Institutionen för systemteknik
Department of Electrical Engineering

Examensarbete

Water Depth Estimation Using Ultrasound Pulses for Handheld Diving Equipment

Examensarbete utfört i Signalbehandling
vid Tekniska högskolan vid Linköpings universitet
av

Katarina Mollén

LiTH-ISY-EX--15/4825--SE

Linköping 2015

Linköpings universitet
TEKNISKA HÖGSKOLAN

Department of Electrical Engineering
Linköpings universitet
SE-581 83 Linköping, Sweden

Linköpings tekniska högskola
Linköpings universitet
581 83 Linköping

Water Depth Estimation Using Ultrasound Pulses for Handheld Diving Equipment

Examensarbete utfört i Signalbehandling
vid Tekniska högskolan vid Linköpings universitet
av

Katarina Mollén

LiTH-ISY-EX--15/4825--SE

Handledare: **Michael Roth**
ISY, Linköpings universitet
Anders Brodin
Aqwary AB

Examinator: **Fredrik Gustafsson**
ISY, Linköpings universitet

Linköping, 3 februari 2015

Avdelning, Institution
Division, Department

Automatic control
Department of Electrical Engineering
SE-581 83 Linköping

Datum
Date

2015-02-03

Språk
Language

☐ Svenska/Swedish
☒ Engelska/English
☐ _____

Rapporttyp
Report category

☐ Licentiatavhandling
☒ Examensarbete
☐ C-uppsats
☐ D-uppsats
☐ Övrig rapport
☐ _____

ISBN
__

ISRN
LiTH-ISY-EX--15/4825--SE

Serietitel och serienummer
Title of series, numbering

ISSN
__

URL för elektronisk version
-

Titel
Title

Skattning av vattendjup med ultraljudspulser för mobil dykarutrustning
Water Depth Estimation Using Ultrasound Pulses for Handheld Diving Equipment

Författare
Author

Katarina Mollén

Sammanfattning
Abstract

This thesis studies the design and implementation of an ultra-sonic water depth sounder. The depth sounder is implemented in a hand-held smart console used by divers. Since the idea of echo sounding is to measure the flight time between transmitting the signal and receiving the echo, the main challenge of this task is to find a time-of-flight (ToF) estimation for a signal in noise. It should be suitable for this specific application and robust when implemented in the device. The thesis contains an investigation of suitable ToF methods. More detailed evaluations of the matched filter, also known as the correlation method, and the linear phase approach are done. Aspects like pulse frequency and duration, speed of sound in water and underwater noise are taken into account.

The ToF-methods are evaluated through simulation and experiments. The matched filter approach is found suitable based on these simulations and tests with signals recorded by the console. This verification leads to the implementation of the algorithm on the device. The algorithm is tested in real time, the results are evaluated and improvements suggested.

Nyckelord
Keywords

Depth sounding, Echo sounding, Underwater ultrasounding, Time of flight estimation, Time delay estimation, Matched filter, Cross correlation, GCC, PHAT, ML

Sammanfattning

Denna rapport behandlar skattning av vattendjup med hjälp av ultraljudspulser och implementation av detta. Djupmätaren implementeras i en handhållen dykarkonsoll. Eftersom grundidén i ekolodning är att mäta tiden mellan att pulsen skickas iväg och att ekot tas emot är en stor del av utmaningen att hitta en lämplig metod för att skatta flykttiden för en signal i brus. Metoden ska passa för detta användingsområde och vara robust. Rapporten tar upp tidigare forskning gjord inom flykttidsestimering. De metoder som utvärderas för implementation är det matchade filtret, också kallad korrelationsmetoden, och linjär fas-metoden. Andra aspekter som avvägs och utreds är pulsfrekvens och pulsvaraktighet, ljudets hastighet och brus under vattnet.

Metoderna för att skatta flykttid utvärderas genom simuleringar. Det matchade filtret bedöms vara lämpligt baserat på dessa simuleringar och experiment med data inspelad med konsollen. Denna verifikation leder till att algoritmen implementeras på konsollen. Den implementerade algoritmen testas i realtid, resultaten utvärderas och förbättringar föreslås.

Abstract

This thesis studies the design and implementation of an ultra-sonic water depth sounder. The depth sounder is implemented in a hand-held smart console used by divers. Since the idea of echo sounding is to measure the flight time between transmitting the signal and receiving the echo, the main challenge of this task is to find a time-of-flight (ToF) estimation for a signal in noise. It should be suitable for this specific application and robust when implemented in the device. The thesis contains an investigation of suitable ToF methods. More detailed evaluations of the matched filter, also known as the correlation method, and the linear phase approach are done. Aspects like pulse frequency and duration, speed of sound in water and underwater noise are taken into account.

The ToF-methods are evaluated through simulation and experiments. The matched filter approach is found suitable based on these simulations and tests with signals recorded by the console. This verification leads to the implementation of the algorithm on the device. The algorithm is tested in real time, the results are evaluated and improvements suggested.

Acknowledgments

I would like to thank everyone at Aqwary AB for giving the opportunity and helping me to conduct this master thesis at their company. I would also like to thank for the support from Linköping University from my supervisor Michael Roth and my examinator Fredrik Gustafsson.

Linköping, February 2015
Katarina Mollén

Contents

Notation

Förkortningar

Förkortning	Betydelse
TOF	Time of flight
SNR	Signal to Noise Ratio
SONAR	Sound Navigation and Ranging
LP	Linear Phase
MF	Matched Filter
OFDM	Orthogonal frequency division multiplexing
TL	Transmission loss
SC	Smart Console
ML	Maximum Likelihood
PHAT	Phase Transform
GCC	Generalized cross correlation

List of Figures

List of Tables

1

Introduction

This chapter introduces the subject and gives an overview of the conducted work. It gives the reader a background of the problem, describes the approach, prior work in the field and the structure of the rest of the report.

The description of the approach is divided into an experiments-part and an implementation-part.

1.1 Background

The company Aqwary has developed a product called “Aqwary Smart Console”. This smart console (SC) is designed to be hand held by divers and to supply them with information underwater. The information consists of for example water depth, air pressure in the tank. It is equipped with a dive computer that calculates the time frames for the diver. It can also communicate with other devices, share its information and send out alarms.

Currently, the SC can measure depth through a pressure sensor. This information is necessary for the diver to be able to calculate time frames. Even though not necessary for a safe dive, enquiry has been made about the possibility to implement an application to measure the distance to the bottom. This information could be interesting to let the user get a better perception of his or her surroundings when it fails to do so visually. Range is difficult to decide visually in water and unclear or dark water can make it even harder. The conventional method to measure depth in water is to use echo sounder. The echo sounder uses ultrasonic pulses to estimate the depth. This means that it is not affected by these factors and in most cases gives a better estimation than a visual attempt. Since the SC already has transducers that can transmit ultrasonic sound waves, the idea is to use these to implement an application that can measure depth.

1.2 Thesis objective

The objective of this thesis is to evaluate the possibility to use the four piezoelectric transducers already integrated with the SC, to implement a ultrasonic depth measurer. The transducers are currently used for communication between devices.

The idea is to send out a pulse and estimate the time it takes for the wave to be reflected in the bottom and returned to the transducer. Once the time is known, the known speed can be used to calculate the distance travelled. That distance should be twice the distance to the bottom, since the wave travels forth and back. This device will be used in groups of divers, each with its own console. A problem for the depth sounder is to find the right echo and to not mix it up with the signal and signal echoes other devices might make.

The goal is to verify if such application is possible or not. This is being done in the software Matlab. A first step is to simulate signals and to try out and evaluate different methods. Then, a suitable method is selected and used on real data. If the results show that it is possible to measure the depth by this method on the console, an application should be created to demonstrate this function.

1.3 Technical specifications of the smart console

The SC has the dimensions 205.3 mm x 86 mm x 42.8 mm and weighs 435 g without batteries. It uses a 536 MHz processor. It can measure tank pressure to estimate gas left in tank and water pressure to estimate depth. The SC also consists of a magnetometer and an accelerometer and uses these to create a digital compass. There are four transducers placed on the SC. Two are placed in the front, aimed forward, and two at each side, angled backwards. Three of them are transmitters and all four are receivers. One of the two transducers placed in the front is only receiver. Their placements are shown in Figure 1.1 and can be seen in Figure 1.2.

The transducers have a center frequency of 40 kHz, an output sound pressure of $\geq 100dB$ and a directivity of 75°. Its operating temperature range is $-30°C$ to $85°C$.

Figure 1.1: *The SC, placement of transducers marked with arrows.*

Figure 1.2: *The SC seen from different angles.*

1.4 Experiments

To be assured that the theories hold and that the application is possible to implement, the problem is first evaluated in Matlab. The first evaluation is made with simulated signals. After this step, the algorithms are applied to real data.

1.4.1 Simulations

To try out the theory, a pulse is synthesized and the echo simulated and corrupted with Gaussian noise. Two different methods for estimation of the time-of-flight (ToF) are tried out and evaluated. These methods are described in Section 2.2. The result from this step is described in Chapter 3. After a suitable method is found, the same calculations are tried out on data collected with the transducers on the SC.

1.4.2 Calculations on real data

Data collected with the SC through the transducers are processed in Matlab. The method with the best results from the simulations is tested to see how it performs with real data.

To obtain the test data, a pulse is transmitted and then the transducers record for a while. Since the required energy of the pulse is unknown, a few different pulse durations are tested, 1 ms, 2 ms and 3 ms.

For this application, only the two transducers in the front are used. It requires the user to aim the device downwards to the bottom whilst collecting data.

The result is described in Chapter 4. It is used as a verification of the possibility to implement a depth-measuring ultrasonic application in the SC.

1.5 Implementation

This part is done after a verification of its possibility has been made by experiments. An application is created, that could be used by the SC. The application is able to handle the algorithm chosen during the experiments whilst getting data from the transducers in real-time. It then displays the result visually to the user. The implementation is in C++ and is described in Chapter 5.

1.6 Related work

Echo sounders have been used since the beginning of World War I. Alexander Behm was granted German patent for the echo sounder 1913, [Höhler, 2002] and is by many regarded as the inventor of the echo sounder. Many articles have been written about its theory. Both the hardware and the processing of the data have been improved over the decades. The main purpose is to improve the accuracy, resolution, range and robustness of the sounder. As sound emitters, explosive cartridges were used in the beginning by the inventor Behm. Nowadays, it is common to use so called transducers. They have experienced a lot of development during the past several decade from being very simple to be able to emit multiple signal frequencies. Their development is described by Bushberg et al. [2011].

Regarding the signal processing, the main research area is the ToF estimation of the transmitted signal. This ToF estimation problem applies to more applications than echo sounding and sonar. Examples of other applications that needs to calculate the ToF is radar, wireless systems, the measuring of spontaneous electrical activity of the brain with electrodes in biomedicine etc. This report focuses on the methods used in echo sounders, where usually a simpler method is used. This as it needs less accuracy than many other applications.

Methods for ToF estimation methods have been widely discussed during the last four decades. In depth sounders a common technique used is the Matched

filter (MF). It was first formalized in 1943 by Dwight O. North, but was security-classified during World War II and was not available to the public until 1963 when it was re-published. The MF is based on the cross-correlation of the pulse and the received signal, [North, 1963]. It is also referred to as the cross-correlation method. Modifications of the MF have been done, e.g. the generalized cross-correlation (GCC) method presented by Knapp and Carter [1976] that introduces a weight function in the frequency domain to avoid the spread of the peak of the cross-correlation generated in the MF. There are different weight function such as the Roth weighting function, Smoothed Coherence Factor (SCOT) the Maximum Likelihood (ML) and the Phase transform (PHAT). The performance using different weighting functions are compared by Yanjie et al. [2014]. The theory of the GCC is explained in Section 2.2.2.

Another method is based on the cross-spectrum between the transmitted and the received signal is suggested by Piersol [1981] and Gustafsson et al. [2010]. This method is in this report called the Linear phase (LP) method and is evaluated in Chapter 3. The LP is explained in Section 2.2.4.

Another method, not based on the cross-correlation, the so called Average square difference function is presented by Jacovitti and Cusani [1987].

1.7 Structure of the work

The thesis starts with an introduction and continues with a theory chapter. The theory chapter is meant to give the reader a bit of background to the experiments and implementation.

The three next chapters, Chapters 3, 4 and 5, describe the design and results of the experiments and the implementation.

The final chapter includes some concluding remarks on the results and suggestions on future work that could be done to improve the implementation.

2

Theory

In this chapter all the theory mentioned in the report is being shortly described.

2.1 Echo sounding

Using ultrasonic sound waves to create a listening device has been done since early 1900s when the sonar was developed. According to Höhler [2002] the increasing interest for underwater sounding came as a result of two big events, namely the sinking of Titanic (1912) and the start of World war I (1914). The war created interest in obtaining information of direction by sounding. The Titanic accident was said to make the German scientist Alexander Behm interested in using ultra acoustics for bathymetry. As sound emitters he used explosive cartridges. He registered a German patent 1913 but marketed his device not before 1920. Since then, echo sounding technology has developed to become more accurate and, in some cases, to cover larger areas.

Echo sounders are a type of sonar. They are either single or multi-beam, [Bremen, 2010]. A multi-beam echo sounder gives information of the topography of the sea floor by covering a wide swath beneath the sounder with a fan of beams. This in contrast to the single-beam echo sounder that only uses one beam to give the depth right below the sounder device. The multi-beam sounder has to keep track the angles of the echoes, compensate for roll- and heave motions and speed of the ship the sounder is mounted on. The single-beam sounder calculates the distance based on the first echo that arrives to the receiver. The difference between the two types of sounders is illustrated in Figure 2.1.

2.1.1 Single-beam echo sounding

The basic principle of a single-beam echo sounder is to use a transducer to send out a pulse and then detect the echo of the sea floor. The distance travelled can then be calculated and divided by two, since the signal travels forth and back. It is most common to find the sounding devices mounted on the hull of a boat. The accuracy varies, depending on the frequency and beam width of the emitted signal and the post-processing of the received signal. The post-processing needs to take into account that the speed of sound varies with temperature, depth and salinity and how accurate ToF estimation method is.

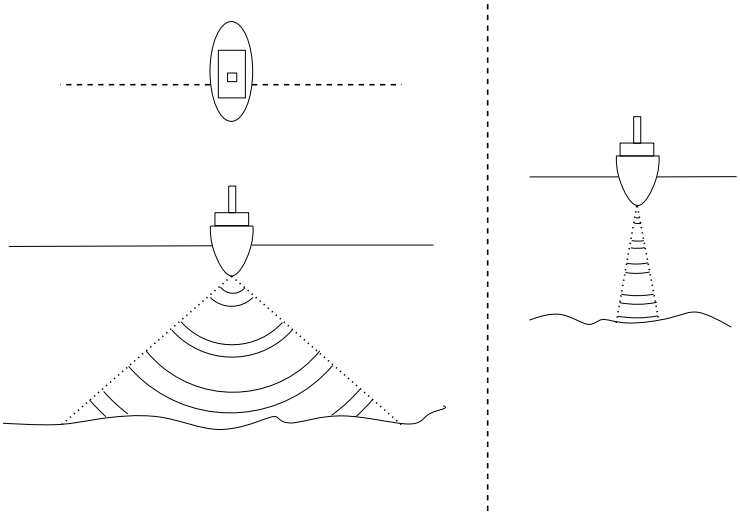

Figure 2.1: Left, top: *Multi-beam echo sounder seen from above.* ***Left, bottom:*** *Multi-beam sounder, seen from the side.* ***Right:*** *Single-beam echo sounder, seen from the side.*

2.2 Time-of-Flight estimation

Applications like sonar and radar are based on finding the time it takes for a signal to travel from a transmitter to a receiver. To do this, the echo of the signal in the received signal must be detected. The received signal is distorted, attenuated and corrupted with noise. There are several methods to do this. A common method is to use a so called matched filter. An extension of this method is the generalized cross-correlation (GCC) which can be applied using different weighting functions. Another approach is the linear phase (LP) method. Other simpler methods exist, such as thresholding, curve fitting and the sliding window method. The system considered is represented in Figure 2.2.

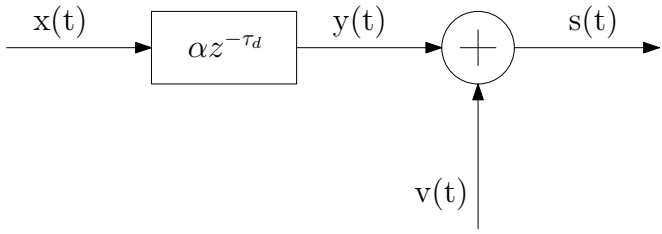

Figure 2.2: *Overview of the system between transmitter and receiver. $x(t)$ denotes the transmitted signal and $y(t)$ the delayed and attenuated version of it. $v(t)$ is noise and $s(t)$ is the returning echo.*

2.2.1 Matched filter

To separate the noise and the echo, you would like the ratio between the noise variance and the signal energy to be high. Then you would be able to separate these two and find at what time the signal arrived, the time-of-flight (ToF). In 1943 Dwight O. North published a paper describing a filter that maximized the signal-to-noise ratio (SNR). This is later called a matched filter, also sometimes known as the North filter. It is described by Turin [1960] and Levanon and Mozeson [2004]. The received signal is filtered and if the SNR is sufficiently high, the output have a peak at the estimated ToF. This because the filter is designed to maximize the SNR. The filter is defined as:

$$h(t) = x^*(-t) \tag{2.1}$$

where $x^*(-t)$ is the complex conjugated, time-reversed emitted signal. The filtered signal $z(t)$ can be calculated through the convolution between $s(t)$ and $h(t)$,

$$z(t) = s(t) \otimes h(t) = s(t) \otimes x^*(-t)\,. \tag{2.2}$$

This is equal to the cross-correlation between $x(t)$ and $s(t)$

$$z(t) = x(t) \star s(t)\,. \tag{2.3}$$

This method is therefore also referred to as the cross-correlation method. The output of the MF can be written as

$$z(\tau) = x(t) \star s(t) = \int_{-\infty}^{\infty} x^*(t)s(t+\tau)\,\mathrm{d}t = \mathcal{F}^{-1}(X^*(f)S(f)) \tag{2.4}$$

where $\mathcal{F}^{-1}$ denotes the inverse Fourier transform and $X(f)$ and $S(f)$ if the Fourier transform of $x(t)$ and $s(t)$ respectively. Since in practice the observation is finite, the cross correlation can only be estimated.

A derivation of this can be found in Appendix A.

In Figure 2.3 a block diagram of the matched filter set-up is shown. $x(t)$ is the emitted signal, $v(t)$ is white Gaussian noise and $z(\tau)$ is the output of the matched

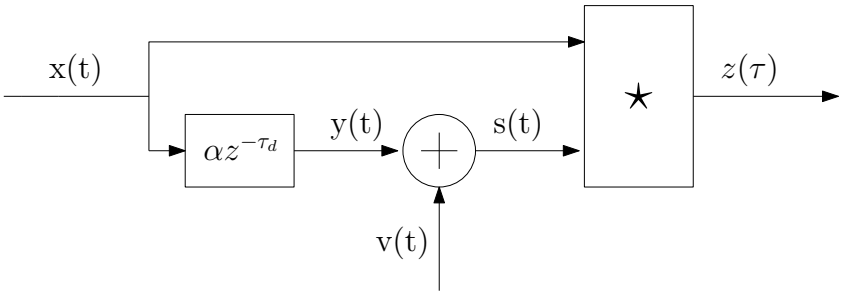

Figure 2.3: *Overview of matched filter. $x(t)$ denotes the transmitted signal and $y(t)$ the delayed and attenuated version of it. $v(t)$ is noise and $s(t)$ is the returning echo. $\star$ denotes a cross-correlator operator.*

filter. y(t) is the attenuated, time-shifted echo as

$$y(t) = \alpha x(t - \tau_d) \tag{2.5}$$

where α is the attenuation-factor and τ_d is the time shift.

The output of the matched filter has, since it maximizes the SNR, a distinguishable peak at the estimated ToF. To accentuate the peak, the output can be low pass filtered and squared.

Example 2.1

In Figure 2.4 the output of the MF can be seen, when the input is a 5-seconds long, delayed chirp pulse, scaled with a factor of 0.6, with noise with variance 0.8 added. With the human eye, a distinction would have been impossible, but the output of the MF shows that it can be located.

Since the filter is $h(t) = x(-t)$ it would be unrealisable in real time since $h(t) \neq 0 \; \forall t < 0$. In order to make it causal, it has to be right shifted as many samples as the chirp pulse is long. As a consequence of this, the output peak is also shifted the same number of seconds to the right. This is why the peak of the output of the matched filter in the example is located at 10 seconds instead of 5.

This filter is the optimal maximizer of the SNR. However, it has drawbacks. Barshan [2000] takes up a few of these. The procedure is time consuming because of the required correlation operation. It also requires the whole echo to be observed before a detection can be made. Because of this, this method might not be suitable to use when the duration of the echo is longer than the ToF. For an echo of duration 0.5 ms this means approximately a minimum target distance of 0.4 m.

An other drawback is that estimation of the time delay can only be made at a sampling moment. That means that there is an error, if the true delay occurred between samplings. If the sampling rate is high enough, this error might be negligible. There are several ways to minimize the error caused by this. One way is to zero-pad both the received signal and the pulse in the frequency domain, see Section 2.2.3.

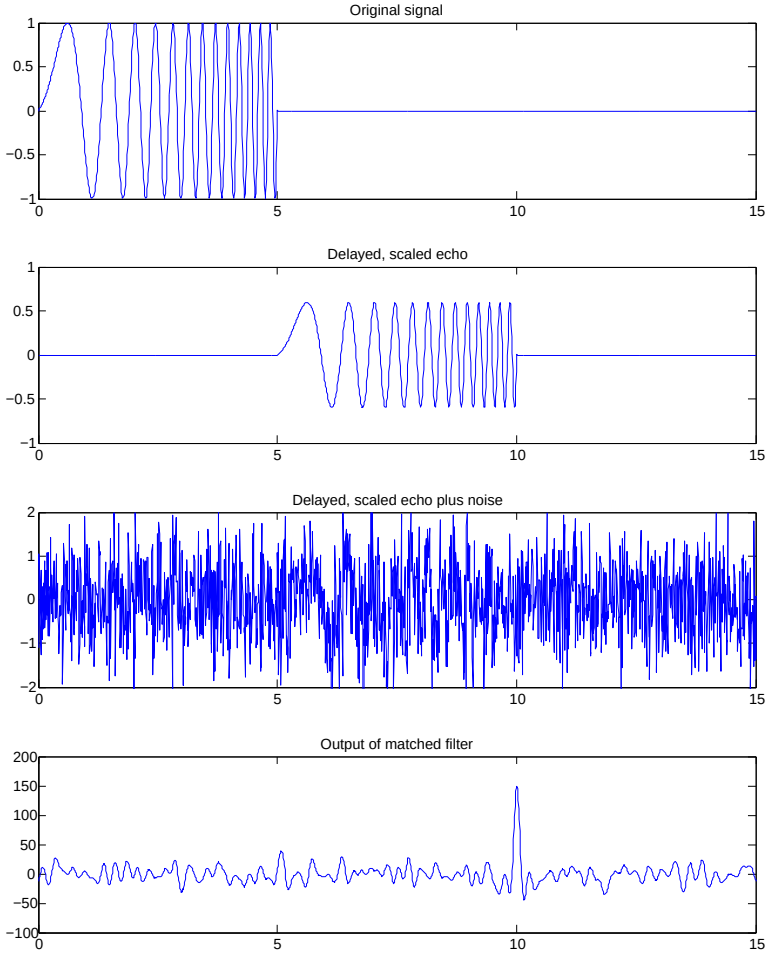

Figure 2.4: *Demonstration of the matched filter. The matched filter is a method used to estimate the ToF. First figure shows the transmitted pulse. In the second figure, it has been time-shifted and attenuated. In the third figure, the pulse has been corrupted with noise, as it will be in the received signal. The last figure shows the cross-correlation between the pulse and the received signal.*

2.2.2 Generalized cross-correlation method

The generalized cross-correlation (GCC) method is an improved version of the MF. It was first introduced by Knapp and Carter [1976]. It introduces a weight function in the frequency domain to avoid the spread of the peak of the cross-correlation generated by the MF. The purpose of the weight function is to accentuate the incoming echo and simultaneously suppress the noise power.

$$z(\tau) = \int_{-\infty}^{\infty} \Psi(f)\Phi_{xs}(f)e^{i2\pi f\tau_d}\,\mathrm{d}f \tag{2.6}$$

where $\Psi(f)$ is the weighting function and Φ_{xs} is the cross-spectrum between $x(t)$ and $s(t)$. Signals $s(t)$ and $x(t)$ are defined as in Figure 2.2.

There are several popular weighting functions. The most common weighting functions to use are the Maximum Likelihood (ML) and the Phase Transform (PHAT) weighting functions. Other functions are e.g. the Roth processor and the Smoothed Coherent Transformation. Difference in their characteristics make them suitable for different environments. Below the PHAT- and the ML-functions are shortly described.

The Phase Transform Weighted

The Phase Transform (PHAT) weighting function sharpens the peak of the cross-correlation by whitening the input signals. It normalizes the amplitude of the spectral density and uses only the phase information to compute the cross-correlation. It is defined as

$$\Psi_{PHAT}(f) = \frac{1}{|\Phi_{xs}(f)|} \tag{2.7}$$

In the ideal case with no noise, then the system

$$s(t) = \alpha x(t - \tau_d) \tag{2.8}$$

would have the frequency response

$$\begin{aligned} H(f) &= \alpha e^{-i\Theta(f)} \\ \Theta(f) &= 2\pi f\tau_d \end{aligned} \tag{2.9}$$

where $\Theta(f)$ is the phase.

The Wiener-Khinchin theorem dictates that the power spectrum is the Fourier transform of the auto-correlation. From this it follows that the cross-spectrum is defined as

$$\begin{aligned} \Phi_{sx}(f) &= H(f)\Phi_{xx}(f) \\ \Phi_{xs}(f) &= H^*(f)\Phi_{xx}(f)\,. \end{aligned} \tag{2.10}$$

$H^*(f)$ is the complex-conjugate of $H(f)$. In our case, eq. 2.10 leads to

$$\Phi_{xs}(f) = \alpha e^{i\Theta(f)}\Phi_{xx}(f) \tag{2.11}$$

Since $x(t)$ and $s(t)$ is assumed to be wide sense stationary, $\Phi_{xx}(f)$ is positive and real-valued for all f. This gives

$$|\Phi_{xs}(f)| = \alpha\Phi_{xx}(f)\,. \tag{2.12}$$

Together with Equations 2.6, 2.7 and 2.11, this gives

$$z(\tau) = \int_{-\infty}^{\infty} e^{i2\pi f\tau_d} e^{i2\pi f\tau}\mathrm{d}f = \delta(\tau - \tau_d)\,. \tag{2.13}$$

This means that in the ideal case, with no corruption of noise, the output of the GCC-PHAT has a Dirac-pulse at the ToF. This helps to sort out other multipath echoes that are not the direct one. If a multipath echo is too close to the echo of interest, the conventional cross-correlation used in the MF gives the two echoes as one peak. If the GCC-PHAT is used, the echoes are separated and two narrower peaks are seen. As this method takes no consideration of the noise, the performance of the PHAT deteriorates with increasing noise level.

According to Donohue et al. [2007], the performance may increase if a parameter, β is added to the expression as

$$\Psi_{PHAT-\beta}(f) = \frac{1}{|\Phi_{xs}(f)|^\beta}\,. \tag{2.14}$$

Donohue et al. [2007] reaches the conclusion that a β between 0.3 and 0.5 achieves good performance, with narrowband signals favouring lower β values. This method is referred to a PHAT-β.

The Maximum Likelihood correlation

The Maximum Likelihood (ML) weighting function gives the ML solution of the ToF estimation problem, [Knapp and Carter, 1976]. The ML-weighting function attenuates the parts of the signal fed to the correlator in the spectral region where the SNR is the lowest. The ML correlation is defined as

$$\Psi_{ML}(f) = \frac{1}{|\Phi_{xs}(f)|}\frac{|\gamma_{xs}(f)|^2}{1 - |\gamma_{xs}(f)|^2} \tag{2.15}$$

where $\gamma_{xs}(f)$ is the so called magnitude-squared coherence function between signals $x(t)$ and $s(t)$ and is defined as

$$|\gamma_{xs}(f)|^2 = \frac{|\Phi_{xs}(f)|^2}{\Phi_{xx}(f)\Phi_{ss}(f)}\,. \tag{2.16}$$

The ML-weighting function gives less weight to the cross spectra estimate where its samples have larger variance and more weight to it where its samples have less variance.

The drawback is that the weighting is frequency dependent and the performance deteriorates in reverberant environments since the spectral characteristics of the signal then are affected.

2.2.3 Sub-sample estimation

In Sections 2.2.1 and 2.2.2, using either the MF or the GCC approach, it is assumed that the time delay is a multiple of the sampling time. If that is not the case, the time delay is rounded to the closest sample. To improve the accuracy of the estimation and make sub-sample estimation, zero-padding of the signals in the frequency domain can be implemented. This introduces an interpolation in the time domain. The interpolation factor is the same factor that the length increases by the zero-padding in the frequency domain. That is, if the signal is zero-padded to its double length in the frequency domain, it is interpolated with an factor 2 in the time domain.

Example 2.2

The original signal, a sinusoid, is sampled at a relatively low rate. The signal is then zero padded in the frequency domain and transformed back to time. The steps and the result can be seen in Figure 2.5.

A side effect that can be seen in this example is that the amplitude of the signal decreases with the factor it has been zero padded with. This can easily be solved by multiplying the new time-domain signal with the zero padding factor.

A zero-padding in the time domain, done by appending zeros at the end of the time-domain-signal, results in a interpolation in the frequency domain.

2.2.4 Linear-phase method

This method uses the fact that a pure time delay filter has an linear phase. It is described in [Gustafsson et al., 2010]. Consider

$$\Phi_{sx}(f) = \alpha e^{-i2\pi f \tau_d} \Phi_{xx}(f) \tag{2.17}$$

where $\Phi_{xx}(f)$ is the spectrum of x(t), $\Phi_{sx}(f)$ the cross-spectra between $s(t)$ and $x(t)$ and $s(t)$ and $x(t)$ as in Figure 2.2. For a background on this relation, see Equations 2.10 and 2.11.

A pure time-delay of the signal can be seen as if the signal passes through a time-delay allpass-filter. This allpass filter have an linear phase. So, in the case of the echo sounder, if the channel that the signal passes through is seen as a filter with an allpass-part, the argument of the cross-spectra has for the frequencies of the emitted signal, an linear part and look like

$$arg(\Phi_{sx}(f)) = \begin{cases} -2\pi f \tau_d + \phi(f) & \text{if } f \in F, \\ \phi(f) & \text{otherwise} \end{cases} \tag{2.18}$$

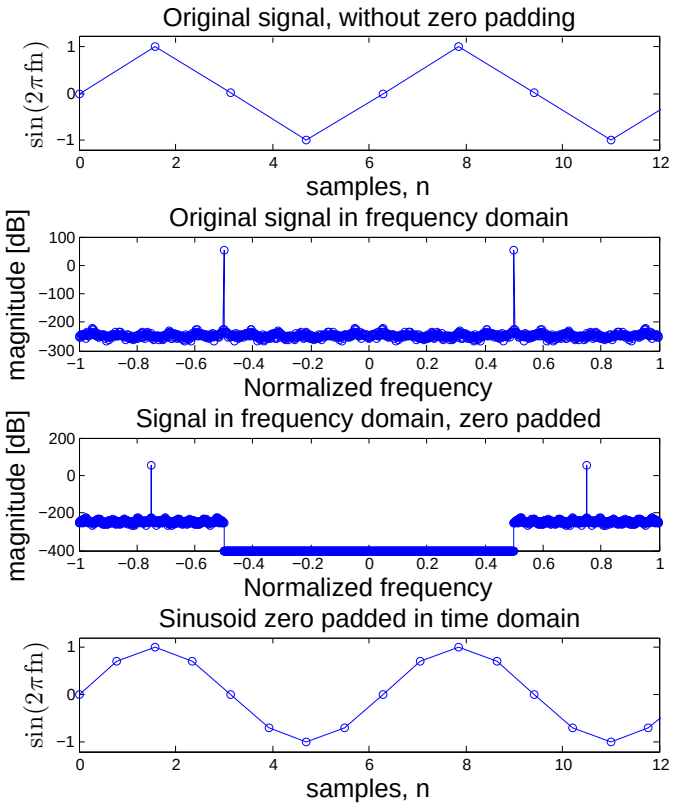

Figure 2.5: *Demonstration of zero padding in the frequency domain*

where $\phi(f)$ is a random number and F is the frequency range of the transmitted signal. If the signal also is of narrow bandwidth B, there is also be an unknown initial phase as

$$arg(\Phi_{sx}(f)) = \begin{cases} \phi_0 - 2\pi f \tau_d + \phi(f) & \text{if } f \in F, \\ \phi(f) & \text{otherwise} \end{cases} \tag{2.19}$$

where $F \in [f_c - B/2, f_c + B/2]$. Observe that the argument has 2π jumps that needs to be unwrapped in order for it to be correct. To estimate $-2\pi f \tau_d$ a line is fitted to the data points that belong to the frequencies of the signal of interest. It is fitted to the data points with the least squares method. The least squares method assumes that the residuals are normally distributed with zero mean and constant variance. The slope of this line is then $-2\pi\tau_d$. If

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \tag{2.20}$$

$$\boldsymbol{Y} = (y_1, y_2, ..., y_n)^T \tag{2.21}$$

$$\boldsymbol{X} = (x_1, x_2, ..., x_n)^T \tag{2.22}$$

where n is the number of observations. Then, using a least-squares-approach, an estimate of β is

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{Y} \tag{2.23}$$

If $\boldsymbol{\beta} = (\phi_0, \tau_d)^T$, $\boldsymbol{Y} = \arg(\Phi_{sx}(\boldsymbol{f}))$, $\boldsymbol{X} = (1\ -2\pi\boldsymbol{f})$ and $\boldsymbol{\epsilon} = \phi(\boldsymbol{f})$, the Equation 2.20 could be written as

$$\begin{pmatrix} \arg(\Phi_{sx}(f_1)) \\ \arg(\Phi_{sx}(f_2)) \\ \vdots \\ \arg(\Phi_{sx}(f_n)) \end{pmatrix} = \begin{pmatrix} 1 & -f_1 \\ 1 & -f_2 \\ \vdots & \vdots \\ 1 & -f_n \end{pmatrix} \begin{pmatrix} \phi_0 \\ \tau_d \end{pmatrix} + \begin{pmatrix} \phi(f_1) \\ \phi(f_2) \\ \vdots \\ \phi(f_n) \end{pmatrix} \tag{2.24}$$

Then an estimate of $\boldsymbol{\beta}$ is

$$\hat{\boldsymbol{\beta}} = \left(\sum_{f \in F} \begin{pmatrix} 1 \\ -2\pi\boldsymbol{f} \end{pmatrix} \begin{pmatrix} 1 & -2\pi\boldsymbol{f} \end{pmatrix} \right)^{-1} \sum_{f \in F} \begin{pmatrix} 1 \\ -2\pi\boldsymbol{f} \end{pmatrix} \arg(\Phi_{sx}(\boldsymbol{f})) \,. \tag{2.25}$$

The variance of the estimator can be obtained as follows:

$$\text{var}(\hat{\boldsymbol{\beta}}) = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\hat{\sigma}^2 \tag{2.26}$$

where $\hat{\sigma}^2$ is a estimate of the noise variance defined as

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} \hat{e}_i^2 \,. \tag{2.27}$$

$\hat{e}$ denotes the residual,

$$\hat{e}_i = y_i - x_{i,1}\hat{\beta} \,. \tag{2.28}$$

Example 2.3

In Figure 2.6 an example of the LP method can be seen. The true ToF is 23 seconds and the OFDM signal has energy in the frequencies [3500, 4000] Hz.

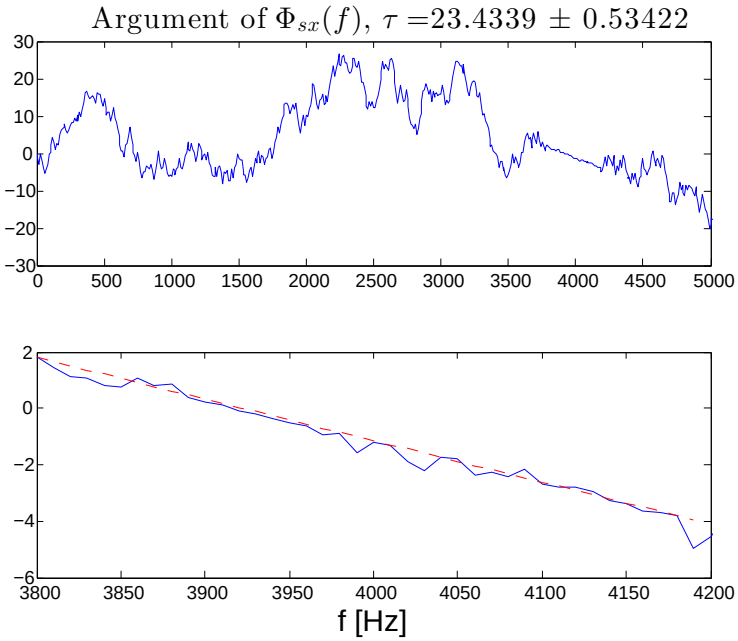

Figure 2.6: *Demonstration of the LP method. The top plot shows the argument of $\Phi_{sx}(f)$ for all frequencies below Nyquist. The bottom plot shows the frequncies corresponding to the signal. The red, dotted, line shows the LS-line fit.*

The method can be preferable to the cross correlation method because it is less affected of small fluctuations between the samples and that it reduces effect of large measurement noise. It also produces sub-sample estimations in contrary to the MF method.

2.2.5 Threshold detection

The simplest method for ToF estimation is the threshold method. This method is discussed in Jackson et al. [2013]. It means that the ToF estimate, $\hat{\tau}_d$ is given when the echo amplitude waveform exceeds a given threshold. It is usually set to about 3-5 times the noise standard deviation. This method is more sensitive to a low SNR than the matched filter. Since the threshold needs to be set to a level above all noise, a rise time between the time of arrival to detection occurs. This time varies for all practical contexts of the method and can not be set as a known offset to compensate for.

2.2.6 Curve fitting

In a article by Barshan [2000] the method of curve fitting is described. A parabola is fitted to the onset of the echo. This is done by setting two thresholds levels, α_1 and α_2, and getting two points of the rising edge of the echo. The parabola, $a_0(t-\tau_d)^2$, that contains these two points is derived by solving the equations

$$\begin{aligned} t_1 &= a_0(t_1-\tau_d)^2 \\ t_2 &= a_0(t_1-\tau_d)^2 \end{aligned} \tag{2.29}$$

eliminating a_0 to get τ_d. A good ratio between α_1 and α_2, according to McMullan et al. [1996] is 2.

2.2.7 Sliding-window method

Barshan [2000] also described the method of sliding-window. A N wide window slides through the received samples, one sample at the time. For every slide, the number of samples exceeding a threshold is being counted. If this number exceeds a second threshold, m, the signal is detected. There are four ways to estimate the ToF once the signal is detected. These are:

1. The ToF estimate is considered to be the first sample of the window that detected the signal.

2. The ToF is estimated to be the first sample that exceeds the first threshold, of the window that detected the signal.

3. The ToF estimate is the center sample of the window.

4. The ToF estimate is the $(N-m)$th sample of the window.

This method has the same drawbacks as the threshold-method, but is a bit more robust to low SNR.

2.3 Transducers

Transducers are used to convert electrical energy to acoustic energy, transmit an ultra sonic signal and then receive an acoustic signal and convert it to an electrical, [Bushberg et al., 2011]. In a piezoelectric transducer, the signal is created by applying an alternating voltage to a piezoelectric crystal. This crystal has the property of changing size when an electrical potential is applied. An alternating voltage causes the crystal to vibrate and create an acoustic signal. If the process is reversed, the crystal can convert an acoustic signal to an electrical, since it generates a voltage when deformed. The amplitude is maximal when the frequency matches the resonant frequency of the crystal.

2.3.1 Center frequency

The signal frequency of a piezoelectric transducer depends on the thickness of the crystal. A thicker crystal produces a lower frequency and a thinner a higher frequency.

When range measuring with ultrasonic sound, the frequency of the transmitted signal matters. The choice of frequency is a trade-off between a few things, [Tetley and Calcutt, 2001]. A low frequency signal is attenuated less when travelling and have stronger penetration capability, but lies in the same frequency spectra as a lot of the ambient noise that occurs in the ocean, see Section 2.8. To get a good SNR, a high frequency signal is therefore most commonly used, usually around 200 kHz for ranges up to 100 m, [International Hydrographic Organization, 2005]. A high frequency also gives better resolution, since the discrimination of two objects placed on a line depends on the wavelength. For longer ranges, a lower frequency must be used, otherwise the echo is attenuated too much to be able to detect. A high center frequency also opens up the possibility for smaller transducers with small beam widths, see Section 2.3.2. Small beam widths lead to higher accuracy.

2.3.2 Transmission beam width

In a transducer, the beam width is defined as the angle of the lobe at a given point. In a lot of transducers this point is set at where the intensity of the beam has dropped down to 70 % of its peak value, [Christ and Wernli, 2014]. 70% of the intensity corresponds to -3 dB.

Jong et al. [2003] discuss the problem that the beam width of the transducer affect the depth estimate. One of the effects that a wide beam width has on the estimation is noticed when an measurement is executed above a slope. The part of the beam that hits the bottom first is reflected first. So that means that the closest part of the slope is the depth estimate and that the echo is somewhat distorted. This is explained in Figure 2.7.

A smaller beam width gives higher precision, but if it is too small there is a risk that the returning echo is so narrow that the transducer misses it.

What beam width the transducer has is dependent on signal frequency and transducer size. The smaller the beam width, the bigger the diameter of the transducer needs to be. The diameter can be reduced for the same beam width if the center frequency is made higher.

The beam width of conventional echo sounders is usually in the order of 30°, [Xu, 2010]. However, since the mid-1980 s there are echo sounders with a beam width less than 5°.

A formula is given by Christ and Wernli [2014] to calculate an approximate beam width and it holds when $L > \lambda$. The beam width, β is calculated as

$$\beta = \frac{\lambda}{L} \tag{2.30}$$

where λ is the wavelength of the signal and L is the diameter of the transducer.

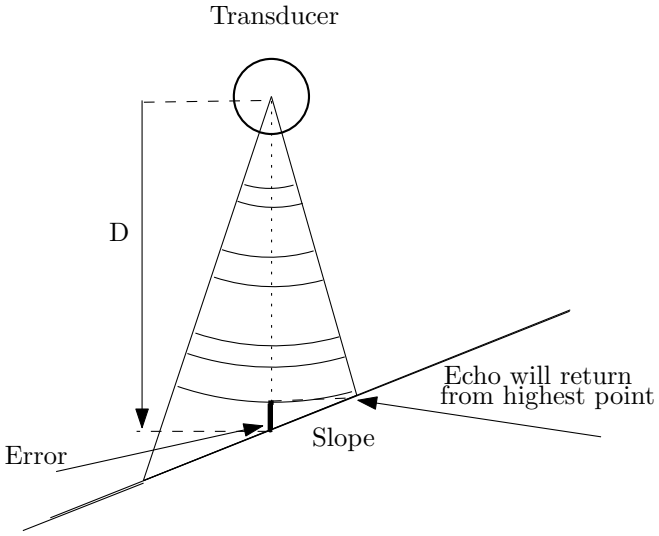

Figure 2.7: *Illustration of possible estimation error, dependent on transducer beam width. The first echo to return to the transducer corresponds to the highest point within the footprint of the beam.*

2.3.3 Ringing

In the rear of the piezoelectric transducer there is a damping block placed, [Bushberg et al., 2011]. The task of this block is to dampen the vibrations after a sound pulse has been emitted. The vibrations that take place after the pulse has been emitted is called ringing. If the damping is high, the ringing will be short, but it will also introduce a broader frequency spectrum. A lighter damped transducer will have a longer ringing, but a narrower spectrum. The ringing is transducer specific and depends also on the dispatched pulse, [Airmar Technology Corporation]. During the ringing no echo can be located and therefore the ringing is one of the factors that decide the minimum distance that can be measured. Figure 4.1 in Chapter 4 illustrates a pulse transmitted from a transducer. The ringing can clearly be seen.

2.4 Multipath propagation

A signal may take multiple paths from the transmitter to the receiver, as shown in Figure 2.8. This is called multipath propagation. The reason it occurs is that the wave field might be reflected or scattered during its travel. As a result, the signal seen at the receiver is a superposition of many wave field components that have taken different paths, [Larsson, 2014]. The possibility of multipath propogation may be needed to be taken into account when making a ToF estimaion, see Section 2.2.

Example 2.4

Figure 2.8 shows one signal taking the straight path down and then up again, travelling the distance $2 * S1$. The other signal is taking another path, resulting in the travelled distance of $S2 + S3 + S4 + S5$. This results in a more attenuated echo arriving later than the first echo and might complicate any attempt to detect the the echo that took the shortest path.

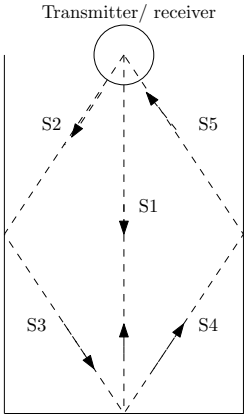

Figure 2.8: *Demonstration of multipath propagation*

2.5 Transmission loss

As the signal travels back and fourth, it loses energy and is attenuated. The biggest factors of transmission loss (TL) are loss due to energy absorption of the water and due to geometrical spread.

2.5.1 Absorption

When the wave is passed through the water, the water absorbs parts of the signal energy. The absorption depends on signal frequency, temperature, salinity and the depth of the water. An simplified model, that only depends on signal frequency is, according to Stojanovic et al. [2008], the Thorp model:

$$\begin{aligned} \alpha &= 0.11\frac{f^2}{1+f^2} + 44\frac{f^2}{4100+f^2} + 2.75 * 10^{-4} f^2 + 0.003 \\ TL &= \alpha R \end{aligned} \tag{2.31}$$

where TL is the TL in decibel, f is the signal frequency in kHz, α is the absorption coefficient given in [db/km] and R is the distance travelled in metres.

2.5.2 Geometrical spread

As the wave travels through water, the intensity of it spreads out on a geometrical surface. This is discussed by Lurton [2010]. This surface is often modelled as a sphere on shorter distances and, underwater if the range is longer, a cylinder. This because the signal is confined between the surface and the floor. The TL due to spherical spread can be calculated through

$$TL = 20log_{10}(R) \tag{2.32}$$

where R is the distance travelled.

2.5.3 Transmission loss dependent on bed sediment reflectivity

According to Tetley and Calcutt [2001], the attenuation due to the reflection in the bottom, changes with the consistency. It also depends on the frequency of the signal. A high frequency signal loses less energy than one with low frequency. For a signal of 24 kHz experiments with an echo sounder were conducted by Tetley and Calcutt [2001] and the results can be seen in Table 2.1. In case of vertical incidence the losses lie around 1-20 dB, [Marage and Mori, 2010]. Hodges [2010] describes several bottom-loss models.

Table 2.1: *TL for differend sea bed consistencies at 24 kHz. Table copied from Tetley and Calcutt [2001], page 24, table 2.1.*

Consistency	TL [dB]
Soft mud	15
Mud/sand	9
Sand/mud	6
Sand	3
Stone/rock	1

2.5.4 Total transmission loss in reflected underwater signals

As the beam in depth sounders is reflected and then returned as an echo to the receiver, the calculations for the TL needs to be done twice. The TL, except for the loss when reflected in sea bed, for a signal reflected in a target R metres away, is then

$$TL = 2 * (20log_{10}(R) + \alpha R) \tag{2.33}$$

where R is the target distance in metres and α is the absorption coefficient in [dB/m].

2.6 Speed of sound underwater

The speed of sound varies in water as a function of temperature, salinity and depth. A simplified model for the speed of sound in sea water was proposed by

Medwin [1975]:

$$c = 1449.2 + 4.6T - 0.055T^2 + 0.00029T^3 + (1.34 - 0.01T)(S - 35) + 0.016D \quad (2.34)$$

where c is the speed of sound in m/s, T is the temperature in °C, S is the salinity in ppt and D is the depth in m. The model is limited to the values in (2.35).

$$\begin{aligned} 0 \leq T \leq 35 \text{ °C} \\ 0 \leq S \leq 45 \text{ ppt} \\ 0 \leq D \leq 1000 \text{ m} \end{aligned} \quad (2.35)$$

This means that the approximation of the speed of sound, at $T = 20$ °C, $S = 35$ ppt and $D = 30$ m is 1522 m/s.

It can be seen in 2.35 that the depth dependency is linear. Table 2.2 shows how the depth affects the speed for a constant temperature and salinity on depths down to 50 m.

Table 2.2: *Speed of sound at different depths, T=15°C, S=35 ppt*

Depth [m]	0	10	20	30	40	50
c [m/s]	1506.8	1507.0	1507.1	1507.3	1507.4	1507.6

The speed is also dependent of the temperature and salinity. Table 2.3 shows the speed of sound for different cases with depth constant at 25 m.

Table 2.3: *Speed of sound [m/s] at different temperatur and salinity levels and depth is 25 m based on equation (2.34)*

T [°C] \ S [ppt]	0	0.5	17	30	33	35	37
4	1422	1422	1444	1460	1465	1467	1470
10	1447	1448	1468	1484	1488	1490	1493
20	1482	1483	1501	1516	1520	1522	1524
30	1510	1510	1527	1541	1544	1546	1548

2.7 Salinity in different waters

According to Dahl [1956], the salinity in the sea lies within the interval 30 – 37 ppt. The salinity in brackish water lies between 0.5 – 30 ppt and fresh water has a salinity of 0.5 ppt or less.

Brackish water arises in waters that have a little contact with the sea and have a large inflow of fresh water. For example, the Baltic Sea is one of the biggest brackish waters in the world.

2.8 Underwater noise

Noise is an important factor of a successful ultrasonic range measurement. One source of noise underwater is noise that is self-inflicted, like scattering from the pulse used for the sounding (reverberation noise) and turbulence around the transducer. It also comes from the ambient noise underwater like animals or shipping movements and electrical devices, [Tetley and Calcutt, 2001], [Marage and Mori, 2010].

2.8.1 Reverberation noise

The noise created by the backscattering of the transmitted signal is called reverberation noise. The amplitudes of the transmitted signal and the reverberation noise are proportional. The further away the target is, the higher is reverberation level, since the width of the beam increases with range and then produces more backscatter.

The characteristics of reverberation noise are

- The amplitude is proportional to the transmitted signal amplitude.
- The amplitude is inversely proportional to the distance of the target.
- The reverberation noise has the same frequency as the transmitted signal.

These characteristics makes it hard to suppress this noise by increasing the signal energy, since the noise energy also would increase. It can not be filtered out, since it has the same frequency as the transmitted signal. The reverberation noise is divided into three classes:

Surface reverberation: Caused by reflection in the water surface. It increases with surface turbulence caused by e.g. rough weather.

Volume reverberation: Caused by matter in the water, e.g. fish and mammals.

Bottom reverberation: Caused by scattering of the signal in the bottom. The level of reverberation depends on type of sea bed. Hard material gives more reverberation.

2.8.2 Ambient noise

The ambient noise is not affected by the transmitted signal amplitude, but stays constant. Examples of ambient noise can be, [National Academies Press, 2003]:

Hydrodynamic noise: Noise caused by the water movement, a result of e.g. in tides, wind, current and storms. Examples of hydrodynamic noise are:

- *Rain on the surface*: Causes noise up to 5000 Hz in the frequency band.
- *Turbulence*: Natural in the water, or around the transducers.

- *Thermal noise*: Caused by microscopic agitation of the environment. Is the dominant contributor above 100 kHz.
- *Surface noise*: Due to waves. Lies between 1 Hz and 100 kHz, dependent on the state of the sea.

Seismic noise: Caused by movement of land underwater, like earthquakes and volcanoes. This is a short and very rare source of ambient noise and lies mostly under 100 Hz.

Ocean traffic: Ships, boats, port activity. Close traffic can cause noise over a wide part of the frequency spectrum, but further away only the low frequencies, below 1000 Hz, are heard.

Biological noise: Caused by marine life, like mammals and fishes. They can create noises between 10 Hz and 100 kHz. Cetacean creates sound that lies in 2-200 kHz. The sounds of the snapping shrimp is also a big source of biological noise and it lies around 1- 100 kHz.

2.9 Effect of human usage of ultrasonic sounds on marine life

During the 20th and the 21st century the question on the human use of ultrasonic sounds in the ocean and its effect on the marine life rose. Several reports were published on how low to mid-frequency sonar affected the marine life, and whales specifically, e.g. [Committee on Low-Frequency Sound and Marine Mammals, 1994] and [Frantzis, 1998].

Low frequency ultrasound (~0-1000 Hz) used to detect large objects on a long range carries such small amount of energy and therefore makes a smaller impact on the marine life. High frequency ultrasound is attenuated very fast and for that reason is estimated to not have a large impact. Studies have found that it is most probably the mid-frequency (~1-10 kHz) usage of sonar that has the largest impact on marine life. This type of active sonar is commonly used by the military to detect threats. The sonar used lies around 2-4 kHz. The usage of this frequency range has been found to correlate with stranding of beaked whales in several studies. In 2013 DeRuiter et al. [2013] showed that beaked whales were highly sensitive to mid range frequencies.

No law has been enforced to limit use of mid-range-frequency sonar, but many navies have their own mitigation measures to limit the impact on marine life. The question has still a strong focus among environmentalists. National Academies Press [2003] calls on all producers of ocean noise, such as marine biologists, military navy and the oil industry, to make greater effort to spread awareness of the problem and encourage further researchers on the subject.

3

Simulation Experiments

Since the idea of the depth measurer is to calculate the depth based on the ToF for the transmitted signal, knowing this time is a main factor to obtain an depth estimate. Why this might be difficult is described in Section 2.2. To find a suitable ToF method, the problem is first investigated by doing simulations in Matlab.

3.1 Methods

Two methods are tested to estimate the ToF. The two methods are MF and LP, see Section 2.2.1 and Section 2.2.4. The methods are implemented in Matlab-code and a simulated signal is used as input. The results of these two methods are compared and analysed. This to draw a conclusion on which one of them is most suitable to use on the real data later.

3.2 Set-up

A signal with center frequency 40 kHz and bandwidth of 4 kHz is created with the orthogonal frequency-division multiplexing- method (OFDM), [Gustafsson et al., 2010]. The signal is seen in Figure 3.1. The code used to generate the signal can be found in Appendix B. This signal is then delayed and corrupted with noise to simulate the received signal, see Figure 3.2. The noise is modelled to be white Gaussian noise with standard deviation σ_v. The variance is then varied so that different levels of SNR could be observed.

To correctly simulate the 40 kHz signal under water, it is necessary to model the attenuation of the signal energy. Since the simulations are supposed to resemble a general case, there is no need to use a more complex method that depends on temperature, salinity and water depth. Therefore, the simpler

model, described in Section 2.5 is used for this purpose.

The simulated received signal consists of an attenuated, time-shifted version of the original signal and white Gaussian noise, as illustrated in Figure 3.2. The problem that needs to be solved here is to find the echo in the noisy signal. When it is found, the estimated ToF can be derived. To solve this, the methods described in the section above are used. The speed of sound in water is approximated to 1500 m/s.

Firstly, a simple case with one echo returning is studied. After this, it is investigated how well the methods can handle the phenomena of multipath signals, see Section 2.4. A returning signal with one main echo and several following, more attenuated, echoes and without noise can be seen in Figure 3.3.

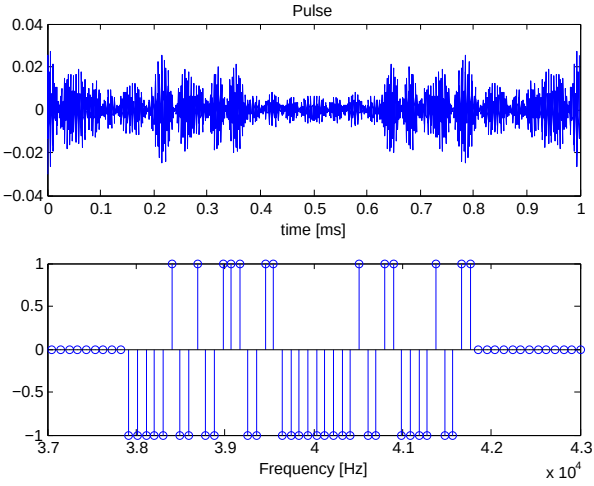

Figure 3.1: *In the simulations a pulse that covers the right frequencies is created using the OFDM method.* ***Top:*** *Pulse in time domain.* ***Bottom:*** *Frequency content of pulse.*

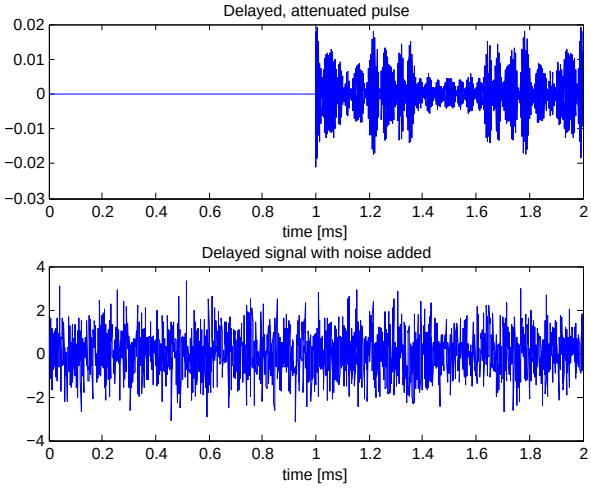

Figure 3.2: *Simulation of returning signal, used as input to the algorithms.* ***Top:*** *Attenuated, delayed echo.* ***Bottom:*** *Attenuated, delayed echo corrupted with noise.*

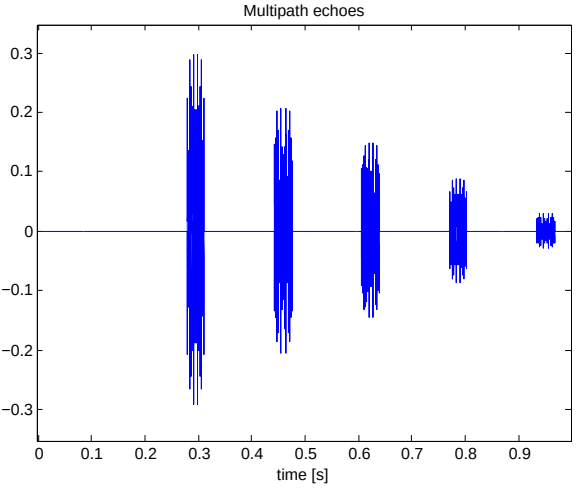

Figure 3.3: *To test the robustness of the ToF estimation methods, a returning signal (here pictured without noise) contaning multipath echoes is simulated.*

3.3 Results

In this section the results from the simulations are presented in form of histograms and boxplots. The standard deviation of each case and how many of the estimates that had a relative error of 5 % or less are presented in tables.

3.3.1 Single echo

In these simulations, the received signal consists of a single time shifted echo corrupted with noise.

The result is documented here in form of histograms at 20 and 40 dB SNR and as boxplots for different SNR. The results from the MF-case can be seen in Figures 3.4 and 3.6. The result from the LP-case can be seen in Figure 3.5 and in 3.6. In Table 3.1 the percentage of estimates with relative error under 5 % and the standard deviation for the different cases are being shown.

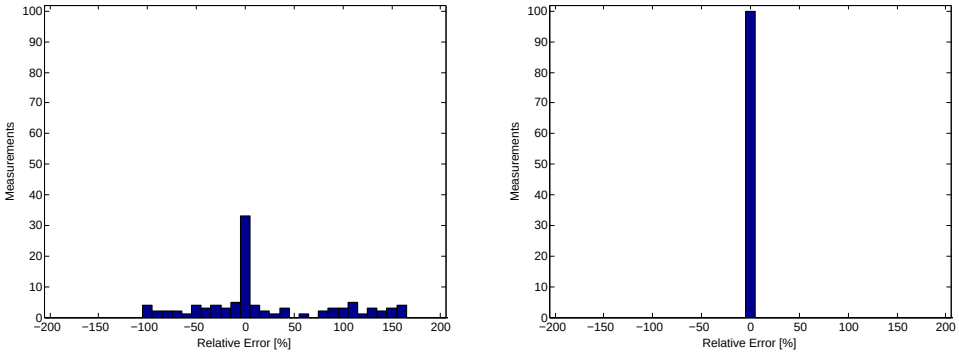

Figure 3.4: *Results from simulation using the MF method. Simulated distance 50 m.* ***Left:*** *SNR 20 dB* ***Right:*** *SNR 40 dB*

3.3.2 Multipath echoes

A possible case that could happen in reality is where the received signal contains the echo from the bottom followed by several multipath echoes, see Section 2.4. To simulate this, a received signal with strong multipath echoes following the first echo is simulated. It can, with the noise removed, be seen in Figure 3.3. This signal is used as the received signal and the MF and LP methods are tried out to see how they handles this case. The results from the MF-case can be seen in Figure 3.7 and the results from the LP-case in 3.8. The results are summarised in Table 3.2 that shows the percentage of estimates with relative error under 5 % and the standard deviation for the different cases.

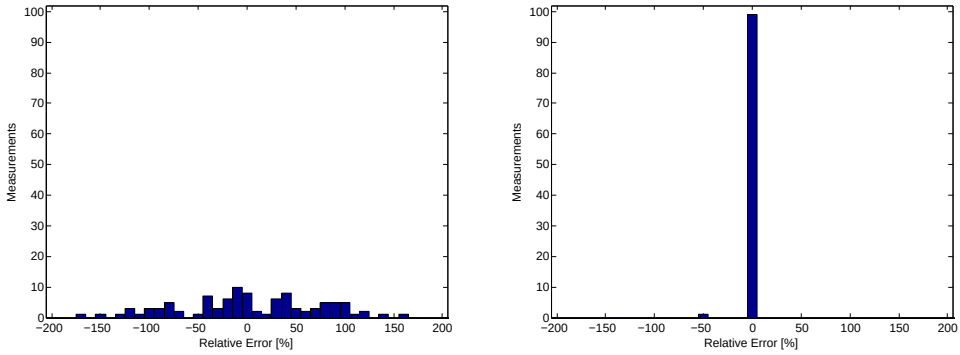

Figure 3.5: *Results from simulation using the LP method. Simulated distance 50 m.* ***Left:*** *SNR 20 dB* ***Right:*** *SNR 40 dB*

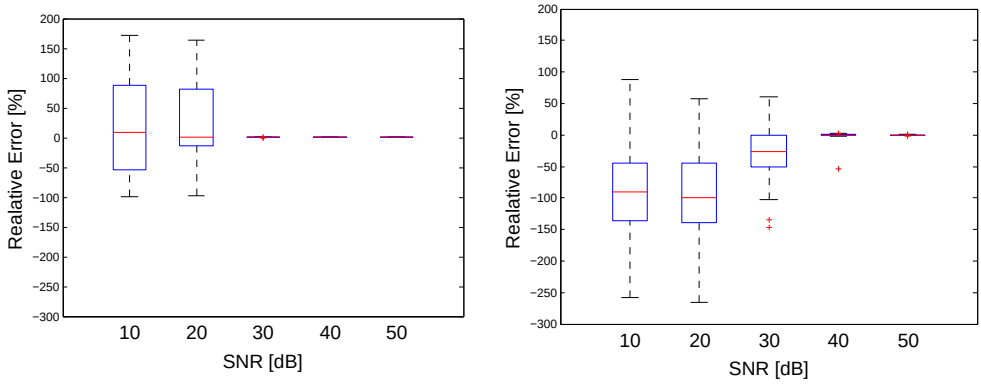

Figure 3.6: *Boxplot of the relative error of the depth measurements. Each data set contains of 100 measurements.* ***Left:*** *MF method.* ***Right:*** *LP method*

Table 3.1: *Result of LP and MF metods, single echo*

Method	SNR [dB]	Relative error < 5% [%]	Standard deviation [m]
MF	20	34	34.58
MF	40	100	0.00
LP	20	6	35.19
LP	40	99	7.34

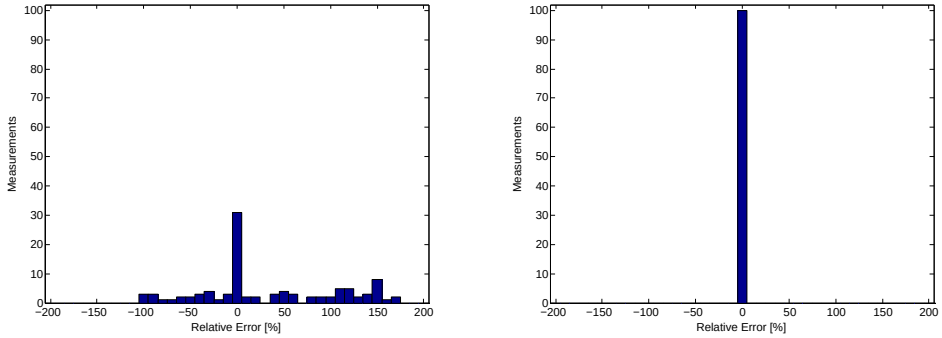

Figure 3.7: *Results from multipath simulation using the MF method.* ***Left:*** *SNR 20 dB* ***Right:*** *SNR 40 dB*

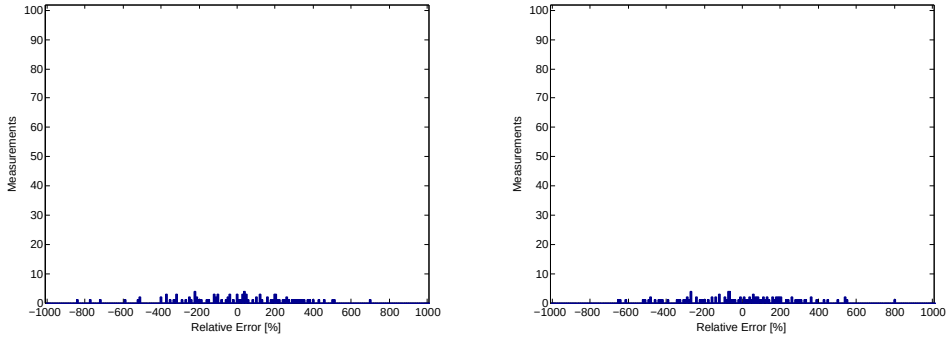

Figure 3.8: *Results from multipath simulation using the LP method.* ***Left:*** *SNR 20 dB* ***Right:*** *SNR 40 dB*

Table 3.2: *Result of LP and MF metods, multipath echoes*

Method	SNR [dB]	Relative error < 5% [%]	Standard deviation [m]
MF	20	32	36.71
MF	40	100	0.01
LP	20	1	145.28
LP	40	0	146.01

3.3.3 Conclusion

When simulating the simpler case, with a single returning echo, the LP-method gives a somewhat better result. In the case with high SNR both of the methods

have almost all of their distance measurements within the 5%- relative error boundaries. The exception is a single measurement from the LP-method that has a relative error of 54%. The low-SNR case shows a large spread for both cases. The MF method has a narrower spread and a peak at 0%. The LP is more evenly spread and does not have such dominant peak at 0%.

The boxplots show that the MF has the median close to 0% and a rather wide spread up to 20 dB SNR. For the cases above the median is at 0% and no spreading. The LP shows spread for cases with SNR up to 30%. The 10 and 20 dB SNR cases have their median far from the desired 0%. The 30 dB SNR case has its median at -25% and is still not adequate, but seems to start to converge towards 0%. The cases with higher SNR have median at 0% and very little spread.

When the case of multipath echoes is simulated, the result from the MF method does not show so much change from the case with a single echo. The LP method on the other hand becomes much more spread out. This is probably a consequence of that the LP method uses all the samples with the signal frequency to estimate the time delay. Since the multipath echoes also have the same frequency, they are not separated from the first echo, causing distortion in the distance measuring. The conclusion that the LP method does not handle multipath echoes very well is drawn.

The MF method seems to converge faster than the LP and at low SNR it might still be possible to distinguish the correct time delay. In addition to this, the MF does not seem affected of multipath echoes contrary to the LP. The final conclusion is that the MF method is the most robust of the two methods compared in the simulations.

3.4 Discussion

The simulations are used to get an initial overview over the problem and to identify a suitable method for the ToF estimation. The model of the channel used is rather simplified and not expected to include all properties of the real channel, but the ones believed most important. Other eventual properties overlooked are believed to be identified when the same algorithm are implemented on real data later.

After the first simulations with a single echo, the result from both methods does not differ too much. The MF shows a slightly better result. The big difference is when simulated multipath echoes are added. The LP method can not handle noise with the same frequency as the transmitted signal. The MF does also pick up these weaker echoes, but the highest peak corresponds to the first, strongest echo. Another reason to not use the LP is also realized later. The pulse that would be used when gathering the real data has most of its energy centred at 40 kHz and only a fraction of it spread out on the rest of the ±4 kHz-bandwidth (see Figure 4.2). This makes the LP-slope shorter and more unreliable since it would not be based on so many samples. The MF-method overall seemed more robust and is therefore chosen as the ToF-estimation method for the rest of the project.

4

Real Data Experiments

The step after the simulations is to gather data with the SC, the same way it would if the functionality were implemented, and to test the method on this data. The calculations are executed in Matlab.

4.1 Methods

Based on the outcome of the simulations, the MF method is chosen to estimate the ToF. The MF method is described in Section 2.2.1.

4.2 Set-up

To gather data, a test-application for the SC is created. It is designed to make the device transmit one pulse, then record for 0.24 s. The pulse has a shape close to a sinus function with center frequency 40 kHz. In Figure 4.1 and Figure 4.2 a recording of this pulse, with duration 1 ms, can be seen in the time domain and the frequency domain respectively. Figure 4.3 shows the autocorrelation of the pulse. When using the MF-method, this would be the ideal output of the matched filter, with no attenuation and no noise.

The pulse is emitted from the left of the two transducers on the front of the device. Two logs of data, that contains of recorded information from both frontal transducers, are extracted. The recording starts directly after the pulse duration. Because of the ringing in the transducers described in 2.3.3 the recorded signals need to be trimmed a few samples in the beginning in order to not interfere with the ToF estimation. The ringing of the 3 ms-pulse can be seen in Figure 4.1. The ringing duration varies a bit with the duration of the pulse and the medium. So how much that needs to be trimmed is investigated for each

pulse duration and for each medium the experiments are executed in. The experiments are executed on land, in a lake and in a swimming pool. The set-up of the three experiments that are conducted are illustrated in Figure 4.4. The speed of sound is approximated to 1500 m/s in water and to 340 m/s in air.

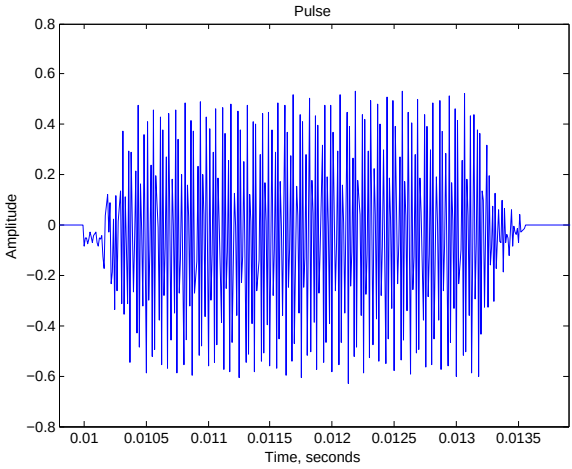

Figure 4.1: *3 ms-pulse in time domain*

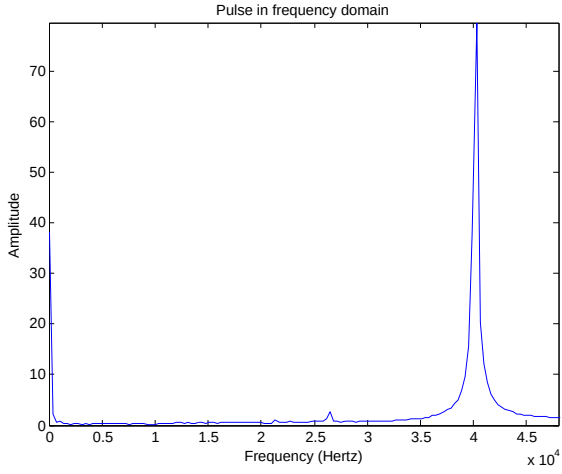

Figure 4.2: *Pulse in frequency domain*

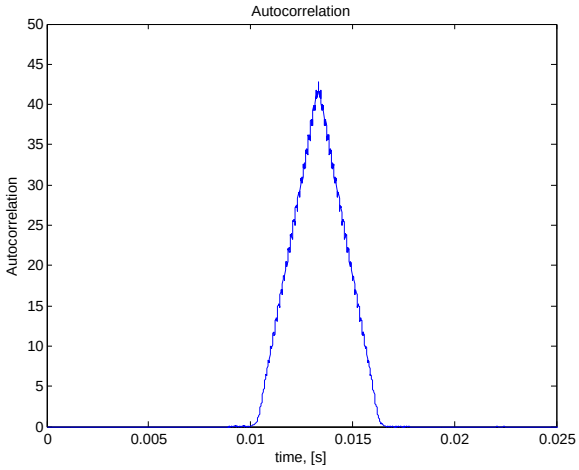

Figure 4.3: *Autocorrelation of pulse*

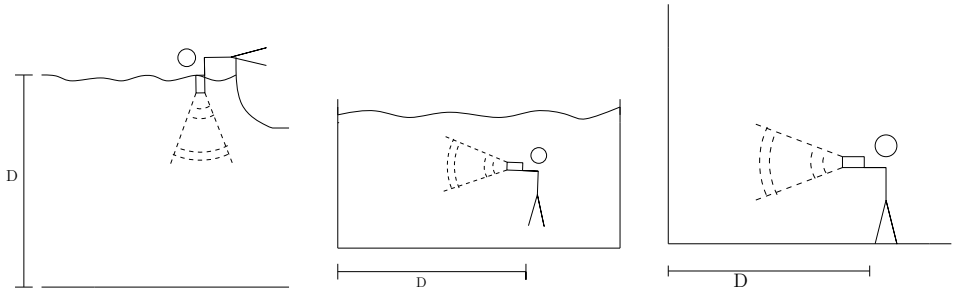

Figure 4.4: Left: *In lake.* ***Middle:*** *In pool.* ***Right:*** *In air.*

First, the pulse had a duration of 1 ms. The data gathered with this pulse duration gives no adequate distance measurement, as can be seen in Section 4.3.1. Therefore, the test-application is changed so that longer pulses can be tested. The difference in results when different durations are used is evaluated to find the best one. Shorter pulses give a better resolution, lower minimum distance for possible measuring and less backscatter, however, it also has a lower signal energy and thereby lower SNR. If the SNR is too low, the MF method would not work.

The minimum distance is dependent on the pulse duration since the whole pulse needs to be transmitted before the recording starts. In 3 ms the signal travels about 4.5 m in water and this means that if the recordings does not start for 3 ms, the shortest distance measurable is 2.25 m in water. The shortest distance measurable with a 3 ms-pulse is a bit longer than that, since also the ringing of the transducers needs to be taken into account. The ringing-problem can be seen in Figure 4.5. From the recordings produced during the experiments,

it is decided that the 200 first samples should cover the ringing. These are trimmed from the recorded signal. 200 samples correspond to 0.0021 s. This means that it takes 0.0051 s before any echo can be detected and the minimum range measurable is 3.82 m.

A few different pulse durations are tried to see what duration gave the best trade-off. The durations tried out are 1, 2 and 3 ms.

The data with pulse duration of 1 ms are gathered from a boat in a lake. The data gathering with pulse durations 2 and 3 ms is done later. Those experiments are executed first on land and then in a swimming pool with the device aimed vertically against the wall.

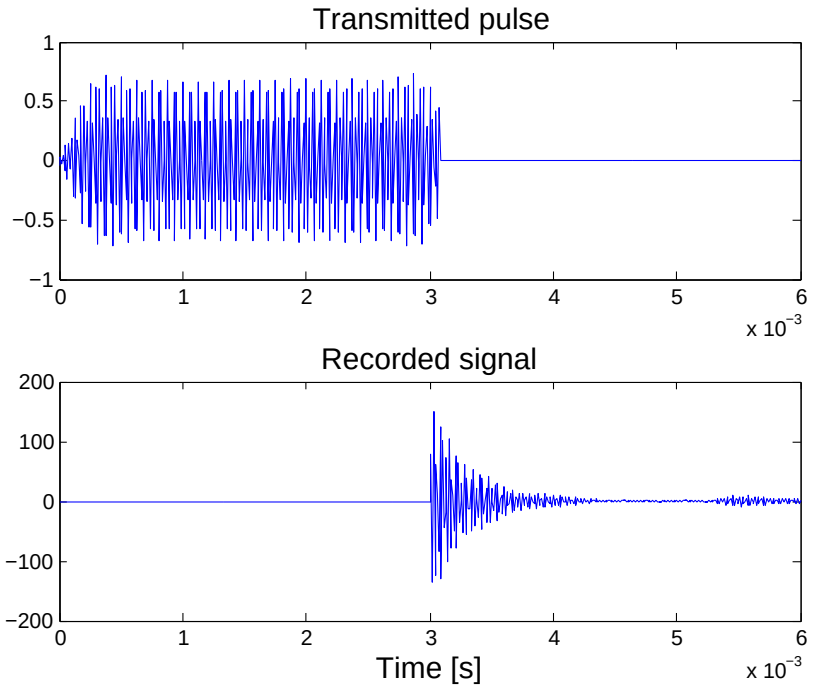

Figure 4.5: *Problem with ringing in the recorded signal visualised.* ***Top:*** *Transmitted pulse.* ***Bottom:*** *Signal recorded in water.*

4.3 Results

In this section, the results are presented. They are divided into a section where the results from the data gathered with pulse duration 1 ms are presented and into a section where the results from the 2 and 3 ms pulse duration results are

presented.

4.3.1 Pulse duration 1 ms

Results from depth measurements made from data gathered with pulse duration 1 ms can be seen in Figure 4.6. The data are gathered in a lake at a depth of 7 m and 20 m. It can be seen that the errors are large. The reason for this is discussed in Section 4.3.4.

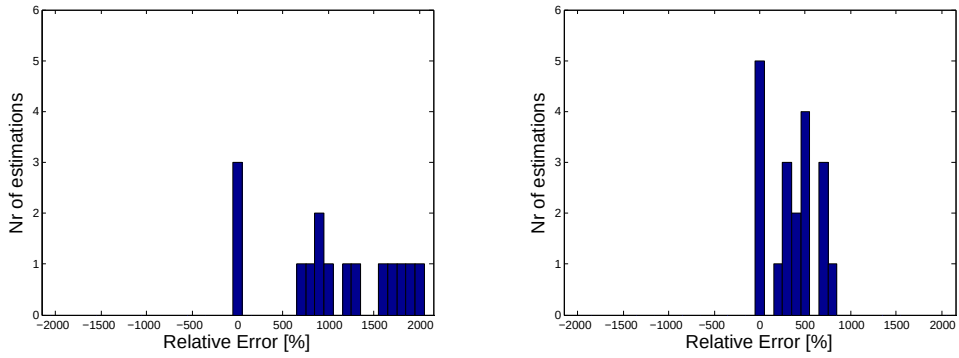

Figure 4.6: *Measurements made with a pulse duration of 1 ms in water.* ***Left:*** *Range 7 m* ***Right:*** *Range 20 m*

4.3.2 Pulse durations 2 and 3 ms

Histograms of the measurements gathered on land, 4 m from a wall, can be seen in Figure 4.7. Since Figure 4.7 has a very wide range, Figure 4.8 shows a more detailed picture of the same data. The data are gathered in water in a swimming pool at a range of 15 m are visualised in Figure 4.9. The results can also be studied in Table 4.1.

Table 4.1: *Results from experiments with pulse durations 2 and 3 ms.*

Pulse duration [ms]	Range [m]	Medium	Relative error < 5% [%]	Standard deviation [m]
2	4	Air	44.44	11.83
3	4	Air	67.39	1.08
2	15	Water	42.86	3.53
3	15	Water	42.86	1.25

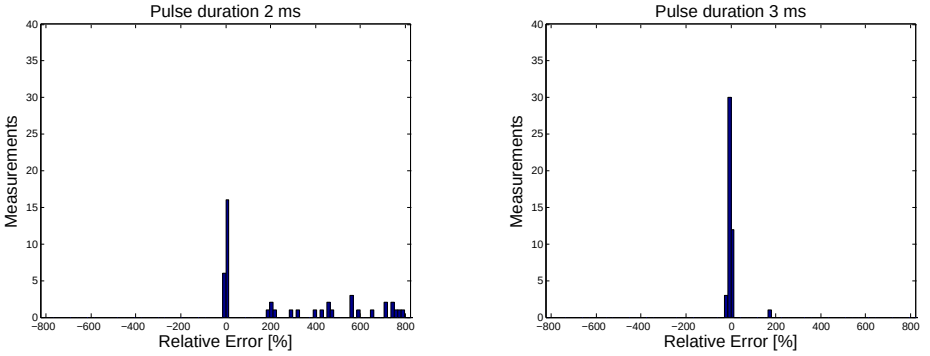

Figure 4.7: *Measurements made on a range of 4 m in air.* ***Left:*** *Pulse duration 2 ms.* ***Right:*** *Pulse duration 3 ms.*

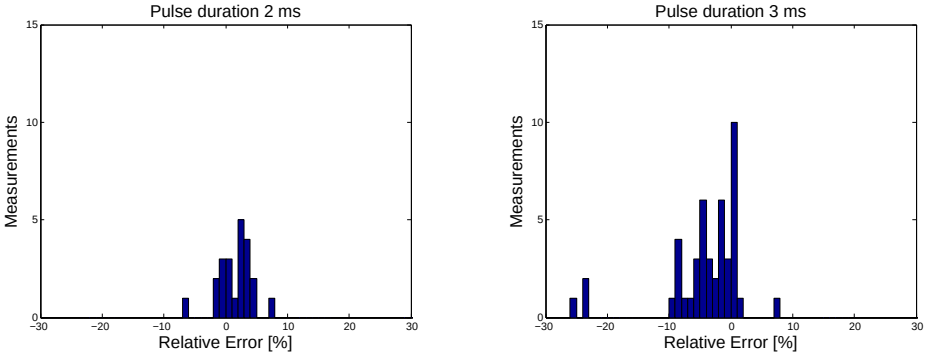

Figure 4.8: *Measurements made on a range of 4 m in air, more detailed.* ***Left:*** *Pulse duration 2 ms.* ***Right:*** *Pulse duration 3 ms.*

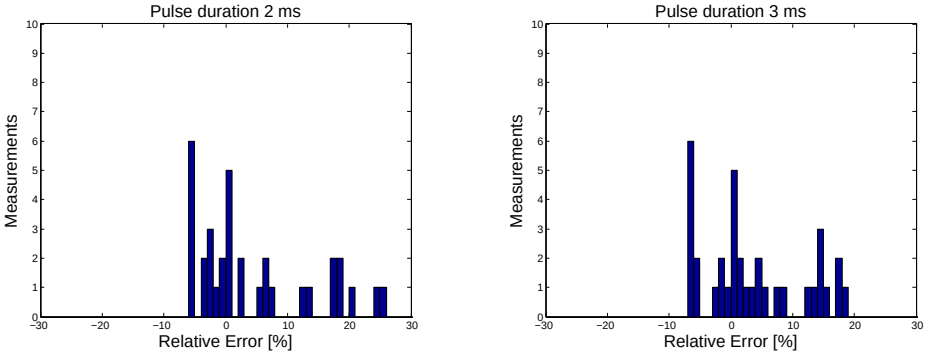

Figure 4.9: *Measurements made on a range of 15 m in water.* ***Left:*** *Pulse duration 2 ms.* ***Right:*** *Pulse duration 3 ms.*

4.3.3 Threshold

An idea to use a threshold approach to sort out outliers is implemented and tried out. The principal idea is to find the maximal value of the output of the MF and place a threshold at 80% of this value. If any other peak than the highest exceed this threshold, the measurement is considered unreliable. In Figure 4.10 a measurement that passes the threshold can be seen to the left and one that does not, to the right.

This approach is tried out on the same data the shown in Figures 4.7 - 4.9. The distance measurements that exceeds the threshold are removed and the result can be seen in Figures 4.11 and 4.13. A more detailed Figure of 4.11 can be found in Figure 4.12. The effect can be most clearly be seen the case when the measurements are gathered on a range of 4 m in air. This is also where the measurements were the most spread before the threshold is implemented. The results can also be studied in Table 4.2.

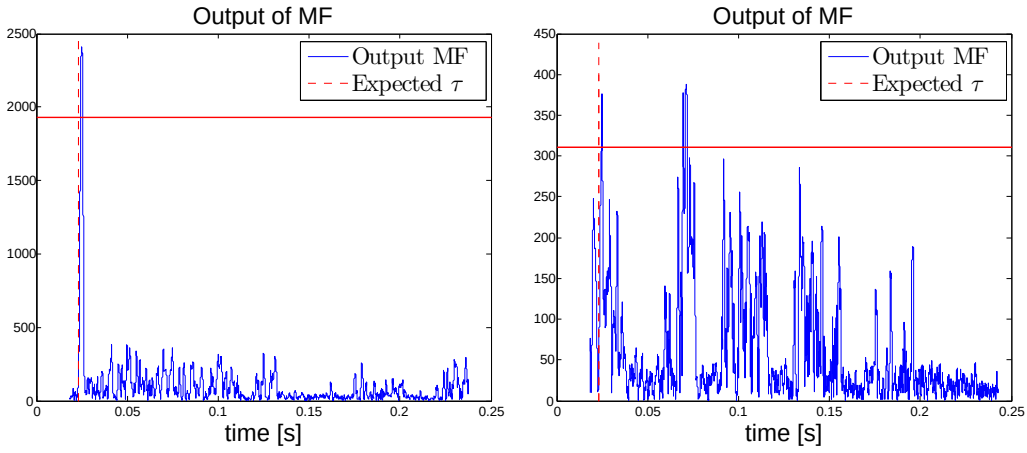

Figure 4.10: *The threshold is set to 80% of the highest peak in the cross-correlation. Cross-correlations containing other peaks that exceeds the threshold are seen as unreliable and the corresponding measurement are sorted out.* ***Left:*** *Example of measurement that would pass the threshold.* ***Right:*** *Example of measurement that would not pass the threshold.*

Table 4.2: *Results with threshold implemented.*

Pulse duration [ms]	Range [m]	Medium	Relative error < 5% [%]	Standard deviation [m]	Passed threshold [%]
2	4	Air	80	8.2298	44.44
3	4	Air	80.95	0.2593	91.30
2	15	Water	44.12	1.4282	97.14
3	15	Water	42.86	1.2472	100

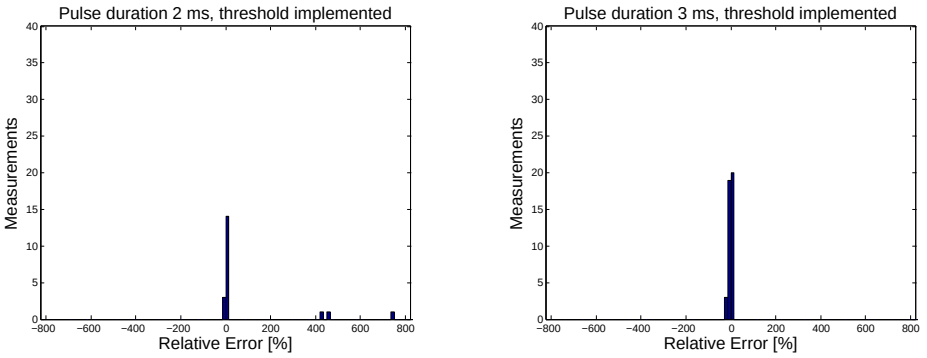

Figure 4.11: *Measurements made on a range of 4 m in air, threshold implemented* ***Left:*** *Pulse duration 2 ms.* ***Right:*** *Pulse duration 3 ms.*

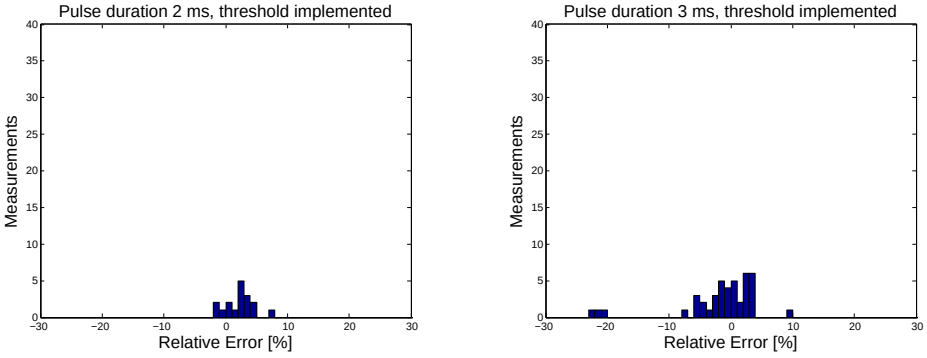

Figure 4.12: *Measurements made on a range of 4 m in air, threshold implemented, more detatiled.* ***Left:*** *Pulse duration 2 ms.* ***Right:*** *Pulse duration 3 ms.*

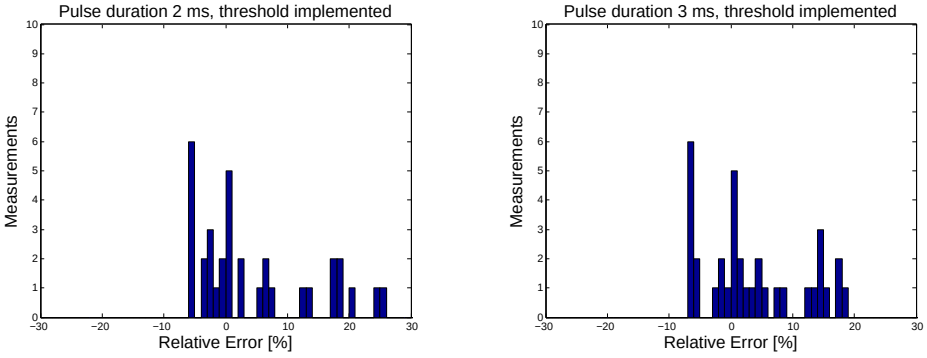

Figure 4.13: *Measurements made on a range of 15 m in water, threshold implemented* ***Left:*** *Pulse duration 2 ms.* ***Right:*** *Pulse duration 3 ms.*

4.3.4 Conclusion

First, it is assumed that a pulse duration of 1 ms was enough to be able to achieve a SNR adequately high. When the calculations are made on the data gathered, it is clear that is not the case. To increase the SNR, longer pulse durations are tried out to see if the measuring of the distance gets better.

The results from the 2 and 3 ms distance measurements are compared to see which of the pulse durations are better to use. Overall, the 3 ms-pulse gives the best results with more measurements with relative error under 5% and smaller standard deviation. To keep in mind is that the tests are conducted in a controlled environment, as a flat wall with no obstacles around. The more energy the pulse has, the more reverberation noise might arise. Other difficulties not foreseen might arise if the data are gathered in a lake or in the ocean as intended for the final application.

4.4 Discussion

Since it is hard to anticipate all the properties of the channel between the transmitted and the received signal it is preferred to gather the data in the same kind of environment the device later will be used in. For the first session of data-gathering, a boat with echo sounder (for measuring the ground truth) is available and the data is gathered in a lake. When calculations on this data are executed, they fail to produce an acceptable estimate. It is then realized that the problem probably is that the pulse duration is too short. 2 and 3 ms-pulses are tried out. At this time, no boat is available and alternative experiment locations are considered. One condition for the location is the possibility to know the true distance to the target. A solution is to make the measurements on land, facing a wall and simply manually measure the distance. Another solution is to gather data it in a pool, where it is easy to measure the distance. The walls of the pool are used as the target to obtain a larger distance than would have been if the bottom had been used. No other opportunity than these two to gather data are presented during the experiments. The pool is considered to be the environment most similar to the real one. One property it lacks is that the transmitted pulse is not reflected truthfully when it is reflected in the hard, smooth pool wall in contrast to a softer lake bottom. The result is that the transmission losses of the signal are less. Also, the pool presents a more confined, smaller area, with hard, smooth walls on all sides. The pool bottom and surface should produce more multipath echoes. This is predicted to produce more and stronger reverberation noise, but is not noticed as a problem whilst calculating the depth measurement.

The data gathered on a shorter distance give worse distance measurements, with many outliers compared to the data gathered on a longer distance. The reason behind this is unknown, but a proposed reason is the difference in medium the gathering is executed in, air and water respectively. That implies that the application works worse in air than in water. To see if that is the case, more data would be needed to be gathered and analysed.

It is decided that for the next step, to implement the functionality in the SC, a pulse duration of 3 ms is used.

5

Implementation

After having analysed the problem by simulations and by doing calculations offline on real data, the next step is to implement the algorithms in the device so that distance measurements could be made in real time. The language used for the implementation of the algorithm is C++.

5.1 Method

For the ToF estimation the MF approach is used. For the theory behind the MF, see Section 2.2.1.

5.2 Design

The application is implemented as a plug-in to the SC. It is connected to the ultrasonic modem by a dBus message system. Figure 5.1 shows how the sensors are connected to the plug-in.

The application tells the ultrasonic modem to transmit a single pulse. The ultrasonic modem then transmits one pulse with duration 3 ms and records with both frontal transducers for 0.1 s and sends the data back to the application. The application waits for 1 s to receive all data. After this, calculations are done on the two data-sets received. The distance measurements are then evaluated by the threshold criterion, see Section 4.3.3. If one of them does not pass, the other one is used. If both fails, no estimation is made from that pulse. If both are good, the estimation from the data from the transducer only recording will be used. This is because that recording usually contains less noise than the other. The cases when no estimation is made will include the cases where the range is too big for echo to return during the recording time, since the recording only will include noise

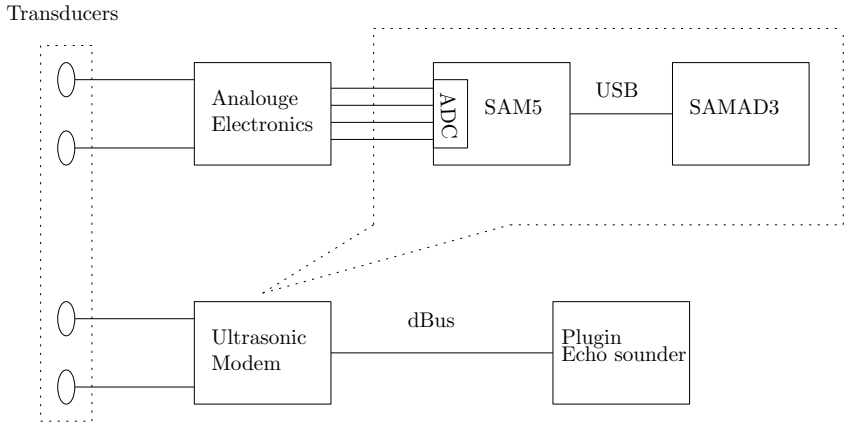

Figure 5.1: *Overview of structure between the transducers and the plug-in.*

and not pass the threshold.

The threshold is implemented in such way, that after the maximum value is found, both the index for this value and the value is stored. The application will then loop through all samples in the array where the cross correlation is stored, besides an interval of 600 samples with center around the maximum value-index. If any of these samples exceeds the threshold, set to 80 % of the maximum value, that measurement will be considered too unreliable.

An aspect that needs to be taken into account is that there is a risk that some data-packages might get lost. In this implementation, those parts of the recorded signal that are not received is put to the DC-offset of the signal. If the lost data contains the returning echo, an error will occur. If it does not, no interference will be noticed in the estimation process. If a recorded signal has lost more than 8 out of 20 packages, that measurement is considered unreliable. There is a risk that even if the signal loses less than 8 packages, the packages lost might carry the data containing the echo. The ToF calculated will then be false. But if the threshold is set so that only signals containing all data would pass, almost all of the signals would be sorted out. The threshold at 8 is chosen so that that risk is small.

Two screen shots of the application can be seen in Figure 5.2. The left figure is showing the application during the data gathering and the right figure the output after a distance measurement has been done. The estimated distance is in the Figure 0.0 m since that is what it shows before the user has made the first measurement attempt. There is also an option to choose either water or air. This option changes the speed of sound that is used for the calculations, so that the application is easy to test both in air and in water.

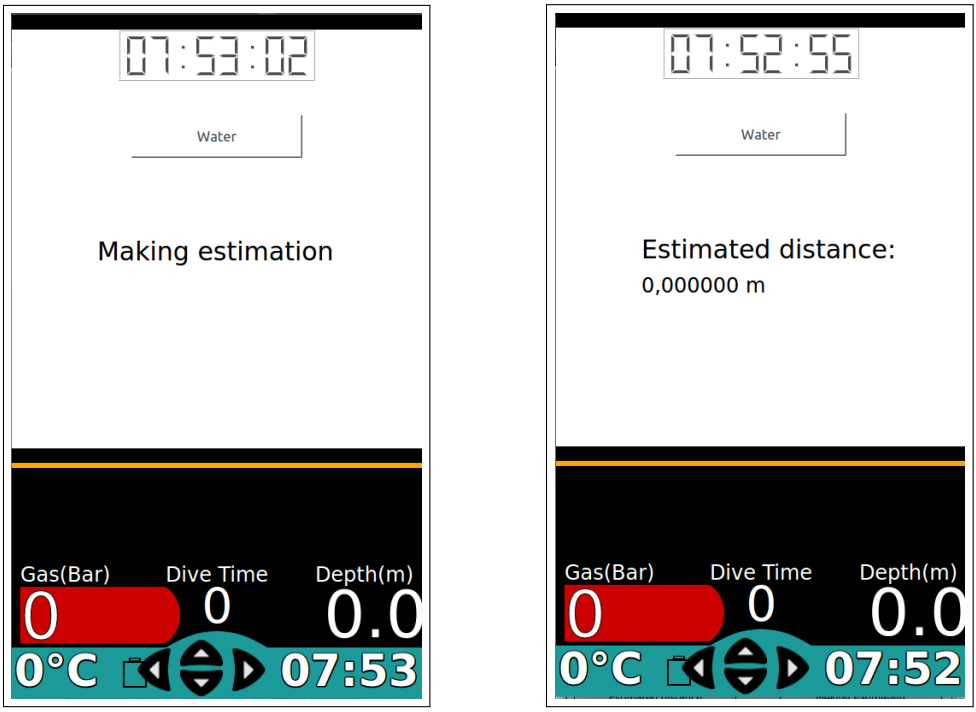

Figure 5.2: *Screenshots from inside the running application.*

5.3 Test set-up

To evaluate the performance of the application, it is tested on land. The device is placed on two different distances from a wall, see Figure 4.4. The distances tried out are 4 m and 8 m.

5.4 Results

The histograms shown in Figures 5.4 and 5.3 displays, together with Table 5.1, the result from the testing of the implementation.

Table 5.1: *Result, implementation.*

Range [m]	Medium	Est. with rel. error < 5% [%]	Percentage failed measurements [%]	Standard deviation [m]
4	Air	75.68	2.70	3.32
8	Air	53.70	33.33	2.05

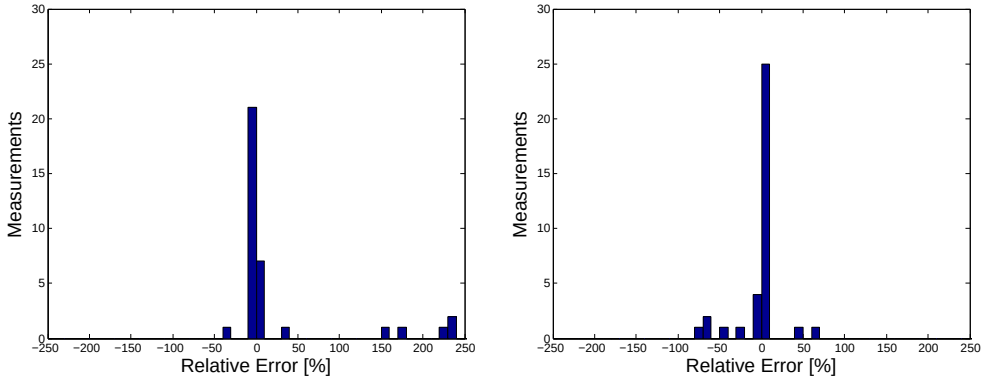

Figure 5.3: *Result from implementation.* ***Left:*** *Measurements gathered on a distance of 4 m in air.* ***Right:*** *Measurements gathered on a distance of 8 m in air.*

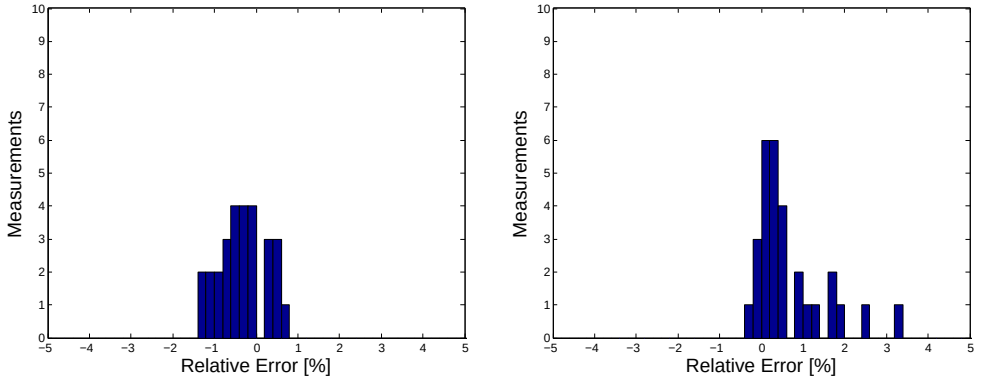

Figure 5.4: *Result from implementation, zoomed in.* ***Left:*** *Measurements gathered on a distance of 4 m in air.* ***Right:*** *Measurements gathered on a distance of 8 m in air.*

5.5 Discussion

Preferably, the tests would have been executed in water, as it is where the application is meant to be used. At this time, there were no possibilities to do so and the tests are executed on land instead. One drawback of this is that the ambient noise on land and in the ocean differs and affects the estimation in different ways. Another one is that the application could not be tested on greater distances in a appropriate way. This because of that large, open areas in front of a suitable, flat wall could not be found. The device could not be higher above the ground than the height of the person gathering the data, so the ground was

always rather close, causing multipath echoes. The further away from the wall the device was placed, the closer in time the multipath echoes would reach the receiver. The close echoes make it hard for the algorithm to handle and might cause error. Therefore, the target distances are chosen to 4 and 8 m. Still, the measurements taken at 8 m had some problems with multipath echoes, as seen in the cross correlation shown in Figure 5.5. Also, a problem with gathering the measurements on land is that the ringing duration seems to be longer in air. Sometimes, traces of it remain after the trimming that is supposed to remove it. These times the algorithm recognises the ringing as the strongest echo and estimates a very short distance. This can be seen in 5.6. Figure 5.7 shows a measurement without problem for comparison. Based on the data gathered in a pool and used in Chapter 4, it is believed that the problems described are not interfering as much as in air while using the application in water.

It can be seen in Figure 5.3 that there exist a few measurements with rather large relative error. These are believed to arise as a consequence of multipath signals. Table 5.1 shows that the experiments made at a distance of 8 m have a larger percentage of failed measurements then the ones made at 4 m. The failed measurements include the ones that did not pass the threshold or lost too many data packages. The reason that that number is higher at range 8 m, compared to 4 m, is that the failed measurements are believed to contain those with interfering multipath echoes that the threshold managed to sort out. And since the 8 m-case is more exposed to multipath echoes, more distance measurements fail.

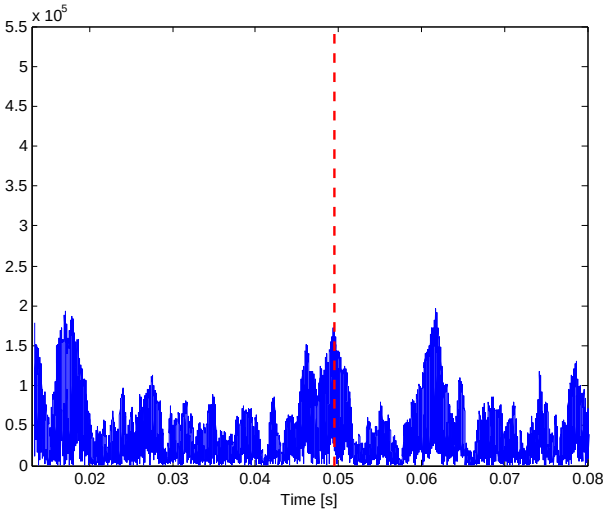

Figure 5.5: *A measurement with multipath echo problem. Red, dashed line shows true ToF.*

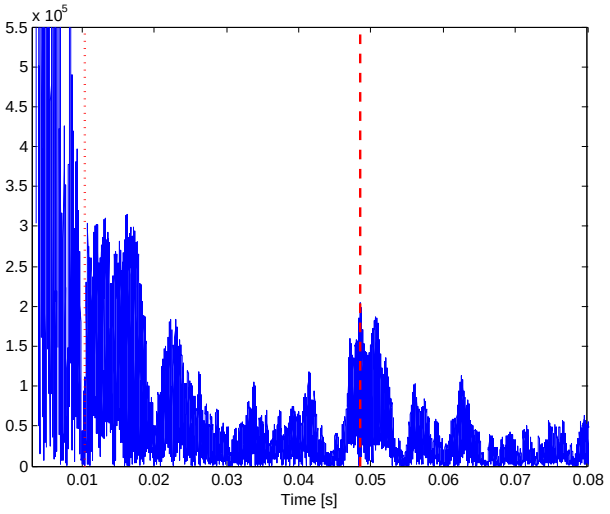

Figure 5.6: *A measurement with ringing problem. Red, dotted line at* $t = 0.0104$ *s shows to where the cross correlation was trimmed before calculating the maximum point. The red, dashed line shows true ToF.*

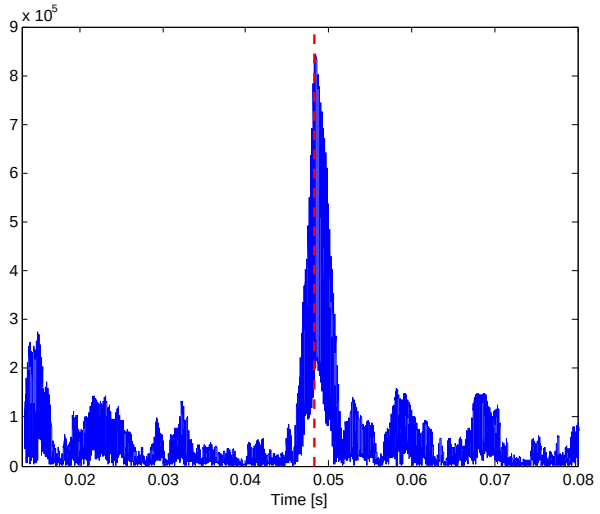

Figure 5.7: *A measurement with little noise and a strong echo at the ToF, producing a good estimation. Red, dashed line shows true ToF.*

6

Concluding remarks

6.1 Remarks on Result

The approach used in the final implementation is the MF-method. To remove outliers a threshold-method is proposed. If there exist other peaks higher than 80% of the highest peak in the cross-correlation, the measurement is considered unreliable.

The MF-method together with the threshold-method is implemented in the SC. The results show that the implementation is capable of making adequate distance measurements but still leaves room for improvement.

The threshold-method sorts out some of the false distance measurements, but not all. It results in outliers rather far away from the true ToF. A way to solve this is to compare the measurement to other measurements taken at the same distance and sort out the outliers. A suggestion on how to implement this is described in Section 6.2.

Sometimes the SC produces false distance measurements with a large relative error. A suggested reason for this is that the transmission beam is rather wide compared to a conventional echo sounder. This is because the transducers on the SC are not chosen with this application in mind, but for communication. This makes the footprint of the beam very big on larger distances. A large footprint results in a poor resolution since the closest distance within the area covered is seen as the depth as explained in Section 2.3.2.

6.2 Suggestions on Future Work

Since the application will be used when diving, test of the application in water should be done. This to be able to draw more truthful conclusions on the

performance and further improvements. It is suggested that this is done before the other ideas on improvement below are evaluated.

The application was first planned to transmit several pulses, sort out the unreliable ones, and take an average of the remaining measurements. This is not implemented due to lack of time as the priority was to get the algorithm to work properly. A proposed implementation approach of this idea is to send out 10 pulses and to record 0.1 s after each. If the numbers of reliable estimates exceed a chosen threshold, use them to calculate an average depth estimate. Outliers that slip past the threshold can also be sorted out when compared to the other measurements. The number of pulses transmitted can be increased if the measurements still seem unreliable. But the more pulses, the longer the data gathering takes. Since the user has to aim the SC towards the target during the transmission and recording, this affects the usability of the application.

Another improvement that can be done is to improve the approximation of the speed of sound used in the calculations. Right now a fixed, averaged speed of sound set to 1500 m/s, is used. As seen in Section 2.6 the speed varies. The closer the speed used is to the true speed of sound, the more accurate distance measurement could be made. As example, Table 2.3 gives that in cold water with low salinity the speed will lie around 1422 m/s. If 1422 m/s is the true speed and the averaged speed used in the post-processing is 1500 m/s, an estimation error of about 2.75 m is made at range of 50 m. To decrease the error made by using a fix speed, an approximation may be made by using the model from Section 2.6. The depth and the temperature is known by the SC, but the salinity of the water has to be averaged. The salinity could also be stepwise constant if the water type is known. The constant values can be for example set to $S = 0.5\ ppt$ for fresh water, $S = 15\ ppt$ for brackish water and $S = 35\ ppt$ for salt water, numbers based on theory from Section 2.7. The water type could for example be provided by the user.

Many studies have shown that the ToF estimation can be improved by pre-filtering the signals as is done in the GCC approach, see Section 2.2.2. Experiments with different weighting function should be executed to see which one of them suits the general environment for the SC the best. As a suggestion, the ML or the PHAT weighting function that is described in Section 2.2.2 could be tried out to see if the result improves. The conventional PHAT is briefly tried out during the simulations, but gives only improved performance compared to the MF in the case of extremely high SNR. This might imply that an improvement of the PHAT-approach is needed. An enhanced approach is suggested by Donohue et al. [2007] and Yanjie et al. [2014]. This method, PHAT-β is described in Section 2.2.2. Varma [2002] proposes another method to improve the ToF estimation made through PHAT. In the report the method is used to perform DOA-estimations, but by using the same principal, it should be applicable to TOA-estimations too. The approach of the method is to not discard the other weaker peaks of the cross-correlation, compiling all possibly sets of time delays and using three different criteria to single out the most probable of all sets. Yanjie et al. [2014] reports also the SCOT weighting function giving promising results.

The maximum range the application can handle is not yet investigated. The transducers record for a time of 0.1 s after the pulse has been transmitted. This implies that the possibility of measuring depths up to 75 m exists. The remaining question is if the echo will have enough energy left after the transmission loss to be detected with current ToF-estimation method. If longer range is desired and the energy remaining in the echo is enough, then the only modification that needs to be done for longer range is to increase the recording time. If the energy is too low, either the SNR needs to be increased or the algorithm needs to be changed or improved. If the SNR is enough, a possibility of using a shorter pulse could be investigated. This would result in shorter transmit duration and lower minimum range. The current minimum range in water is 3.82 m, depending on the pulse duration and the ringing time of the transducers.

When integrating the application with the existing communication, that uses the same transducers and signal frequency, the duration of the transmission might also be of interest. Currently, the application uses 0.13 s to transmit and record one pulse. An idea to integrate the application with existing communication is to use the time hatch every unit has to communicate, to transmit pulses. This would make sure no other unit interferes with the depth measuring. The problem is that the time frame of the hatch only allows one single pulse to be transmitted and recorded. This collides with the plan to make the measurements more robust by transmitting more pulses during the same depth measuring.

A

Derivation of matched filter

Let us say that signal $x(t)$ is sent out. The received signal will then be

$$s(t) = \alpha x(t - \tau_d) + v(t) \tag{A.1}$$

where $v(t)$ is white Gaussian noise with variance $N_0/2$ and $\alpha x(t - \tau_d)$ is the time-shifted, possibly scaled, echo of the emitted signal. Define $y(t)$ as

$$y(t) = \alpha x(t - \tau_d)\,. \tag{A.2}$$

Let $s(t)$ pass through a filter, $h(t)$. Then the output is

$$z(t) = y_0(t) + v_0(t)\,. \tag{A.3}$$

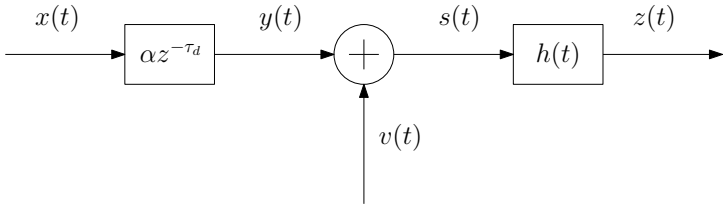

Figure A.1: *The implementation of the MF, $h(t)$, where $x(t)$ is the transmitted signal and $v(t)$ is noise*

A.1 Maximize the SNR

What you would like is to maximize the SNR of the output of the matched filter at time τ_d. A derivation of the matched filter based on this has been done by

Levanon and Mozeson [2004] and Turin [1960]. To maximize the SNR is to maximize

$$SNR = \frac{|y_0(\tau_d)|^2}{\overline{v_0(t)^2}} . \tag{A.4}$$

If $Y(f)$ is the Fourier transform of $y(t)$ and $H(f)$ is the Fourier transform of $h(t)$, then

$$y_0(\tau_d) = \int_{-\infty}^{\infty} H(f)Y(f)e^{j2\pi f\tau_d} df \tag{A.5}$$

and the mean-squared value of the noise is

$$\overline{v_0(t)^2} = \frac{N_0}{2} \int_{-\infty}^{\infty} |H(f)|^2 df . \tag{A.6}$$

Substituting equations (A.5) and (A.6) into (A.4) yields

$$SNR = \frac{2\left| \int_{-\infty}^{\infty} H(f)Y(f)e^{j2\pi f\tau_d} df \right|^2}{N_0 \int_{-\infty}^{\infty} |H(f)|^2 df} . \tag{A.7}$$

To continue, the Schwartz inequality is used,

$$\left| \int_{-\infty}^{\infty} A(x)B(x)dx \right|^2 \leq \int_{-\infty}^{\infty} |A(x)|^2 dx \int_{-\infty}^{\infty} |B(x)|^2 dx \tag{A.8}$$

that holds if $A(x) = KB^*(x)$, where K is a arbitrary constant and $B^*(x)$ is the complex conjugate of $B(x)$. Choosing

$$A(x) = H(f) \tag{A.9}$$

and

$$B(x) = Y(f)e^{j2\pi f\tau_d} \tag{A.10}$$

yields

$$SNR \leq \frac{2}{N_0} \int_{-\infty}^{\infty} |Y(f)e^{j2\pi f\tau_d}|^2 df = \frac{2}{N_0} \int_{-\infty}^{\infty} |Y(f)|^2 df = \frac{2E}{N_0} \tag{A.11}$$

where E is the energy of the input signal, defined as

$$E = \int_{-\infty}^{\infty} y^2(t)dt = \int_{-\infty}^{\infty} Y^2(f)df . \tag{A.12}$$

When using Schwartz inequality (A.8), it is required that

$$A(x) = KB^*(x) \tag{A.13}$$

holds. If $A(x)$ and $B(x)$ is set as in Equation (A.9) and (A.10), it follows that

$$H(f) = KY^*(f)e^{-j2\pi f\tau_d} . \tag{A.14}$$

The result is the frequency response of the matched filter. To get the impulse response, the inverse Fourier transform is applied,

$$h(t) = Ky^*(\tau_d - t) . \tag{A.15}$$

If $h(t)$ is expressed in $x(t)$ (according to (A.2)) and $K = \frac{1}{a}$ is choosen, then

$$h(t) = x^*(-t) . \tag{A.16}$$

This is the matched filter. To be realizable, $h(t)$ needs to be equal to 0 for $t \leq 0$ The output $z(t)$ is

$$z(t) = s(t) \otimes x^*(-t) = \int_{-\infty}^{\infty} s(m)x^*(-(t-m))dm . \tag{A.17}$$

Since in reality, the emitted signal is real, the complex conjugate could be omitted. Also, the filter needs to be zero for all times smaller than 0. It needs be right shifted until it fulfils this. To remember, when this is done, is that the peak in $z(t)$ will be shifted equally many samples to the right.

$$z(t) = s(t) \otimes x(-t-T) = \int_{-\infty}^{\infty} s(m)x(-(t-m)-T)dm \tag{A.18}$$

where T is the right shift of $x(-t)$. So, in order to do a matched filtering, the convolution between the received signal and the time-reversed emitted signal is calculated. The result will be a peak at $t = \tau_d$ and the time-of-flight estimate can be obtained.

A.2 Least-square method

Silvia and Elliott [1987] derives the matched filter using a least square approach.

Let us say that the signal $x(t)$ is transmitted. The received signal is then

$$s(t) = \alpha x(t - \tau_d) + v(t) \tag{A.19}$$

where $v(t)$ is white Gaussian noise with variance $N_0/2$ and $\alpha x(t - \tau_d)$ is the time-shifted, possibly scaled, echo of the emitted signal. $y(t)$ is defined as

$$y(t) = \alpha x(t - \tau_d) . \tag{A.20}$$

The quantity

$$J(\tilde{\tau}_d) = \int_{t_p}^{t_p+t_0} (s(t) - \alpha x(t-\tilde{\tau}_d))^2 dt \tag{A.21}$$

is formed, where t_0 is the observation time and t_p is the transmitted pulse length. Let $\tilde{\tau}_d$ vary until $J(\tilde{\tau}_d)$ is a minimum. The $\tilde{\tau}_d$ that produces the minimum is called $\hat{\tau}_d$ and is an estimate of the ToF.

To find this estimate, differentiate (A.21) with respect to $\tilde{\tau}_d$

$$\frac{dJ}{d\tilde{\tau}_d} = \int_{t_p}^{t_p+t_0} 2((s(t) - \alpha x(t-\tilde{\tau}_d))(\alpha \frac{dx(t-\tilde{\tau}_d)}{d\tilde{\tau}_d}) dt \,. \tag{A.22}$$

Then set this equal to zero

$$\begin{aligned}
&\frac{dJ}{d\tilde{\tau}_d} = 0 \Rightarrow \int_{t_p}^{t_p+t_0} \alpha s(t) \frac{dx(t-\tilde{\tau}_d)}{d\tilde{\tau}_d} - \alpha^2 x(t-\tilde{\tau}_d) \frac{dx(t-\tilde{\tau}_d)}{d\tilde{\tau}_d} dt = 0 \Rightarrow \\
&\int_{t_p}^{t_p+t_0} s(t) \frac{dx(t-\tilde{\tau}_d)}{d\tilde{\tau}_d} dt = \int_{t_p}^{t_p+t_0} \alpha x(t-\tilde{\tau}_d) \frac{dx(t-\tilde{\tau}_d)}{d\tilde{\tau}_d} dt \Rightarrow \\
&\int_{t_p}^{t_p+t_0} s(t) \frac{dx(t-\tilde{\tau}_d)}{d\tilde{\tau}_d} dt = \alpha \int_{t_p}^{t_p+t_0} \frac{d}{d\tilde{\tau}_d} \left[\frac{1}{2} x^2(t-\tilde{\tau}_d)\right] dt \,.
\end{aligned} \tag{A.23}$$

If it is assumed that $t_p \leq \tau_d \leq t_0$, then it can be said that the right hand side of A.23 is approximately equal to zero,

$$\int_{t_p}^{t_p+t_0} \frac{d}{d\tilde{\tau}_d} \left[\frac{1}{2} x^2(t-\tilde{\tau}_d)\right] dt \approx \frac{d}{d\tilde{\tau}_d} \int_{t_p}^{t_p+t_0} \left[\frac{1}{2} x^2(t-\tilde{\tau}_d)\right] dt \approx \frac{1}{2} \frac{d}{d\tilde{\tau}_d}(E_x) = 0 \,. \tag{A.24}$$

E_x is the energy contained in the transmitted signal and is constant for all τ_d. Now it is possible to rewrite (A.23) to

$$\int_{t_p}^{t_p+t_0} s(t) \frac{dx(t-\tilde{\tau}_d)}{d\tilde{\tau}_d} dt \approx 0 \Rightarrow \frac{d}{d\tilde{\tau}_d}[z(\tilde{\tau}_d)] \approx 0 \tag{A.25}$$

where

$$z(\tilde{\tau}_d) = \int_{t_p}^{t_p+t_0} s(t) x(t-\tilde{\tau}_d) dt \,. \tag{A.26}$$

So, to the solution to (A.25), where $\tilde{\tau}_d = \hat{\tau}_d$ is the maximum value of (A.26). This is the same result as was given by the derivation where the SNR was maximized, see section A.1.

B

Generation of OFDM pulse

This chapter includes the code that is used to generate the OFDM pulse used in the simulation experiments. The code origins from [Gustafsson et al., 2010].

```
N=1000;                      % Number of samples in pulse
fs=96000;                    % Sampling frequency
f=(0:N-1)'/N*fs;             % Frequency interval
t=(0:N-1)/1000;              % Time interval
f1=38016; f2=41952;          % OFDM frequency interval
n=(f1/fs*N:f2/fs*N-1/fs)';   % Freq bins in OFDM interval
c=sign(randn(size(n)));      % Generate random code
X=zeros(N,1);                %
X(n)=c; X(N-n+2)=c;          % OFDM principle
x=ifft(X);

% --- Plot signals -----
subplot(2,1,1)
plot(t,x)
title('Signal')
subplot(2,1,2)
stem(f,X)
axis([37000 43000 -1 1])
```

Bibliography

Airmar Technology Corporation. *Overview for Applying Ultrasonic Technology*. URL `http://airmartechnology.com/uploads/AirPDF/Intro_Overview.pdf`. (accessed October 03, 2014). Cited on page 20.

B. Barshan. Fast processing techniques for accurate ultrasonic range measurements. *Journal of the Acoustical Society of America*, 11(1):45–50, 2000. Cited on pages 10 and 18.

Joe Bremen. *Ocean Globe*. Environmental Systems Research Institute Inc., first edition, 2010. Cited on page 7.

Jerrold T. Bushberg, Seibert, J. Anthony, Leidholdt Jr Edwin M., and John M. Boone. *The Essential Physics of Medical Imaging*. Lippincott Williams & Wilkins, third edition, 2011. Cited on pages 4, 18, and 20.

Robert D. Christ and Robert L. Wernli. *The ROV Manual: A User Guide for Remotely Operated Vehicles*. Elsevier Ltd., second edition, 2014. Cited on page 19.

Committee on Low-Frequency Sound and Marine Mammals. *Low-Frequency Sound and Marine Mammals: Current Knowledge and Research Needs*. Washington, D.C. : National Academy Press, first edition, 1994. Cited on page 25.

Erik Dahl. Ecological salinity boundaries in poikilohaline waters. *Oikos*, 7(1): 1–21, 1956. Cited on page 23.

Stacy L. DeRuiter, Brandon L. Southall, John Calambokidis, Walter M. X. Zimmer, Dinara Sadykova, Erin A. Falcone, Ari S. Friedlaender, John E. Joseph, David Moretti, Gregory S. Schorr, Len Thomas, and Peter L. Tyack. First direct measurements of behavioural responses by cuvier's beaked whales to mid-frequency active sonar. *Biology Letters*, 9(4), 2013. Cited on page 25.

Kevin D. Donohue, Alvin Agrinsoni, and Jens Hannemann. Audio signal delay estimation using partial whitening. In *IEEE SoutheastCon 2007*, 2007. Cited on pages 13 and 52.

A. Frantzis. Does acoustic testing strand whales? *Nature*, 392(6671):29–29, 1998. Cited on page 25.

Fredrik Gustafsson, Lennart Ljung, and Mille Millnert. *Signal Processing*. Studentlitteratur AB, 1:3 edition, 2010. Cited on pages 5, 14, 27, and 61.

Richard P. Hodges. *Underwater acoustics: analysis, design and performance of sonar*. John Wiley & Sons, first edition, 2010. Cited on page 22.

Sabine Höhler. Depth records and ocean volumes: Ocean profiling by sounding technology, 1850-1930. *History & Technology*, 18(2):119, 2002. Cited on pages 4 and 7.

International Hydrographic Organization. *Manual on hydrography*. International Hydrographic Bureau, first edition, 2005. Cited on page 19.

J.C. Jackson, R. Summan, G.I. Dobie, S.M. Whiteley, S.G. Pierce, and G. Hayward. Time-of-flight measurement techniques for airborne ultrasonic ranging. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 6(2):343–55, 2013. Cited on page 17.

G. Jacovitti and R. Cusani. An efficient technique for high correlation estimation. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 35(5):654–660, 1987. Cited on page 5.

C.D. de Jong, G. Lachapelle, S. Skone, and I.A. Elema. *Hydrography*. VSSD, second edition, 2003. Cited on page 19.

C. Knapp and G. Clifford Carter. The generalized correlation method for estimation of time delay. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 24(4):320–327, 1976. Cited on pages 5, 12, and 13.

Erik G. Larsson. *Signals, Informations and Communication*. LiU-Press, first edition, 2014. Cited on page 20.

Nadav Levanon and Eli Mozeson. *Radar signals*. John Wiley and Sons, Inc., first edition, 2004. Cited on pages 9 and 56.

Xavier Lurton. *An Introduction to Underwater Acoustics- Principles and Applications*. Springer, second edition, 2010. Cited on page 22.

Jean-Paul Marage and Yvon Mori. *Sonar and Underwater Acoustics*. John Wiley and Sons, Inc., first edition, 2010. Cited on pages 22 and 24.

W.G. McMullan, B.A. Deianghe, and J.S. Bird. A simple rising-edge detector for time-of-arrival estimation. *IEEE Transactions on Instrumentation and Measurement*, 45(4):823–827, 1996. Cited on page 18.

Herman Medwin. Speed of sound in water: A simple equation for realistic parameters. *Journal of the Acoustical Society of America*, 58(6):1318–1319, 1975. Cited on page 23.

National Academies Press. *Ocean Noise and Marine Mammals*. National Academies Press, first edition, 2003. Cited on pages 24 and 25.

Dwight O. North. An analysis of the factors which determine signal/noise discrimination in pulsed-carrier systems. *Proceedings of the IEEE*, 51(7):1016–1027, 1963. Cited on page 5.

A. Piersol. Time delay estimation using phase data. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 29(3):471, 1981. Cited on page 5.

M.T. Silvia and Douglas F. Elliott. *Handbook of Digital Signal Processing*. Academic Press, Inc., first edition, 1987. Cited on page 57.

Milica Stojanovic, Daniel E. Lucani, and Muriel Medard. On the relationship between transmission power and capacity of an underwater acoustic communication channel. 2008. Cited on page 21.

Laurie Tetley and David Calcutt. *Electronic Navigation Systems*. Butterworth-Heinemann, third edition, 2001. Cited on pages 19, 22, and 24.

George L. Turin. An introduction to matched filters. *IRE Transactions on Information Theory*, 6(3):311–329, 1960. Cited on pages 9 and 56.

Krishnaraj Varma. Time-delay-estimate based direction-of-arrival estimation for speech in reverberant environments. Technical report, 2002. Cited on page 52.

Guochang Xu. *Sciences of Geodesy - I*. Springer, first edition, 2010. Cited on page 19.

Li Yanjie, Wang Ming, Guo Xian, Li Xiaomao, Hu Xiangchao, Qian Hai, and Zhu Baoliang. Time delay signal processing research in sound source localization technology. *Applied Mechanics & Materials*, 556-562:5159–5162, 2014. Cited on pages 5 and 52.