Aircraft	1013	1811	2824
non-Aircraft	74	2750	2824
Totals	1087	4561	5648

Table B.3: Confusion matrix on HOG features by Quadratic Distance Classifier trained on Noisy Aircraft Signal, having error rate $33.37\,\%$

True Labels	**Estimated Labels**		
	Aircraft	non-Aircraft	Totals
Aircraft	1322	1502	2824
non-Aircraft	551	2273	2824
Totals	1873	3775	5648

Table B.4: Confusion matrix on HOG features by K nearest neighbour classifier trained on Noisy Aircraft Signal, having error rate $36.35\,\%$

HOG Results with Mixed Aircraft Signal for Training

In the following the PCA has reduced the features matrix dimension to 50. The 5-fold, 100 iterations cross validation errors for the two trained classifiers are

$$\text{QDC Cross Validation Error} = 23.12\,\%$$
$$\text{KNNC Cross Validation Error} = 27.95\,\%$$

When applying the two trained classifier on the test data, the QDC returns the confusion matrix seen in Table B.5 And the KNNC returns the confusion matrix seen in Table B.6

True Labels	**Estimated Labels**		
	Aircraft	non-Aircraft	Totals
Aircraft	1004	1820	2824
non-Aircraft	102	2722	2824
Totals	1106	4542	5648

Table B.5: Confusion matrix on HOG features by Quadratic Distance Classifier trained on Mixed Aircraft Signal, having error rate $34.03\,\%$

True Labels	**Estimated Labels**		
	Aircraft	non-Aircraft	Totals
Aircraft	1090	1734	2824
non-Aircraft	450	2374	2824
Totals	1540	4108	5648

Table B.6: Confusion matrix on HOG features by K nearest neighbour classifier trained on Mixed Aircraft Signal, having error rate $38.67\,\%$

Conclusion of HOG Classification

Starting with the first trained classifier, the cross validation error rate was $17.50\,\%$ for the QDC and $21.59\,\%$ for the KNNC and with that trained classifier, the error rates on test data became $34.17\,\%$ for QDC and $39.08\,\%$, which is very poor results. It did improve with noise aircraft training data, but not much. The error rate on test data was $33.37\,\%$ for QDC and $36.35\,\%$ for the KNNC. For the mixed training data, the error rates did not improve, and the conclusion which can be done in the experiment is that, the best training data for HOG is the noisy aircraft data highly similar to test data, but the assumption about high similarity of any aircraft was not obeyed, and therefore the error rates became poor.

Appendix C

Classification by Pitch and Related Features

We use the estimator proposed in as in Subsection 3.4.4 with the initial frequency range $50\,\mathrm{Hz}$ to $120\,\mathrm{Hz}$. We high pass the input signal, so frequencies below $50\,\mathrm{Hz}$ are removed. We also down sample by a factor of 6, which means going from $16\,\mathrm{kHz}$ to $2.67\,\mathrm{kHz}$. The upper limit of harmonics for the MAP order estimation is 10, the maximum iterations in the HCM iterative chirp pitch estimator is 10 and convergence criteria 2e-3. The features used for classification was

- f_0
- b, the chirp rate
- $\hat{L}$, from MAP-order
- Stationary number of complex harmonics, 4, normalized, in range 0 to 1.
- Stationary number of complex harmonics, 4, normalized and summed giving one value.

Which in total gives 8 features.

Training and Test Data

The HCM pitch estimator will be extreme time consuming if using a signal with length 1 second, equal 2667 samples after down-sampling. Therefore, for this test, we have buffered the original aggregated signals into blocks of $100\,\mathrm{ms}$ with $67\,\%$ overlap. The test and training data can be seen in Table C.1 It is clear that, the measurement features matrix is huge, not because the number

Names	Number of $100\,\mathrm{ms}$ Intervals
cleanAirCraftSignal_for_training	9454
noisyAirCraftSignal_for_training	9454
mixedAirCraftSignal_for_training	9454
noiseSignal_for_training	9454
noisyAirCraftSignal_for_testing	28243
noiseSignal_for_testing	28243

Table C.1: Overview of testing and training data, and the number of $100\,\mathrm{ms}$ intervals in each.

of features, but because of the number of measurements and it was a problem to apply the non-parametric KNNC, which is the reason why only QDC will be shown in this Appendix.

Pitch and Related Features Results with Clean Aircraft Signal for Training

The 5-fold, 100 iterations cross validation errors for the two trained classifiers are

$$\text{QDC Cross Validation Error} = 25.27\,\%$$

When applying the two trained classifier on the test data, the QDC returns the confusion matrix seen in Table C.2

True Labels	**Estimated Labels**		
	Aircraft	non-Aircraft	Totals
Aircraft	12185	16058	28243
non-Aircraft	3915	24328	28243
Totals	16100	40386	56486

Table C.2: Confusion matrix on Pitch and Related Features by Quadratic Distance Classifier trained on Clean Aircraft Signal, having error rate $35.36\,\%$

Pitch and Related Features with Noisy Aircraft Signal for Training

The 5-fold, 100 iterations cross validation errors for the two trained classifiers are

$$\text{QDC Cross Validation Error} = 35.10\,\%$$

When applying the two trained classifier on the test data, the QDC returns the confusion matrix seen in Table C.3

True Labels	**Estimated Labels**		
	Aircraft	non-Aircraft	Totals
Aircraft	13008	15235	28243
non-Aircraft	4562	23681	28243
Totals	17570	38916	56486

Table C.3: Confusion matrix on Pitch and Related Features by Quadratic Distance Classifier trained on Noisy Aircraft Signal, having error rate $35.05\,\%$

Pitch and Related Features with Mixed Aircraft Signal for Training

The 5-fold, 100 iterations cross validation errors for the two trained classifiers are

$$\text{QDC Cross Validation Error} = 30.52\,\%$$

When applying the two trained classifier on the test data, the QDC returns the confusion matrix seen in Table C.4

True Labels	Estimated Labels		
	Aircraft	non-Aircraft	Totals
Aircraft	12703	15540	28243
non-Aircraft	4302	23941	28243
Totals	17005	39481	56486

Table C.4: Confusion matrix on Pitch and Related Features by Quadratic Distance Classifier trained on Mixed Aircraft Signal, having error rate $35.13\,\%$

Conclusion of Pitch and Related Features

It is clear both from cross validation error rates and the error rates on experiment test data, that this approach, is not superior compared to MFFC from experiment in Section 4.2 or Cyclic Analysis from experiment in Section 4.3. In this experiment the best cross validation error rate was $25.27\,\%$, obtained from clean data set. The best error rate on experiment test data was $35.05\,\%$, obtained when classifier was trained on noisy aircraft signal and tested on similar data. However, observing the error rates from the two other trained classifier, they were within $0.36\,\%$, which indicates the training does not changes the performance of a classifier based on chirp pitch and related features. If the features in this experiments should be working directly, further investigation are needed on how to use them directly. It might be sufficient to detect the center of a signal where the sound pressure level is at it maximum and within a small time window around this position, the features can be estimated and might give a better results. This gives all kinds of problem like how to find the center, how often to find center, threshold for maximum sound pressure level and lastly a truck passing by are believed to give very similar pitch, chirp and amplitudes.

We believe the features cannot stand alone and they are needed to be used in a model, where they can be used for estimating significant unique features. This is exactly the reason why, we found the flight parameter estimation interesting, see experiment in Section 4.5, since the HCM chirp pitch estimator in this case could provide the IF directly, which is crucial for the flight parameters estimator to work.

Appendix D

CD Content

Observe, it has not been possible to share audio files and the used MFCC implementation, due to the NDA with Brüel and Kjær. Neither has it been possible to share saved -mat files, for the same reason, and because of the size of these.

The CD contains

- A/V - Product "video-final.mp4"
  - It is a movie made with "Gource" which is visualization tools used in this case for visualizing my Git repository
- Report "master.pdf" - report in digital form
- MATLAB - Used MATLAB scripts and functions in report.
  - anc_lib : it is a library, which here is cleaned and only shared the functions used in the report
  - RMSE_test : The HCM chirp pitch estimator RMSE test and the flight parameter RMSE test. All done on synthetic signal
  - non_executable : It contains scripts, which cannot be ran without the audio data base protected by the NDA with Brüel and Kjær
  - signalModelVer1.m : a script for producing a synthetic signal when an aircraft is passing by
  - HCM_IF_verus_Doppler_model_Flight_Parm_estimation_test.m : It will produce a synthetic signal, using HCM chirp pitch estimator for estimating IF and it uses the flight parameter estimator to estimate the flight parameters, which are all known values. Lastly, the Doppler IF model will be compared with the estimated IF