Department of Electrical and Computer Engineering

FACULTY OF ENGINEERING OF PORTO UNIVERSITY

Interferometric Synthetic Aperture Sonar System supported by Satellite

Thesis Submitted in Candidature for the Degree of Doctor of Philosophy

July 2009

Sérgio Rui dos Santos Barbosa Oliveira da Silva

Thesis Identification

Title:	Interferometric Synthetic Aperture Sonar System supported by Satellite
Keywords:	Sonar, synthetic aperture, interferometry, height mapping, surface autonomous vehicle.
Start:	December 2004
Duration:	4 years
Candidate information:	
Name:	Sérgio Rui dos Santos Barbosa Oliveira da Silva
e-Mail:	srui@fe.up.pt
Curriculum Vitæ	http://www.fe.up.pt/~srui/cv.pdf
Advisor information:	
Name:	Sérgio Reis Cunha
e-Mail:	sergio@fe.up.pt
Curriculum Vitæ	http://www.fe.up.pt/~sergio/cv.pdf
Educational Establishment:	Faculdade de Engenharia da Universidade do Porto
Research Institution:	Faculdade de Engenharia da Universidade do Porto

Statement of Originality

The work presented in this thesis was carried out by the candidate. It has not been presented previously for any degree, nor is it at present under consideration by any other degree awarding body.

Candidate:

_____

(Sérgio Rui dos Santos Barbosa Oliveira da Silva)

Advisor:

_____

(Sérgio Reis Cunha)

Statement of Availability

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organizations.

Candidate:

_____

(Sérgio Rui dos Santos Barbosa Oliveira da Silva)

Acknowledgments

In the first place I would like to give thanks to my supervisor, Prof. Sérgio Reis Cunha, for having always supported me since my early undergraduate years to the completion of this thesis. Without his quick-witted spirit and original thinking, this work would not have been possible.

Without the help of Prof. Anibal Matos and Nuno Cruz, the sonar would not have such a good moving platform nor would I have been able to test the sonar in a real world scenario.

I would also like to thank Prof. Henrique Salgado for providing me a fully equipped electronics laboratory where I was able to conduct my work and Prof. José Carlos Alves for giving me the opportunity to learn and work with FPGA systems.

I would like to acknowledge FCT (Foundation for Science and Technology) for giving me the scholarship and supporting my trips to conferences and also CIMAR (Centre of Marine and Environmental Research) for providing funds that where used to purchase essential equipment for my work.

"If you cause your ship to stop, and place the head of a long tube in the water and place the outer extremity to your ear you will hear ships at a great distance from you."

Leonardo da Vinci, 1490

Summary

The work presented here describes a high-resolution interferometric synthetic aperture sonar system to map submerged floors from small ships and surface vehicles. This system was developed with operation in shallow water environments like rivers, lakes and estuaries in mind.

Synthetic aperture is a technique that enables range independent high-resolution sonar imaging by coherently combining echoes gathered using the displacement of the sonar platform. If two geometrically displaced images are obtained, the construction of three-dimensional topographic maps is possible through the use of interferometric techniques.

While high along-track resolution can be obtained by exploiting synthetic aperture techniques, finer resolution in the cross-track direction can only be achieved using large bandwidths, thus forcing the use of high frequency signals. This, in the other hand, imposes strict requirements on the allowed phase errors due to the motion of the sonar platform and the instability of the medium instability when using synthetic aperture techniques.

The approach followed in this work was to evaluate synthetic aperture imaging performance when directly using only one receiver and directly the position estimation given by the vehicle navigation system. For this, the use of a surface craft presents advantages that help overcome the problems related to tight sonar platform motion control and necessary navigation precision for the formation of the motion-corrected high-frequency synthetic aperture sonar images. Also, a small autonomous surface vehicle is advantageous in terms of cost of operation, maintenance and ease of deployment.

Mapping of shallow water areas is an important task for many commercial and scientific applications like river navigability, infrastructure maintenance or natural resource monitoring. These tasks can be done efficiently with a system such as the one presented in this work.

Sample data obtained during test trials will be used to demonstrate how synthetic aperture images can be obtained with this system and subsequently used to generate bathymetric data.

Sumário

O trabalho aqui apresentado, descreve um sistema de sonar de abertura sintética interferométrico de alta resolução para uso em pequenas embarcações ou veículos de superfície. Este sistema foi desenvolvido tendo como objectivo o seu uso em áreas submersas com águas pouco profundas tais como rios, lagos ou estuários.

A abertura sintética é uma técnica que permite obter imagens de sonar de alta resolução independentemente da distância aos alvos, por combinação coerentemente dos ecos que são obtidos sucessivamente usando o deslocamento da plataforma do sonar. Se forem obtidas duas imagens completas da mesma área a partir de posições diferentes é possível também, através do uso de técnicas de interferometria, fazer uma reconstrução tridimensional completa da topografia desta mesma área.

Enquanto que uma melhoria na resolução na direcção do deslocamento da plataforma pode ser obtida explorando técnicas de abertura sintética, uma resolução mais fina na distância ao alvo só pode ser conseguida transmitindo sinais de maior largura de banda, forçando o uso de frequências elevadas. Isto, por outro lado, impõe restrições duras para os erros de fase devidos à instabilidade do movimento da plataforma e do meio de propagação durante a aquisição dos ecos, podendo impedir a correcta combinação coerente dos ecos aquando do processamento da abertura sintética.

A abordagem seguida neste trabalho foi de explorar técnicas de abertura sintética usando apenas um receptor e as estimativas de posições obtidas directamente a partir do sistema de navegação do veículo. Para isto, o uso de um veículo de superfície apresenta vantagens que ajudam a ultrapassar os problemas relativos à elevada precisão da informação de navegação necessárias à formação de imagens de abertura sintética. Por outro lado, um pequeno veículo autónomo de superfície apresenta também vantagens em termos de custos de operação, manutenção e rapidez de intervenção.

O mapeamento de áreas submersas é uma tarefa importante para muitas áreas de aplicação comercias e científicas tais como a navegabilidade de rios, manutenção de infra-estruturas submersas ou a monitorização de recursos naturais. Estas tarefas podem ser realizadas de uma forma eficiente com um sistema como o apresentado neste trabalho.

Exemplos obtidos a partir de dados recolhidos durante missões de teste serão usados para ilustrar a geração de imagens sonar de abertura sintética com este sistema e o seu uso para a obtenção de mapas batimétricos.

Sommaire

Le travail exposé décrit un système de sonar à ouverture synthétique de haute définition sur bases interférométrique installe sur un véhicule de surface. Ce système a été développé en vue d'une opération dans les environnements d'eau peu profonds comme les fleuves, les lacs et les estuaires.

L'ouverture synthétique est une technique qui permet obtenir des images de sonar de haute définition indépendant de la distance des cibles, par la combinaison cohérente des échos recueillis successivement en utilisant le déplacement de la plateforme sur le quelle le sonar est installe. Si deux images sont obtenues de la même zone à partir des positions différentes, il est possible grâce à l'utilisation de techniques interférométriques, de faire une reconstruction trois-dimensionnelle de la topographie de cette zone.

Bien qu'une amélioration de la résolution dans le sens du déplacement de la plateforme puisse être atteinte en utilisant les techniques d'ouverture synthétique, une résolution plus fine pour distance de la cible ne peut être obtenue que par la transmission de signaux de bande passante élevée, ce qui oblige l'utilisation d'hautes fréquences. Ceci, en outre, impose des restrictions sévères sur les erreurs de phase qui sont dues à l'instabilité du mouvement de la plateforme et du milieu de propagation pendant l'acquisition de l'écho, qui rendent impossible la combinaison cohérente des échos au cours de la transformation d'ouverture synthétique.

L'approche suivie dans ce travail était d'explorer les techniques d'ouverture synthétique en utilisant un seul récepteur et les estimations des positions obtenues directement à partir du système de navigation du véhicule. Pour cela, l'utilisation d'un véhicule de surface a des avantages qui aident à surmonter les problèmes liés à la haute précision des informations de navigation nécessaires à la formation d'images d'ouverture synthétique. En outre, un petit véhicule autonome de surface présente aussi des avantages en termes de coût de fonctionnement, d'entretien et de la rapidité d'intervention.

La cartographie des régions peu profondes d'eau est une tâche importante pour beaucoup d'applications commerciales et scientifiques comme la navigabilité du fleuve, la maintenance d'infrastructure ou le contrôle des ressources minérales. Ces tâches peuvent être remplies efficacement avec un système comme celui présenté dans ce travail.

Des données recueillies au cours de missions de teste seront utilisées pour démontrer la création d'images de sonar à ouverture synthétique avec ce système, ainsi que leur utilisation dans la création des cartes bathymétriques.

Contents

List of Figures

Chapter 1: Introduction

1.1. Background and Motivation

Today a good percentage of our planet is known and well mapped. Synthetic aperture techniques used in space and airborne systems have greatly aided to obtain this information. Nevertheless, our planet is mostly covered by water and the level of detail of knowledge about this segment is still very far away from that of the land segment.

The presented system focus is the mapping and characterization of shallow water environments, such as rivers, estuaries, lakes and dams. Here, applications such as bottom topography, river navigability and inspection, dredging and aggregate extraction surveillance, infra-structure maintenance and object finding can play an important role in the study of these areas, and is of a major interest to the related fields of science, economy and ecology ([de Groot, S.J. (1996)]).

Topography mapping of the riverbeds is of high importance for a large set of reasons. Very few riverbeds have been scanned with sub-meter accuracy. Furthermore, these are dynamic environments and periodic studies are valuable for river navigability assessment, for instance. Also, dynamic sand deposits that form in rivers and overgrowth of submerged plants have an impact on the ecosystem and should therefore be identified and mapped ([Coffey and Hollingsworth (1972)]). These procedures must be performed on a regular basis and can be efficiently accomplished with interferometric synthetic aperture sonar. Surveillance of harbors is another important application for this system. Here, it is important to map objects like sunken containers, rocks and sand dunes that can be hazardous to ships using the docks ([Whalen, J.E. et al (2005)]).

Synthetic aperture is a technique that enables high resolution through the coherent processing of consecutive displaced echo data. Instead of using one static large array of transducers, it uses the along-track displacement of the sensors to synthesize a large virtual array. The resolution thus obtained is in the order of the transducer size and, most importantly, independent of the range between sensor and target. While a modern high frequency real-aperture sonar system can have a beam width below 1º, this translates into a resolution of half a meter at a range of just 25m. A synthetic aperture system using the same transducer can obtain a resolution of about 5cm across the whole range. Moreover the transducers used for synthetic aperture can be much simpler and so a good deal less expensive. Because there is no need to

have a small real aperture, the frequency employed can be considerably lower, which enables longer reach due to the better propagation of lower frequencies in water.

This potential resolution increase comes at the cost of algorithm complexity in the image formation and tight tolerances with respect to the sonar platform deviations from known path and velocity. The maximum velocity limit is related to the maximum phase shift between adjacent echoes that can still lead to their coherent combination. As this is a function of the pulse repetition rate and sensor size, this limit imposes very slow survey speeds whose problematic must be mitigated.

The most used platform for synthetic aperture sonar is the tow-fish. Good designs enable smooth motion, but the inability to use satellite navigation technology leads to expensive solutions that integrate high grade inertial navigation units and data extracted from the sonar array itself. This only works for arrays with a high count of elements that operate at the nominal or above the nominal pulse repetition frequency. The sonar can also be mounted on the hull of a ship, providing access to high precision GPS navigation that can be integrated with data from moderate cost inertial systems to further refine the navigation solution ([Legris, M. et al (2004),]). Nevertheless a ship is seldom easy to maneuver and presents considerable operation and maintenance costs of the sonar system itself. A surface vehicle arises as an interesting solution for these problems. It can be used as standalone or with the support of a ship. It enables the use of GPS and inertial navigation units efficiently. Moreover, its path and velocity can be controlled easily improving motion stability. The operation and maintenance costs are low and its availability is very high. It can be used both for sporadic missions and for regular security check of harbors, river navigability assessment, infrastructure inspection, etc.

The position of the sonar must be known to a 1/8 of the wavelength for proper synthetic image formation. Traditional synthetic aperture image formation techniques are derived from the assumption that the sonar motion is linear (typically for methods that operate in the frequency domain) and treat deviations from this straight path as motion errors. A newer approach uses time-domain methods (such as back-projection) that don’t rely on the assumption that the sonar follows any particular path. Instead, the information obtained by the navigation system is used at each sonar sampling position to form the virtual array. Here, only the position uncertainties are considered as errors.

High frequency systems require navigation precision bellow the centimeter level. This level of precision is not feasible to be obtained with the navigation systems of today. Therefore, the image formation starts with the available navigation solution and then a global auto-focus algorithm (that searches for an optimum measurement of image quality) refines the image

formation parameters to mitigate the unavailable necessary navigation precision. Instead of using redundancy in the data (that comes at a high cost), global auto-focus algorithms parameterizes the image formation process enclosing navigation errors and medium fluctuations. Because of the large number of parameters that result from this, an efficient process of focusing synthetic aperture images is necessary.

1.2. State of the art

An active sonar system enables underwater imaging through angular and range echo discrimination. When a single beam is used to illuminate a swath as the sonar platform moves, it is said that the sonar is a side-scan. In these systems a single echo line is obtained at each time with the angular discrimination being given by the beam width. Thus a narrow beam is desirable for high angular discrimination or along-track resolution. Typical beam widths are in the order of 1º ([Brissette, M. B. (2006); Denny, J. F. et al (2006)]). In these systems the along-track resolution is dependent of the range and a large array has to be used to obtain suitably low beam widths at the desired range. This type of sonar enables high area coverage speed. Alternatively, several narrow beams can be used to spatially sample the swath obtaining range and intensity information for each angular position. This system is called multi-beam sonar ([Pratson, L. F.; Edwards, M. H. (1996)]). In this case the footprint of each beam is also dependent of the range. This type of sonar requires expensive and complex hardware to achieve a high number of sampling narrow beams. The area coverage speed is also limited by the area covered by the beams.

Synthetic aperture enables a high resolution/high area coverage binomial not possible with other sonar techniques. Instead of using a long physical array, a large virtual array is synthesized through the coherent combination of the echoes in the along track dimension of a side-scan sonar. Range independent along-track resolution is in this way obtained. Moreover the obtained along-track resolution is not influenced by the frequency of the signals employed and is in the order of the transducer physical dimensions. Lower frequency signals can thus be employed to extend the sonar range. Also because of the processing gain, the necessary transmitting power is lower when compared to its real aperture counterpart.

While it is possible to apply synthetic aperture techniques to multi-beam sonar, these have not been of great dissemination and the focus has been on obtaining high resolution sonar systems with large area coverage speed ([Asada, A.; Yabuki, T. (2001)]).

It is possible to obtain height estimation through the use of interferometric techniques on side-scan sonar images or synthetic aperture sonar images ([Saebo, T.O. (2007); Silva, S. et al (2008 c)]). Multi-beam sonar systems obtain these height measurements directly.

Synthetic aperture sonar systems can operate in strip-map or spot-light mode. In spot-light mode the beam is steered so a single zone is illuminated as the sonar platform moves, while in strip-map mode an area is sequentially sampled. Since sonar applications normally strive for area-coverage, strip-map is the more used synthetic aperture mode.

Synthetic aperture techniques are of common use in radar systems (typically know as SAR, [Tomiyasu, K. (1978)]). Here the signal propagation velocity through the medium is much higher and the wavelength is long in comparison to radar platform motion uncertainties and medium phase fluctuations. Moreover the scene width is short when compared to the centre range which enables the use of simplifying approximation in the image formation and motion compensation algorithms. Also the bandwidth to centre frequency ratio is small and typically narrow beam widths antennas are used. This means that SAR algorithms are not suitable for direct application in sonar data.

Medium stability for synthetic aperture sonar system is a major issue, but it has been proved not to be a restraining factor for the application of this technique ([Gough, P. T. et al (1989)]).

Most active synthetic aperture sonar systems are still in the research and development stage ([Douglas, B. L.; Lee, H. (1993); Neudorfer, M. et al (1996); Sammelmann, G. S. et al (1997); Nelson, M. A. (1998); Chatillon, J. et al (1999)]) but this technology is starting to emerge as a commercially advantageous technology ([Hansen, R.E. et al (2005); Putney, A. et al (2005); Legris, M.; Jean, F. (2007)]). A large part of these systems rely on complex and costly underwater vehicle configurations with multiple receivers, making them not attractive cost-wise. This is because one needs to know the position with a very high degree of precision to correctly focus a synthetic array. Since it is not possible to use a high precision navigation system, such as a GPS, in the underwater environments, these systems have to rely on positioning schemes that use sonar data itself. This usually means having to use a complex array with multiple receivers, higher pulse repetition frequency or greatly limit the area coverage speed. Inertial navigation systems are of hardly any utility in standalone mode and must be corrected by other navigation sources. These underwater vehicles can be simply towed by a ship or be autonomous.

Surface autonomous vehicles are easier to operate and maintain. Also, it is possible to use readily available high precision differential GPS navigation system. Other navigation data

sources can also be used effectively to enhance the navigation solution. A simple transducer array can be used since it is no longer mandatory to have a multiple element system for navigation. This makes high resolution synthetic aperture sonar, obtained directly from navigation estimates, a cost-effective possibility.

1.3. Problem statement

Synthetic aperture techniques can be used to obtain centimeter level resolution in the along-track and cross-track dimensions. To obtain this level of resolution in the cross-track direction, one needs to use large bandwidths and, thus, high frequency signals. This makes the position accuracy issue even more problematic for synthetic aperture sonar system since the necessary accuracy is directly related to the wavelength of the signal used.

Furthermore a sonar platform is subject to several undesirable motions that negatively affect the synthetic aperture performance. Normally, a synthetic aperture sonar platform would move through a straight path with constant velocity. But this is seldom the case. Low frequency oscillations around the desired path and high frequency motion in heave, sway and surge directions adversely affect the sonar platform. This is due to water surface motion, currents and platform instability.

Tow-fish vehicles avoid the water surface induced motion instability, but at a cost of a more complex vehicle navigation. For an autonomous under-water vehicle it is difficult to maintain a path reference since there are no wide spread high accuracy navigation sources.

An autonomous boat offers automatic position and velocity following capabilities through the use of a highly available differential GPS system (DGPS-RTK). Furthermore, the sonar sampling position and attitude is known to a high precision level which can than be integrated in the image formation algorithms.

The synthetic aperture sonar is then thus only limited by the unknown portion of the platform motion and medium phase fluctuations. Typical DGPS-RTK systems can operate with an error in the centimeter level. This mean that a synthetic aperture image with a wavelength of about 8 cm can be directly formed using only this position estimation. Furthermore, in this way, the obtained images can be easily and accurately integrated with other geographic information systems. Nevertheless, shorter wavelengths require auto focus procedures to be applied to the data for successful image formation.

Traditional synthetic aperture image formation algorithms treat deviations from a linear path as motion errors. They do not integrate well large deviation from a theoretical linear,

constant velocity path, and so new algorithms had to be developed to cope with this information ([Callow, H. J. (2003)]).

Auto-focus algorithms, including the particular case of micronavigation, are essential to successfully produce high-resolution synthetic aperture images. Their role is to mitigate phase fluctuations due to medium stability and more importantly reduce the navigation precision requirements to acceptable levels.

Most auto-focus algorithms require large pulse repetition frequencies (PRF), several receiver transducers or both. In practical terms is not always possible to use high pulse repetition frequencies due to range ambiguity. With several receiver transducers the PRF can be lower, but the system cost is considerably higher not only due to the transducer array itself but also because of the data acquisition and signal processing system increased complexity.

1.4. Author Contributions

With this work it was possible to study sonar imagery in general and in particular study synthetic aperture sonar imagery. A large emphasis was given in broad band and broad beam systems single element systems.

An autonomous surface vehicle based synthetic aperture sonar system operated directly from navigation data was developed. Moreover a demanding centimeter level range resolution was targeted which lead to the use of high frequency signals. Due to the short wavelengths, this imposes very high constrains to the image formation errors that in turn lead to the investigation of new motion compensation techniques and auto-focus algorithms suitable for this system.

The complex motion dynamics and error constrains in the image formation rendered typical SAS algorithms unsuitable, instead a back-projection algorithm capable of efficiently using the available position and attitude information was implemented. Furthermore, a new algorithm for synthetic aperture processing of sonar images was developed: the arithmetic image reconstruction algorithm.

The performance requirements of the auto-focus procedures lead to a development of the sub-band algorithm and, furthermore, the study and development of a variable and dimension reduction algorithm for fast search of the optimum image generation parameter solution space.

The bathymetric capabilities of this system were studied, highlighting interferometric height mapping.

1.5. Publications

This work lead to the publication of the several papers in conference proceedings, a journal paper and participation on a book has editor and contributor. The bibliographic references are the following:

1. Cruz, N.; Matos, A.; Cunha, S.; Silva, S. (2007), "ZARCO - An Autonomous Craft for Underwater Surveys", Proceedings of the 7th Geomatic Week, Barcelona, Spain, February, 2007.

2. Silva, S.; Cunha, S.; Matos, A.; Cruz, N. (2007 a), “An In-SAS System For Shallow Water Surveying”, Proceedings of the 7th Geomatic Week, Barcelona, Spain, February, 2007.

3. Silva, S.; Cunha, S.; Matos, A.; Cruz, N. (2007 b), "An Autonomous Boat Based Synthetic Aperture Sonar”, Proceedings of the Oceans 2007 MTS/IEEE Conference, vol., no., pp.1-7, September, 2007.

4. Silva, S.; Cunha, S.; Matos, A.; Cruz, N. (2008 a), "An Algebraic Approach to Synthetic Aperture Sonar Image Reconstruction", Proceedings of the Oceans 2008 MTS/IEEE Conference, September, 2008.

5. Silva, S.; Cunha, S.; Matos, A.; Cruz, N. (2008 b), "Sub-Band Processing of Synthetic Aperture Sonar Data", Proceedings of the Oceans 2008 MTS/IEEE Conference, September, 2008.

6. Silva, S.; Cunha, S.; Matos, A.; Cruz, N. (2008 c), "Shallow Water Height Mapping with Interferometric Synthetic Aperture Sonar", Proceedings of the Oceans 2008 MTS/IEEE Conference, September, 2008.

7. Silva, S. et al. (2009 a), "Advances in Sonar Technology", In-Tech, ISBN 978-3-902613-48-6, February, 2009.

8. Silva, S.; Cunha, S.; Matos, A.; Cruz, N. (2009 b), "Shallow Water Surveying Using Experimental Interferometric Synthetic Aperture Sonar", Marine Technology Society Journal, vol. 43, no. 1, pp. 50-63, 2009.

9. Silva, S.; Cunha, S.; Matos, A.; Cruz, N. (2009 c), "An Autonomous Surface Vehicle InSAS for Maintenance and Surveillance Applications", Proceedings of the SMART 2009 Conference, Porto, Portugal, July, 2009.

1.6. Thesis Outline

After contextualization and problem description, this thesis proceeds with the study of signal and sonar fundamentals. Here important concepts for array theory and pulse compression will be highlighted and used later in the development of the image formation algorithms. A description of the developed system follows describing the hardware architecture and navigation system. Its capabilities, shortcomings and engineering tradeoffs are explained and, more importantly, their impacts on the sonar system characteristics are put in evidence.

The chapter about the sonar model describes the sonar imaging problem geometry. The image formation algorithms are presented next with the solutions for inclusion of navigation data described. Image enhancing chapter describes the image filtering procedures to reduce noise and speckle. The auto-focus methods developed are explained in the chapter that follows. In the Bathymetry chapter, height mapping methods are explored using both interferometry and volumetric back-projection.

Results obtained during test trials in the Douro River will be presented for the synthetic aperture and bathymetric capabilities. The conclusion and future work chapter points out the achievements but also the problems that were found with this system, pointing some future paths for research work.

Chapter 2: Sonar & Signal fundamentals

2.1. Acoustic Wave Propagation

The water medium is a time and spatially varying propagation environment whose characteristics pose significant challenges to the development of sonar imaging systems.

When compared to electromagnetic propagation through the atmosphere, acoustic propagation through water is characterized by significant frequency dependent attenuation and relative slow speed of propagation. These characteristics are present in all water environments. Sound waves propagate in a medium through alternating compression and rarefaction effects. These are the changes in pressure that are then converted to electrical signals by a transducer (receiving transducer - hydrophone). Electrical signal can also be converter to pressure changes by electromechanical actuation (transmitting transducer - projector). Typical frequencies used for underwater acoustic systems range from 10Hz to 1Mhz. The propagation of sound in the underwater realms at frequencies lower than 10Hz is usually not possible without penetrating deep into the bottom layers, while frequencies above 1MHz are rarely used because the attenuation due to distance is very high.

Assuming that the compression laws govern the particles deformation (adiabatic process), that the particles mass is constant (principle of conservation of the mass) and that the particle motion is small (so that the particle speed and pressure can be decomposed in a constant value and a linear term), it is possible to arrive at the following expression for the single dimension x ([Kinsler, L. et al (1982)]):

$$\frac{\partial p}{\partial x} = -\rho \frac{\partial u}{\partial t} \tag{2.1}$$

Here, p is the instantaneous fluid local pressure, ρ is the mass density and u is the particle velocity. From this expression it is possible to deduce the sound wave equation for the general three dimensional space:

$$\nabla^2 p - \left(\frac{\rho}{K}\right)\frac{\partial^2 p}{\partial^2 t} = 0 \tag{2.2}$$

K is the bulk modulus of the fluid, expressed in Pa. This second order differential equation can be recognized as the wave equation making the ratio (ρ/K) the square of the propagation speed. The speed of sound can thus be defined as:

$$c = \sqrt{\frac{\rho}{K}} \tag{2.3}$$

Substituting c in the equation we get:

$$\nabla^2 p - \frac{1}{c^2}\frac{\partial^2 p}{\partial^2 t} = 0 \tag{2.4}$$

Considering sinusoidal particle velocities, the solution for the above equation is:

$$p(r,k) = \frac{A}{r}e^{\pm jk} \tag{2.5}$$

Where k is defined as $k = 2\pi/\lambda = \omega/c$ for the particles sinusoidal motion frequency ω, A is the pressure amplitude and r the distance to the source.

The power associated with this pressure is given by:

$$I = \frac{RMS\{p\}^2}{c\rho} \quad (W) \tag{2.6}$$

Pressure changes are measured as sound pressure level (SPL). Typical units are: RMS acoustic pressure in micropascal (dB re 1μPa); RMS acoustic pressure in a specific bandwidth (dB re 1μPa/octave); spectral density: (dB re 1μPa^2/Hz).

The linearity assumption taken above for the mass density is just an approximation. In water, especially with air bubbles, the change in density due to a change in pressure is not exactly linearly proportional. The effect is greater at high source levels. Because of the non-linearity there is a dependence of sound speed on the pressure amplitude so that large changes travel faster than small ones. As a consequence, for a sinusoidal wave inputs intermodulation products are generated. Thus a sinusoidal waveform gradually becomes a saw-tooth one with a steep rise and a gradual tail.

Spreading loss, absorption loss, and scattering loss are the three primary mechanisms which attenuate underwater acoustic signals. Spreading losses are due to the expansion of the fixed amount of transmitted energy over a large surface area as the signal propagates away

from its source. At relatively short ranges, the increasing surface area is represented by the surface of a sphere so signal energy decays due to spreading loss at a rate of r^{-2}, where r is the range from the source.

However, typical underwater mediums (like oceans or rivers) are bounded from above by the surface and, at the frequencies and ranges typically of interest for sonar, it is effectively bounded from below by the sea floor. Thus, at some range from the source the acoustic signal can no longer spread vertically and the nature of the spreading changes from spherical to cylindrical spreading. This transition typically occurs at ranges much greater than water depth ([Jensen, J.; Kuperman, W.; Porter, M.; Schmidt, H. (2000)]). In the cylindrical spreading region, signal energy decay due to spreading loss is at a rate of r^{-1}.

A second mechanism of signal loss results from the conversion of the energy in the propagation signal into heat. This mechanism is referred to as absorption loss. In water the absorption loss of acoustic signals is strongly frequency dependent and increases with the increasing frequency ([Clay, C. S.; Medwin, H. (1977)]). Signal decay due to absorption loss is proportional to $e^{-\alpha(f)r}$, where $\alpha(f)$ is an increasing function of frequency. The acoustic signals are absorbed by the medium due to its viscosity (fresh and salt water) and ionic relaxation phenomena (salt water). In sea water the dominant ions are due to the boric acid and magnesium sulfate.

Two characteristics of spreading and absorption loss are worth noting. First, at short ranges the spherical spreading loss dominates the absorption loss. Second, even at short ranges (approximately 400m) the absorption loss at 100kHz exceeds that at 25kHz by close to 15dB. The practical impact of the frequency dependence of absorption loss is that the propagation channel is effectively bandwidth limited and the available bandwidth is a decreasing function of range.

The relatively slow speed of propagation of sound through water ($c \approx 1500$m/s) is also a factor that differentiates it from electromagnetic propagation ($c \approx 3 \times 10^8$m/s).

This relatively slow propagation speed results in high Doppler spreads or shifts. A Doppler shift is an apparent frequency change in the received signal caused by a sound source that is in movement in relation to the receiver. Received signals suffer Doppler shifts resulting from propagation path length fluctuations due to platform motion or scattering from the moving sea surface.

The Doppler shift (f_d) of a received signal is given, approximately for a sinusoidal signal, $v << c$ and large source receiver distances compared to the wavelength, by ([Rayleigh, J. W. S.; (1945)]):

$$f_d = f_0 \frac{v}{c} \tag{2.7}$$

Where f_0 is the original signal frequency and v is the rate of change in propagation path length (can be due to the platform velocity of surface reflection effects). Thus even at modest values of $v = 2m/s$ and $f_0 = 25kHz$, the Doppler shift of this signal would be approximately 33Hz (>0.1%). Similar Doppler spreads have been reported resulting from a difference in rates of fluctuation of the lengths of two propagation paths ([Preisig, J. (2005)]). These Doppler spreads and shifts result in a reduction in coherence time or apparent increase in the rate of channel fluctuation.

In most environments and at the frequencies of interest for sonar signals, the sea or river water can be modeled as a waveguide with a reflecting surface and bottom and a spatially variant sound speed. Both the water surface and bottom are reflecting and scattering boundaries. Due to the large impedance difference between the air and the water, the water surface is normally assumed as perfect reflector below 1kHz, but behaves as a scatter above 1kHz. The seabed reflections are complex and depend on the nature of the materials, types of existing sediments layers, frequency and grazing angle. Bottom scattering strengths may be considered roughly constant at frequencies up to 10kHz and grazing angles up to 10º ([Waite, A. D. (2005)]). At frequencies above 10kHz, the strength increases at about 3dB per octave for the smother (mud, sand) bottoms (where the roughness is small compared to a wavelength), but for rough bottoms (shingle, pebbles, rock), the scattering strength is independent of frequency. Scattering strengths vary considerably for the type of material present at the bottom, ranging from -45dB for mud, to -35 dB for sand and shingle and -28dB for pebbles and rocks ([Waite, A. D. (2005)]).

The reflections of acoustic signals from the water surface and bottom and the refraction of signals by the spatially varying sound speed in the water column results in multiple propagation paths from each source to receiver. This multipath results in a delay spread in the often time-varying impulse response of the channel, leading to reduced range discrimination ability. Sound fluctuations can be seen in amplitude and phase of the received signals and are

of temporal and spatial nature. These are mainly due to interference phenomena caused by reflections in the water surface and changes in propagation conditions.

The refraction of signals by the sound speed fluctuation not only gives rise to multipath, but can result in the formation of "shadow zones" ([Clay, C. S.; Medwin, H. (1977)]). These are areas where there is little propagation of signal energy. Thus it could be difficult to receive signals from a source (or target) located in a shadow zone. This effect is typical of long range deep water channels, but the same phenomenon can be found in shallow waters (in the order of 100m depth) and shorter ranges (in the order of 3km). In these environments, the vertical movement of masses of water results in vertical movement of the sound speed structure of the water column. This phenomenon gives rise to variations in the location of shadow zones, even for the case of a stationary source and receiver, and has been studied in [Porter, M. B., and the KauaiEx Group (2004)].

The rough water surface gives rise to a spreading in delay for each surface bounce path that can reduce spatial correlation of scattered signals, and can result in very high intensity and rapidly fluctuating arrivals in the channel impulse response. Surface focusing results from the fact that waves moving over the sea surface can act as downwardly facing curved mirrors that reflect the sound down into the water column and focus it at predictable locations.

The speed of sound in the water is a function of the density and bulk modulus and so depends on the temperature, depth and salinity. The speed of sound at 10ºC, zero depth and for a salinity of 35ppm, is 1495 m/s. The speed of sound changes approximately linearly for salinity and depth but has a strong non linear dependence with temperature (where T is temperature, S is the salinity and D is the depth):

$$c(D,S,T) = 1448.96+4.591T+1.340\,(S-35)+1.630\cdot 10^{-2}D \quad (m/s) \tag{2.8}$$

A variety of empirical formulas exist for calculation of the sound speed, the equation above is for the speed of sound in water, this is a simplification of the Mackenzie equation ([Mackenzie, K.V. (1981)]), where only the first order terms where kept. The range of validity of the full Mackenzie equation is: temperature 2 to 30ºC, salinity 25 to 40 parts per thousand, depth 0 to 8000m. More recently the UNESCO equation has been adopted which has a broader range of validity, but the sound speed is given as a function of pressure and not depth and so it is necessary to proceed with the corresponding conversion. The interested reader is referred to [Wong, G. S. K.; Zhu, Z. (1995)] for more details.

The thermocline is a thin but distinct layer, in which temperature changes more rapidly with depth than it does in the layers above or below. In the ocean, the thermocline may be thought of as an invisible blanket which separates the upper mixed layer from the calm deep water below. Depending largely on season, latitude and turbulent mixing by wind, thermoclines may be a semi-permanent feature of the body of water in which they occur, or they may form temporarily in response to phenomena such as the radiative heating and cooling of surface water during the day and night. Factors that affect the depth and thickness of a thermocline include seasonal weather variations, latitude, and local environmental conditions, such as tides and currents. Thermocline can also be seen in river waters and with lower depths than in the ocean.

Bubbles generated by braking waves at the sea surface can have a major influence on high frequency acoustic propagation in both open ocean and near shore regions. Layers of bubbles near the surface can result in a significant attenuation of surface scattered signals. In one experiment, the impact of scattering of surface bubble layer was estimated to be an attenuation of the surface scattered signal by 3 dB per surface bounce. For bubble densities characteristic of wind speeds up to around 6m/s, no bubble induced losses were reported. Above this level, bubble induced losses increase as a function of wind speed with almost total signal loss at wind speeds of approximately 10m/s. Bubble clouds injected down into the water column also significantly attenuate propagation signals with rates as high as 26 dB/m being reported ([Farmer, D. M; Deane, G. B., Vagle, S. (2001)]). The injection of bubbles by a breaking wave in shallow water can result in a sudden channel outage.

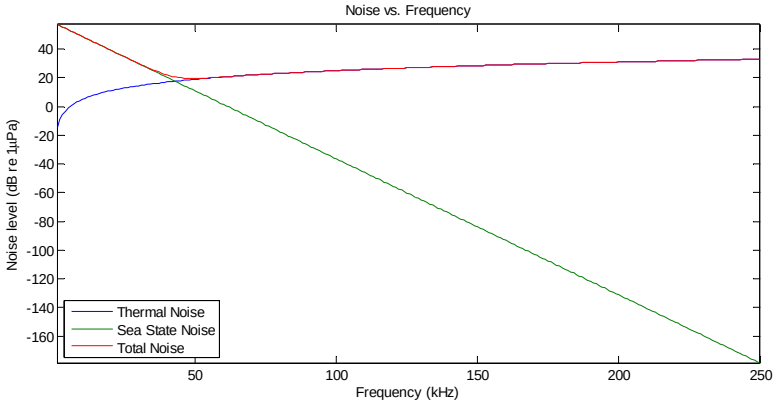

Figure 1: Noise level as a function of frequency.

There are several important natural sources of ambient noise in the ocean at frequencies of interest for acoustic imaging. These include breaking waves and bubbles, thermal noise (molecular excitation), noise from the water agitation (ambient noise), vessel produced noise

biological sources and rain. Ambient noise has been studied extensively with a common theme that the power spectral density of the noise decreases with increasing frequency ([Jensen, J.; Kuperman, W.; Porter, M.; Schmidt, H. (2000), Farmer, D. M; Lemon, D. (1984)]). Bubbles are one primary source of ambient noise in the open ocean in the 3 to 30kHz band, with the noise having a roll off of approximately 5dB per octave. Noise levels increase with wind speed up to a point after which the noise level is reduced due to absorption by the surface layer of bubbles. The frequency dependence of the ambient noise should be one of the factors considered when selecting frequency bands for underwater acoustic imaging. Thermal noise is dominant at high frequencies (above 50kHz) while sea state noise dominates the lower frequency spectrum (Figure 1). In this case the noise level is approximately given by ([Waite, A. D. (2005)]):

$$N_{Thermal} = -15 + 20\log_{10} f \quad \text{(dB)} \tag{2.9}$$

2.2. Sonar equation

As described in the previous section, the propagation of sound in the medium is dependent on many factors such as the sound speed profile, sound absorption, reflection in the surface and bottom, etc. For short ranges and high frequency (single hop, as typical in side scan sonar applications), the propagation model can be formed using ray theory. In particular here forward it will be assumed that for the short ranges and shallow depth of the application scenarios of synthetic aperture sonar, the sound waves will propagate in straight lines, and its propagation can be correctly approximated by the sonar equation, which relates the received signal with the signal generated or reflected by the target. It is a simple budget notation and is useful for sonar system design. Reverberation effects are due to scattering and usually add to the background noise and will be treated as such.

Passive sonar detects signals radiated by the target, for these systems, the sonar equation is ([Waite, A. D. (2005)]):

$$SE_P = (SL + DI_R - PL) - N - DT \quad (dB) \tag{2.10}$$

For active sonar systems, the targets are insonified by a transmitting transducer, enabling the detection of the object. In this case the sonar equation is ([Waite, A. D. (2005)]):

$$SE_A = (SL + DI_T + DI_R + TS - 2PL) - N - DT \quad (dB) \tag{2.11}$$

The source level (SL) is the sonar acoustic pressure level. For an omnidirectional projector it is referred to a standard range (1m) from its acoustic centre. At 1m, the acoustic centre of an omnidirectional source is surrounded by a sphere of a surface area $4\pi r^2 = 12.6m^2$. If the omnidirectional acoustic pressure produced by the sensor is P in Watts, then the source intensity at 1m is $P/12.6W/m^2$ and the source level becomes ([Waite, A. D. (2005)]):

$$SL = 10\log_{10}\left(\frac{I_1}{I_R}\right) = 10\log_{10}\left(\frac{P/12.6}{0.67\cdot 10^{-18}}\right) = 10\log_{10} P + 170.8 \quad (dB) \tag{2.12}$$

DI_T is the transmitter directivity and DI_R is the receiver directivity and is given by the relation between the intensity along the axis of the beam pattern and the intensity if the source was omnidirectional.

$$DI = 10\log_{10}\left(\frac{I_{dir}}{I_{omni}}\right) \quad (dB) \tag{2.13}$$

The acoustic path loss is the quantity PL ([Waite, A. D. (2005)]):

$$PL = 10\log_{10}\left(\frac{I_S}{I_R}\right) = 20\log_{10}\left(\frac{p_S}{p_R}\right) \quad (dB) \tag{2.14}$$

The path loss depends on the propagation conditions. In deep waters the propagation is typically assumed to be spherical, while in shallow waters the propagation is taken to be cylindrical as the surface of the water and bottom form a wave guide. An approximation to the path loss equation for spherical losses is given by:

$$PL = 20\log_{10} r + \alpha(f)r \quad (dB) \tag{2.15}$$

And for cylindrical spread losses:

$$PL = 10\log_{10} r + \alpha(f)r \quad (dB) \tag{2.16}$$

The acoustic pressure absorption is taken into account by the parameter $\alpha(f)$ that has an empirical relation with frequency (f in kHz) given by ([Waite, A. D. (2005)]):

$$\alpha(f) = 0.05 f^{1.4} \quad (dB/km) \tag{2.17}$$

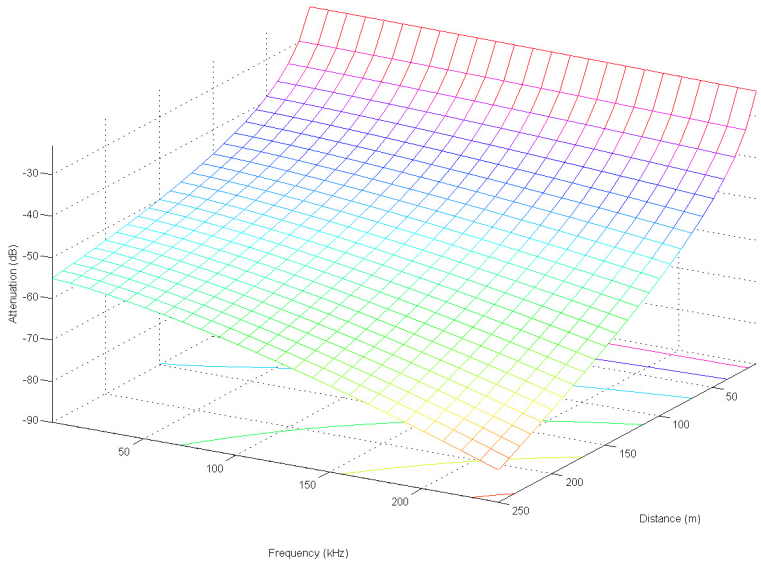

Figure 2: Attenuation as a function of frequency and distance.

Figure 2 shows the attenuation suffered by an acoustic signal as a function of frequency and distance. As it can be seen, the underwater sound channel behaves effectively as a low pass filter.

TS is the target strength, which is given as the ratio between the reflected (I_R) and incident (I_I) acoustic pressure on a target as ([Waite, A. D. (2005)]):

$$TS = 10\log_{10}\left(\frac{I_R}{I_I}\right) \quad (dB) \tag{2.18}$$

Target reflections depend on its impedance in relation to the water. This is the target ability to reflect acoustic signals (usually a negative numeric value in *dBs*). In relation to the target cross section, the target strength can also be given as:

$$TS = 10\log_{10}\left(\frac{BS}{4}\right) \quad (dB) \tag{2.19}$$

The backscattering cross section (*BS*), relates the mean incident wave power (I_I) with the backscattered wave intensity (I_B) at midpoint by the equivalent volume as:

$$BS = 4R^2 \frac{I_I}{I_B} \quad (m^2) \tag{2.20}$$

The noise level N is due to noise from the water environment but also has a component due to the sonar electronics (transducer, amplifier and analog-to-digital converter).

For correct sonar processing (being it synthetic aperture or target qualification in real aperture systems), the minimum signal to noise ratio (SNR) necessary is given by DT .

2.3. Array Theory

2.3.1. Transducers

Transducers operating in the transmit mode are usually assembled as arrays in order to increase the source level of the transmitted acoustic pulse. The directivity of a projector array concentrates the transmitted sound in a given direction.

In the horizontal plane the transmissions may be omnidirectional or, with a corresponding increase in source level, they may take place over some smaller sector. In the vertical plane there is clearly no point in transmitting omnidirectionally, so a typical array is designed to make the vertical beam width between 3º and 30º to the 3dB points.

The required source level might not be available from one projector. A single projector requires a large surface area to avoid cavitation, and at higher frequencies this is not a simple engineering solution: directivity would be sacrificed or at best compromised. Finally, the reliability of an array of projectors is much better because several, perhaps even half, of the projectors can fail before performance is critically impaired. A multiple transducers system is, nevertheless, more complex and more expensive. Synthetic aperture tries to overcome the limitations of systems that use few transducers at the expense of more signal processing complexity.

Transducers operating in the receive mode are assembled as arrays to improve the response of the array in a desired direction, thereby increasing the signal-to-noise ratio and indicating the direction of a signal source. The directivity of a hydrophone array is due to all signals impinging upon the array from the same direction being in phase, and therefore reinforcing one another.

2.3.2. Beam-Patterns

The beam pattern of an array plots the array response against angle, describing the relative response of a beam formed from an array to signals and noise from all directions. Mathematically, the beam pattern of an array is identical for both transmit and receive. The improvement in the signal-to-noise ratio due to the array is known as the array gain. A special

case of array gain, that considers the signal coherent and the noise is incoherent, is designated by directivity index. This parameter is simpler to calculate and will be normally be a satisfactory measure of the increase in the signal-to-noise ratio due to the array.

2.3.3. Circular Transducer

Before starting our study of multiple transducer array we will first see how sound is projected by a single circular transducer. For this we will consider that it can be modeled by a flat circular piston of radius a made of thin and rigid material. This piston is completely surrounded by an infinite mass of water (and so the sound does not suffer reverberation) and is vibrating with a simple harmonic motion $u(t) = u_0 e^{jwt}$. The axes of coordinates are centered in the piston, and the radiated sound field has symmetry of revolution with the same axis as the piston.

The pressure in an area $dS = 2\pi x dx$ of the piston is ([Kinsler, L. et al (1982)]):

$$dp(r,\theta,t) = -\rho j\omega u_0 e^{j\omega\left(t-\frac{r}{c}\right)} = \rho j\omega u_0 e^{j\omega t - kr} dS \tag{2.21}$$

Where r/c is the delay due to the distance r to the source. Integrating over the surface of the piston we get (and assuming cylindrical propagation):

$$p(r,\theta,t) = j\frac{\rho c u_0}{2\pi}\int_S \frac{e^{j(\omega t - kr)}}{r} dS \tag{2.22}$$

The integral is taken between 0 and the piston radius a. The filed along the acoustic axis ($\theta = 0$) is relatively simple to calculate. The integral becomes:

$$p(r,0,t) = j\frac{\rho c u_0 k}{2\pi}\int_0^a \frac{e^{j(\omega t - kr)}}{\sqrt{z^2 + x^2}} 2\pi x dx \tag{2.23}$$

Solving the integral we find that the amplitude of the pressure through the z axis is:

$$p(z,0) = 2\rho c u_0 \left|\sin\left\{\frac{1}{2}kr\left[\sqrt{1+\left(\frac{a}{r}\right)^2} - 1\right]\right\}\right| \tag{2.24}$$

Since the transducers are usually small, this expression can be simplified considering $a/r << 1$ and that $\sqrt{1+x}$ can be approximated by the first order Taylor expansion:

$$p(z,0) = 2\rho c u_0 \left| \sin\left\{ \frac{a^2}{4} \frac{k}{r} \right\} \right| \tag{2.25}$$

The sine can also be approximated by a linear term considering that $k/r << 1$ which is true for normal sonar operating conditions since the wavelengths are in the order of centimeters and the closest range in the order of meters:

$$p(z,0) = \frac{\rho c u_0}{2} a^2 \frac{k}{r} \tag{2.26}$$

This approximation is what is called the far field approximation. In the near field ($a/r << 1 \cap k/r << 1$) the pressure amplitude varies rapidly between 0 and $2\rho c u_0$ as is suggested by the sine. As the range increases, the pressure amplitude follows a $1/r$ dependence as it would be expected since the cylindrical propagation is being considered here.

In the far field, the circular transducer directivity is given by:

$$DI(\theta) = \left| \frac{2J_1(ka\sin\theta)}{ka\sin\theta} \right| \approx \left| 2\mathrm{sinc}(ka\sin\theta) \right| \tag{2.27}$$

Where J_1 is the first order Bessel function. If the radius of the transducer is large compared to wavelength ($ka >> 1$) the first lobe is narrow and there are several side lobes. On the other hand, if the wavelength is large compared to the size of the transducer ($ka << 1$), only the major lobe will be present.

Taking the approximation for the far field, the null-to-null beam-width ($\theta_{null-to-null}$) is (where *D* is the piston diameter):

$$\theta_{null-to-null} = \sin^{-1}\left(\frac{\lambda}{r}\right) \approx \frac{\lambda}{r} = \frac{2\lambda}{D} \tag{2.28}$$

2.3.4. Real Arrays

Receiving and transmitting arrays are perfectly dual, and so without lost of generality we will analyze the case of the transmitting array as it is more intuitive. If S_1 and S_2 are two point sources (forming a dipole) vibrating in phase and equal amplitudes (p_0), a hydrophone placed in the far field in the direction θ will receive two sound waves of equal amplitude (since S_1 and

S_2 are almost equidistant from the hydrophone), but of different phase corresponding to the distance as illustrated Figure 3 ([Waite, A. D. (2005)]):

$$\Delta d = d \sin\theta \tag{2.29}$$

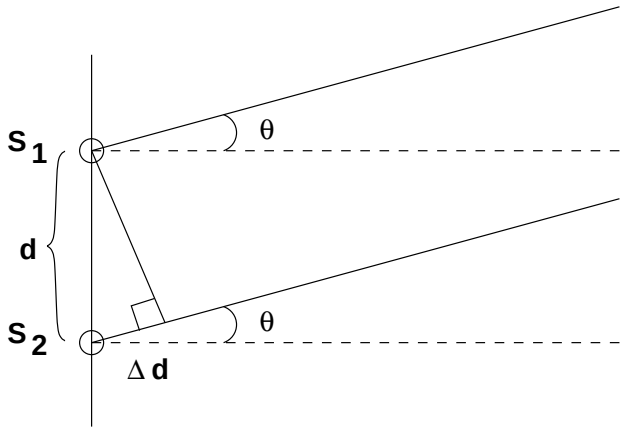

Figure 3: Dipole array.

The phase difference is:

$$\varphi = \frac{2\pi d}{\lambda} \sin\theta \tag{2.30}$$

The sound pressure at the hydrophone is the vector sum of the pressure p_1 and p_2, of the same amplitude and with a phase difference φ:

$$p(\theta) = p_1(\theta) + p_2(\theta) = p_0 + p_0 e^{j\varphi} \tag{2.31}$$

For a single source S_1, the pressure is constant whatever the direction. When a second source S_2 is present, the pressure varies between 0 and $2p_0$ as a function of θ:

$$p(\theta) = p_0\left[1 + e^{j\varphi}\right] = 2p_0 \cos\left(\frac{\pi d}{\lambda}\sin\theta\right) e^{j\frac{\pi d}{\lambda}\sin\theta} \tag{2.32}$$

At the axis of symmetry of the dipole (where $\theta = 0$), the pressure is $p(0)$ and so we can write:

$$\frac{p(\theta)}{p(0)} = \cos\left(\frac{\pi d}{\lambda}\sin\theta\right) e^{j\frac{\pi d}{\lambda}\sin\theta} \tag{2.33}$$

The intensity varies as the square of the absolute value of the pressure, as a result:

$$\frac{I(\theta)}{I(0)} = \cos^2\left(\frac{\pi d}{\lambda}\sin\theta\right) \tag{2.34}$$

And in logarithmic form:

$$10\log_{10}\left(\frac{I(\theta)}{I(0)}\right) = 20\log_{10}\left(\cos\left(\frac{\pi d}{\lambda}\sin\theta\right)\right) \tag{2.35}$$

The 3dB beam-width (the width in degrees at the half-power points of the dipole response) for a distance between dipoles d can be obtained by:

$$20\log_{10}\left(\cos\left(\frac{\pi d}{\lambda}\sin\theta\right)\right) = 3dB \Rightarrow \cos\left(\frac{\pi d}{\lambda}\sin\theta\right) = \frac{\sqrt{2}}{2} \Leftrightarrow \theta = \sin^{-1}\left(\frac{\lambda}{4d}\right) \tag{2.36}$$

When the spacing is exactly $d = \lambda/2$, nulls appear at 90^o and -90^o. As the spacing is increased secondary (side lobes) appear which are soon unacceptable, and when $d = \lambda$ they totally change the beam pattern so that the major response now occur at the previous null angles.

Considering now that the dipole of element is at the centre of a large sphere of radius r, the sound flux passing through an area of $rd\theta$ around the centre is then $I(\theta, 2\pi r\cos\theta)rd\theta$.

If $\langle I\rangle$ is the average intensity at the surface of the sphere, the total sound flux passing through its complete surface is: $\langle I\rangle 4\pi r^2$, and therefore ([Waite, A. D. (2005)]):

$$\langle I\rangle = \frac{1}{4\pi r^2}\int_{-\pi/2}^{\pi/2} I(\theta, 2\pi r\cos\theta)rd\theta \tag{2.37}$$

Defining the directive index (DI) as $DI = 10\log_{10} K$, where K is the directivity factor we get:

$$\frac{1}{K} = \frac{\langle I\rangle}{I(0)} = \frac{1}{4\pi r^2}\int_{-\pi/2}^{\pi/2}\frac{I(\theta, 2\pi r\cos\theta)rd\theta}{I(0)} \tag{2.38}$$

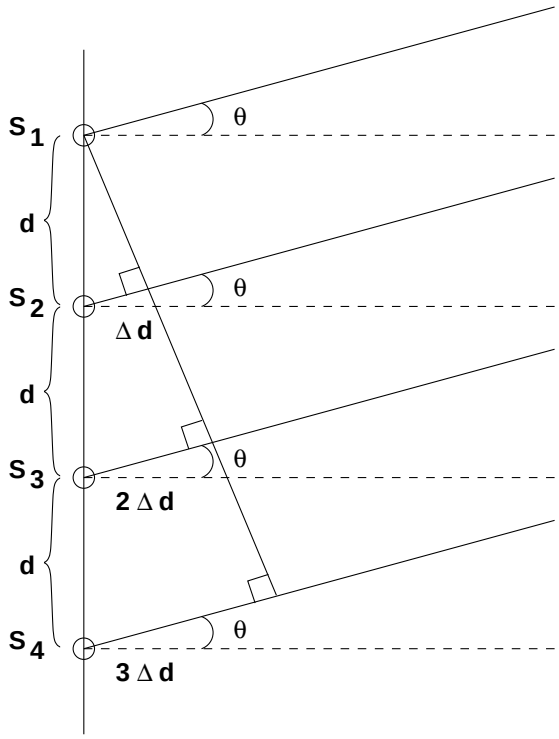

Figure 4: Array of n hydrophones.

Substituting:

$$\frac{I(\theta)}{I(0)} = \cos^2\left(\frac{\pi d}{\lambda}\sin\theta\right) \tag{2.39}$$

We obtain:

$$\frac{1}{K} = \frac{\langle I \rangle}{I(0)} = \frac{1}{2}\int_{-\pi/2}^{\pi/2} \cos^2\left(\frac{\pi d}{\lambda}\sin\theta\right)\cos\theta d\theta \tag{2.40}$$

And thus integrating and changing the variable, we get:

$$\frac{1}{K} = \frac{1}{2}\left[1 + \frac{\sin\left(\frac{2\pi d}{\lambda}\right)}{\frac{2\pi d}{\lambda}}\right] \tag{2.41}$$

And,

$$DI = 10\log_{10} K = 3 - 10\log_{10}\left[1 + \frac{\sin\left(\frac{2\pi d}{\lambda}\right)}{\frac{2\pi d}{\lambda}}\right] \tag{2.42}$$

When the elements are separated by $d = \lambda/2$, this reduces to $DI = 3 \quad dB$. When the elements are separated by $d = 0$, then $DI = 0 \quad dB$. We would expect both of these results intuitively: two elements have twice (3 dB) the power as one; and when the two become one ($d = 0$), the dipole is omnidirectional and $DI = 0 \quad dB$.

The simple dipole can be is extended to an array of n hydrophones of spacing d (Figure 4), the beam pattern is then given by ([Waite, A. D. (2005)]):

$$p(\theta) = p_0(\theta) + p_0(\theta)e^{j\varphi} + \ldots + p_0(\theta)e^{j(n-1)\varphi} = np_0\sum_{k=1}^{n} e^{j(k-1)\varphi} = np_0\frac{1-e^{jn\varphi}}{1-e^{j\varphi}} \tag{2.43}$$

Using the complex trigonometric identities it is possible to simplify the above equation obtaining:

$$p(\theta) = np_0\frac{\sin\left(\frac{n\pi}{\lambda}\sin\theta\right)}{\sin\left(\frac{\pi}{\lambda}\sin\theta\right)}e^{j(n-1)\frac{d\pi}{\lambda}\sin\theta} \tag{2.44}$$

If the individual beam patterns of each of the transducer are not omnidirectional and thus have a dependency in amplitude and phase with θ (for example an array of circular transducers), but are approximately equal ($p_0(\theta)$), it is trivial to see that the resulting array will have the following beam pattern:

$$p_0(\theta) = np_0(\theta)\frac{\sin\left(\frac{n\pi}{\lambda}\sin\theta\right)}{\sin\left(\frac{\pi}{\lambda}\sin\theta\right)}e^{j(n-1)\frac{d\pi}{\lambda}\sin\theta} \tag{2.45}$$

The main lobes of an array can be steered by introducing phase or time delays in series with the elements. To place the maximum of the beam pattern at the angle θ_s the delay must be:

$$\frac{n\pi}{\lambda}\sin\theta_s - d_s = 0 \Rightarrow d_s = \frac{\pi}{\lambda}\sin\theta_s \tag{2.46}$$

When this is done the beam width and side lobes structure of the array is changed. The beam pattern of the line array is then modified to yield:

$$p(\theta) = np_0 \frac{\sin\left(n\pi\left(\frac{d}{\lambda}\sin\theta - \frac{d}{\lambda}\sin\theta_s\right)\right)}{\sin\left(\pi\frac{d}{\lambda}\left(\sin\theta - \frac{d}{\lambda}\sin\theta_s\right)\right)} e^{j\left((n-1)\frac{d\pi}{\lambda}\sin\theta - \frac{d\pi}{\lambda}\sin\theta_s\right)} \tag{2.47}$$

The directivity of an array of n elements spaced $d = \lambda/2$ is given by:

$$DI = 10\log_{10} n \tag{2.48}$$

For a continuous line of length L, the pressure in the far field is found by integration to be:

$$p(\theta) = p_0(\theta)\frac{\sin\left(\frac{L\pi}{\lambda}\sin\theta\right)}{\sin\left(\frac{L\pi}{\lambda}\sin\theta\right)} e^{j\frac{L\pi}{2\lambda}\sin\theta} = p_0(\theta)\,\mathrm{sinc}\left(\frac{L\pi}{\lambda}\sin\theta\right) e^{j\frac{L\pi}{2\lambda}\sin\theta} \tag{2.49}$$

Since this expression reduces to a sinc function ($\mathrm{sinc}(x) = \mathrm{sinc}(x)/x$), it is possible to see that there will be nulls in the beam pattern where the argument of the sinc function equals $\pm n\pi, n = 1, 2, 3...$ and peaks at 0 and $\pm\left((n-1)\pi + \frac{\pi}{2}\right), n = 1, 2, 3...$

Except for the simple dipole with elements spacing less than or equal to $d = \lambda/2$, all the above beam patterns have significant side lobe levels. Strong signals will be detected through the side lobes of adjacent beams as well as, correctly, in the main lobe of the beam at the bearing of the signal. The resultant bearing ambiguity and additional, false signals complicate all further processing. The aim, therefore, should be to produce the narrowest possible main lobe consistent with some reasonable side lobes.

In the case of a continuous line, the directivity index is:

$$DI = 10\log_{10}\left(\frac{2L}{\lambda}\right) \tag{2.50}$$

By using amplitude shading or weighting, meaning varying the amplitudes of the signals applied across the array (in transmit) or coming from the array (in receive), it is possible

to reduce the side lobes, but always at the expense of widening the main lobe and reducing the directivity somewhat:

$$p(\theta) = w_1 p_1(\theta) + w_1 p_2(\theta) + \ldots + w_n p_n(\theta) = p_0(\theta) \sum_{k=1}^{n} w_k e^{j(k-1)\varphi} \tag{2.51}$$

Reducing the directivity means that the source level is reduced as a consequence. In an active system it is the combined, transmit and receive, beam pattern which is important and it is hardly ever desirable to further reduce the side lobes at the expense of source level. A better solution is to adequately shade the receiver elements only.

All shading functions widen the main lobe by some factor. Mathematical procedures may be used to determine the amplitude shading coefficients which yield the narrowest main lobe for a specified level of side lobes ([Wettergren, T.A. et al (2005)]).

Note the similarity between amplitude shading used to reduce bearing side lobes in a beam pattern and pulse shaping used to reduce frequency side lobes in the frequency spectrum of a pulse as will be described next.

2.3.5. Synthetic Arrays

Synthetic arrays are formed when instead of having n physical transducer; a common swath is sampled at n different positions with distances between them d_i with the same transducer. For static scenes, a synthetic array should be equivalent to having a real array with the same number of elements ($d_0 = 0$):

$$p(\theta) = p_0(\theta) + p_0(\theta) e^{j\frac{2\pi d_1}{\lambda}\sin\theta} + \ldots + p_0(\theta) e^{j\frac{2\pi d_{n-1}}{\lambda}\sin\theta} = n p_0 \sum_{k=1}^{n} e^{j\frac{2\pi d_{k-1}}{\lambda}\sin\theta} \tag{2.52}$$

Synthetic arrays can be processed much the same ways as real arrays, by combining the received echoes with the correct delay and shading ([Heering, P. (1984)]). Nevertheless, by doing so, the information available in the echoes is not being correctly explored. It will be shown in the next chapters, that by using specific algorithms, it is possible to obtain an image from a synthetic array that has a resolution independent of the range: a very different result from that obtained using common array techniques where the resolution has a range dependency given by:

$$\delta_{XT} = \sin\theta_{3dB} r \tag{2.53}$$

2.4. Pass-band, Base-band and Analytical Signals

In signal processing, it is common to use the complex signals notion ([Oppenheim A. V. et al (1999)]). In the implementation sense, a complex signal is nothing more than the combination of two real signals (real and imaginary parts). However, the convenience of the complex signal notation in frequency domain analysis has made the complex signal concept an essential ingredient in theoretic signal processing.

Another important ingredient in signal processing is the band-pass (frequency spectrum centered at the carrier frequency) signal concept, and the possibility to model a (real) band-pass system with a complex base-band model (frequency spectrum centered at 0). In this context the complex signal concept becomes mandatory. The signals transmitted by sonar systems are band-pass signals and thus the interest on this representation.

2.4.1. Analytical signals

A signal which has no negative-frequency components is called an analytic signal ([Oppenheim A. V. et al (1999)]). Therefore, in continuous time, every analytic signal $z(t)$ can be represented using only the positive frequency portion of its Fourier transform as:

$$z(t) = \frac{1}{2\pi}\int_0^{\infty} Z(\omega) e^{j\omega t} d\omega \tag{2.54}$$

Where $Z(\omega)$ is the complex coefficient (setting the amplitude and phase) of the positive-frequency complex sinusoid $e^{j\omega t}$ at frequency ω.

Any real sinusoid $A\cos(\omega t + \phi)$ may be converted to a positive-frequency complex sinusoid $Ae^{j(\omega t+\phi)}$ by simply generating a phase-quadrature component $A\sin(\omega t + \phi)$ to serve as the imaginary part:

$$Ae^{j(\omega t+\phi)} = A\cos(\omega t + \phi) + jA\sin(\omega t + \phi) \tag{2.55}$$

The phase-quadrature component can be generated from the in-phase component by a simple quarter-cycle time shift. For more complicated signals which are expressible as a sum of many sinusoids, a filter can be constructed by shifting each sinusoidal component by a quarter of cycle. This is called a Hilbert transform filter. Let $H_t\{x\}$ denote the output at time t of the Hilbert-transform filter applied to the signal x. Ideally, this filter has magnitude 1 at all

frequencies and introduces a phase shift of $-\frac{\pi}{2}$ at each positive frequency and $+\frac{\pi}{2}$ at each negative frequency. When a real signal $x(t)$ and its Hilbert transform $y(t) = H_t\{x\}$ are used to form a new complex signal $z(t) = x(t) + jy(t)$, the signal $z(t)$ is the (complex) analytic signal corresponding to the real signal $x(t)$. In other words, for any real signal $x(t)$, the corresponding analytic signal $z(t) = x(t) + jH_t\{x\}$ has the property that all negative frequencies of $x(t)$ have been taken out.

Consider the positive and negative frequency components at the particular frequency ω_0:

$$\begin{aligned} x_+(t) &= e^{j\omega_0 t} \\ x_-(t) &= e^{-j\omega_0 t} \end{aligned} \tag{2.56}$$

Applying a –90 degrees phase shift to the positive-frequency component (phase shifts can be impressed on a complex sinusoid by multiplying it by $e^{\pm j\frac{\pi}{2}} = \pm j$), and a +90 degrees phase shift to the negative-frequency component:

$$\begin{aligned} y_+(t) &= e^{-j\frac{\pi}{2}} e^{j\omega_0 t} = -je^{j\omega_0 t} \\ y_-(t) &= e^{+j\frac{\pi}{2}} e^{-j\omega_0 t} = je^{-j\omega_0 t} \end{aligned} \tag{2.57}$$

And adding the signals together we get:

$$\begin{aligned} z_+(t) &= x_+(t) + jy_+(t) = e^{j\omega_0 t} - j^2 e^{j\omega_0 t} = 2e^{j\omega_0 t} \\ z_-(t) &= x_-(t) + jy_-(t) = e^{-j\omega_0 t} + j^2 e^{-j\omega_0 t} = 0 \end{aligned} \tag{2.58}$$

With the negative frequency component has been canceled. There is also a gain of 2 at positive frequencies.

In the case of a real sinusoid:

$$x(t) = 2\cos(\omega_0 t) = e^{j\omega_0 t} + e^{-j\omega_0 t} \tag{2.59}$$

Applying the ideal phase shifts, the Hilbert transform is:

$$y(t) = e^{j\omega_0 t - j\frac{\pi}{2}} + e^{-j\omega_0 t + j\frac{\pi}{2}} = -je^{j\omega_0 t} + je^{-j\omega_0 t} = 2\sin(\omega_0 t) \tag{2.60}$$

The analytical signal is then:

$$z(t) = x(t) + jy(t) = 2\cos(\omega_0 t) + j2\sin(\omega_0 t) = 2e^{j\omega_0 t} \tag{2.61}$$

By Euler's identity, in the sum $z(t) = x(t) + jy(t)$, the negative-frequency components of $x(t)$ and $jy(t)$ cancel out, leaving only the positive-frequency component. This happens for any real signal $x(t)$, not just for sinusoids as in this case.

2.4.2. Pass-band and base band signals

Since all the information in a signal transmitted by a sonar is contained in a bandwidth B around a frequency ω_c, the band-pass signal transmitted by a sonar can be described as:

$$s(t) = \mathrm{Re}\left\{s_b(t)e^{j\omega_c t}\right\} = I(t)\cos(\omega_c t) - Q(t)\sin(\omega_c t) \tag{2.62}$$

$I(t)$ and $Q(t)$ are the real and complex components of the base-band version of the signal:

$$s_b(t) = I(t) + jQ(t) \tag{2.63}$$

The band-pass signal can also be written to put in evidence the modulus and phase of its constituting base-band signal as:

$$s(t) = \sqrt{I(t)^2 + Q(t)^2}\cos\left(\omega_c t + \tan^{-1}\left(\frac{-Q(t)}{I(t)}\right)\right) \tag{2.64}$$

To recover the base-band signal from the pass-band signal we can first apply the Hilbert transform to the signal and add to the original version:

$$\begin{aligned} s'(t) = I(t)\cos(\omega_c t) - Q(t)\sin(\omega_c t) + jH\left\{I(t)\cos(\omega_c t) - Q(t)\sin(\omega_c t)\right\} = \\ = \left[I(t) + jQ(t)\right]e^{j\omega_c t} \end{aligned} \tag{2.65}$$

The base-band signal is now obtained by:

$$s_b(t) = s'(t)e^{-j\omega_c t} = I(t) + jQ(t) \tag{2.66}$$

In practice, it is also possible to obtain the base-band signal by multiplying the pass-band signal by $e^{-j\omega_c t}$ and filtering the result, thus removing the cosine and sine terms with $\omega_c t$ dependency.

2.4.3. Sampling of pass-band signals

Considering again the signal transmitted by a sonar:

$$s(t) = I(t)\cos(\omega_c t) - Q(t)\sin(\omega_c t) \tag{2.67}$$

Remember that this signal has bandwidth B and has its spectrum centered at f_c, but let B represent the signal bandwidth plus some guard frequency space that can be adjusted. If the signal is sampled at a sample time give by:

$$T_S = \frac{1}{2B} \tag{2.68}$$

And make:

$$f_c = (2l-1)\frac{B}{2} \tag{2.69}$$

We get:

$$\begin{aligned} &s(T_s n) = I(T_s n)\cos(\omega_c T_s n) - Q(T_s n)\sin(\omega_c T_s n) = \\ &= I(T_s n)\cos\left(\frac{\pi}{2}(2l-1)n\right) - Q(T_s n)\sin\left(\frac{\pi}{2}(2l-1)n\right) \end{aligned} \tag{2.70}$$

Dividing this expression in n odd and even we get:

$$s(T_s n) = \begin{cases} I(T_s n)(-1)^{\frac{n}{2}}, n\ even \\ Q(T_s n)(-1)^{l+\frac{n+1}{2}+1}, n\ odd \end{cases} \tag{2.71}$$

This means that the signal doesn't have to be sampled at $f_s = 2(f_c + B/2)$, instead if it is sampled at a frequency given by the above formulas it is possible to correctly recuperate the in phase and quadrature components of the base-band signal. In reality the in phase signal is

obtained multiplying $s(T_s n)$ by $(-1)^{\frac{n}{2}}$ when n is even and by 0 when is n odd. In the same way, the quadrature component is obtained multiplying $s(T_s n)$ by $(-1)^{l+\frac{n+1}{2}+1}$ when n is odd and by 0 when n is even.

The reason for this can be seen intuitively. When sampling at a lower sampling rate (but chosen to respect the above equation) several spectral replicas of the signal will emerge, one of which will be cantered at DC and thus having its spectral content between $-f_s/2$ and $+f_s/2$. This means that the sampling and down-conversion of a pass-band signal can be done in one step, lowering the sampling frequency requirements and the hardware complexity.

2.5. Pulse Compression

2.5.1. Matched Filters

Pulse compression is an example of matched filtering. Two-dimensional matched filters are commonly used in image processing and will be used throughout to explain how cross-track and along-track target discrimination is achieved.

A matched filter is obtained by correlating a known signal, or template, with an unknown signal to detect the presence of the template in the unknown signal ([Haykin, S. (2000)]). This is equivalent to convolving the unknown signal with a time-reversed version of the template (cross-correlation). The matched filter is the optimal linear filter for maximizing the signal to noise ratio (SNR) in the presence of additive stochastic noise. Matched filters are commonly used in radar and sonar, in which a known signal is sent out, and the reflected signal is examined for common elements of the out-going signal.

The matched filter is the linear filter, h,

$$y(t) = \int_{-\infty}^{+\infty} h(t-\tau)\,x(\tau)\,d\tau = F^{-1}\{H(f)\,X(f)\} \tag{2.72}$$

that maximizes the output signal-to-noise ratio $SNR = |s|^2 \Big/ E\{|v|^2\}$ ([Haykin, S. (2000)]), where s is the filtered signal and v is the noise and E denotes the expected value. In the discrete form we have:

$$y[n] = \sum_{k=-\infty}^{+\infty} h[n-k]\,x[k] \tag{2.73}$$

If $s(t)$ is the transmitted signal and $x(t)=s(t)+n(t)$ is the signal plus noise, the output of the filter is:

$$y(t)=\int_{-\infty}^{+\infty}h(t-\tau)x(\tau)d\tau=\int_{-\infty}^{+\infty}h(t-\tau)s(\tau)d\tau+\int_{-\infty}^{+\infty}h(t-\tau)n(\tau)d\tau=\int_{-\infty}^{+\infty}h(t-\tau)s(\tau)d\tau+v(t) \tag{2.74}$$

The expected value of $v(t)$ is $E\{v(t)\}=0$ and the energy (variance) is:

$$E\{v^2(t)\}=E\left\{\int_{-\infty}^{+\infty}h(t-\tau)n(\tau)d\tau\int_{-\infty}^{+\infty}h(t-\tau')n(\tau')d\tau'\right\}=\frac{N_0}{2}\int_{-\infty}^{+\infty}\left|h(t-\tau)\right|^2d\tau \tag{2.75}$$

N_0 is the noise amplitude level. The signal energy is:

$$E\{s^2(t)\}=\left(\int_{-\infty}^{+\infty}h(t-\tau)s(\tau)d\tau\right)^2 \tag{2.76}$$

And so:

$$SNR=\frac{\left(\int_{-\infty}^{+\infty}h(t-\tau)s(\tau)d\tau\right)^2}{\frac{N_0}{2}\int_{-\infty}^{+\infty}\left|h(t-\tau)\right|^2d\tau} \tag{2.77}$$

Using the Cauchy-Schwarz inequality (* denotes complex conjugate):

$$\left(\int_{-\infty}^{+\infty}g_1(\tau)g_2^*(\tau)d\tau\right)^2\le\int_{-\infty}^{+\infty}\left|g_1(\tau)\right|^2d\tau\int_{-\infty}^{+\infty}\left|g_2(\tau)\right|^2d\tau \tag{2.78}$$

Knowing that the equality is achieved when:

$$g_1(\tau)=\alpha g_2(\tau) \tag{2.79}$$

It is possible to find a maximum limit for the SNR:

$$\frac{\left(\int_{-\infty}^{+\infty} h(t-\tau)s(\tau)d\tau\right)^2}{\frac{N_0}{2}\int_{-\infty}^{+\infty}\left|h(t-\tau)\right|^2 d\tau} = \frac{\int_{-\infty}^{+\infty}\left|h(t-\tau)\right|^2 d\tau \int_{-\infty}^{+\infty}\left|s(t-\tau)\right|^2 d\tau}{\frac{N_0}{2}\int_{-\infty}^{+\infty}\left|h(t-\tau)\right|^2 d\tau} = \frac{N_0}{2}\int_{-\infty}^{+\infty}\left|s(t-\tau)\right|^2 d\tau \tag{2.80}$$

If we know make:

$$h(t) = \alpha s^*(\tau)d\tau \tag{2.81}$$

We get:

$$\frac{\left(\int_{-\infty}^{+\infty} s^*(t-\tau)s(\tau)d\tau\right)^2}{\frac{N_0}{2}\int_{-\infty}^{+\infty}\left|s(t-\tau)\right|^2 d\tau} = \frac{N_0}{2}\int_{-\infty}^{+\infty}\left|s(t-\tau)\right|^2 d\tau \tag{2.82}$$

And so the matched filter is the complex conjugate of the signal that was sent. It is also useful to think of the matched filter in the context of the inner product as it is more intuitive. Lets the output, y , the inner product of the filter and the observed signal such that:

$$y[n] = \sum_{k=-\infty}^{+\infty} h^*[k]x[k] = h^H x = h^H s + h^H v = y_s + y_v \tag{2.83}$$

We can derive the linear filter that maximizes output signal-to-noise ratio by invoking a geometric argument. The intuition behind the matched filter relies on correlating the received signal (a vector) with a filter (another vector) that is parallel with the signal, maximizing the inner product. This enhances the signal. When we consider the additive stochastic noise, we have the additional challenge of minimizing the output due to noise by choosing a filter that is orthogonal to the noise.

Formally defining the problem, we seek a filter, h , such that we maximize the output signal-to-noise ratio, where the output is the inner product of the filter and the observed signal x . Again the observed signal consists of the desirable signal s and additive noise v : $x = s + v$. Defining the covariance matrix of the noise and noting that this matrix has Hermitian symmetry ($A = A^H$, where ^{H} denotes Hermitian transpose):

$$R_v = E\{vv^H\} \tag{2.84}$$

Let us call our output, $y[n]$, the inner product of our filter and the observed signal such that

$$y[n] = \sum_{k=-\infty}^{+\infty} h^*[k]x[k] = h^H x = h^H s + h^H v = y_s + y_v \tag{2.85}$$

We now define the signal-to-noise ratio, which is our objective function, to be the ratio of the power of the output due to the desired signal to the power of the output due to the noise:

$$SNR = \frac{|y_s|^2}{E\left\{|y_v|^2\right\}} = \frac{\left|h^H s\right|^2}{E\left\{\left|h^H v\right|^2\right\}} \tag{2.86}$$

We wish to maximize this quantity by choosing h. Expanding the denominator of our objective function, we have

$$E\left\{h^H v\right\} = E\left\{\left(h^H v\right)\left(h^H v\right)^H\right\} = h^H E\left\{vv^H\right\}h = h^H R_v h \tag{2.87}$$

The *SNR* now becomes:

$$SNR = \frac{\left|h^H s\right|^2}{h^H R_v h} \tag{2.88}$$

Rewriting this expression with some matrix manipulation, exploiting the Hermitian symmetry of the covariance matrix R_v, we obtain:

$$SNR = \frac{\left|\left(R_v^{\frac{1}{2}}h\right)\left(R_v^{\frac{-1}{2}}s\right)\right|^2}{\left(R_v^{\frac{1}{2}}h\right)^H\left(R_v^{\frac{-1}{2}}h\right)} \tag{2.89}$$

To find an upper bound on this expression we again recognize a form of the Cauchy-Schwarz inequality:

$$\left|a^H b\right|^2 \le \left(a^H a\right)\left(b^H b\right) \tag{2.90}$$

This is to say that the square of the inner product of two vectors can only be as large as the product of the individual inner products of the vectors. This concept returns to the intuition behind the matched filter: this upper bound is achieved when the two vectors a and b are parallel.

Expressing the upper bound of the using this geometric inequality above:

$$SNR = \frac{\left|\left(R_v^{\frac{1}{2}}h\right)\left(R_v^{\frac{-1}{2}}s\right)\right|^2}{\left(R_v^{\frac{1}{2}}h\right)^H\left(R_v^{\frac{-1}{2}}h\right)} \leq \frac{\left[\left(R_v^{\frac{1}{2}}h\right)^H\left(R_v^{\frac{1}{2}}h\right)\right]\left[\left(R_v^{\frac{1}{2}}s\right)^H\left(R_v^{\frac{1}{2}}s\right)\right]}{\left(R_v^{\frac{1}{2}}h\right)^H\left(R_v^{\frac{-1}{2}}h\right)} = s^H R_v^{-1} s \tag{2.91}$$

This upper bound can be achieved making,

$$R_v^{\frac{1}{2}}h = \alpha R_v^{\frac{-1}{2}}s \tag{2.92}$$

And thus:

$$h = s^H R_v^{-1} s \tag{2.93}$$

Where α is an arbitrary real number. To prove that this expression for h is the optimal filter, we substitute in the expression for the output SNR reminding that R_v has Hermitian symmetry:

$$SNR = \frac{\left|\left(R_v^{\frac{1}{2}}h\right)\left(R_v^{-\frac{1}{2}}s\right)\right|^2}{\left(R_v^{\frac{1}{2}}h\right)^H\left(R_v^{-\frac{1}{2}}h\right)} = \frac{\left|\left(\alpha R_v^{-\frac{1}{2}}s\right)\left(R_v^{-\frac{1}{2}}s\right)\right|^2}{\left(\alpha R_v^{-\frac{1}{2}}s\right)^H\left(R_v^{-\frac{1}{2}}h\right)} = \frac{\left|s^H R_v^{-1}s\right|^2}{s^H R_v^{-1}s} = s^H R_v^{-1}s \tag{2.94}$$

We often choose to normalize the expected value of the power of the filter output due to the noise to unity. That is, we constrain

$$E\left\{\left|y_v^{\,2}\right|\right\} = 1 \tag{2.95}$$

This constraint implies a value of α, for which we can solve:

$$E\left\{\left|y_v^{\,2}\right|\right\} = \alpha^2 s^H R_v^{-1} s = 1 \tag{2.96}$$

Yielding,

$$\alpha = \frac{1}{\sqrt{s^H R_v^{-1} s}} \tag{2.97}$$

Giving the normalized filter:

$$h = \frac{1}{\sqrt{s^H R_v^{-1} s}} R_v^{-1} s \tag{2.98}$$

If we care to write the impulse response of the filter for the convolution system, it is simply the complex conjugate time reversal of h.

This solution was intuitively expected for both deductions. The output of the filter will be maximum if it corresponds to a correlation, for which the template is the signal itself.

2.5.2. Pulse Signal

Matched filters are often used in signal detection as is the case with sonar. As an example, suppose that we wish to judge the distance of an object by reflecting a signal off it. The simplest signal a pulse sonar can transmit is a sinusoidal pulse of amplitude A and carrier frequency f_0, truncated by a rectangular function of width T. The pulse is transmitted periodically although, without loss of generality, we will consider only a single pulse $s(t)$. If we assume the pulse to start at time $t = 0$, the signal can be written the following way, using the complex notation:

$$s(t) = A\text{rect}\left(\frac{t - T/2}{T}\right) e^{j2\pi f_0 t} \tag{2.99}$$

With the rectangular function defined as:

$$\text{rect}(t) = \begin{cases} 1 & , -T/2 \le t \le T/2 \\ 0 & , t < -T/2 \cup t > T/2 \end{cases} \tag{2.100}$$

And the received signal will be an attenuated and phase-shifted version of the transmitted signal plus added noise. To judge the distance of the object, we correlate the received signal with a matched filter, which, in the case of white (uncorrelated) noise, is another pure-tone sinusoid. When the output of the matched filter system exceeds a certain threshold, we conclude with high probability that the received signal has been reflected off the

object. Using the speed of propagation and the time that the reflected signal is first observed, an estimate of the distance of the object can be made. Changing the shape of the pulse in a specially-designed way, the signal-to-noise ratio and the distance resolution can be even improved after matched filtering: this is a technique known as pulse compression.

Additionally, matched filters can be used in parameter estimation problems. Returning to the previous example, we may desire to estimate the speed of the object, in addition to its position. To exploit the Doppler Effect, we would like to estimate the frequency of the received signal. To do so, we may correlate the received signal with several matched filters of sinusoids at varying frequencies. The matched filter with the highest output will reveal, with high probability, the frequency of the reflected signal and help determine the speed of the object. This method is, in fact, a simple version of the discrete Fourier transform (DFT). The DFT takes an N-valued complex input and correlates it with N matched filters, corresponding to complex exponentials at N different frequencies, to yield N complex-valued numbers corresponding to the relative amplitudes and phases of the sinusoidal.

To determine the range resolution which can be obtained with such a signal, we will consider the attenuated and time-shifted copy of the original transmitted signal (there might also be a Doppler effect but will not be considered here): the return signal, $e(t)$. There is also noise in the incoming signal, both on the imaginary and the real channel, which we will assume to be white and Gaussian (this generally holds in reality); we write $n(t)$ to denote that noise. To detect the incoming signal, matched filtering is used. Like seen before, this method is optimal when a known signal is to be detected among an additive white Gaussian noise. In other words, the cross-correlation of the received signal with the transmitted signal is computed. This comes down to convolving the incoming signal with a conjugated and mirrored version of the transmitted signal. This operation can be done either in software or with hardware. If write $se(t)$ for this cross-correlation, we get:

$$se(t) = \int_0^{+\infty} s^*(\tau) e(t-\tau)\, d\tau \tag{2.101}$$

If the reflected signal comes back to the receiver at time t_v and is attenuated by a factor K, this yields:

$$e(t) = \begin{cases} KAe^{j2\pi f_0 (t-t_v)} & , t_v \le t < t_v + T \\ n(t) & , t \ge t_v + T \end{cases} \tag{2.102}$$

Since we know the transmitted signal, we obtain:

$$se(t) = KA^2\Lambda\left(\frac{t-t_v}{T}\right)e^{j2\pi f_0(t-t_v)} + n'(t) \tag{2.103}$$

The function $\Lambda(t)$ is the triangular shaped pulse defined as:

$$\Lambda(t) = \begin{cases} 2(t+T/2)/T & ,-T/2 \le t \le 0 \\ -2(t-T/2)/T & ,0 \le t \le T/2 \\ 0 & ,t < -T/2 \cup t > T/2 \end{cases} \tag{2.104}$$

The result of the cross-correlation between the noise and the transmitted signal, $n'(t)$, remains a white noise of same characteristics as $n(t)$ since it is not correlated to the transmitted signal.

If two pulses come back (nearly) at the same time, the cross-correlation is equal to the sum of the cross-correlation of the two elementary signals. To be able to distinguish two different triangular envelopes, it is necessary that the times of arrival of the two pulses must be separated by at least T so that the maxima of both pulses can be uniquely identified. If this condition is false, both triangles will be mixed together and impossible to separate.

Since the distance travelled by a wave during T is cT, and since this distance corresponds to a round-trip time, we get that the range resolution with a sinusoidal pulse is:

$$\delta_{XT} = \frac{cT}{2} \tag{2.105}$$

As a conclusion, to augment the resolution, the pulse length must be reduced.

The instantaneous power of the transmitted pulse is $P(t) = |s(t)|^2$. The energy put into that signal is:

$$E = \int_0^{+\infty} P(t)\,dt = A^2T \tag{2.106}$$

Similarly, the energy in the received pulse is $E_r = K^2A^2T$. If σ is the standard deviation of the noise, the signal-to-noise ratio (SNR) at the receiver is:

$$SNR = \frac{E_r}{\sigma} = \frac{K^2A^2T}{\sigma} \tag{2.107}$$

The SNR augments with the pulse duration, if other parameters are fixed. This goes against the resolution requirements, since generally one wants a large resolution.

2.5.3. Chirp Signal

As seen above, it is not possible with a simple modulated rectangular, to obtain simultaneously high resolution and high energy. It is also possible to explore the autocorrelation properties of coded signals to transmitted a long enough signal, but after matched filter presents a sharp peak: hence the name pulse compression.

In radar and sonar applications, linear chirps (linear frequency sweep) are the most typically used signals to achieve pulse compression. The pulse being of finite length, the amplitude is a rectangle function. If the transmitted signal has duration T, begins at $t=0$ and linearly sweeps the frequency band Δf centered on carrier f_0, it can be written:

$$s(t) = A\,rect\left(\frac{t-T/2}{T}\right)e^{j2\pi\left(f_0-\frac{\Delta f}{2}+\frac{\Delta f}{2T}t\right)t} \tag{2.108}$$

The phase of the chirped signal (that is, the argument of the complex exponential), is:

$$\phi(t) = 2\pi\left(f_0-\frac{\Delta f}{2}+\frac{\Delta f}{2T}t\right)t \tag{2.109}$$

Thus the instantaneous frequency is (by definition):

$$f(t) = \frac{1}{2\pi}\frac{d\phi(t)}{dt} = f_0-\frac{\Delta f}{2}+\frac{\Delta f}{T}t \tag{2.110}$$

Which is the intended linear ramp going from $f_0-\Delta f/2$ f0 at $t=0$ to $f_0+\Delta f/2$ at $t=T$.

It can be shown [Hein, A. (2003)] that the cross-correlation function of $s(t)$ with the received $s^r(t)$ signal is:

$$se(t) = \frac{TA}{2}\Lambda\left(\frac{t}{T}\right)\mathrm{sinc}\left(\pi\Delta f t\Lambda\left(\frac{t}{T}\right)\right)e^{j2\pi f_0(t-t_v)}+n'(t) \tag{2.111}$$

The maximum of the cross-correlation function of $s(t)$ is reached at 0. Around 0, this function behaves as the $\mathrm{sinc}(x)$ term. The $-3\,dB$ temporal width of that cardinal sine is approximately $T'=1/\Delta f$. Everything happens as if, after matched filtering, a simple pulse of

duration T' had been used with the corresponding time resolution. For the common values of Δf , T' is smaller than T , thus obtaining the desired pulse compression.

Since the cardinal sine can have strong side lobes, a common practice is to filter the result by a window (Hamming, Hann, etc). In practice, this can be done at the same time as the adapted filtering by multiplying the reference chirp with the filter. The result will be a signal with slightly lower maximum amplitude, but the side lobes will be filtered out, which is more important.

The distance resolution reachable with a linear frequency modulation of a pulse on a bandwidth Δf is:

$$\delta_{XT} = \frac{c}{2\Delta f} \tag{2.112}$$

Ratio $T/T' = T\Delta f$ is the pulse compression ratio. It is generally greater than 1 (usually its value ranges from 20 to 30).

The energy of the signal does not vary during pulse compression. However, it is now located in the main lobe of the cardinal sine, whose width is approximately T'. If P is the power of the signal before compression, and P' the power of the signal after compression, we have:

$$PT = P'T' \tag{2.113}$$

This then yields:

$$P' = T'/PT \tag{2.114}$$

Besides, the power of the noise does not change through intercorrelation since it is not correlated to the transmitted pulse (it is totally random). As a consequence, after pulse compression, the power of the received signal can be considered as being amplified by $T\Delta f$. This additional gain can be injected in the sonar equation.

Although other types of chirp pulses exist (like quadratic), linear chirps provide good correlation properties and low sensitivity to Doppler shifts. ([Collins, T.; Atkins, P. (1998)] and [Collins, T.; Atkins, P. (1999)]). The linear chirp is orthogonal with a chirp that has the same symmetric frequency rate.

2.5.4. Pseudo-Random Sequence Signals

There are other means to modulate the signal. Phase modulation is a commonly used technique; in this case, the pulse is divided in N time slots of duration T/N for which the phase at the origin is chosen according to a pre-established convention and thus occupying a bandwidth of $\Delta f = N/T$. The signal is thus simply given by ([Frerking, M. (1994)]):

$$s(t) = real\left\{e^{j\pi c(t)} e^{j2\pi f_0 t}\right\} = \cos\left(2\pi f_0 t + \pi c(t)\right) \tag{2.115}$$

Where $c(t)$ is the sequence of random integers. The precise way of choosing the sequence of $\{0,1\}$ integers is done according to a pseudo random sequence. Barker codes ([Barker, R. H. (1953)]) are an example of these type of sequences. As with a linear chirp, pulse compression is achieved through intercorrelation. The advantages of the Barker codes are their simplicity (a π de-phasing is a simple sign change), but the pulse compression ratio is lower than in the chirp case. The compression is also very sensitive to frequency changes due to the Doppler Effect if that change is larger than $1/T$.

If similar resolution characteristics when compared to chirp are wanted, the signal should be constituted by the orthogonal codes in phase and quadrate components ([Frerking, M. (1994)]):

$$s(t) = real\left\{\left[c_1(t) + jc_2(t)\right] e^{j2\pi f_0 t}\right\} = c_1(t)\cos\left(2\pi f_0 t\right) - c_2(t)\sin\left(2\pi f_0 t\right) \tag{2.116}$$

Where $c_1(t)$ and $c_2(t)$ are two orthogonal pseudo random codes of the same length (for example Gold codes [Gold, R. (1959)]). The reason for this is that if we use only one of the components the signal will have a spectrum that is symmetrical in relation to the origin. Thus the bandwidth available will not be efficiently used. With a code in the phase and quadrate components of the signal, the bandwidth will no longer be symmetrical and the resolution thus achieved will be $c/2\Delta f$.

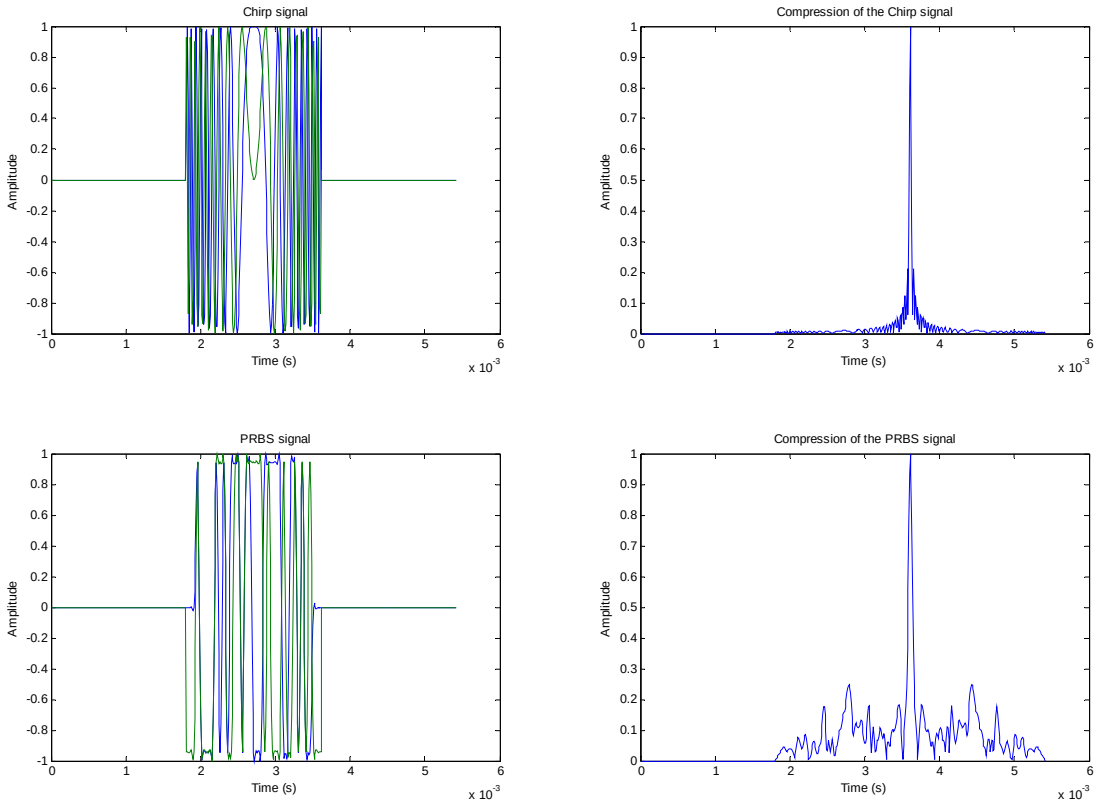

Figure 5: Chirp and pseudo-random pulses.

Pseudo-random sequence pulses provide a mean to extend the set of orthogonal pulses by using the properties of the pseudo random codes. This type of pulses are interesting in sceneries where it is desired (and possible) to transmit a pulse before the echoes of the previous one are completely received or situations where multiple sonar devices are operating with the same frequency (multiple boats).

Figure 5 presents an example of a chirp and pseudo-random phase signal of the same bandwidth (and thus theoretical time resolution) before and after compression.

Nevertheless dynamic range limitations make the simultaneous transmission and reception of orthogonal signals very difficult for today's electronic sonar front end devices.

Chapter 3: System description

3.1. System Overview

The complete system setup ([Silva, S. (2007a)]) is divided into two major subsystems: the sonar device transported by a surface craft and a control base station (Figure 6).

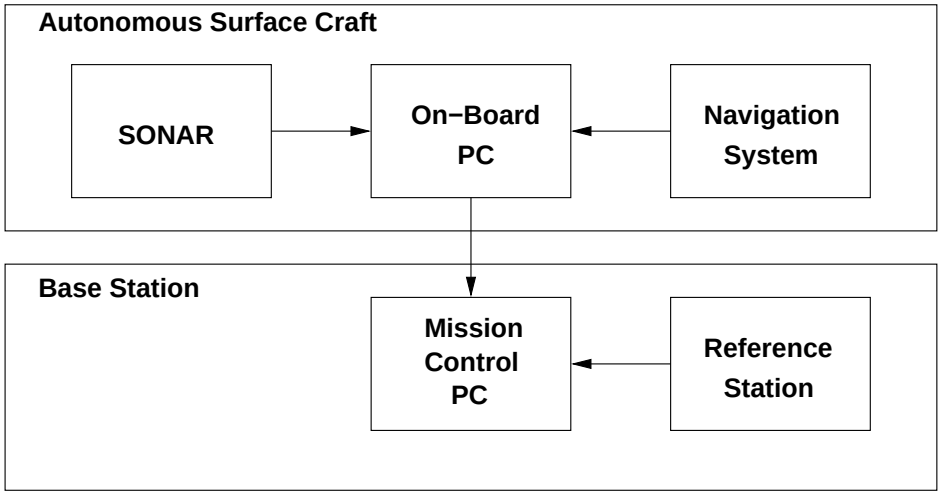

Figure 6: Sonar system overview.

The sonar platform is the autonomous surface craft shown in Figure 7 ([Cruz, N. (2007)]). This is a small catamaran like craft, proving high direction stability, smooth maneuverability and several hours of unmanned operation, which can be manually commanded from the base station or fulfill a pre-defined mission plan. It was built using commonly available components to lower cost and simplify maintenance. It has two independent thrusters for longitudinal and angular motion that provide high maneuverability at low speeds and a maximum speed of 2 m/s. Its size is suitable for, together with the navigation system, executing profiles and other maneuvers with sub-meter accuracy. Because of its size and modular construction, this setup is easily portable and has low deployment time. The boat itself is modular and easy to assemble in site without the need of any special tools. The boat carries three GPS receivers, for position and attitude calculation, a digital compass and an inertial navigation system. The data from the digital compass and inertial system is integrated with the GPS data to provide a refined navigation solution. It also embodies an on-board computer for system control, as well as for acquisition and storage of data.

Figure 7: Autonomous surface craft based synthetic aperture sonar.

The communication to the boat is obtained using an Ethernet radio link (Wi-Fi). The boat and base station antennas were studied as to minimize the effect of the water reflective surface and maximize radio reach which is in the kilometer range.

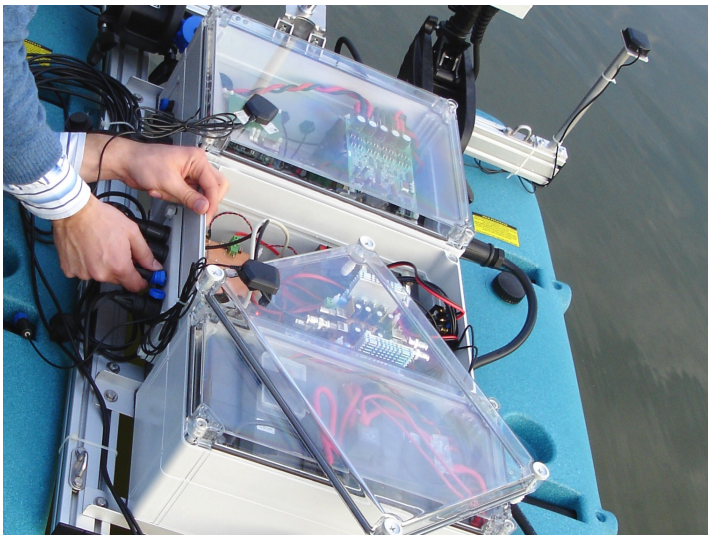

Figure 8: Autonomous craft in preparation for a mission.

The sonar system is carried by the boat as payload (Figure 8) and transducers are placed at the front of the vessel, rigidly coupled to the boats structure. Compared to a solution based on a tow-fish, this sonar platform was mainly chosen because of the possibility to use a GPS positioning system. The velocity and position following is of great importance too. Low frequency errors such as deviations from the desired path or inconsistent platform velocity are in this way limited. Other surface crafts could be used, but it would still have to provide some degree of motion control.

Having the sonar placed at the surface of the water limits its use to shallow water realms such as rivers, lakes, dams, harbors, etc, because of the maximum possible range. These are an important part of the synthetic sonar possible application scenarios. Placing the sonar on a surface craft also makes the system more vulnerable to undesirable heave, roll and pitch motion that has to be measured by the boats navigation system and integrated in the sonar image formation algorithms.

Figure 9: Autonomous craft and base station before a mission.

The base station is constituted by a portable PC that runs the boat and sonar control software, a GPS reference station and a high speed digital radio link to the boat (Figure 9). The purpose of the base station is to control the mission and provide real-time visualization of the sonar data.

Figure 10: Transducer array and support system.

3.2. Sonar system

The synthetic aperture sonar system is itself constituted by two major components: the transducer array (Figure 10) and the digital signal processing system for signal generation and acquisition (Figure 11). The hardware for the sonar systems was designed and developed within the frame of this PhD work.

3.3. Transducer array

The floating platform transports the acoustic transducer array, placed beneath the waterline. The transducer array is formed by two displaced sets of 4 transducers separated by 0.5 m. Although only one transducer is necessary for transmission and reception of the sonar echoes, the vertical arrangement enables interferometric height mapping. The horizontal set of transducers is for future use where image enhancement through the use of micro-navigation techniques will be studied. Each individual set can be regulated for an angle suitable for the illumination of the swath at the expected underwater surface depth.

The transducers operate at a centre frequency of 200kHz, corresponding to a wavelength of 0.75cm. As appropriate for synthetic aperture operations, their real aperture is large (approximately 18 degrees), but have a strong front-to-back lobe attenuation ratio which is fundamental to minimize the reflections on the boat structure and water surface.

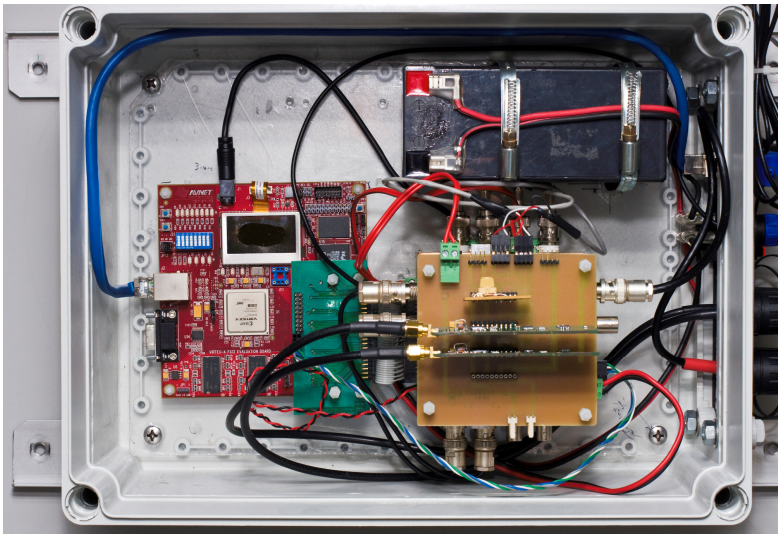

Figure 11: Sonar signal generation and acquisition system.

The effective transducer diameter is 5cm, which allows for synthetic images with this order of magnitude of resolution in the along-track direction.

The usable bandwidth of the transducers (and of the signals employed) is explored thought the use of amplitude and phase compensation to obtain the highest possible range resolution from the system, thus enabling the use of bandwidths of 20 to 40 kHz with the current signal acquisition configuration.

3.4. Signal Generation and Acquisition System

The sonar signal acquisition and generation system is made up of four principal components: the digital processing and control system, the power amplifier, the low noise amplifier, and a GPS receiver for time reference (Figure 12).

This system is tailored for a high resolution interferometric synthetic aperture sonar and so special attention was paid to the different analogue components in terms of bandwidth, amplitude response and phase linearity.

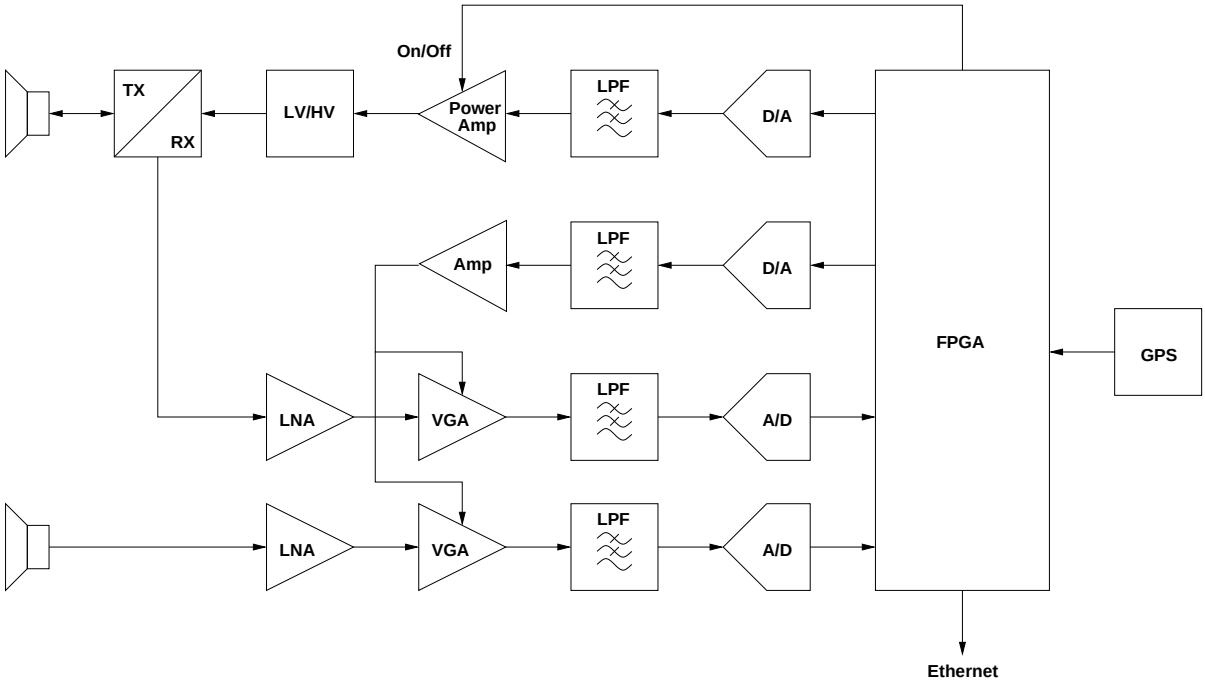

Figure 12: Sonar schematic system overview.

3.4.1. Linear PA

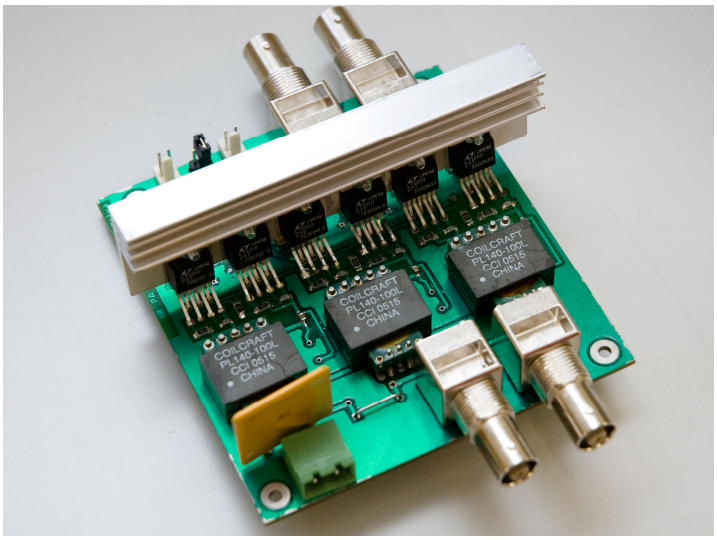

Figure 13: Linear PA

The power-amplifier (PA) is a linear amplifier (Figure 13) which can have a bandwidth as high as 10 MHz, and has a continuous output power rate of 50 W (RMS). A trade-of was made between output power and bandwidth. Because we are interested in using our system in shallow waters, an output power of 50 W was found to be adequate for this 200 kHz sonar system. In exchange it was possible to build a power-amplifier with very flat amplitude and linear phase response in addition to low distortion (THD < -60 dB) in the spectral band of interest. The power amplifier only operates through the transmission period: this saves energy and lowers the electric noise levels during echo reception. Traditionally, switch amplifier are use to drive the transducers. For high frequency sonar systems, pulse width modulation or sigma-delta modulation require very high switching speeds that are difficult to obtain with the output stage of an output amplifier. If only phase modulated signals are used, a switching amplifier can be used with lower switching speed requirements. A linear amplifier design enables the use of amplitude modulated signals and typically provides low noise and spectral interference. The transmitted signal can be of any kind: linear chirp, logarithmic chirp or pseudo-random sequence. It can also have any suitable windowing function applied. At this moment, a linear chirp of 30 kHz bandwidth is being used for system tests. The system is prepared to output a pulse rate between 1 to 15 Hz, because this system is meant to ne used in shallow water surveys, short ranges enable higher pulse rates. Nevertheless, because of the broad beam-width of the transducer and its effective diameter, a pulse rate of 15 Hz still imposes very low moving speeds to the autonomous craft (0.16 m/s) that in turn has to couple with higher precision trajectory tracking constrains.

3.4.2. Low-Noise Amplifier

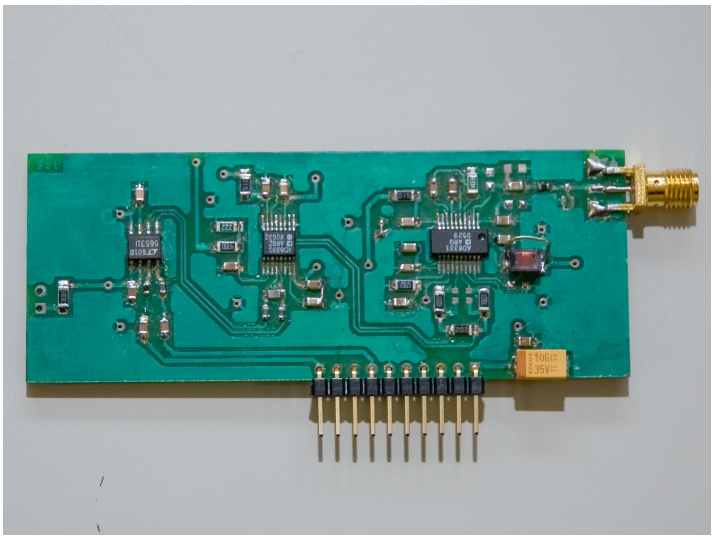

Figure 14: Low-Noise Amplifier.

Each receiving channel has a low noise amplifier and a controllable gain amplifier to handle the high dynamic range of echo signals (Figure 14). The pre-amplifier system has a bandwidth of 10 MHz, a noise figure of 2 dB and a 50-115 dB gain range. It can also recover rapidly from overdrive caused by the transmitting pulse.

To reduced aliasing, a linear phase analogue filter is used. This anti-aliasing filter is interchangeable and can have a cut-off frequency of, for example, 650 kHz or 2.3 MHz.

3.4.3. Digital-to-Analog/Analog-to-Digital Conversion

The conversion between the digital domain and analog domain is made by a high speed digital-to-analog and analog-to-digital conversion board with two output channels and two input channels (Figure 15).

Both the D/A and A/D have 12 bit resolution and are capable of 200 MSamples/s and 40 MSamples/s respectively. The 12 bit resolution of the A/D (effective 72 dB SNR) was found to be high enough to cope with the necessary front-end performance that is dominated by the transducers noise.

This board is capable of full duplex operation although the sonar operates typically in pulsed mode, switching between the generation mode and the acquisition mode. Also for this reason, the dynamic range necessary is moderate and can be compensated by the low-noise amplifier gain control.

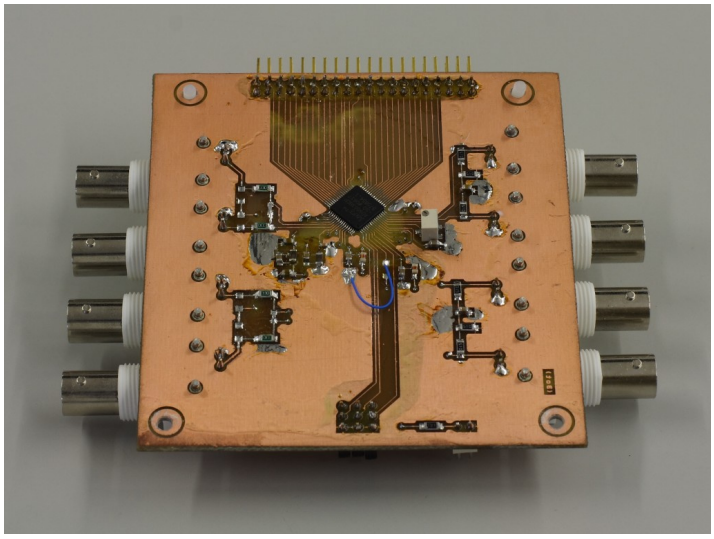

Figure 15: Digital-to-Analog/Analog-to-Digital Conversion Board.

3.4.4. Time reference system

A GPS board provides the time reference for the Sonar (Figure 16). This is constituted by a simple L1 GPS receiver that outputs the GPS time in binary format through a serial port together with a pulse-per-second signal (PPS). The binary time information enables a large correction to the clock while the pulse signal enables fine adjustment of the disciplined clock frequency. It also eliminates residual random delays ([Silva, S. et al (2001)]) in the communication path of the time information (processing, parsing and serial interface).

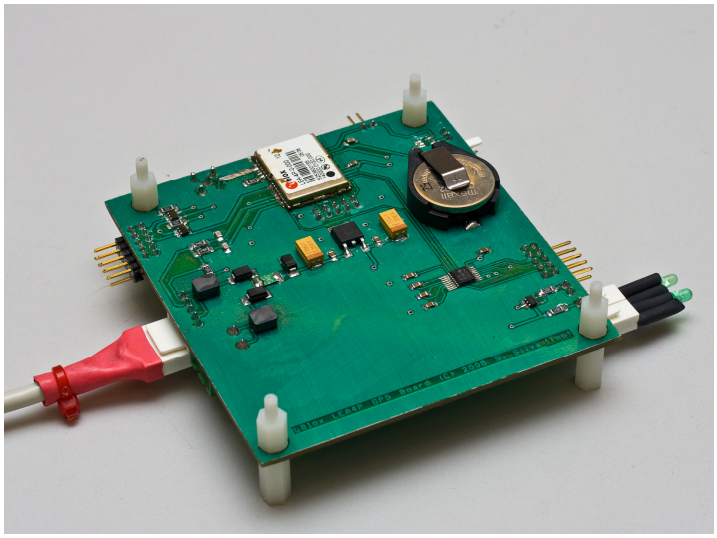

Figure 16: GPS Time reference board.

3.4.5. Base-band conversion and data storage

The sonar signal processing system uses a direct to digital system architecture ([Mitola, J. (1997)]). This means that only the front-end elements of the sonar, like the power amplifier and the low-noise amplifier, use analogue electronics and all other functions such as signal generation, frequency down/up-conversion, filtering and demodulation are performed in the digital domain.

This greatly simplifies the electronic hardware at the cost of more complex digital signal processing. But these in turn are powerfully performed in a field programmed array system (FPGA), providing a low cost solution.

A FPGA is a device that enables the implementation of customized digital circuits. This device can perform several tasks in parallel and at high speed. The system complexity stands, therefore, in the digital domain, enabling more flexible and higher quality signal acquisition and processing through this implementation. The use of this technology results in a low power consumption system that fits a small box, compatible with the autonomous surface craft both in size and energy consumption.

The FPGA system is responsible for the generation of complex acoustic signals, for controlling the transmitting power amplifiers and the adaptive gain low noise receiving amplifiers, for demodulating and for match filtering the received signals with the transmitted waveform (Figure 17).

The first task of the FPGA is to control the digital-to-analog (D/A) and analog-to-digital interface (A/D). At each transmission pulse trigger, the FPGA reads the signal wave form from memory (that is stored in base-band, decimated form), converts the signal to pass-band (frequency up conversion), interpolates and filters the signal, finally supplying it to the D/A.

During the receiver stage, the samples are read from the A/D, converted to base-band (frequency down-conversion), filtered, decimated and finally placed in a FIFO memory.

Each transmit pulse is time-stamped using a real-time clock implemented in the FPGA system that is corrected using the time information and pulse-per-second trigger from a GPS receiver. This enables precise correlation with the navigation data.

Interpolation, decimation and filtering is done using CIC (Cascaded Integrate and Comb) filters, which are a class of FIR (Finite Impulse Response) filters whose coefficients are all 1 or -1, therefore providing a simple mean of implementation in a FPGA ([Hogenauer, E. B. (1981)]).

The processor embedded in the FPGA bridges the low level hardware and the control PC, providing an Ethernet access to the sonar for data readout and issue of command instructions.

The results are then supplied to an embedded computer for storage and acoustic image computation. The base station at a nearly shore accesses this data through a high-speed digital radio-link making the surveyed data available.

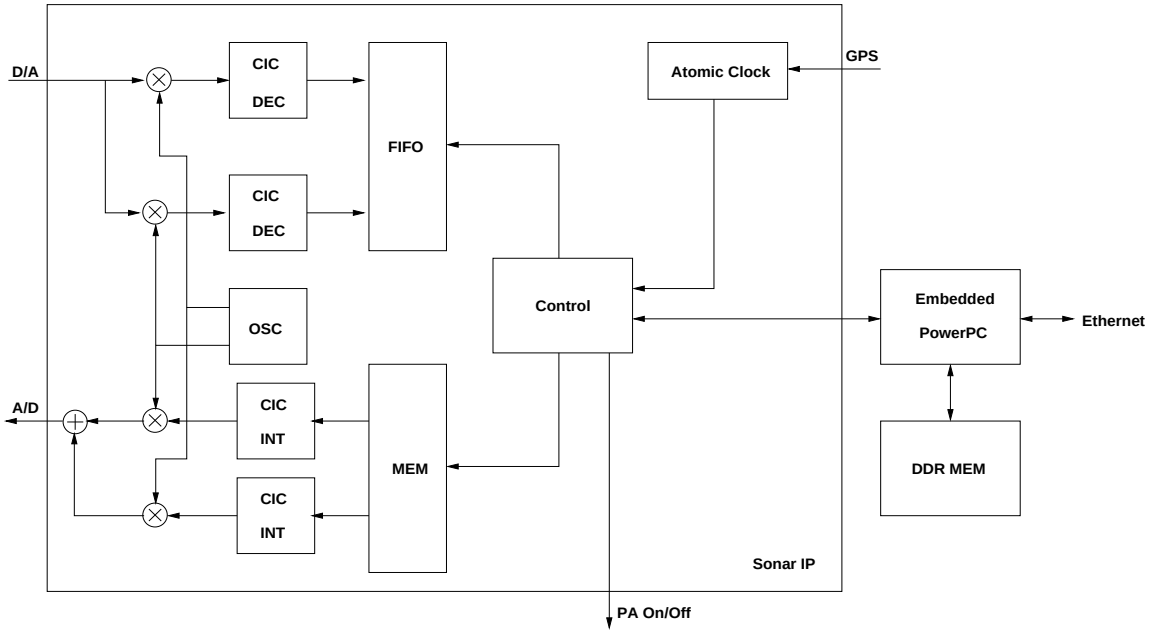

Figure 17: FGPA system detail.

3.5. Navigation system

3.5.1. Overview

The quality of the images depends heavily on the ability to correct for the sonar motion. The followed approach is to compensate for these with information from a high accuracy GPS-IMU navigation system.

In detail, the navigation system is composed by the following elements (Figure 18): an IMU (Inertial Measurement Unit) with three gyroscopes and three accelerometers; a digital magnetic compass; two low cost L1 GPS receivers capable of supplying carrier phase measurements, and a L1+L2 GPS system capable of real time kinematic carrier phase based differential processing (RTK-DGPS). An L1+L2 geodesy grade GPS receiver installed in a static nearby position operates as a reference station.

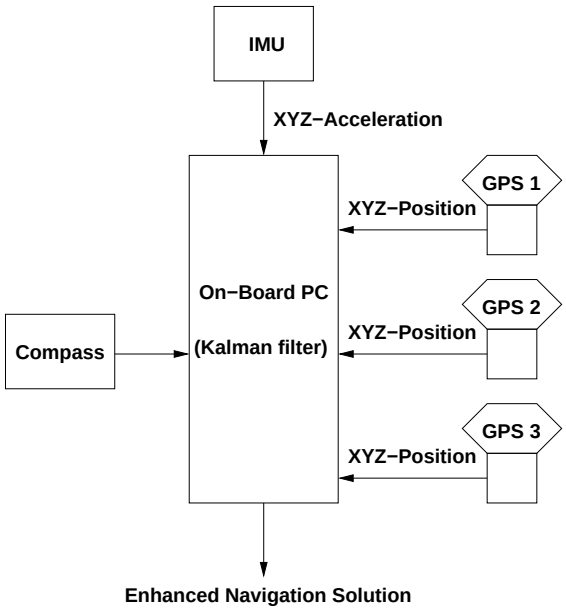

Figure 18: Autonomous vehicle navigation sensor integration schematic.

The receiver installed on the boat is capable of receiving the differential corrections from the reference station and supplying centimeter level accuracy position estimates through processing of this information.

All data from these devices are logged for post-processing, besides being used in real-time. The carrier phase measurements of the two auxiliary GPS receivers are processed in differential mode to obtain absolute estimates of the boat heading. These are used to calibrate the compass. The inertial measurements are integrated and combined with the L1+L2 GPS positions and calibrated compass measurements in a Kalman filter to produce a full position and attitude navigation solution. The independent heading estimates are needed due to the fact that heading errors are loosely coupled to absolute positions when most of the vehicle motion consists of straight lines (with low levels of horizontal acceleration). The accuracy of the navigation system is better than 1cm in position, 0.02º in roll and pitch and 0.05º in heading. As the synthetic aperture processing depends on the displacement information of the received sonar echoes, the sonar motion influences its performance considerably.

3.5.2. Navigation data fusion

Starting with the angular accelerations ($\left[\Delta\theta^i \quad \Delta\omega^i \quad \Delta\kappa^i\right]^T$) given by the IMU, the attitude of the vehicle at the instant i is given by:

$$\begin{bmatrix}\theta^i\\ \omega^i\\ \kappa^i\end{bmatrix}=\begin{bmatrix}\theta^{i-1}\\ \omega^{i-1}\\ \kappa^{i-1}\end{bmatrix}+M\left(\begin{bmatrix}\theta^{i-1} & \omega^{i-1} & \kappa^{i-1}\end{bmatrix}^T\right)\left(\begin{bmatrix}\Delta\theta^i\\ \Delta\omega^i\\ \Delta\kappa^i\end{bmatrix}-\begin{bmatrix}\Delta\theta_B^i\\ \Delta\omega_B^i\\ \Delta\kappa_B^i\end{bmatrix}\right) \tag{3.1}$$

The attitude vector is roll, pitch and yaw ($\begin{bmatrix}\theta^i & \omega^i & \kappa^i\end{bmatrix}^T$). The attitude is thus given by the previous attitude value plus the angular change measured by the IMU, corrected by the estimated gyro bias ($\begin{bmatrix}\Delta\theta_B^i & \Delta\omega_B^i & \Delta\kappa_B^i\end{bmatrix}^T$) and translated to the ECEF referential (Earth Fixed Earth Centered) through a matrix function of the previous attitude $M\left(\begin{bmatrix}\theta^{i-1} & \omega^{i-1} & \kappa^{i-1}\end{bmatrix}^T\right)$ vector ([Farrell, J. L., Jr. (1976); Kuipers, J. B. (1999)]).

The vehicle accelerations ($\begin{bmatrix}a_x^i & a_y^i & a_z^i\end{bmatrix}^T$) in the ECEF referential are given by the measured linear accelerations by the IMU ($\begin{bmatrix}a_{B_x}^i & a_{B_y}^i & a_{B_z}^i\end{bmatrix}^T$) corrected by the bias gravitational acceleration (g) and affected by the rotation matrix ($R^{-1}\left(\theta^i,\omega^i,\kappa^i\right)$) which is a function of the vehicle attitude ([Farrell, J. L., Jr. (1976)]):

$$\begin{bmatrix}a_x^i\\ a_y^i\\ a_z^i\end{bmatrix}=R^{-1}\left(\theta^i,\omega^i,\kappa^i\right)\left(\begin{bmatrix}\hat{a}_x^i\\ \hat{a}_y^i\\ \hat{a}_z^i\end{bmatrix}-\begin{bmatrix}a_{B_x}^i\\ a_{B_y}^i\\ a_{B_z}^i\end{bmatrix}\right)-\begin{bmatrix}0\\ 0\\ g\end{bmatrix} \tag{3.2}$$

And the velocity vector ($\begin{bmatrix}v_x^i & v_y^i & v_z^i\end{bmatrix}^T$) is just the accumulation of the IMU correct acceleration measurements multiplied by the elapsed time Δt:

$$\begin{bmatrix}v_x^i\\ v_y^i\\ v_z^i\end{bmatrix}=\begin{bmatrix}v_x^{i-1}\\ v_y^{i-1}\\ v_z^{i-1}\end{bmatrix}+\begin{bmatrix}a_x^i\\ a_y^i\\ a_z^i\end{bmatrix}\Delta t \tag{3.3}$$

Finally, with the velocity vector it is possible to obtain the vehicle positions as ($\begin{bmatrix}x^i & y^i & z^i\end{bmatrix}^T$) given by:

$$\begin{bmatrix}x^i\\ y^i\\ z^i\end{bmatrix}=\begin{bmatrix}x^{i-1}\\ y^{i-1}\\ z^{i-1}\end{bmatrix}+\begin{bmatrix}v_x^i\\ v_y^i\\ v_z^i\end{bmatrix}\Delta t \tag{3.4}$$

So in this way the measurements of the IMU can be used to calculate the relative position of the vehicle (Figure 19).

Nevertheless, the IMU accelerometer and gyroscope biases are random variables with time the values of which have to be estimated. Failing to do so, results in errors that accumulate rapidly due to the double integration.

To estimate the IMU bias values, a Kalman filter is usually used since it enables real-time operation ([Berdjag, D.; Pomorski, D. (2004)] and [Welch, G.; Bishop, G. (2006)]). Another more simple solution is possible when working with offline data: adjust the bias estimates so the error between the calculated positions and the positions given by the GPS is minimized. This concept will be further explored in the chapter about auto-focus of sonar images.

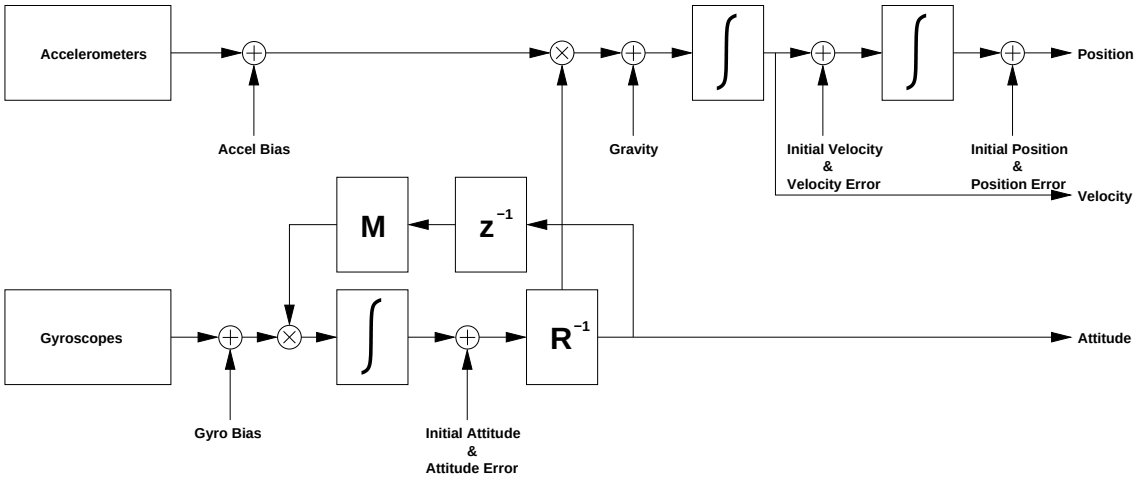

Figure 19: Navigation dynamics flow diagram.

Using the position information obtained from the DGPS system and the heading data from the digital compass (that is calibrated using the heading information also from the GPS system), the bias of the IMU are calculated using the Kalman filter (Figure 20). The filter is constrained assuming that the boat has a mean constant pitch and roll, the mean height is constant and the mean velocity in steady state is the one set by the user. By using this approach the corrected IMU measurements give the high frequency components of the vehicle position and attitude with the low frequency components being given by the DGPS system and digital compass.

Since the relation between the accelerations and vehicle position is non-linear (because of the rotation matrix), the extended version of the Kalman filter has to be used. The time update (state and error covariance projection ahead) equations are then:

$$\begin{cases} \hat{x}_k^- = f\left(\hat{x}_{k-1}, u_{k-1}\right) \\ P_k^- = A_k P_{k-1} A_k^T + W_k Q_{k-1} W_k^T \end{cases} \tag{3.5}$$

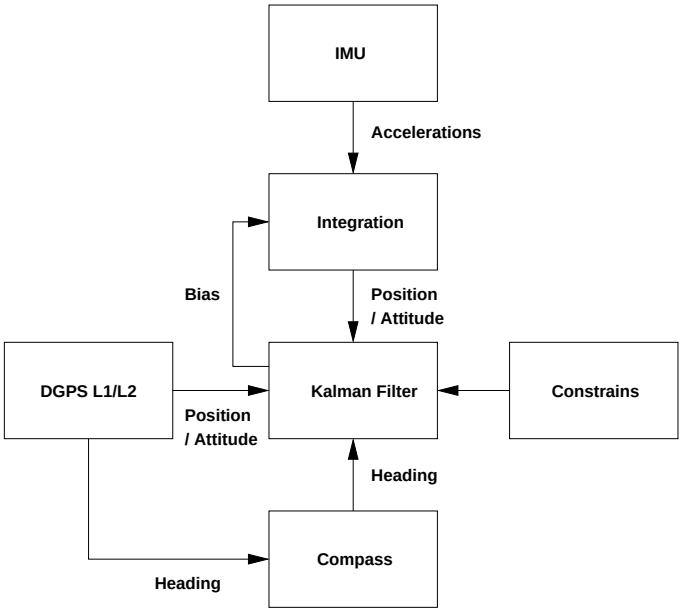

Figure 20: Kalman filter flow diagram.

The vector $\hat{x}_k$ is the concatenation of the attitude, velocity and position error vectors plus the IMU sensor biases, while u_k are the concatenated measured accelerations and angular velocities changes. The function $f\left(\hat{x}_k, u_k\right)$ calculates the vehicle position and attitude as described above. Since this function is non-linear the error covariance P_k^- is calculated using the Jacobian matrix A_k of partial derivatives of f with respect to x and the Jacobian matrix W_k of partial derivatives of f with respect to the measurement noise w. Q is the noise covariance matrix.

The measurement update (Kalman gain K_k, measurement corrected estimation $\hat{x}_k$ and error covariance P_k) equations are:

$$\begin{cases} K_k = P_k^- H_k^T \left(H_k P_k^- H_k^T + V_k R_k V_k^T\right)^{-1} \\ \hat{x}_k = \hat{x}_k^- + K_k \left(z_k - h\left(\hat{x}_k^-, 0\right)\right) \\ P_k = \left(I - K_k H_k\right) P_k^- \end{cases} \tag{3.6}$$

The observation function is in this case selects the position and attitude elements from the vector $\hat{x}_k^-$ that are compared to the measurements z_k and enforces the chosen constrains. H_k is the Jacobian matrix of the observation function with respect to x. The vector z_k is the concatenation of the DGPS positions and digital compass heading.

3.6. Control system

To achieve the best results, the boat trajectories must be executed with minimal deviation from the predefined paths at constant speed over ground and with minimal roll and pitch motions to allow for good correlation characteristics in surveillance and monitoring missions. The speed must be kept below the maximum survey velocity, even when facing river currents, in order not to create aliasing artifacts in the reconstructed image.

To accomplish these goals, a control system that automatically drives the vehicle along user specified trajectories at given speeds was developed (Figure 21). The control system is organized in two independent control loops, one for the velocity and the other for the horizontal position of the vehicle.

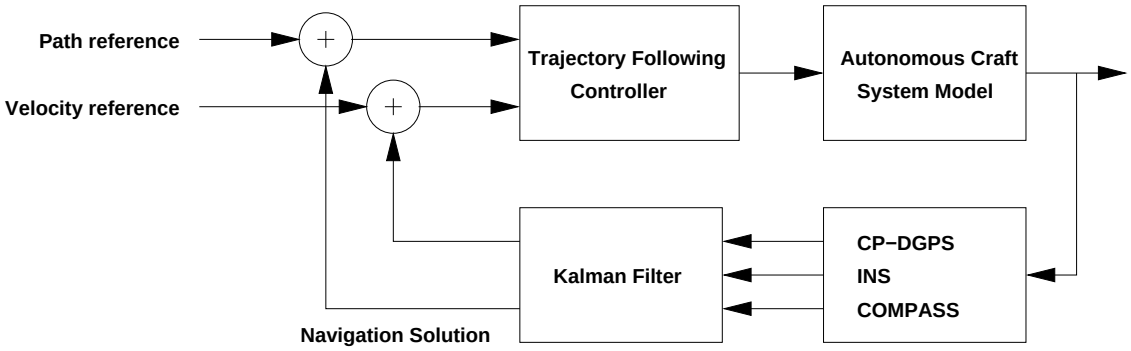

Figure 21: Integration between the navigation system and motion control system.

The velocity loop determines the common mode actuation of the boat while the horizontal plane loop determines the differential mode actuation. These values are then combined to produce the commands for the starboard and port thrusters.

The velocity loop is based on a proportional plus integral controller that assures that the velocity of the boat is, in steady state, the one defined by the user. The controller parameters are tuned to assure a smooth motion by rejecting high frequency noise from the navigation sensors. Different controller parameters can be used to obtain the best behavior at different velocity ranges.

The horizontal plane loop implements a line tracking algorithm. It is composed by an outer controller that computes a heading reference based on the cross track error (distance of

the boat to the desired straight line) and an inner controller that drives the vehicle heading to the given reference. This two stage control loop assures zero cross track error in steady state regardless of the water current. The tuning of the controller parameters takes a dynamic model of the boat into account. Drift due to currents and boat model imperfections are estimated inside these control loops ([Cruz, N. (2007)]).

Relevant to the performance of the automatic control system of the boat, and in general to its behavior, is the quality of the real time estimates of the vehicle position, attitude and velocities. The whole system has already been tested in operational scenarios with great success. In particular, it has been observed that the vehicle follows intended trajectories with average cross tracking errors below 10cm when following 50m lines and a velocity error less than 10% for a target velocity of 0.5 m/s (in calm river waters).

Figure 22 presents a histogram of the cross tracking error for a 50 m straight line described by the autonomous boat while collecting SAS data.

In Figure 23 it is illustrated the path following ability of the autonomous boat, where it was programmed to execute two parallel profiles with a distance of 1 m of each other.

The autonomous surface craft presents an adequate navigation and motion controlled platform for the development of this synthetic aperture sonar system.

Figure 24 shows the boat velocity graph, measured by the navigation system, during the execution of a profile.

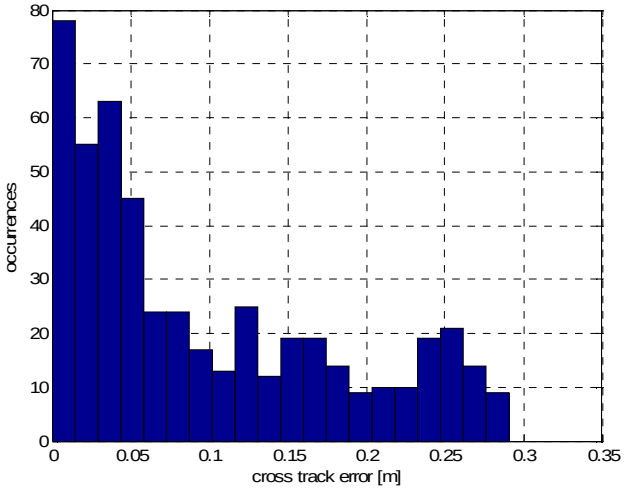

Figure 22: Line tracking performance.

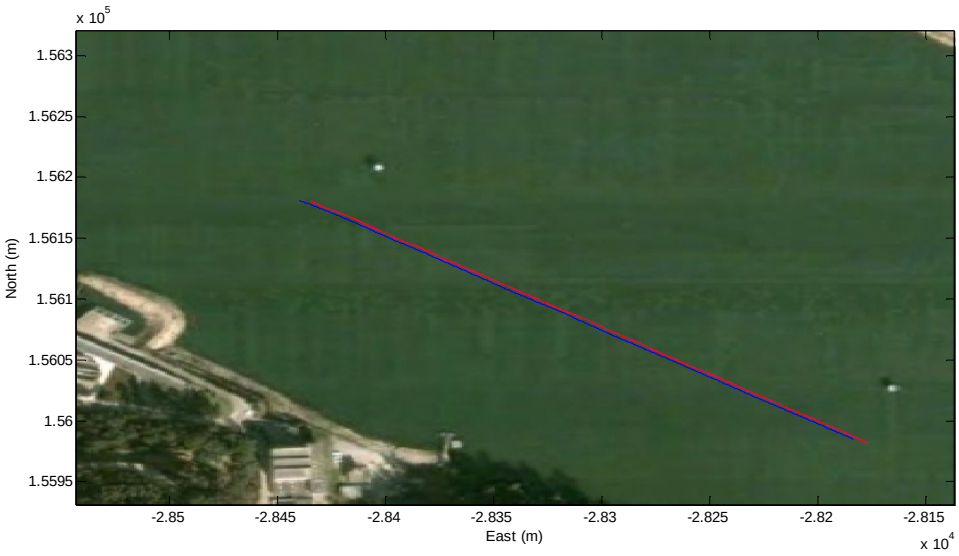

Figure 23: Example path.

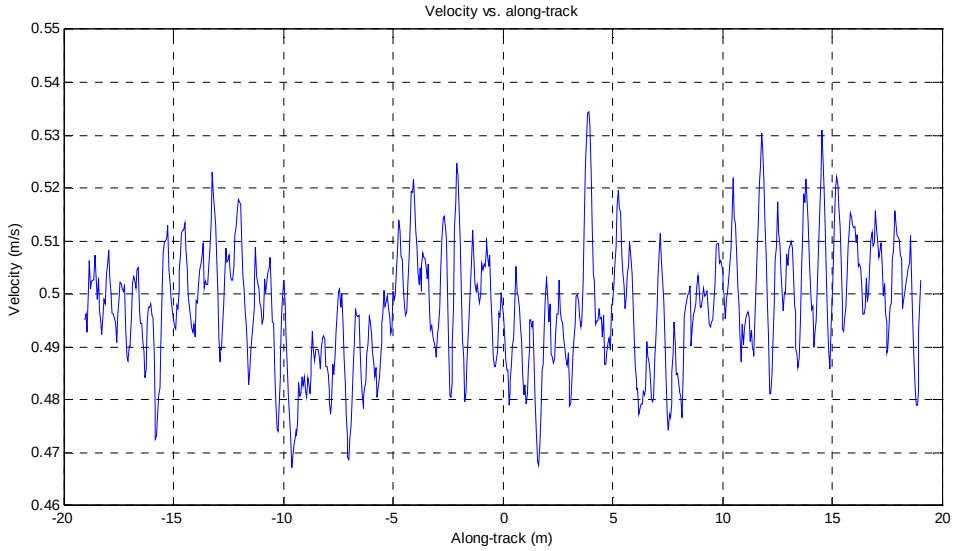

Figure 24: Autonomous surface craft velocity during a mission.

Chapter 4: Sonar model

4.1. System geometry

Side looking sonar is an imaging technique that provides 2D reflectance maps of the acoustic backscatter energy from the submerged floor. These maps are characterized by an intensity variation that is proportional to the acoustic backscattered signal strength. Reconstructing the image from this reflection maps and explore the data as to achieve the best possible resolution requires a good knowledge of the imaging process model.

For this system, the sonar transducers are located beneath the boat pontoons, rigidly connected to its structure. The boat is programmed to follow a series of straight lines. As the platform moves with velocity v, the projector transducer ensonifies strips on the floor at each sampling instant (τ), corresponding to the sonar position $\left[x(\tau) \quad y(\tau) \quad z(\tau)\right]$, and the hydrophone transducer listens for the echoes of the targets (Figure 25). Each echo contains undistinguished information of an area corresponding to the beam pattern ensonified area. This is known as a strip-map configuration. Track and across-track differentiation is obtained by the use of matched filters.

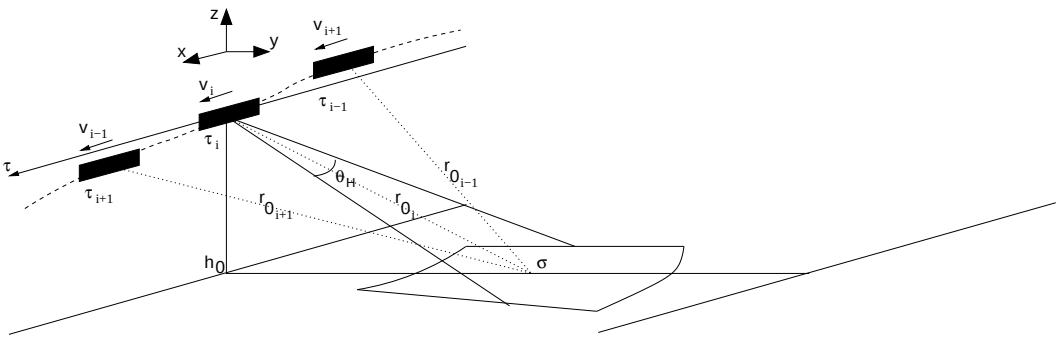

Figure 25: System geometry.

In Figure 25, r_{0_i} is the slant range distance at each sampling position for a target σ in the center of the ensonified swath. The transducers are set so that a swath of necessary size in range is covered (Figure 26). The swath has the dimensions given by the transducer horizontal (θ_V) and vertical (θ_H) beam widths for a certain height above the floor (h_0) and elevation angle (θ_E). The echoes are obtained from r_{min} to r_{max}. Bellow r_{min} the sonar is not recording data (Nadir). Above r_{max} it is assumed that the echoes don't have enough strength and can be

ignored. Since the echoes are obtained as function of range and along-track displacement, the image coordinate system can be by denoted by (x, y_s), where y_s is the slant range given by:

$$y_s = \sqrt{y^2 + \left(z - z(\tau)\right)^2} \tag{4.1}$$

The angle θ_G is given by:

$$\theta_G = \tan^{-1}\left(\frac{z - z(\tau)}{y}\right) \tag{4.2}$$

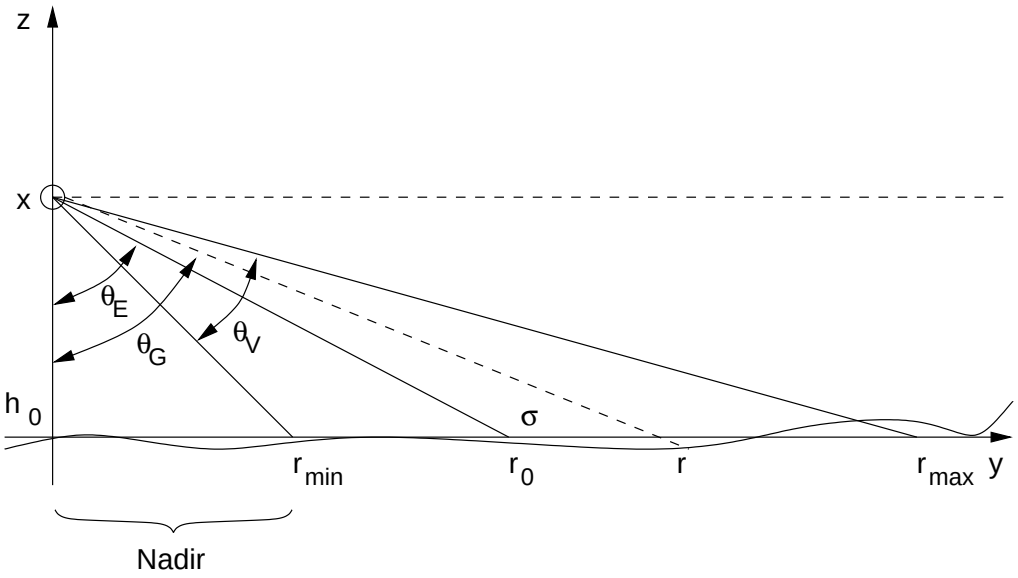

Figure 26: System geometry (side view).

The complex reflectivity distribution of the target field is given by $ff(x, y_s)$ and is obtained from the 3D reflectivity map as:

$$ff(x, y_s) = \int_{-\pi}^{+\pi} ff\left(x, y_s, z(\tau) + y_s \tan\theta_G\right) d\theta_G \tag{4.3}$$

Reconstruction of an image without correcting for the ground-plane mapping causes range variant image compression and shifting. Reconstruction onto the original spatial coordinates is only possible if an elevation map is known or estimated by means of a bathymetry method ([Banks et al (2001)]).

The 2D reflectivity maps can be described by a complex function with the along-track (x) and slant-range (y_s) dimensions as variables:

$$ff(x, y_s) = \left| ff(x, y_s) \right| e^{\phi\phi(x, y_s)} \tag{4.4}$$

The function $\phi\phi(x, y_s)$ is a highly random phase function and its Fourier transform is extremely broad and stretches to all points of the spectrum. In the frequency domain, $ff(x, y_s)$ is the convolution of $\left| ff(x, y_s) \right|$ with $e^{\phi\phi(x, y_s)}$. This means that $\left| ff(x, y_s) \right|$ gets modulated to all parts of the frequency spectrum. For this reason, any window of the complete Fourier transform domain is able to reconstruct the magnitude of the reflectivity map to the resolution determined by the size and shape of the window.

The time delay of each echo is a measure of the distance to the target. Since the platform is moving during the reception of the pulses, the total distance a pulse travels is:

$$r(\tau, x, y_s) = \left\| \sigma - T_{TX}(\tau) \right\| + \left\| T_{RX}(\tau) - \sigma \right\| \tag{4.5}$$

Where $\sigma = [x \quad y \quad z]$ is the vector containing the target coordinates, T_{TX} is the position of the sonar during the transmission and T_{RX} the position of the sonar during the reception of the echoes that for the simple linear trajectory case is ($u = v\tau$):

$$T = [u \quad 0 \quad 0] \tag{4.6}$$

And so, the time of flight is related to the range by:

$$t_f(\tau, x, y_s) = \frac{2}{c} r(\tau, x, y_s) \tag{4.7}$$

If we assume that during the transmission and reception the sonar is essentially static (stop-and-hop approximation), the above expression for the range simplifies to:

$$r(\tau, x, y, z) = 2 \left\| \sigma - T(\tau) \right\| = \frac{2}{c} \sqrt{(x - v\tau)^2 + y_s^2} \tag{4.8}$$

$T(\tau)$ is assumed to be the position at the instant of transmission (mono-static). This approximation assumes that the displacement that occurs during the transmission and reception time is negligible. This is the common assumption that is good enough for SAR data models, and is usually, but not always, sufficient for SAS.

Stop-and-hop approximation decouples the time of flight from the along-track position parameter. In reality this approximation is only valid for small vehicle velocities when

compared to the propagation speed. The movement of the sonar during the transmission and reception causes a temporal Doppler shift and a difference between receiver and transmit positions. The received echoes can be thought of having a frequency contraction for the returns forward the sonar and an expansion for the targets behind the sonar. Ignoring this effect results in a slight geometrical error (image skew) and some minor blurring ([Bonifant, W. W. (1999)]). The continuous movement of the sonar can be modeled and corrected for as explained in [Hayes, M. P. (1989)].

In practice, however, the sonar platform motion has perturbations in its motion, and so the sonar model must account for the irregular along-track sampling positions. Therefore the polar coordinate model must contain a broader expression for the echo travel time with the estimated position of the transducers and a rough estimation of the bottom height (a flat bottom is enough for a first approximation):

$$t_f(\tau, x, y, z) = \frac{2}{c}\sqrt{\left(x - x(\tau)\right)^2 + \left(y - y(\tau)\right)^2 + \left(z - z(\tau)\right)^2} \tag{4.9}$$

Moreover, different transducer geometries mean that the travel time will not be exactly the simple double of the distance to reach the target (Figure 27), but a function of the distance from the transmitting transducer to the target and back to the receiver.

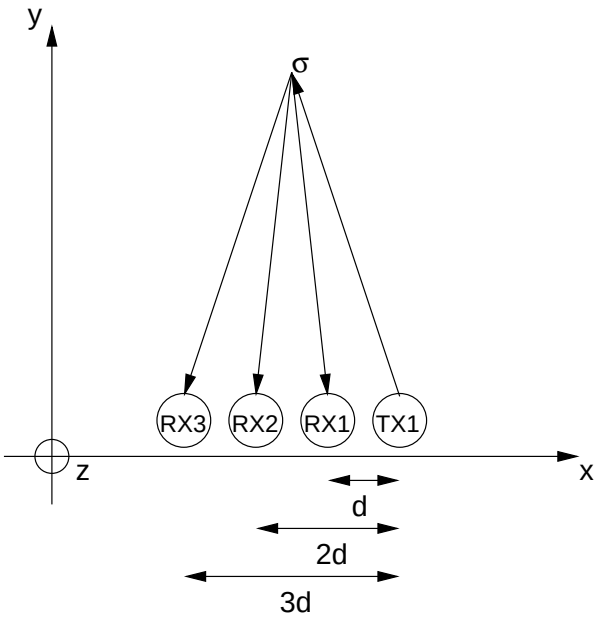

Figure 27: Multiple receiver geometry.

For this, an approximation can be made and the transmitter and receiver treated as if it were a single co-located transducer placed midway between the two. By doing this a multiple

receiver array can be modeled as a single-receiver sonar taking samples at the positions given by the phase center. This implies a correction for each along-track sample of the received echo data by half of the distance from the transmitter to the receiver ($kd/2, k = 1, 2, 3\ldots$) by phase adjustment or interpolation.

An error is committed when doing this approximation that is more severe in the near-filed and for large displacements between the receiver and the transmitter elements ([Bellettini, A.; Pinto, M. A. (2002)], [Bonifant, W. W. (1999)], [Wilkinson, D. R. (2001)], [Banks, S. M.; Griffiths, H. D. (2002)]).

A number of assumptions are inherent to this model. First, the model ignores the effects of medium turbulence, refraction and multipath, in that the signal is assumed to have travelled along a straight path, and the propagation delay is proportional to the range from target to platform. Second, the model assumes that each target's reflectivity is constant and does not change with the viewing angle.

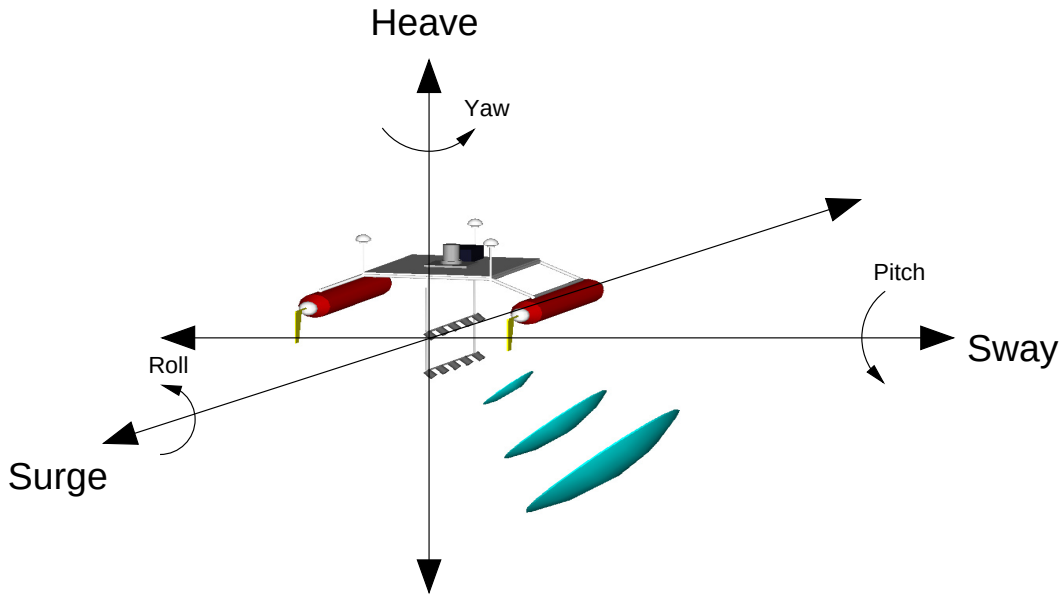

Figure 28: Motion perturbations.

Unknown motion components prevent the echo signals to be coherently combined results in severe degradation of the reconstructed image ([Cutrona (1975)], [Johnson, K. A. (1992)]). It is typically assumed that the position must be known to better than a tenth of a wavelength of the signal used. The surface vehicle is subject to perturbations in its position (surge, sway and heave) and attitude (roll, pitch and yaw) as exemplified in Figure 28.

Also medium coherence in time and space is fundamental. Non homogeneities in the medium cause ray bending and propagation speed fluctuation, corrupting the assumed model ([Jensen, J.; Kuperman, W.; Porter, M.; Schmidt, H. (2000)]).

Medium stability as been shown to be good enough for SAS imagery, with temporal stability within the π/4 limit over a period of a minute ([Gough, P. T.; Hayes, M. P. (1989)]).

Platform motion errors limit short-range SAS and medium coherence limit long-range SAS ([Chang, E.; Tinkle, M. D. (2001)]).

The motion and medium coherence restrictions impose difficult barriers to SAS systems that have to be overcome using sophisticated processing techniques.

4.2. Data model

The complex pass-band transmitted signal is (where $p_b(t)$ is the base band representation of the transmitted signal):

$$p_m(t) = p_b(t)e^{j2\pi f_0 t} \tag{4.10}$$

The received echoes will be the super imposition of the response of multiple target individual responses. The received signal due to one target is a time delayed and attenuated version of the transmitted signal (supposing no sources of distortion) convolved with the combined transmitter and receiver beam patterns $a(\tau, x, y, z)$.

$$a(\tau, x, y, z) = \frac{\alpha(f, r(\tau, x, y, z))}{r(\tau, x, y, z)^{\varepsilon}}\left[a_{RX}(\tau, x, y, z) * a_{TX}(\tau, x, y, z)\right] \tag{4.11}$$

The attenuation of the signal due to propagation loss (cylindrical, $\varepsilon = 1$ or spherical propagation $\varepsilon = 2$) and other loss mechanism is agglutinated in this function. The echo received for one target at each sampling position is:

$$ee_m(\tau, t) = a(\tau, x, y, z) * p_b\left(t - t_f(\tau, x, y, z)\right)e^{j2\pi f_0\left(t - t_f(\tau, x, y, z)\right)} \tag{4.12}$$

This assumes that the echo is not significantly influenced by Doppler effects, be it the ones due to platform motion or medium motion (river waters). Moreover, Doppler tolerant waveforms should be used if possible ([Rihaczek, A.W. (1966)], [Urkowitz, H.; Bucci, N.J. (1992)]).

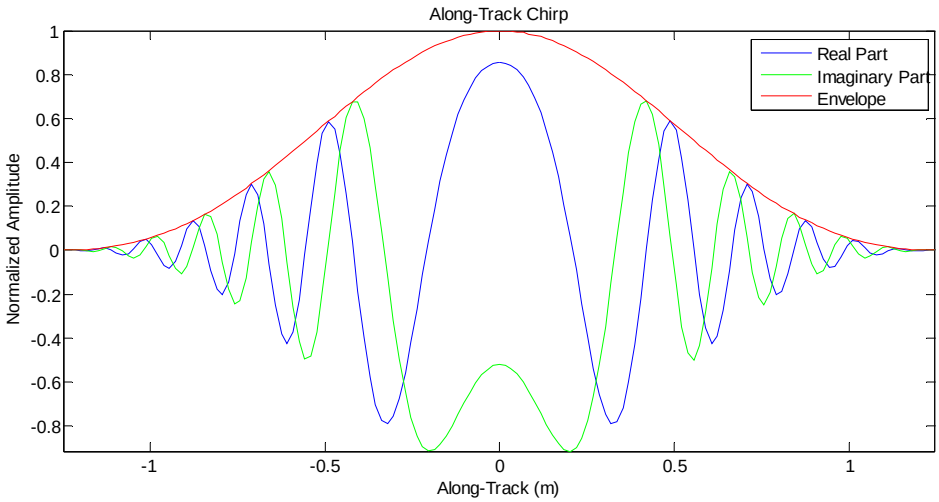

Figure 29: Along-Track Chirp.

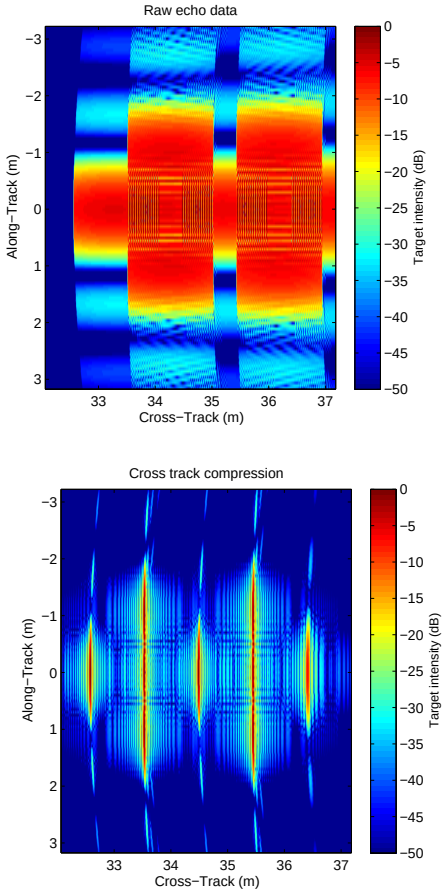

Figure 30: Uncompressed (top) and cross-track compressed (bottom) echo data.

So that it is possible to work with a minimal number of data samples to allow efficient implementation via digital signal processing, the data is demodulated and decimated at the earliest stage. The carrier frequency is removed from the data via phase multiplication by $e^{-j2\pi f_0 t}$ and filtering. It is straight forward to show that the resulting baseband expression for $ee_m(\tau,t)$ is then:

$$ee_b(\tau,t) = a(\tau,x,y,z) * p_b\left(t - t_f(\tau,x,y,z)\right) e^{-j2\pi f_0 t_f(\tau,x,y,z)} \tag{4.13}$$

The term $e^{-j2\pi f_0 t_f(\tau,x_\sigma,y_\sigma,z_\sigma)}$ is a hyperbolic frequency chirp in the τ coordinate. Assuming a simple linear motion, $r(\tau,x,y,z) = 2\sqrt{(x-v\tau)^2 + y^2 + z^2}$, we obtain the frequency chirp depicted in Figure 29, for which the described sonar parameters were used.

The range matched filter is quite simple and obtained by correlating the received echoes with the transmitted pulse:

$$ss_b(\tau,t) = \int ee_b(\tau,t') p_b(t'-t) dt' \tag{4.14}$$

An example sample data with 9 targets is shown in Figure 30. Here it is possible to see the effect of the pulse compression in the received echo data.

Taking the Fourier transform of $ee_b(\tau,t)$ in the t dimension we get:

$$eE_b(\tau,k) = P(f_t) \iint a(\tau,x,y,z)\, ff(x,y_s) e^{-j2\pi f_0 t_f(\tau,x,y,z)} dxdy \tag{4.15}$$

Using the principle of the Stationary Phase ([Hein, A. (2003)]), the Fourier transform in the τ dimension is given by:

$$EE_b(k_u,k) = P(k_u) A(k_u,k) \iint ff(x,y_s) e^{-jy\sqrt{4(k+k_0)^2 - k_u}} e^{-jk_u x} dxdy \tag{4.16}$$

$k_u = \frac{k_\tau}{v} = \frac{2\pi f_\tau}{v}$ and $k = \frac{2\pi f}{c}$ are the wave number related to the along-track and cross-track sampling times respectively. By making the following relation between the time and spatial wave numbers:

$$\begin{cases} k_x = k_u \\ k_y = \sqrt{4(k+k_0)^2 - k_u} \end{cases} \tag{4.17}$$

We get:

$$EE_b(k_u,k) = P(k_u)\,A(k_u,k)\,FF(k_x,k_y) \tag{4.18}$$

4.3. Resolution and sampling requirements

In the cross-track dimension, since pulse-compression is used, the range resolution will be given by the signal bandwidth and not by its duration. Because of this, it is possible to use longer pulses to reduce the peak transmitted power, maintaining the total signal energy. Since this system operates in shallow waters, the pulse maximum length is limited by the nearest distance of interest, taking into account that the receiver cannot receive data while the transmitter is operating.

The resolution in the cross-track (δ_{XT}) is given by the ability to distinguish the peaks of the two closely separated pulse compressed signals:

$$\Delta t = \frac{2\delta_{XT}}{c} \Rightarrow \delta_{XT} = \frac{c\Delta t}{2} \tag{4.19}$$

From the previous chapter we know that Δt is $1/\Delta f$ for a chirp signal.

The resolution in the along-track (δ_{AT}) for a side-scan system is given by the horizontal beam-width, furthermore the resolution is dependent of the distance to the target r (Figure 31):

$$\delta_{AT} = r\cos\theta_H \tag{4.20}$$

The resolution in the ground plane (δ_{XT_g}) is given by:

$$\delta_{XT_g} = \frac{\delta_{XT}}{\cos\theta_g} \tag{4.21}$$

To make the resolution higher it is necessary to decrease θ_H using higher frequency signals, larger diameter transducers or longer arrays (bigger window of observation of $ff(x,y_s)$ in the frequency domain).

In side scan sonar imagery the along-track position must be sampled with a minimum rate (Pulse Repetition Frequency - PRF) given by the resolution in this direction as to avoid gaps:

$$PRF \geq \frac{\delta_{AT}}{v} \tag{4.22}$$

On the other side, the maximum pulse repetition rate is given by the time that is necessary to wait for the furthest target echo to arrive:

$$PRF \leq \frac{2r_{\max}}{c} \tag{4.23}$$

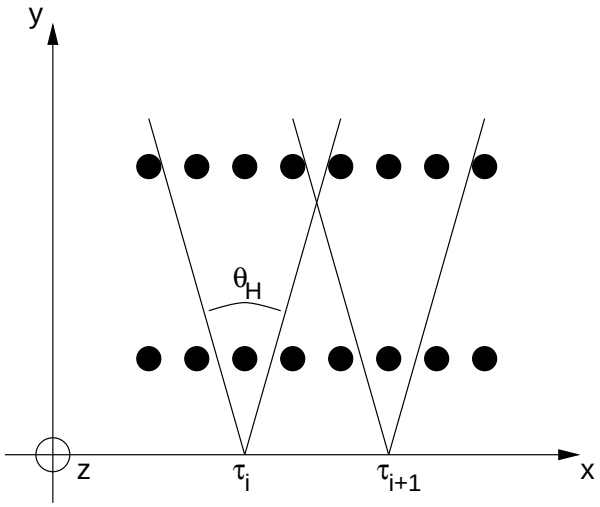

Figure 31: Strip-map sonar along-track resolution.

In synthetic aperture sonar along-track echoes are combined regarding their acquisition positions to form a virtual array which in turn is used to synthesize the image ([Gough, P.T. (1998)]). The maximum synthetic-array length is defined as:

$$L_{\max} = r_{\max} \cos\theta_H \tag{4.24}$$

Figure 32 represents how a single target echo is viewed by the sonar as it traverses the aperture. The target describes a hyperbole thought the sampling positions which fall in the main beam width of the combined transmitter and receiver radiation diagram. The along-track length that contains the target response is the synthetic aperture length, $L_{\max}$.

To correctly synthesize an image without aliasing artifacts in the along-track dimension, it is necessary to sample the swath with an interval of $D/2$ (considering the use of only one transducer for transmission and reception). This constraint together with the maximum pulse repetition frequency (PRF) imposes a very slow survey speed ([McHugh, P

(1998), Putney, P. (2005)]). For that reason automatic motion control is of high importance for this system.

Consider the target of Figure 32, the target will be seen by the sonar during the time it is inside the aperture (3dB lobe) of the transducer, which is approximately given by:

$$\theta_{3dB} \approx \frac{\lambda}{D} \tag{4.25}$$

The array spacing from Nyquist spatial sampling and classical array theory is λ/2 (two way equivalent 2π phase shift), which means that for angles of arrival of a wave-front the inter-element phase difference must be less than 2π ([McHugh, P. (1998)]).

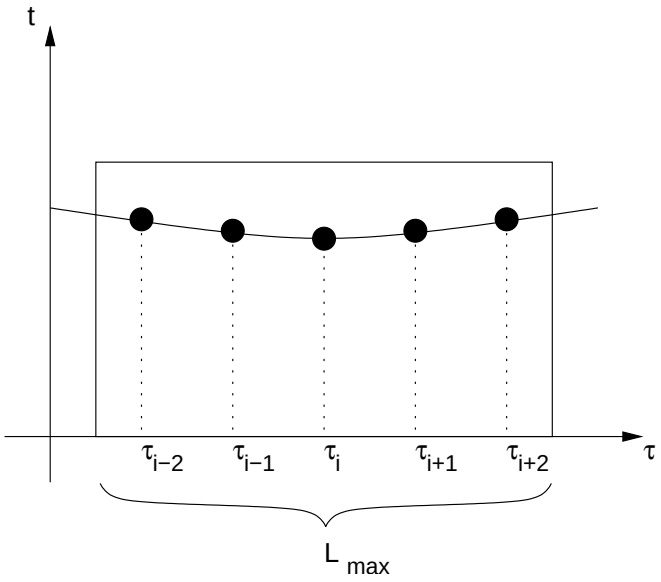

Figure 32: Hyperbolic target signature.

This is also true for motion errors. To correctly form a synthetic aperture the platform position must be known within 1/8 of a wavelength so the echoes can be coherently combined with negligible image deterioration ([Tomiyasu, K. (1978), Fornaro, G. (1999), Cutrona, L. J. (1975)]).

Another perspective would be to consider the array spacing to be given by a Pulse repetition Frequency (PRF) that is at least equal the maximum Doppler shift experienced by a target. The Doppler shift f_D is related to the radial velocity v_r by:

$$f_D = \frac{2v_r}{\lambda} = \frac{2v\sin\theta(\tau)}{\lambda} \tag{4.26}$$

The maximum radial velocity is obtained at the beam edge and so the lower bound for the PRF is:

$$PRF \geq \frac{2v \sin \theta_{3dB}}{\lambda} \approx \frac{2v\lambda / D}{\lambda} = 2\frac{v}{D} \tag{4.27}$$

The along-track resolution is independent of the range and wavelength. This results from the fact that for a transducer with a fixed diameter D, the synthetic aperture length D_{SA} will be given, approximately, by:

$$D_{SA} \approx 2r_0\theta_{3dB} = 2r_0\frac{\lambda}{D} \tag{4.28}$$

Remembering that r_0 is the distance to the center of the scene.

This than gives the classical synthetic aperture along-track resolution δ_{AT} formula:

$$\delta_{AT} \approx r_0\theta_{SA} = r_0\frac{\lambda}{D_{SA}} = r_0\frac{\lambda}{2R_0\frac{\lambda}{D}} = \frac{D}{2} \tag{4.29}$$

The image formed in this way has thus a cross-track resolution of $\delta_{XT} = c\Delta f/2$ and an along-track resolution of $\delta_{AT} = D/2$ (where c is the speed of sound, BW is the transmitted signal bandwidth and D is the transducer diameter), but more importantly the along-track resolution is independent of the distance between the sonar and the target ([Gough, P. T.; Hawkins, D. (1998)]).

We see here that the phase relations that enable the synthetic array formation are tightly related to the wavelength of the signal and the effective synthetic array length. Normally these two values are interconnected due to the transducers real aperture width, but can be explored to mitigate some of the problems inherent to synthetic aperture.

The aliasing artifacts are therefore not related to the presence of targets in some particular spatial concentration, but are yes related to the phase relation between sampling positions. The sampling constrain chosen for a synthetic aperture sonar system must target a along-track ambiguity to signal ratio (AASR) and peak to grating lobe peak ratio (PGLR) in the reconstructed image. The dynamic range can be estimated through the PGLR measure. An under-sampling can be amenable when using high bandwidth signals as in the case for the transmitting pulses employed in this particular system ([Gough, P. T.; Hawkins, D. (1998)]).

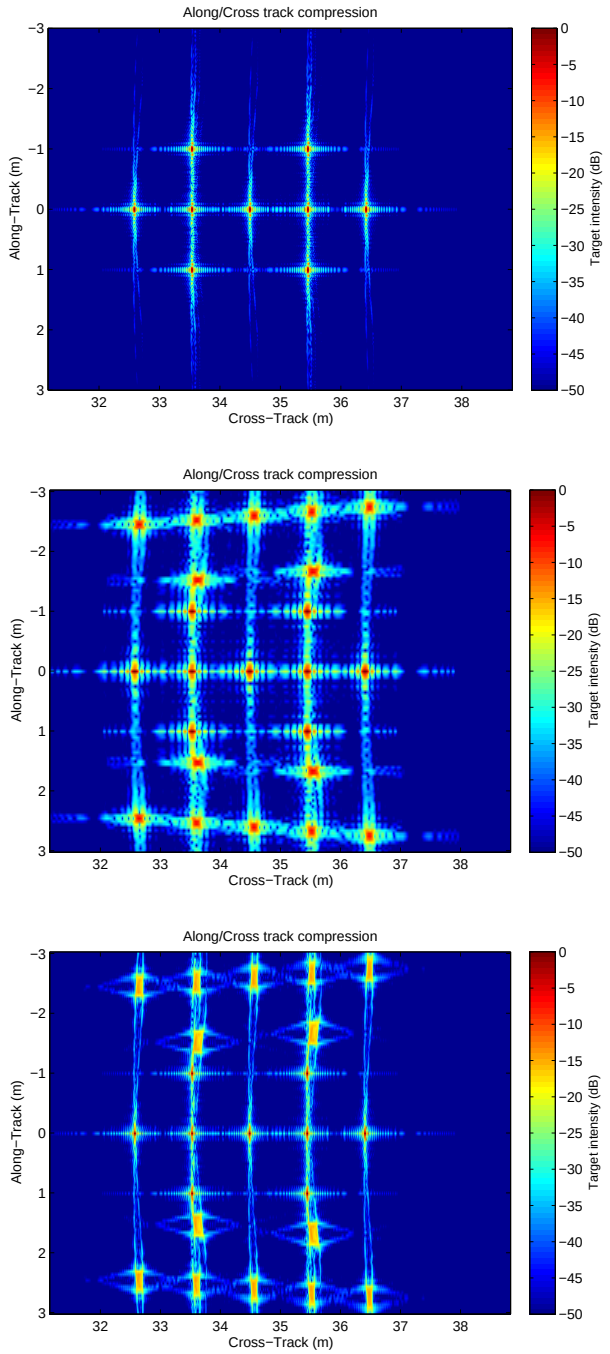

Figure 33: Effects of bandwidth in the synthetic image reconstruction: sampled correctly (BW = 30kHz); 10 times under-sampled (BW = 10kHz); 10 times under-sampled (BW = 30kHz).

This is because when using high bandwidth signals the energy in the point target response side-bands is spread across the along track dimension instead of concentrating on specific points. Figure 33 shows the result of reconstructing an under-sampled synthetic

aperture for the case of using a transmitted signal with 10 kHz and 30 kHz. It can be seen that the image using higher bandwidth signal benefits from the smeared side-lobes.

On the other side, AASR does not improve by increasing the bandwidth. Increasing the bandwidth only causes smearing of the grating lobes. Since the total energy is maintained, the value for AASR is maintained ([Chatillon, J. et al (1992)], [Chatillon, J. et al (2000)]).

To truly overcome the along-track sampling constrains it is necessary to use an array of multiple elements in this direction ([Heering, P. (1982)]). By doing so, an increase of the PRF by a factor equal to the number of elements is possible. For commercial and military applications, this is typically the only way to obtain a suitable area coverage ratio. The increase of the number of elements in the array causes a substantial increase in hardware complexity and data storage (more channels). The algorithms for synthetic aperture must also account for the presence of multiple receivers.

Chapter 5: Image Formation

5.1. Explicit Matched Filtering Algorithm

Suppose a scene containing a single point reflector. The reflectivity function is a single Dirac impulse in the center of the image as given by:

$$ff(x, y_s) = \delta(0, r_0) \tag{5.1}$$

As the sonar platform traverses the aperture following a simple linear motion, the changing range to a given to a given target will follow a hyperbolic locus, or range migration, in space-time, given by:

$$r(\tau, x, y_s) = \frac{c}{2} t_f(\tau, x, y_s) = \sqrt{(x - v\tau)^2 + (y_s)^2} \tag{5.2}$$

The obtained echoes are given by:

$$ee_b(\tau, t) = a(\tau, x, y, z) * p_b\left(t - t_f(\tau, x, y_s)\right) e^{-j2\pi f_0 t_f(\tau, x, y_s)} \tag{5.3}$$

The pulse compression in cross-track is independent from the along-track processing and so the pulse compressed echo data is:

$$ss_b(x, y_s) = ee_b(\tau, t) * p_b(t) = \delta\left(t - t_f(\tau, x, y_s)\right) e^{-j2\pi f_0 t_f(\tau, x, y_s)} \tag{5.4}$$

Since the locus is range-variant, the image inversion problem is two-dimensional and non-separable. Based in the echo data, we must compress the 2-D signature of each target into a single, correctly-located impulse. An exact algorithm uses the targets locus equation at each location, thereby exactly inverting the system equation.

The reconstructed image point is thus:

$$ff(x, y_s) = \iint \delta\left(t - t_f(\tau, x, y_s)\right) e^{-j2\pi f_0 t_f(\tau, x, y_s)} e^{j2\pi f_0 t_f(\tau, x, y_s)} d\tau dt \tag{5.5}$$

Extending this reasoning to the whole scene, each point in reconstructed the image is obtained by applying a matched filter to the baseband data of $ss_b(\tau, t)$ using a point spread function based on each target location that is given by:

$$h(\tau,t)=\delta\left(t-t_f\left(\tau,x,y_s\right)\right)e^{j2\pi f_0 t_f(\tau,x,y_s)} \tag{5.6}$$

These operations require interpolators due to the sampled nature of the data and can be implemented in either spatial-time, using convolution operations, or in the frequency domain.

Implementation of multiple receiver and transmitter configurations is straight forward with this algorithm as is the integration of sonar positioning information.

The first SAS processing systems used extensively this method. Nevertheless, military and commercial applications demanded faster algorithms, more efficient and where real-time image processing could be utilized.

Obviously, even with FFT operations, these methods are very computationally intensive, since a matched filter must be computed for each pixel location:

$$ff\left(x,y_s\right)=\iint eE_b(u,k)P_b^*\left(k\right)e^{j2k_0 r(\tau,x,y_s)}dudk \tag{5.7}$$

If it is assumed that for ranges that are large compared to the aperture length, the hyperbola can be approximated by a range invariant parabola obtained for the center of the scene. This simplifies the processing since there is no need to explicitly compute the matched filter for each target and the reconstructed image is simply the correlation between the pulse-compressed echo data and the range invariant parabola (Figure 34).

The along-track Fourier transform of the above equation gives the fast correlation implementation, by using Weyl's identity ([Cumming, I. G.; Wong, F. H. (2005)], [Callow, H. J. (2003)]) and noting that this is only correct around the center range of the scene (y_{s_0}) because of the space-variant nature of the point-spread function:

$$ff\left(x,y_s\right)=\iint EE_b(k_u,k)P_b^*(k)e^{j\left|y_{s_0}\right|\sqrt{4k_o^2-k_u^2}}e^{jk_u y_s}dk_u dk \tag{5.8}$$

For SAS systems this is usually not a very good approximation, since the scene size is usually of the same order of magnitude as the distance to its center. As the matching of the range invariant hyperbola will be lower for ranges that differ from the chosen one (y_{s_0}), blur will occur and limit the depth-of-focus (set of ranges with an acceptable measure of resolution).

If the scene is divided in several blocks, the depth of focus can be somewhat extended, but then again at higher computational effort. Algorithms such as the range-Doppler or the Wave-number algorithm implement forms of fast correlation, but take a different approach for extending the depth-of-focus.

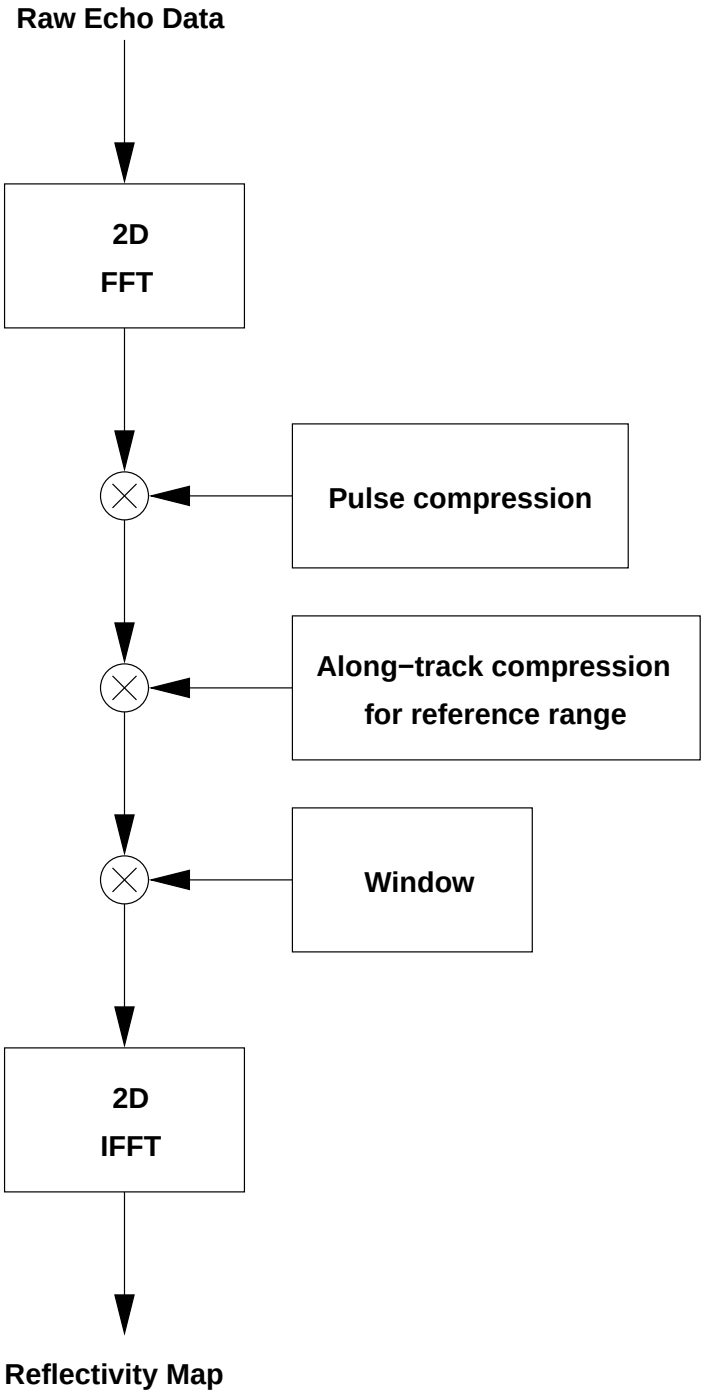

Figure 34: Simplified explicit matched filtering algorithm signal flow chart.

Typically, algorithms that use the transfer function of each target response are designated matched filter system.

Figure 35 shows a scene with 9 point targets obtained through the use of the application of the explicit matched filtering algorithm.

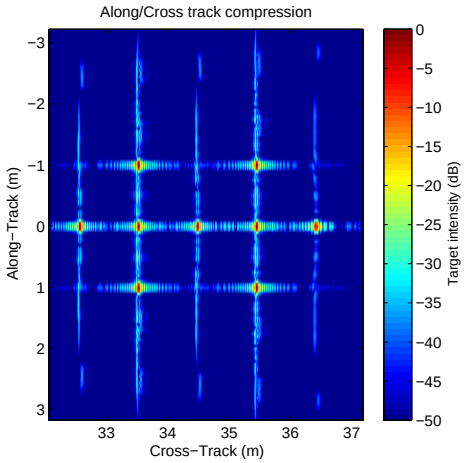

Figure 35: Explicit matched filtering output.

5.2. Range-Doppler Algorithm

The range-Doppler algorithm aims at extending the depth of focus obtained using along-track compression with a single point-spread function by using a coordinate change. This is a popular algorithm for SAR ([Cumming, I. G.; Wong, F. H. (2005)]).

Starting with the pulsed compressed data $ss_b(u,t)$, a 1-D Fourier transform is performed to the data in the along-track direction. The data is then represented in the Doppler (k_u) and slant-range (y_s) domain:

$$Ss_b(k_u,t) \approx Ff(k_u, y_{s_0})\delta\left(t - \frac{2}{c} y_{s_0} C_s(k_u)\right) e^{j|y_{s_0}|\sqrt{4k_o^2 - k_u^2}} \tag{5.9}$$

The locus of a point target placed at y_{s_0} in the range-Doppler domain is give by:

$$y_{s_0} C_s(k_u) \tag{5.10}$$

With the curvature factor $C_s(k_u)$ given by:

$$C_s(k_u) = \frac{1}{\sqrt{1 - \left(\frac{k_u}{2k_0}\right)^2}} - 1 \tag{5.11}$$

Range and along-track decoupling is performed by the coordinate change:

$$T\{\cdot\}=\begin{cases} t=\dfrac{2}{c}y_{s_0}\left[1-C_s\left(k_u\right)\right] \\ k_x\left(k_u,t\right)=k_u \end{cases} \tag{5.12}$$

After the coordinate transform we obtain:

$$T\left\{Ss\left(k_u,t\right)\right\}=Ff\left(k_x,y_s\right)e^{j\left|y_{s_0}\right|\sqrt{4k_o^2-k_u^2}} \tag{5.13}$$

This step is implemented by an interpolation. This straightens the targets signature in the range-Doppler domain, making possible to apply an along-track matching filter given by:

$$Qq(k_x,y_s)=e^{j\left|y_{s_0}\right|\sqrt{4k_o^2-k_u^2}} \tag{5.14}$$

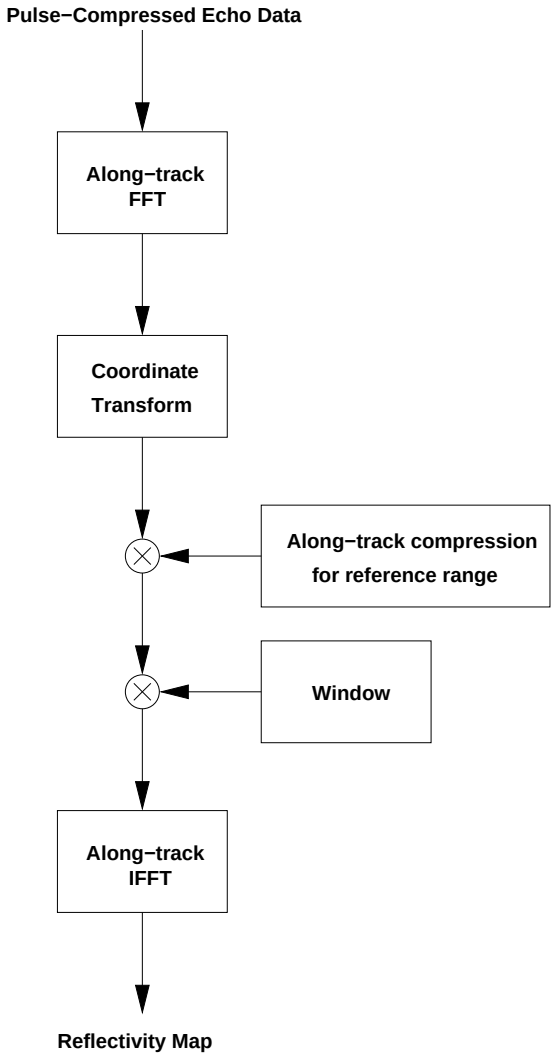

Figure 36: Range-Doppler algorithm signal flow chart.

This operation is then followed by a tapered window $W(k_x)$ which may be any one of a number of 1D window functions that will ensure low side lobes in the image along-track direction. The concept here is that the coordinate transformation puts a time advance into the domain that is both delay time and spatial wave-number dependent. This decouples the rows and the columns of the range-Doppler matrix. Phase only multiplication is followed which removes the Fresnel-like dependence of the reflected echoes in the range-Doppler domain, producing data corresponding to a pseudo-Fourier transform of the diffraction limited image.

The final algorithm can then be represented as (Figure 36):

$$ff(x, y_s) = F_x^{-1}\left\{W(k_x)Qq(k_x, y_s)T\left\{F_u\left\{ss_b(u,t)\right\}\right\}\right\} \tag{5.15}$$

The main disadvantages of this algorithm are for one the required interpolation between two sets of sampling grids, secondly it degrades as the scene extends away from the center. With large beam widths, the spatial Doppler may not be completely decoupled from the range which results in additional linear frequency modulated chirp at large wave numbers. For wide-beam or broad-band systems, the use of this range-Doppler algorithm produces images with degradation. This results form the narrow-band assumption ($k \approx k_0$) made in the derivation of the algorithm. The narrow-band assumption neglects the spreading in the range direction of the range-migration corrected data. This spreading is noticeable for non-zero k_u values. This means that the pulse compression appropriate for $k_u = 0$ is no longer appropriate at the larger wave numbers.

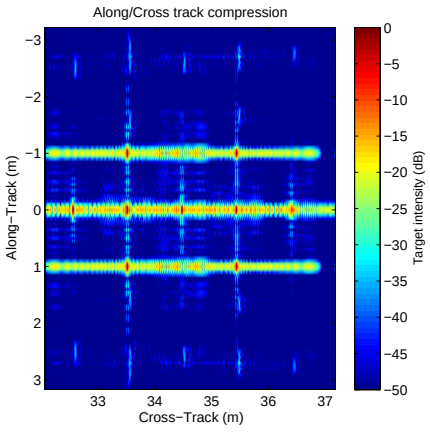

Figure 37: Range-Doppler algorithm output.

To compensate this, a secondary range compression (SRC) was developed ([Curlander, J. C.; McDonough, R. N. (1996)]). SRC provides a k_u dependent compression in the range direction to give improved imagery.

Figure 37 shows a scene with 9 point targets obtained through the use of the range-Doppler algorithm.

5.3. Wave-number Algorithm

The essence of the wavenumber algorithm is a coordinate transform from (k_u, k) to the (k_x, k_y) domain.

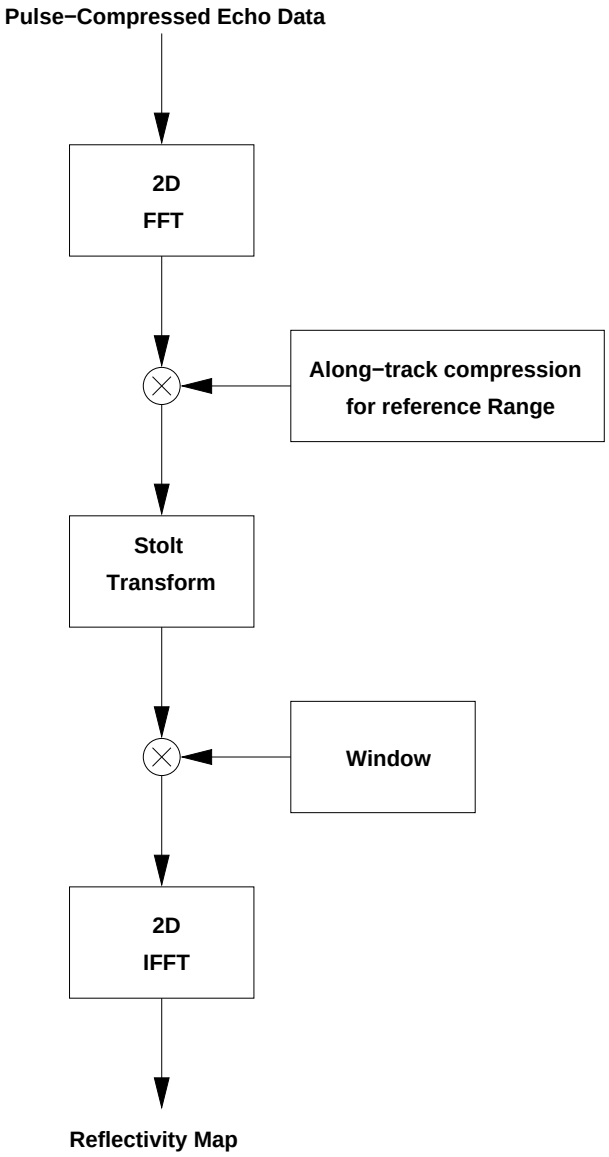

Figure 38: Wave-number algorithm signal flow chart.

The 2-D Fourier transform is taken from the pulse compressed echo data $ss(u,t)$. By applying a matched filter for a target at a reference range y_{s_0}, oscillatory terms dependent of y_{s_0} are removed, easing the coordinate transform implementation by interpolation:

$$FF\left(k_x,k_{y_s}\right)=S^{-1}\left\{SS\left(k_u,k\right)e^{j\left|y_{s_0}\right|\sqrt{4k_o^2-k_u^2}}\right\} \tag{5.16}$$

$S^{-1}\{\cdot\}$ is the Stolt inverse coordinate transform given by ([Cumming, I. G.; Wong, F. H. (2005)], [Hawkins 1996]):

$$S^{-1}\{\cdot\}=\begin{cases} k_x=k_u \\ k_{y_s}=\sqrt{4k^2-k_u^2}\end{cases} \tag{5.17}$$

And the Stolt mapping, $S^{-1}\{\cdot\}$:

$$S\{\cdot\}=\begin{cases} k_x=k_u \\ k=\frac{1}{2}\sqrt{k_{y_s}{}^2+k_u^2}\end{cases} \tag{5.18}$$

The coordinate transform is the mapping of 2-D arcs of constant (k_u,k) centered around the wave-number origin onto a regular grid $\left(k_x,k_y\right)$.

The final image is obtained by 2-D filtering and inverse Fourier transform of the processed data.

Although the later derivation was made for pass-band data, typically SAS data is recorded in base-band. Moreover the Stolt coordinate transform requires an accurate interpolation. By using base-band data, rapidly varying phase data is avoided. Furthermore, doing so, the input data is reentered on y_{s_0} and the output data on y_0.

The equation describing the algorithm taking into account the modulation and offsets is:

$$FF\left(k_x,k_{y_s}\right)=S_b\left\{SS_b\left(k_u,k\right)e^{j(k_u+k_0)y_{s_0}}e^{-j2(k+k_0)y_0}\right\} \tag{5.19}$$

And the inverse base band Stolt transform is:

$$S_b^{-1}\{\cdot\} = \begin{cases} k_x = k_u \\ k_{y_s} = \sqrt{4(k+k_0)^2 - k_u^2} - k_{y_0} \end{cases} \tag{5.20}$$

Since typically $y_{s_0} = y_0$, $k_{y_0} = 2k_0$ and the above expression simplifies to:

$$FF\left(k_x, k_{y_s}\right) = S_b\left\{SS_b\left(k_u, k\right)e^{-j(k_u+k_0)y_0}\right\} \tag{5.21}$$

Without the interpolation step (coordinate change) this algorithm is equivalent to use a match filter in the along-track direction for a reference range (Figure 38). By doing the coordinate change, the depth-of-focus is extended. Thus the image is always well focused at the reference range (even for large spatial bandwidths) and image distortion from inaccurate interpolation only appears at the edges of the image. Near range imaging requires even more accurate interpolation due to the non-linear and space-variant nature of the coordinate change.

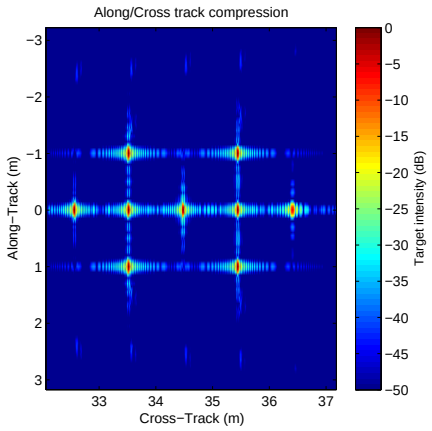

Figure 39: Wave-number algorithm output.

Figure 39 shows a scene with 9 point targets obtained through the use of the wave-number algorithm.

5.4. Chirp Scaling Algorithm

The chirp-scaling algorithm avoids the interpolation of the wave-number algorithm by using the first three terms in the Taylor expansion of the Stolt transform. Doing so, the time-shifting and scaling properties of linear-FM chirp are then exploited to replace the interpolation ([Cumming, I. G.; Wong, F. H. (2005)]).

Therefore assuming that the transmitted pulse is a linear chirp of the form:

$$p(t) = e^{j\pi Kt^2} \tag{5.22}$$

With $K = 2\frac{f_2 - f_1}{T}$. The along-track Fourier transform is taken, giving:

$$Ee(k_u, t) \approx Ff(k_u, t) rect\left[\frac{t - \frac{2}{c} y_s C_s(k_u)}{T}\right] e^{jK_s(k_u,x)\left[t - \frac{2}{c} y_s C_s(k_u)\right]} e^{j|y_s|\sqrt{4k_0^2 - k_u^2}} \tag{5.23}$$

Here the effects of the transducer pattern and amplitude terms have been ignored. The function $K_s(k_u, x)$ is the new k_u dependent received chirp rate:

$$K_s(k_u, x) = \frac{1}{1/K - K_{SRC}(k_u, x)} \tag{5.24}$$

And $K_{SRC}(k_u, x)$ is the secondary range compression range correction, due to the imaging geometry:

$$K_{SRC}(k_u, x) = \frac{8\pi x}{c} \frac{k_u^2}{\left(4k_0^2 - k_u^2\right)^{3/2}} \tag{5.25}$$

The along-track Fourier transform was solved using two-approximations. The first being the Principle of Stationary Phase. The second approximation is a quadratic Taylor series approximation of the Stolt transformation give by ([Callow, H. J. (2003)]):

$$\sqrt{4(k + k_0)^2 - k_u^2} \approx \sqrt{4k_0^2 - k_u^2} + \frac{2k_0}{\sqrt{4k_0^2 - k_u^2}} k + \frac{1}{2}\left[\frac{4}{\sqrt{4k_0^2 - k_u^2}} - \frac{(8k_0)^2}{\sqrt[3]{4k_0^2 - k_u^2}}\right] k^2 \tag{5.26}$$

With:

$$\beta = \frac{1}{C_S(k_\tau) - 1} = \sqrt{1 - \left(\frac{k_u}{2k_0}\right)^2} \tag{5.27}$$

We get:

$$\sqrt{4(k + k_0)^2 - k_u^2} \approx 2k_0\beta + \frac{2k}{\beta} + \frac{k^2}{k_0} - \frac{k^2}{8k_0\beta^3} \tag{5.28}$$

The algorithm proceeds by making all targets have the same phase signal as a target at a reference range y_{s_0} by applying a chirp-scaling phase multiplication to the range-Doppler data:

$$Mm(k_u,t) = Ee(k_u,t)\Psi\psi_1(k_u,t) \tag{5.29}$$

With $\Psi\psi_1(k_u,t)$ given by:

$$\Psi\psi_1(k_u,t) = e^{j\pi K_S(k_u,x)C_s(k_u)(t-t_0(k_u))^2} \tag{5.30}$$

As mentioned this operation equalizes all the scatters range migration phase term to that of the reference scatter at range y_{s_0}. The curvature factor $C_s(k_u)$ ([Gough, P. T.; Hawkins, D. (1998)]), mentioned already within the description of the range-Doppler algorithm, describes the spatial wave number dependent part of the signal trajectory and the time locus of the reference range is given by $t_0(k_u)$, the time shift of the reference focus locus:

$$t_0(k_u) = \frac{2}{c} y_{s_0}\left(1+C_s(k_u)\right) \tag{5.31}$$

To apply matched-filtering to the reference range, a Fourier transform is taken in the cross-track dimension:

$$MM(k_u,k) = F_t\left\{Mm(k_u,t)\right\} \tag{5.32}$$

Pulse compression and SRC are applied at this step through the following phase multiplication:

$$NN(k_u,k) = MM(k_u,k)\Psi\Psi_2(k_u,k) \tag{5.33}$$

Where the phase multiplication term $\Psi\Psi_2(k_u,k)$ is given by:

$$\Psi\Psi_2(k_u,k) = e^{j\frac{c^2k^2}{4\pi K_s(k_u)\left[1+C_s(k_u)\right]}} e^{j2ky_{s_0}C_s(k_u)} \tag{5.34}$$

The first term is the cross-track pulse compression including SRC for the modified chirp rate, and the second term is the bulk range curvature correction. Additionally this along-track and cross-track coordinates have been decoupled, and $NN(k_u, k)$ is simply given by:

$$NN\left(k_x, k_{y_s}\right) = NN\left(k_u, 2k\right) \tag{5.35}$$

The data is then windowed across the processing bandwidth and inverse Fourier transformed in the k_u domain, making possible to perform now the same along-track compression that was used in the range-Doppler algorithm:

$$Ff\left(k_x, y_s\right) = F_{k_{y_s}}^{-1}\left\{NN\left(k_x, k_{y_s}\right)\right\}\Psi\psi_3\left(k_x, y_s\right) \tag{5.36}$$

The $\Psi\psi_3(k_x, y_s)$ phase term encloses $Qq(k_x, y_s)$ and some additional phase compensation terms:

$$\Psi\psi_3\left(k_x, y_s\right) = Qq\left(k_x, y_s\right)e^{-\frac{4\pi}{c^2}K_s(k_x)C_s(k_x)\left(1+C_s(k_x)\right)\left(y_s - y_{s_0}\right)^2} \tag{5.37}$$

Finally the reconstructed image is obtained by inverse Fourier transform in the k_x domain. This algorithm uses only phase multiplications and Fourier transforms (Figure 40) which can be implemented efficiently ([Gimeno, E.; Lopez-Sanchez, J. M. (2001),]).

The substance of this algorithm is to recognize that the equivalent of the t and k_u time advance in the range-Doppler algorithm implicit in $T\{\cdot\}$ can be accomplished by a phase multiplication of the uncompressed range-Doppler data with the provision that the transmitted pulse is a linear chirp.

A complete mathematical derivation of the chirp scaling algorithm can be found in [Callow, H. J. (2003)]. Other derivations can be found in [Gough, P. T.; Hawkins, D. (1998)] and [Gimeno, E.; Lopez-Sanchez, J. M. (2001)].

The above process lacks an interpolation step, thus removing some of the limitations in the computational efficiency of the seismic migration algorithm. The dominant computational burden of the seismic migration algorithm is the interpolation step, which determines both efficiency and accuracy. This is a trade-off, and typical implementations use an interpolation kernel with several samples to favor efficiency ([Callow, H. J. (2003)]).

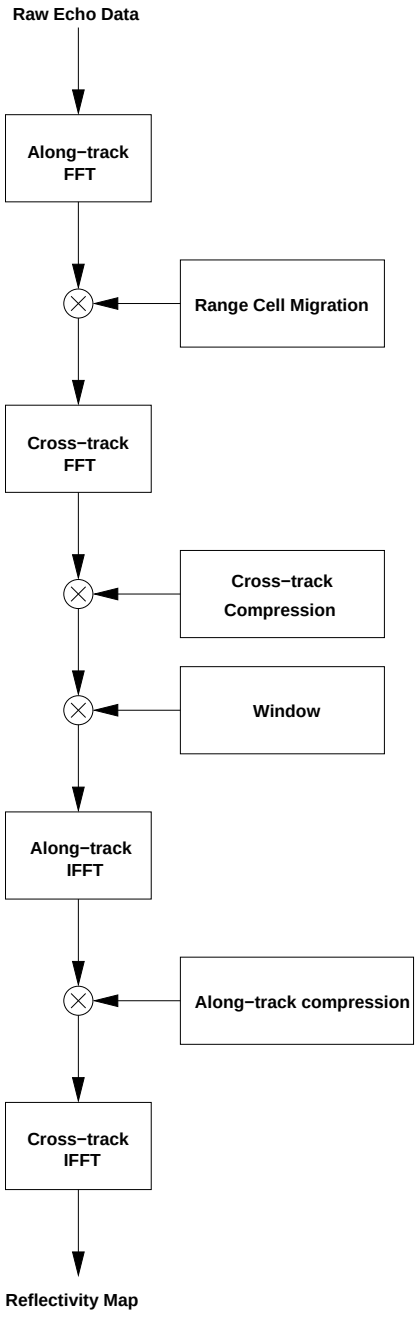

Figure 40: Chirp-Scaling algorithm signal flow chart.

The chirp scaling algorithms eliminates the need for an interpolation step by approximating the Stolt transformation.

The chirp scaling algorithm has, however, some limitations of its own. First, there are a number of phase errors inherent to the process. The Principle of Stationary Phase is invoked in several steps and the chirp scaling multiplier neglects the range dependence of K_S. Additionally, the derivation of the expression for the range-Doppler data used a quadratic

Taylor series expansion, and doing so, approximated the Stolt interpolation. The effect of this approximation is discussed in [Callow, H. J. (2003)].

In [Raney, R. K. et al (1994)], all of the intrinsic phase errors are described as small with the range invariance assumption cited as the dominant error. Another phase error correction method is proposed in [Hawkins, D. W.; Gough, P. T. (1997)], based on the theoretical phase function of an image point target.

Although this algorithm can only be used with uncompressed data obtained form the transmission of a linear chirp, a different version of the algorithm was proposed by [Hawkins, D. W. (1996)], [Hawkins, D. W.; Gough, P. T. (1997)] that allows computational saving and the use of arbitrary waveforms.

Figure 41 illustrates the result of the Chirp scaling algorithm in a scene with 9 point targets.

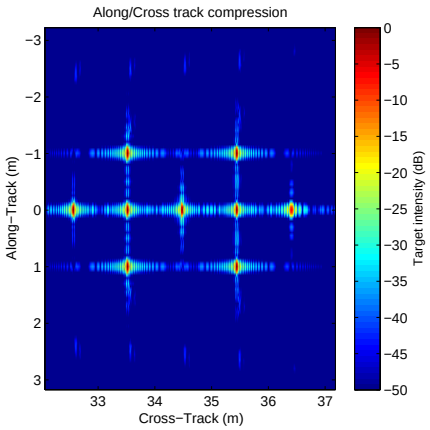

Figure 41: Chirp-Scaling algorithm output.

5.5. ISFFT Algorithm

We have seen that cross-track compression is an easy task, since all the parameters of the model are known and, more relevantly, because, the cross-track compression is not variant in any of the coordinates. Implementing azimuth compression through match filtering is more difficult, because of the different cross-track hyperbolas under the influence of each point-target: cross-track variance of the spread function.

The chirp-scaling algorithm avoids the burdensome non-linear interpolation of a coordinate change by using and the time scaling properties of the chirps that are applied in a sequence of multiplications and convolutions. Nevertheless, the chirp scaling algorithm is limited in use to the processing of uncompressed echo data that is obtained by the transmission

of chirps signals. The approach followed by the inverse scaled fast Fourier transform algorithm (ISFFT) ([Hein, A. (2003)]) interprets the raw data spectrum as a scaled and shifted replica of the scene spectrum. This scaling can then be removed during the inverse Fourier transformation if the normal IFFT is substituted by a scaled IFFT. This scaled IFFT can be realized by a chirp multiplication in the time and frequency domain. The obtained algorithm is computationally efficient and phase preserving (e.g. fit for interferometric imagery). This algorithm was developed for SAR imagery and applied to SAS within the scope of this work ([Silva, S. (2007 a)]).

Consider a complete scene raw echo data that can be obtained by calculating the response for one target and then substituting the single received backscatter by a set of backscatters with a delay described by $t_f(\tau, R(\tau)) = 2R(\tau)/c$, where $R(\tau) = y_{s_0} + y_s$ and integrating these scatters through the entire scene using the superimposition integral and performing some approximations steps. In SAR data one can approximate $\sqrt{y_{s_0} + y_s} \approx \sqrt{y_{s_0}}$, because the mean range is many orders of magnitude higher than the scene range. In SAS data however, the average distance is of the order of the scene distance. This is reflected as amplitude weighting dependent range presented on the scene spectrum. Since this weighting is dependent only on range, its compensation can be included in the range compression step or prior to synthetic aperture image formation. The spectrum of an entire SAS scene in the equivalent low-pass domain is then:

$$FF'(k_u, k) = \sqrt{y_{s_0}}\, e^{-jy_{s_0}\sqrt{4(k+k_0)^2 - k_u^2}} \iint FF\left(k_u, y_{s_0} + y_s\right) e^{-jy_s\sqrt{4(k+k_0)^2 - k_u^2}}\, dy_s \tag{5.38}$$

Which in turn can be written has:

$$FF'(k_u, k) = \sqrt{y_{s_0}}\, FF\left(k_u, k_r(k_u, k)\right) \tag{5.39}$$

By making $k_r(k_u, k) = \sqrt{4(k+k_0)^2 - k_u^2}$. Because of the dependency of $k_r(k_u, k)$ on both k_u and k it is not possible to obtain the original backscattering spectrum back by conventional means (without some kind of coordinate change or interpolation). To compress all point targets it is necessary to treat all of the individually.

Consider $k_r(k_u, k)$ a non-linear distortion of the k axis depending on both k_u and k and take the following Taylor series approximation (similar to the one made in the Chirp algorithm):

$$k_r\left(k_u,k\right)\approx k_r\left(k_u,k_0\right)+\dot{k}_r\left(k_u,k_0\right)\left(k-k_0\right)+\frac{1}{2}\ddot{k}_r\left(k_u,k_0\right)\left(k-k_0\right)^2+\ldots \tag{5.40}$$

Taking only the first order term, we get:

$$k_r\left(k_u,k\right)\approx 2k_0\beta+\frac{2k}{\beta} \tag{5.41}$$

The spectrum of the complete backscattering coefficient may be expressed as:

$$FF\left(k_u,k_r\left(k_u,k\right)\right)\approx FF\left(k_u,2k_0\beta+\frac{2k}{\beta}\right) \tag{5.42}$$

This approximation was found to be valid for SAR data, but is also valid for SAS data, where the propagation velocity is much lower, the center frequency is also lower and the center frequency to bandwidth ratio is quite higher than in SAR. As can be seen by the error surface depicted in Figure 42, this assumption is applicable still.

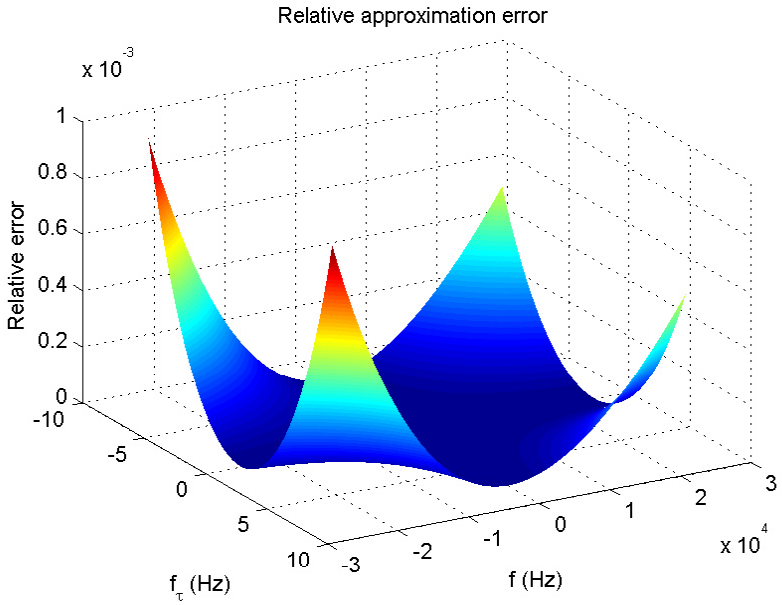

Figure 42: Approximation error in the frequency scaling Taylor expansion.

Using this approximation, we can write the inverse cross-track Fourier transform of the scene spectrum as:

$$Ff\left(k_u, y_{s_0}+y_s\right) \approx \frac{2}{\beta} e^{j2k_0\beta y_{s_0}} e^{j2k_0\beta y_s} \int FF\left(k_u, 2k_0\beta + \frac{2k}{\beta}\right) e^{-j2y_{s_0}(k+k_0)} e^{-j2ky_{s_0}/\beta} e^{-j2ky_s/\beta} dk \tag{5.43}$$

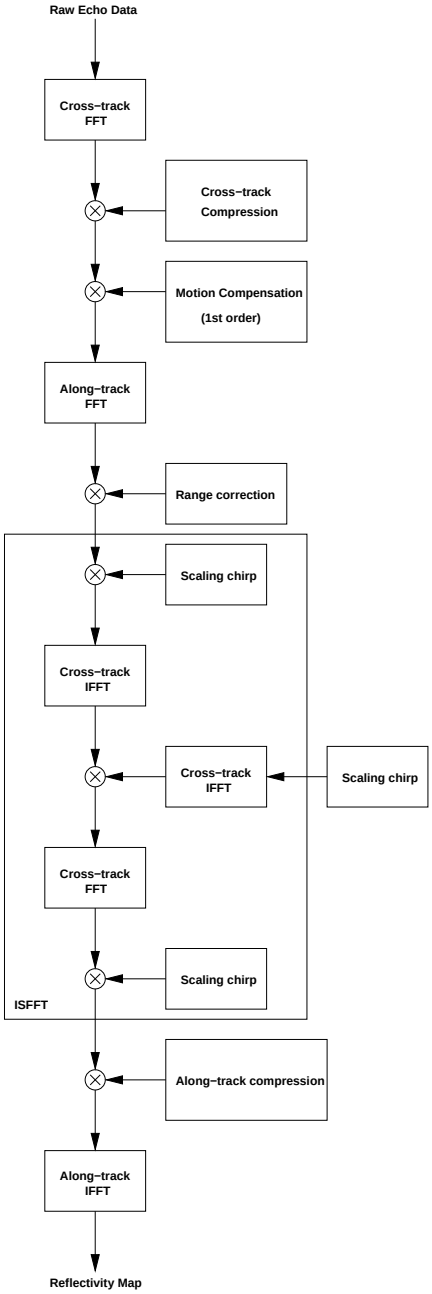

Figure 43: ISFFT algorithm signal flow chart.

In this equation we can identify the Fourier transform of $FF(k_u, k)$ scaled in the frequency axis by $2/\beta$ and shifted by $2k_0\beta$. The original backscattering can now be obtained using the scaled Fourier transform [Hein, A. (2003)], noting that:

$$\frac{1}{|a|}s(t/a)=\int_{-\infty}^{+\infty}S(af)e^{j2\pi ft}df=F^{-1}\{S(af)\}=e^{j\pi at^2}\left[\left\{S(af)e^{j\pi f^2}\right\}e^{-j\pi tf^2}\right] \tag{5.44}$$

This results in:

$$Ff\left(k_u,y_{s_0}+y_s\right)\approx\frac{2}{\beta}e^{j2k_0\beta y_s}e^{-j2y_{s_0}k_0(1-\beta)}\frac{\beta}{2}Ff\left(k_u,y_{s_0}+y_s\right)e^{-j2k_0\beta y_s}=Ff\left(k_u,y_{s_0}+y_s\right) \tag{5.45}$$

$Ff\left(k_u,y_s\right)$ can thus be found by the use of the inverse-scaled Fourier transform, and after these steps, the desired synthetic aperture image is obtained by inverse Fourier transform in the k_u dimension.

Figure 43 illustrates the algorithm signal flow, putting in evidence the inverse scaled Fourier transform step, and motion compensation factor that will be discussed in the following section.

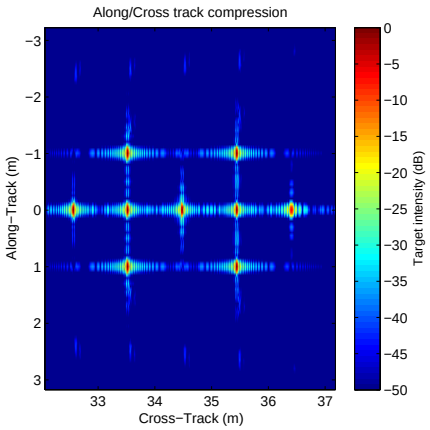

Figure 44: ISFTT algorithm output: chirp.

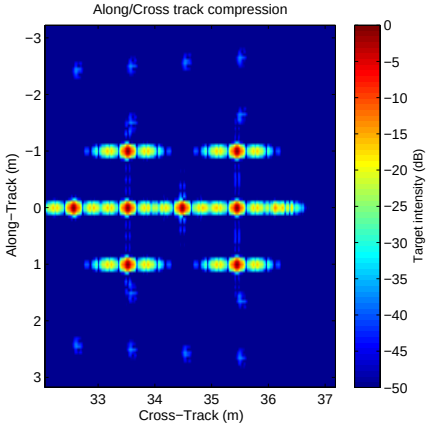

Figure 45: ISFTT algorithm output: pseudo-random sequence.

An example of a SAS image obtained through this algorithm can be seen in Figure 44 and Figure 45, where 9 point targets obtained by simulation using the parameters of this sonar system of the received echo signals is shown. Both hypotheses for transmission signals are tested: the chirp signal and the pseudo-random sequence signal.

5.6. Back-projection

This class of synthetic aperture imaging algorithms, although quite computationally expensive in comparison with frequency domain algorithms, lends itself very well to non-straight acquisition trajectories and, therefore, to the inclusion of known motion deviations from the expected path ([Hunter, A. J. (2003), Shippey, G. (2005), Silva, S. (2007 b), Silva, S. (2008 a)]).

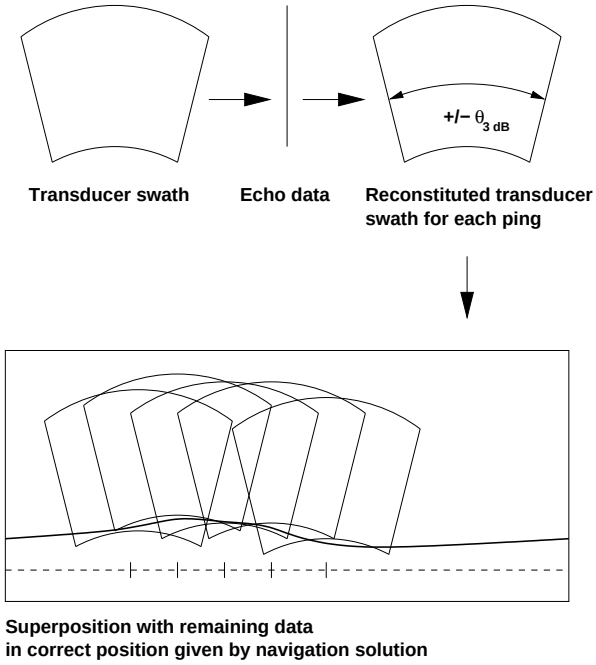

Figure 46: Back-projection schematic.

The back-projection algorithm enables image reconstruction for any desired path, limited only by sonar position uncertainties, estimation of the bottom topography and interpolator quality. It does not rely on the simple data resampling motion compensation algorithms or approximations to a linear acquisition path ([Callow, H. J. (2003)]). Instead, it considers that each point in one echo is the summation of the contributions of the targets in the transducer aperture span with the same range.

To reconstruct the reflectivity map, each echo is spread in the image at the correct coordinates (back-projected) using the known transducer position at the time of acquisition (Figure 46). The back-projection algorithm can be formulated as:

$$ff(x,y_s)=\int_{\tau_1}^{\tau_2} ss_b\left(\tau,t-\frac{2}{c}r(\tau,x,y,z)\right)e^{j2k_0r(\tau,x,y,z)}r^{\alpha}(\tau,x,y,z)w(\tau)d\tau \quad (5.46)$$

Where $r(\tau,x,y,z)$ is the distance between the transmitter to location (x,y,z) and back to the receiver through the span, k_0 is the transmitted signal center wave-number, $r^{\alpha}(\tau,x,y,z)$ compensates for the along-range attenuation, $w(\tau)$ is a weighting factor that compensates the transducer aperture attenuation within the 3 dB swath size ($[\tau_1,\tau_2]$). As a first approximation, the value of the constant α equals 1, this relates to the path loss of a cylindrical propagation model that is usually assumed in shallow waters ([Hunter, A. J. (2006)]).

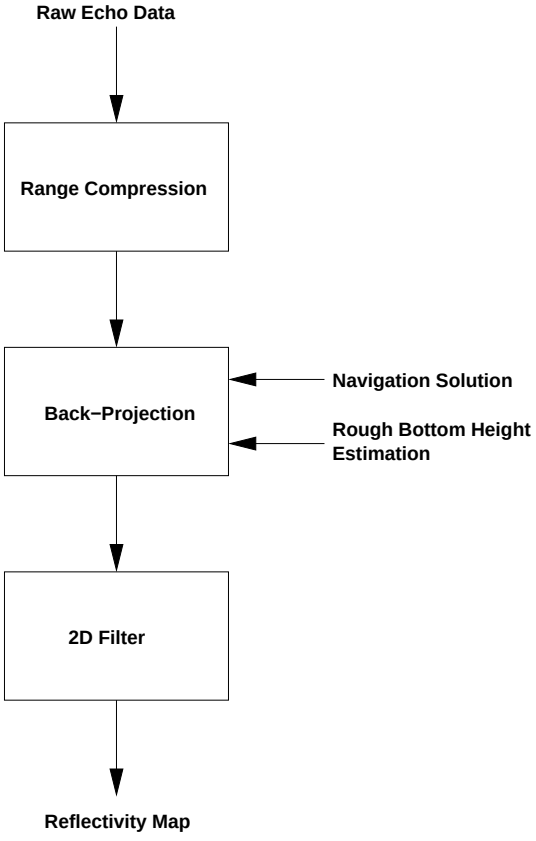

Figure 47: Back-projection algorithm signal-flow chart.

This formulation enables motion compensation for sway, heave and also surge. This last component is of special importance because although the ping-rate is constant the autonomous craft velocity is subject to variations. Yaw, roll and pitch corrections can also be

easily introduced. After the transducer position is calculated, these parameters only affect the weighting factor. The cross-track pulse compression is independent of the along-track pulse compression in the back-projection algorithm. Therefore, the signal processing chain can be divided in two main steps: cross-track pulse compression and along-track aperture compression (Figure 47). Furthermore, there are no constraints in using specific transmitted waveforms, such as a chirp or a random phase signal. An additional step of filtering can be carried out to remove artifacts such as speckle ([Chanussot, J. et al (2002)]) or improve the image sharpness ([Hungerbühler, N. (1998)]). A sharpening filter is, however, not suitable for cases were the signal-to-noise rate is low as it emphasizes the high frequency components of the spectrum and is only consequential for correctly focused images.

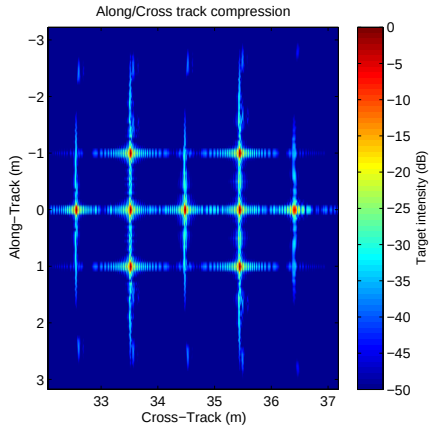

Figure 48: Back-Projection algorithm output.

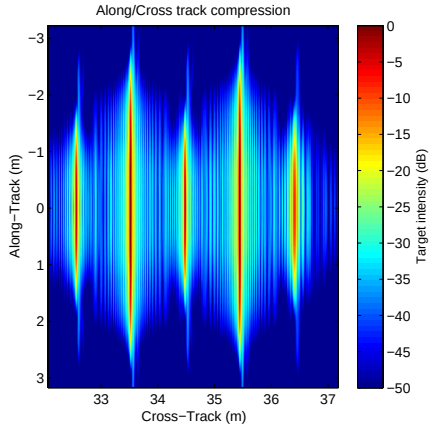

Figure 49: Back-Projection (non-coherent) algorithm output.

Note that only pixels ensonified by the ping have to be calculated. If the points to be considered are only those included in the 3 dB transducer aperture span, a substantial performance gain can be obtained for long survey tracks. If the transducer aperture is large

enough compared to the roll and yaw values and the temporal Doppler effect can be ignored ([Hawkins, D. W. (2004)]), the back-projection of each echo becomes symmetrical to the axis defined by the acquisition position and a small simplification can be made through which it is possible to calculate only half of the affected points. Nevertheless, this algorithm is no match for the frequency domain algorithms in terms of processing speed. On the other hand, frequency domain algorithms require small errors while following straight lines and are therefore not suitable for this application ([Fornaro, G. (1999), Silkaitis, J. M. (1995)]).

Apart from the described algorithm that can be regarded a coherent back-projection, it is also possible to employ an incoherent back-projection algorithm ([Foo, K.Y.; Atkins, P.R.; Collins, T. (2003)]):

$$ff(x,y_s)=\int_{\tau_1}^{\tau_2}\left|ss_b\left(\tau,t-\frac{2}{c}r(\tau,x,y,z)\right)\right|r^{\alpha}(\tau,x,y,z)w(\tau)d\tau \tag{5.47}$$

This algorithm trades along-track resolution for processing speed and a considerable gain in robustness to unknown platform motion and medium fluctuations. Also, the along-track sampling rate constrain is also lightened and can be higher than without creating aliasing artifacts. The along-track resolution becomes instead of $\delta_{AT}=D/2$:

$$\delta_{AT_{IC}}=\frac{4.5\delta_{XT}D}{\lambda} \tag{5.48}$$

As a by-product, the typical speckle in sonar images is greatly reduced.

The main advantages of the back-projection algorithm are related to its image quality and ease of incorporation of motion compensation factors at the cost of processing speed. This is true for implementations using standard CPUs. However, it is possible to use the back-projection algorithm in an elegant parallel implementation. Each echo back-projection can be computed separately and then summed together to reconstitute the image. This leads to a fast implementation using modern multiple processor system with a speed gain close to the number of processors used.

Figure 48 represents the 9 target reconstructed image using coherent back-projection algorithm and Figure 49 the reconstructed image with the incoherent back-projection algorithm. The resolution gain of the later compared to cross-track only compressed image is small although the target hyperboles are know straight.

5.7. Fast Back-projection

Real-time operation of synthetic aperture sonar and iterative parameter search for auto-focus algorithms, require computationally efficient implementations of the image synthesis algorithms. Back-projection enables flexibility in the way the sonar parameters and available information data is used, but suffers from computational inefficiency.

Looking into the back-projection summation in discrete along-track sampling time, it is trivial to see that the summation can be divided into sub-summations:

$$ff(x,y_s)=\sum_{i=1}^{N} ss_b\left(\tau_i, t-t_f\left(\tau_i,x,y,z\right)\right)e^{j2k_0 r(\tau_i,x,y,z)}=\sum_{i=1}^{N/P}\sum_{j=1}^{P} ss_b\left(\tau_{i+j}, t-t_f\left(\tau_{i+j},x,y,z\right)\right)e^{j2k_0 r(\tau_{i+j},x,y,z)} \quad (5.49)$$

Each summation has a lower along-track resolution that the total combined summations. This can be explored to achieve a processing gain by decimating the along-track image dimension to fewer points to which it is necessary to calculate the back projected pixel.
Several implementation of fast back projection exist, that trade off accuracy and quality in the image synthesis through a depth-of-filed narrowing approximation to achieve processing speed gains. They can be mainly divided in to fast polar back projection algorithms ([Shippey, G.; Banks, S.; Phil, J. (2005)]) and fast factorized back projection implementations ([Banks, S. M.; Griffiths, H. D. (2002)], [Hunter, A. J. ; Hayes, M. P. ; Gough, P. T. (2003)], [Ulander, L.; Hellsten, H.; Stenstrom, G. (2000)], [Ulander, L.; Hellsten, H.; Stenstrom, G. (2003)], [Callow, H.J.; Hansen, R.E.; Sæbø, T.O. (2006)]).

The difference lays in the way data is decimated. One obtains sub-summations arranged in a polar grid and decimates the angular positions. The other progresses through recursive steps were the image is divided in successive higher resolutions representations and fewer back-projected beams (one beam for each sub-summation in each step).

The approach followed in this work tried to maximize the image quality and input parameter flexibility by decimating only the along-track coordinates of each sub-summation image, obtaining the final image by interpolating and summation of the individual sub-images.

5.8. Algebraic image reconstruction

Algebraic image reconstruction algorithm ([Silva, S. (2008 a)] is an algorithm derived from back-projection image reconstruction with roots in medical imaging.

The integral expression used for the Explicit Matched filter is normally implemented as a summation, due to the discrete nature of the recorded echo data. The samples are interpolated to their correct position given by $t_f(\tau, x, y, z)$ and summed to form the synthetic aperture image.

To get a better insight into the algebraic reconstruction algorithm, seeing how interpolation can be implemented in matrix annotation is helpful.

A nearest neighbor interpolation can be written in matrix format in the following way:

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \end{bmatrix}' = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \tag{5.50}$$

Here the samples $[y_1, y_2, y_3, y_4, y_5, y_6]$ are obtained from $[x_1, x_2, x_3, x_4]$. The linear interpolation in following a similar manner can also be written in matrix notation:

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \end{bmatrix}' = \begin{bmatrix} a_{11} & 0 & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 \\ 0 & a_{32} & 0 & 0 \\ 0 & 0 & a_{43} & 0 \\ 0 & 0 & a_{53} & a_{54} \\ 0 & 0 & a_{63} & a_{64} \end{bmatrix} \times \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \tag{5.51}$$

With this in mind, the received echo data can be regarded as a combination of the original image pixels, and the value of a single pixel (y_j) can be given as:

$$y_j = \sum_i H_{ij} x_i = H_j x \tag{5.52}$$

Here x is swath reflectivity arranged in vector manner, column by column, and H_j a column vector whose non null elements contain the correct complex weights of the embraced swath. Extending the equation to the complete echo scene we get:

$$y = Hx \tag{5.53}$$

Where y is the echo data image also arranged in a vector, column by column, and H is a matrix with dimension $(m \times n)^2$. Here m is the number of along-track samples and n the

number of cross-track samples. This is a very large matrix, but also very sparse. The reconstructed image can thus be given by:

$$x = H^{-1} y \tag{5.54}$$

Instead of inverting the large matrix H which is only possible in special cases and generally will not provide the desired goal of obtaining the best signal to noise ratio, an approximation can be made which leads to:

$$x = H' y \tag{5.55}$$

H' is the transpose of the matrix H. This is equivalent to the discrete summation that implements the explicit matched filter. The matrix H can be regarded as the map of the elements of the image that contributed to one echo.

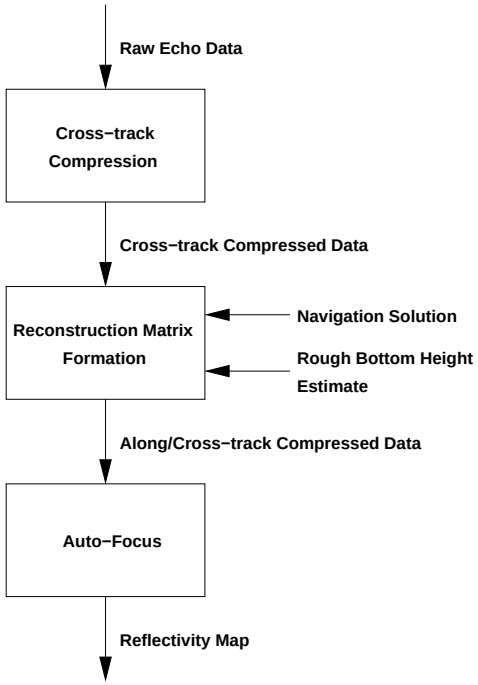

Figure 50: Arithmetic Reconstruction Signal Flow.

Its elements value and position are calculated through the navigation solution data and transducer parameters as dictated by the sonar model. So the navigation data is gracefully included in the image formation process, without any approximation regarding specific paths or transducer location being made.

Figure 50 shows the flow diagram of this sonar synthetic aperture processor. A solution to the above equation can also be obtained using a data fitting method by minimizing iteratively the image x that best fits the model H, knowing y:

$$\min_x \|Hx - y\| \tag{5.56}$$

This, however, does not provide results as good as the auto-focus method described in the following section.

5.8.1. Auto-Focus

As will be discussed in the following chapter, auto-focus is an important step in image reconstruction. The use of precise navigation data is fundamental for the reconstruction of synthetic aperture images. Nevertheless, in the case of high frequency sonar systems, it is very difficult to obtain the necessary navigation precision of $1/8$ of the wavelength even with the use of an integrated DGPS/IMU system. The algebraic reconstruction algorithm can be easily extended as to include an iterative auto-focus procedure.

One of the major advantages of the arithmetic reconstruction algorithm is to lend itself very well to estimation of the correct image formation parameters through the use of a global auto-focus algorithm.

Global auto-focus algorithms have several advantages. There is no need for any particular image characteristic such as a strong target or any statistical assumption on the type of clutter. It is also possible to obtain good results with along-track under-sampled data and only one transducer.

For efficient implementation of the auto-focus algorithm, instead of changing the elements of the matrix H, which is computationally very demanding, an auxiliary matrix of with complex weights is used (Figure 51):

$$x = H'Wy \tag{5.57}$$

The matrix W is diagonal matrix with only m distinct elements: one for each along-track echo. To further simplify this process only the phase of the complex weights is changed. Therefore the non-null elements of W are complex number of unitary absolute value and adjustable phase as a function of the image quality metric.

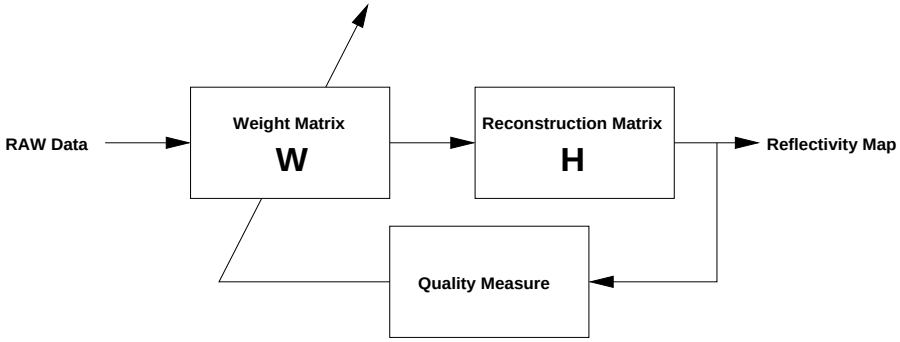

Figure 51: Algebraic reconstruction algorithm auto-focus procedure.

This naturally assumes that the matrix H is formed correctly through the use of navigation data to within one wavelength of the transited signal used.

The role of the auto-focus algorithm is thus to search the best set of phase adjustments that maximize some image quality parameter, within one wavelength. Since the major burden is in calculating the matrix H, the auto-focus iterations are very fast.

Several image quality measurements can be used, such as contrast or entropy, but it was found that entropy measurement and, in particular, quadratic entropy, is a more robust measure of image quality. The main advantage of quadratic entropy is that no specific assumption of the data probability density function is used. The entropy is calculated using an probability density function estimated through the available data using a Panzer window method as described in [Liu, W.; Pokharel, P. P.; Principe, J. C.. (2006)].

Figure 52 shows the cross-track compressed image, algebraic algorithm reconstruction output and output after auto-focus for the data collected in a test mission. The hyperbole obtained shows the strong reflectivity of an artificial target with little waviness, which is due to the good motion control of the autonomous vehicle. Without auto-focus the synthetic aperture reconstruction is not very successful.

The error of the navigation system is too high for a successful image synthesis with a signal of the frequency used in this system. Although it is possible to distinguish a target, the resolution is poor and the target is smeared in the along track direction. Using the described auto-focus step, after 25 iterations the image in the right was obtained. Here the image is perfectly reconstructed. The image quality limitation, in this case, is the side-lobes created by under-sampling in the along-track.

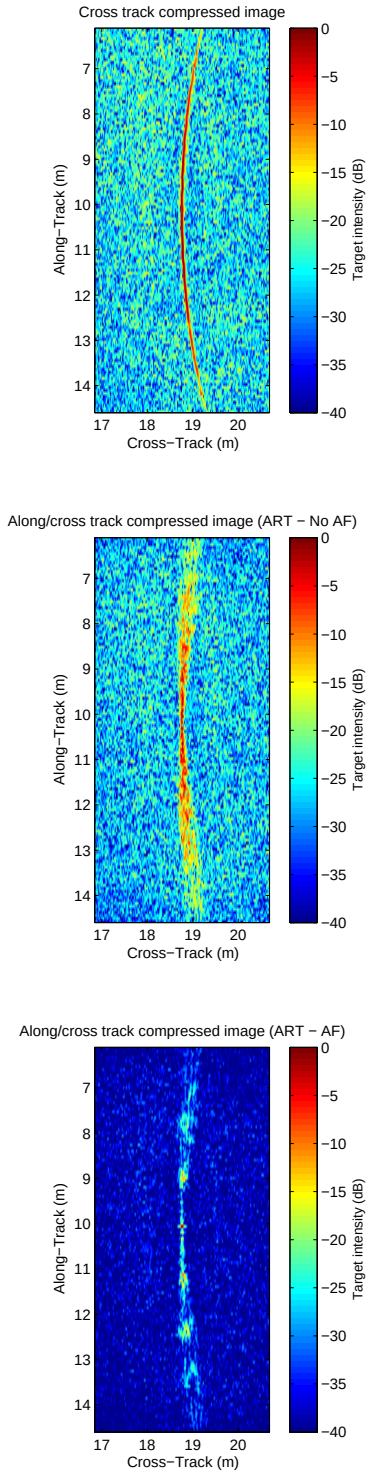

Figure 52: Algebraic algorithm output in the several steps: cross-track compression only; algebraic image reconstruction without AF, algebraic image reconstruction with AF.

5.8.2. Extension to a Large Data Set

Real life data sets might lead to reconstruction matrixes too big to be efficiently computed. In this case, the original echo data can be subdivided into several blocks with suitable overlap. For each block a matrix is created and auto-focus is performed on the respective data. The adjacent blocks use the previous overlapping auto-focus parameters to initialize their own auto-focus step.

5.9. Motion compensation

When deviations from the predefined linear path are known, it is possible integrate this information in the image formation algorithm, avoiding severe image defocusing.

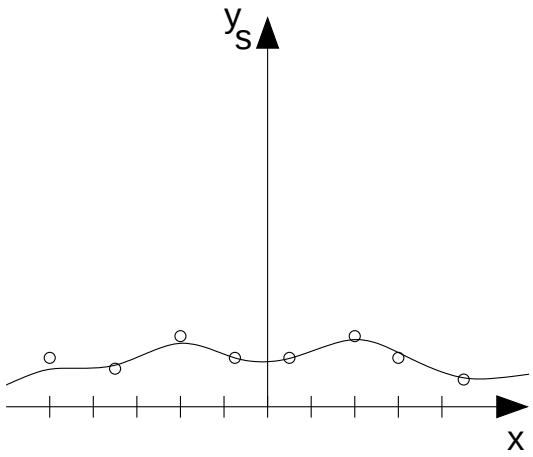

Figure 53: Representation of the sampling point variance.

Time domain reconstruction algorithms are able to compensate for arbitrary path trajectories ([Ulander et al (2000, 2001); Silva, S (2007 b)]). On the other hand frequency domain reconstructions algorithms assume a linear path and constant velocity, therefore any deviation from this assumption must be treated as a motion error. Compensation for these errors is not straightforward and might not even be possible to reduce to a suitable level, especially for single receiver and wide-beam systems.

Typically the echo data is realigned in a regular grid assuming a linear trajectory (Figure 53), before reconstruction through the frequency domain algorithm ([Banks, S. M.; Griffiths, H. D. (2002); Hawkins, D. W.; Gough, P. T. (1997); Hawkins, D. W. (1996); Silva, S. (2007 a); Wilkinson, D. R. (2001)]).

The received echoes model including the motion deviations from the linear path is:

$$ee_b(\tau,t) = \iint ff(x,y_s)\, p_b\left(t - \frac{2}{c}\sqrt{(x-x(\tau))^2 + (y-y(\tau))^2 + (z-z(\tau))^2}\right) \tag{5.58}$$

If we assume $z(\tau)$ and $x(\tau)$ is very small (thus assuming that the sway error as the dominant one), the above expression simplifies to (using the slant range notation):

$$ee_b(u-x(\tau),t) = \iint ff(x,y_s)\, p_b\left(t - \frac{2}{c}\sqrt{(x-x(\tau))^2 + (y_s-y_s(\tau))^2}\right) \tag{5.59}$$

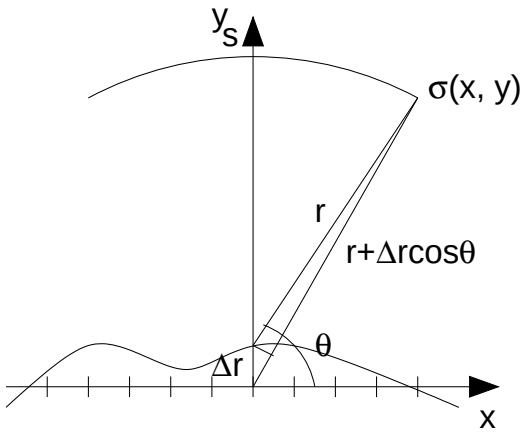

Figure 54: Correction of motion error.

This assumes an additional approximation: the amplitude of the variation must be small ($y_s(\tau) << y_s$). Furthermore for narrow-beam systems, $\cos(\theta) \approx 1$ and is possible to write the regular grid aligned echo data as:

$$ee_{b_{corrected}}(\tau,t) = ee_b\left(\tau, t - 2/cy_s(\tau)\right) \tag{5.60}$$

These motion-errors are dependent on the angle to the target, this dependency is more pronounced in wide-beam systems (Figure 54). This has to be taken into account in high-resolution wide-beam systems. The error of this approximation is:

$$\varepsilon \approx 2y_s(\tau)\left[1-\cos(\theta)\right] \tag{5.61}$$

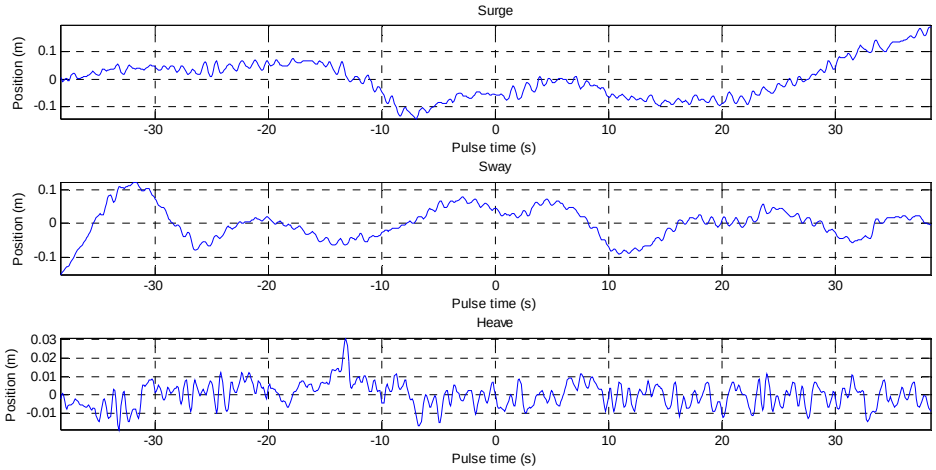

Figure 55: Autonomous craft position deviation during imaging mission.

It is not possible, in this particular sonar platform and especially for high resolution sonar system using short wavelength signals, to neglect heave and surge. During the boat motion, sway and surge are kept low by the boat motion control system, but heave is an important component in opposition to tow-fish based sonar systems, because of the waves on the water surface (Figure 55). Yaw variations can also be larger than in tow-fish based systems.

If we suppose small values for these motion errors we get:

$$ee_{b_{corrected}}(\tau,t)=ee_b\left(\tau-\frac{x(\tau)}{v},t-\frac{2}{c}\sqrt{\left(x-x(\tau)\right)^2+\left(y-y(\tau)\right)^2+\left(z-z(\tau)\right)^2}\right) \tag{5.62}$$

In the frequency domain:

$$EE_{b_{corrected}}(k_u,k)=EE_b(k_u,k)e^{jk_u x(\tau)}e^{j2ky_s(\tau)} \tag{5.63}$$

These corrections are applied through interpolation of the echo data or directly as phase shifts in the Fourier domain as is the case in the developed ISFFT algorithm.

In [Madsen, S. N. (2001)], an approach using space-variant filtering on local image patches is used to perform wide-band single-receiver motion correction.

For multiple receiver systems, wide-beam motion compensation complexity is reduced because many samples are collected for each sonar sampling position ([Callow, H. J. (2003)], [Silva, S. et al. (2009 a)]). In these systems the direction to target, for all targets, can be

estimated for each sample position. Furthermore, all receivers suffer the same sway error for a given ping.

By Fourier transform of the multiple-receiver data (along the receivers) to get a beam-space representation of the data, wide-beam motion correction can be applied ([Callow, H. J. (2003)]). The direct application of wide-beam techniques provides motion compensation for targets independently of their angle from bore-sight.

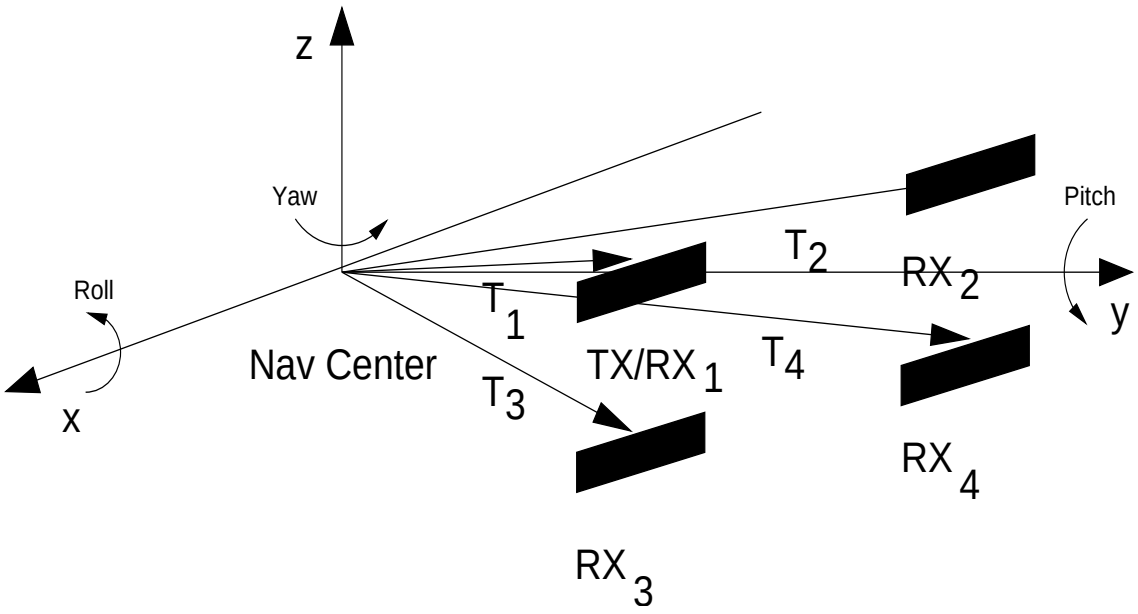

Figure 56: Transducer position calculation.

Attitude variations must also be accounted for in the calculation of the transducer positions. Knowing the arm between the boat navigation center (reference for the navigation system) and each transducer (Figure 56), and also the roll, pitch and yaw angles, the positions of the transducers can be correctly calculated by:

$$T = R \times B + N \tag{5.64}$$

Where R is rotation matrix associated to the roll, pitch and yaw angles, B is the arm vector between the transducer and the center of navigation of the boat, N is the absolute position vector given by the navigation system and T is the transducer position vector. The correct position of each transducer is then used by the image formation algorithm to synthesize the image.

The transducers attitude also influences the swath footprint. During image synthesis this information should be used to calculate the correct weights of the combined echoes.

Figure 57 shows a simulated boat positions sequence through a nominal linear path with the typical motion deviation suffered by this system and Figure 58 and Figure 59 the reconstruction of the simulated image using a frequency domain algorithm (ISFFT) and a time-domain algorithm (back-projection). Although it was possible to reconstruct to a high degree with the motion compensated ISFTT algorithm even in a presence of such high surge, heave and sway deviations (in relation to the system center wavelength), it is clear that back-projection performs much better in integrating the navigation data into the image reconstruction. Back-projection and algebraic image reconstruction, as a time-domain reconstruction algorithms, enable perfect motion-compensation at the cost of processing time. Motion induced phase errors cause the primary limitation of short-range (<100m) SAS imaging and as such must be kept bellow $\lambda/8$ ([Hayes, M. P.; Gough, P. T. (1992)]) for correct image reconstruction.

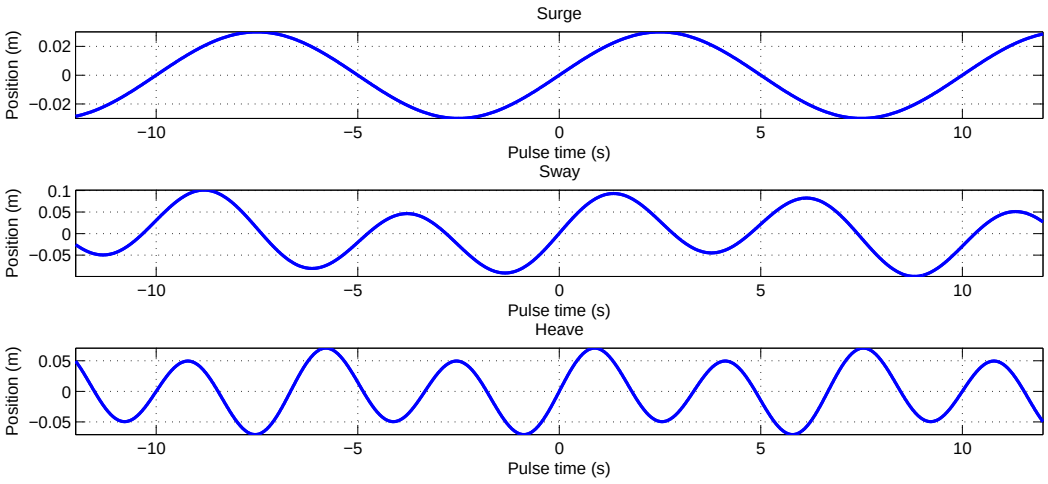

Figure 57: Simulated position of the autonomous surface craft.

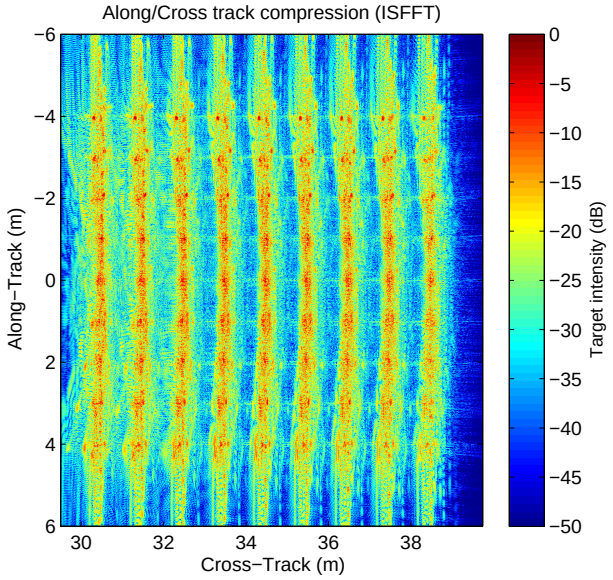

Figure 58: Reconstruction of the simulated image using ISFFT.

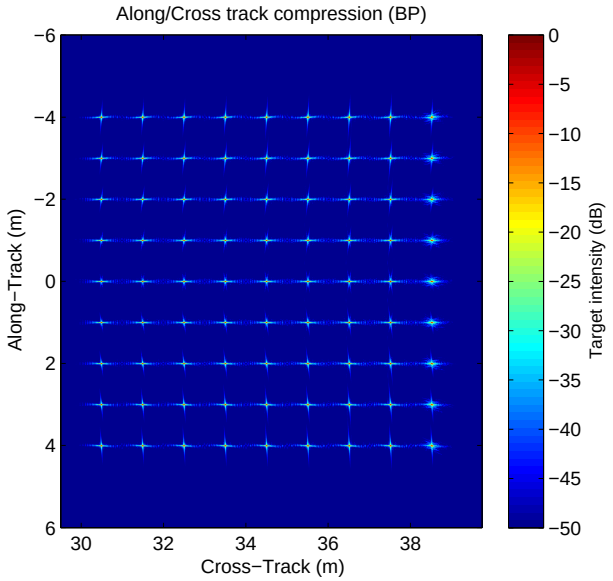

Figure 59: Reconstruction of the simulated image using back-projection.

Chapter 6: Image Enhancing

6.1. Auto-focus

Unknown motion components of the sonar platform and medium instability corrupt the phase of the echo signals leading to image blurring. Furthermore, the use of high frequency signals and the long total path length of the synthetic aperture mean that the sonar platform will deviate from its known position several wavelengths.

These phase distortion are one of the main obstacles to the widespread of synthetic aperture sonar. Autofocus addresses these problems of unknown phase distortions, by estimating them through the use of redundant information contained in the echo data or by means of optimization of a set of parameters under a certain suitable metric. The autofocus problem can be thought of a typical system estimation problem: estimate the unknown system using a random noise input. The system in this case is the unknown distortion phase caused by the path-derivation and random noise in the imaged submerged scene. Autofocus algorithms therefore exploit the redundancy of the phase distortion across the scene to estimate motion errors.

In creating the synthetic aperture, one assumes that the propagation medium is homogeneous and does not have variations in the speed of sound. When in presence of sound speed gradients, the sonar pulses travel along bent and twisted paths and the correct sound profile must be estimated and correct for to prevent blurred images ([Hansen, R. (2007)]). Moreover, medium instability leads to significant random phase delays. Medium induced phase-distortions are the limiting factor in long-range synthetic aperture sonar ([Chang and Tinkle (2001)]).

It was found that the accuracy of even the best current IMU (Inertial Measurement Unit), operating in standalone mode, is not enough to allow for diffraction-limited imagery of SAS ([Bellettini, A.; Pinto, M. A. (2002); Pinto et al (2002); Wang et al (2001)]). Navigation systems based on integration of GPS and inertial measurement units enable direct image formation as long as the wavelength is lower than the maximum RMS error of the navigation by an order of magnitude ([Silva, S. (2009 a)]).

Unknown path deviation of only a fraction of the wavelength ($< \lambda/8$) can cause the reconstructed synthetic aperture image to blur. The blurring caused by motion errors is dependent in the amplitude and how rapidly are the errors varying.

Motion induced phase errors are common to all targets in the sonar beam for a given along-track sampling position. In opposition medium fluctuations caused distortions vary significantly, different areas of the received echo image can have different types of distortion. This limits the area of the image with common phase errors and thus the size for the patch that can be auto-focused assuming common phase errors for each along-track sample.

We will define the procedure of autofocus has the one that enables the estimation of a set of parameters used in the synthetic image formation process that leads to an enhancement of a specific image quality measurement. These parameters can be directly connect to the sonar platform path, as is the case with the position or attitude estimation or be an arbitrary phase error as the one cause by the unknown portion of the platform position and medium instability.

Autofocus algorithms can be used to estimate the sonar path with the help of other sources of navigation data such as an IMU as is the case described for a multiple receiver system in [Pinto et al (2002)]. In this case it is common to refer to the auto-focus procedure as micro-navigation.

Micro-navigation algorithms such as RPC, DPCA and shear average ([Callow, H. J. (2003)], [Silva, S. et al (2009 b)]), exploit redundancy in the echo collection. In contrast, autofocus algorithms require certain statistical properties in the scene of interest to work correctly. When using multiple receiver system it is possible to perform DPCA navigation as described in [Silva, S. et al (2009 b)].

6.2. Echo-correlation auto-focus

Echo-correlation based auto-focus techniques rely on redundancy in the collected data to provide reliable motion estimates. SAS data must be oversampled in the along-track for the algorithms to operate (which is not possible with the described sonar system). Extensions relying on typical scene-based auto-focus information can remove this requirement at the expense of accuracy and algorithmic simplicity ([Callow, H. J. (2003)]).

Shear average is a simple-echo correlation auto-focus technique (probably better described as micro-navigation) that uses the redundancy from adjacent recorded echo signals to obtain a sway information estimate. This requires that the samples must be taken closer that the limit of $D/2$ imposed for SAS. Moreover, the submerged scene must translate into a random image of uniform amplitude meaning that the pulse compressed echo data of adjacent along-track echoes must be delta-correlated. The sway (δ_{y_s}) is thus estimated as:

$$\delta_{y_s} = \frac{\lambda_0}{4\pi} \sum unwrap\left\{ w(f)\, sS_b(\tau_i, f)\, sS_b^*(\tau_{i+1}, f) \right\} \tag{6.1}$$

In the presence of a strong scattering target, the assumption of delta-correlation breaks down and the path estimation is biased. The influence of strong targets can be ameliorated by applying a weighting function ($w(f)$) to the echo data ([Callow, H. J.; Hayes, M. P.; Gough, P. T. (2001)]).

Also, the amount of phase distortion that can be measured is limited by the amount of phase estimate that can be correctly unwrapped. Methods for dealing with this problem have been proposed ([Callow, H. J.; Hayes, M. P.; Gough, P. T. (2001)]).

The RPC algorithm, also known as DPC and the DPCA algorithm is similar to shear average in operation and is used for multiple receiver systems ([Silva, S. et al (2009 b), Groen, J.; Sabel, J. C. (2002)]).

A variation of the shear-average algorithm does non-coherent estimation of adjacent pulse-compressed sonar pulse displacement by calculating the sub-pixel position of the correlation maximum. Large path-deviations can be estimated since there are no issues with phase wrapping and the correlation length is extended, making this algorithm useful for under-sampled swaths. The sway displacement between adjacent echoes can be calculated as:

$$\delta_{y_s} = \sum \max d\left\{ F^{-1}\left\{ sS_b(\tau_i, f)\, sS_b^*(\tau_{i+1}, f) \right\} \right\} - y_{s_0} \tag{6.2}$$

6.3. Phase gradient auto-focus

Phase gradient auto-focus (PGA) selects point-like targets from every cross-track position in the synthetic aperture image generated using any known navigation information and priory phase error estimation. The residual phase error is estimated along the sonar path at those points and combined into an estimate of the total along-track phase error.

For the phase estimation only strong targets are used, the image is windowed and shifted so the targets occupy the same location in the synthetic aperture (Doppler history) which removes any linear phase error trend. The phase error estimation is then integrated and converted to sway displacement. This phase estimation is then used to create a new synthetic aperture image and the phase error estimation repeated until a suitable lower phase error bound is achieved.

Phase gradient auto-focus use narrow beam and narrow band assumptions that are not suitable for application in this described system. Moreover is not suitable for large motion

errors (the estimated errors are modeled as phase only) containing significant components in all coordinate directions (sway, heave and surge) or highly under-sampled swaths.

The interested reader is referred to [Callow, H. J. (2003)] for an excellent description and comparison of several PGA and PCA (Phase Curvature Auto-Focus) algorithms for both spot-light and strip-map synthetic aperture sonar.

6.4. Global optimization auto-focus

Global optimization provides a promising framework in which to perform SAS autofocus. This is an auto-focus technique that iteratively tries to maximize an image likelihood or metric function.

The problem this type of autofocus algorithms faces is in choosing a suitable image metric and avoid the problem of local solutions in the optimization process (solutions other than the optimum due to several local minima in the optimization surface).

In [Fortune et al (2001)], a contrast-based image metric has been used and a local Fourier magnitude uniformity (wavefront sensing) proved inadequate in [Callow, H. J.; Hayes, M. P.; Gough, P. T. (2002)].

The computational burden of this optimization type autofocus is high, making it difficult to apply and study.

This auto-focus technique is suitable for motion and medium induced phase errors. Several error parameters can be estimated simultaneously: heave, sway, surge and medium distortions such as sound propagation speed.

The approach followed in this work was to enhanced the synthetic aperture sonar images through an auto-focus algorithm that is based on a global optimization technique ([Silva, S. (2007 b)]) that uses the already high quality navigation solution to further refine the trajectory estimation and in turn, improve the obtained image. In fact, the algorithm itself can be regarded as a micro-navigation algorithm since the image quality measurement is exploited to better estimate the boat's position. Given the accuracy of the navigation system, the purpose of the auto-focus system is to correct unknown motion components below 1 cm level, elevating the image quality to the standards needed for high precision applications, such as interferometry.

Starting with the navigation solution provided by the GPS/IMU system an optimization algorithm (Nelder-Mead Simplex or Quasi-Newton method) is used to search throughout the solution space to find the position parameters that maximize the image quality. If the boat

movement has very low power spectral density at the higher frequencies, one can use the decimated positions or the coefficients of a fitting polynomial as optimization parameters. Otherwise all the coordinates must be used in the search algorithm resulting in a cumbersome optimization problem where it is convenient to fix the attention to an area of interest of the image in question.

The image quality estimate is the quadratic entropy measurement ($\hat{H}_2(x)$):

$$\hat{H}_2(x) = -\log IP(x) \tag{6.3}$$

This is a measure of the image x sharpness. A low entropy value means a sharper image. To calculate the quadratic entropy one needs to estimate the image information potential $IP(x)$. Instead of making the assumption that the image intensity has a particular probability density function (uniform, Gaussian or negative exponential [Ruderman, Daniel L. (1994)]), the probability density function is estimated thought a Parzen window method using only the available data samples ([Liw, W. et al (2006)]):

$$IP(x) = \frac{1}{N^2}\sum_{j=1}^{N}\sum_{i=1}^{N} k_\sigma(x_j - x_i) \tag{6.4}$$

Where $k_\sigma(x - x_i)$ is the Gaussian kernel defined as:

$$k_\sigma(x - x_i) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-x_i)^2}{2\sigma^2}} \tag{6.5}$$

The parameter σ controls the width of the Gaussian kernel, large values are necessary for scarce data while small values lead to better estimation results. Because this method of estimation requires a computationally intensive calculation of the sum of Gaussians, this is implemented through the Improved Fast Gaussian Transform described in [Yang, C. (2003)].

We know beforehand that the solution cannot be very far from the initial position estimate. So a similarity measure is used between the estimated path $p(x, w)$ (which is a function of the input data x and the adjustment parameters w) and the measured path y. The total objective cost also includes this path deviation cost which acts as a regularization factor of the image cost function and eases convergence of the solution. For this similarity measure the correntropy function described in [Liw, W. et al (2006)] was used, as it does not assume any particular characteristics about the path of the boat:

$$\min P(w) = -\frac{1}{N}\sum_{i=1}^{N} k(p(x_i, w) - y_i) \tag{6.6}$$

A schematic of the autofocus procedure is shown in Figure 60. Performing an image optimization in this way does not rely on strong targets or any other particular image characteristic or statistical assumption. It also has a good degree of robustness to along-track under-sampling.

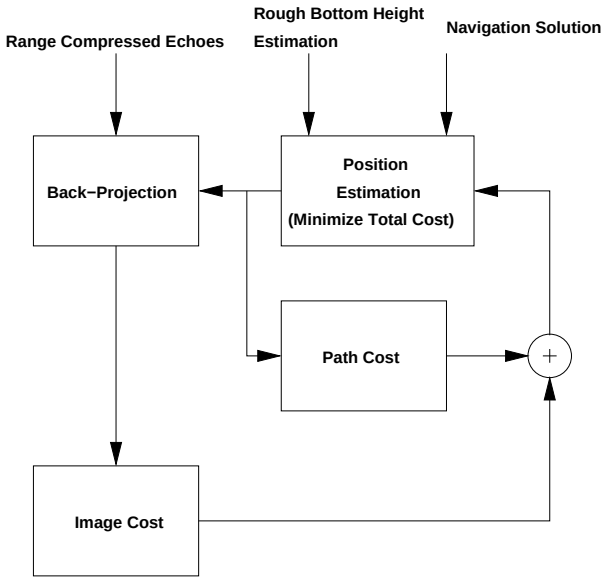

Figure 60: Auto-focus block diagram.

6.5. Delta displacement

The back-projection algorithm uses the available navigation solution to place the echo data in the correct image coordinates. Assuming that navigation errors are small compared to the range to the center of the image, further adjustments to the echo data placement can be made using only a small adjustment to the mean range (Figure 61). So, each along-track sampling position has it space coordinates and a variable, Δr, that affects all the data that in the cross-track coordinates. This adjustment is responsible for the correct integration in the along-track direction.

These can also be regarded as phases only adjustment. In this case no change in range direction is made and only a small change to the phase of the echoes is applied before image reconstruction. Doing so it is possible to correct motion uncertainties induced phase errors

(although only those smaller than a wavelength) and also phase errors due to medium fluctuations and others.

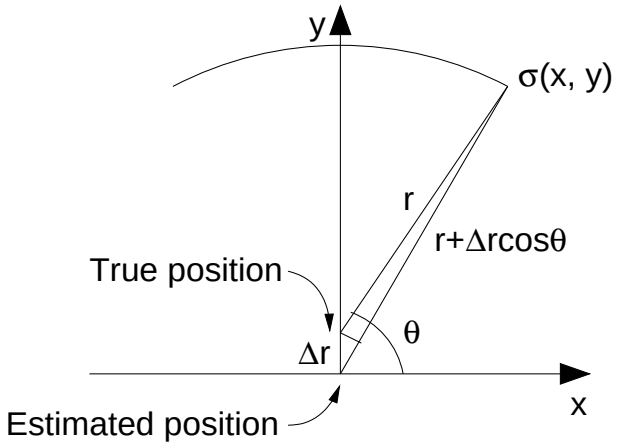

Figure 61: Position adjustment approximation for auto-focus.

6.6. Dimension reduction through PCA

To form the synthetic aperture image we have to know the coordinates and attitude of the transmitter transducer and receiver transducer to a very high level of precision. The large dimensionality of the auto-focus adjustment parameters makes the optimization process very burdensome. This becomes even a greater problem if instead of using just small range or phase adjustment, the complete coordinate vector of each sampling position is used.

The coordinates have a very high correlation between them, since transmitting transducer is mechanically coupled to the receiving transducer and the attitude relates the position of the transducers. So it is possible to obtain a representation space with smaller dimensionality that is still able to correctly represent the entire set of transducer coordinates with minimum error.

Using the strong correlation between dimensions in the navigation data, a method known as Principal Component Analysis ([Jolliffe, I. T. (2002)]) can be used to reduce its representation set ([Silva, S. et al (2009 c)]). From the position data, any linear variation term is removed and the covariance matrix calculated. The higher energy Eigen values and corresponding Eigen vectors are chosen to represent the data with a resultant reduced dimensionality. To reconstruct the data, the chosen Eigen vectors and covariance matrix are then used, and the linear variation term added back to the data.

Starting with the navigation data for the position of the transmitting and receiving transducer (T), the data is written as the variation from a mean straight line (T_L):

$$T_L = \begin{bmatrix} x_{TX} \\ y_{TX} \\ z_{TX} \\ x_{RX} \\ y_{RX} \\ z_{RX} \end{bmatrix} - \begin{bmatrix} v\tau \\ 0 \\ 0 \\ v\tau \\ 0 \\ 0 \end{bmatrix} \tag{6.7}$$

After this we have to subtract the mean value from each of the data dimensions:

$$T_O = \begin{bmatrix} x_{TX} \\ y_{TX} \\ z_{TX} \\ x_{RX} \\ y_{RX} \\ z_{RX} \end{bmatrix} - \langle T_L \rangle \tag{6.8}$$

By calculating the covariance matrix C of this data, a new representation of the data can be obtained. The covariance matrix can be written with the help of its Eigen vectors and Eigen values:

$$C = VDV^{-1} \tag{6.9}$$

Where V and are the Eigen vectors of C and D is a diagonal matrix of the Eigen values of C. The matrix T_O can thus be expressed in the space defined by the Eigen vector of the covariance matrix, by doing so the greatest variance by any projection of the data comes to lie on the first coordinate, the second greatest in the second coordinate and so on. If we now only take only a subset of the Eigen vectors (W) of higher energy content, we obtain a compact representation of T_O, T_C:

$$T_C = WT_O \tag{6.10}$$

The estimation of the receiver and transmitter positions is obtained by adding back the mean and linear horizontal position variation to $\hat{T}$:

$$\hat{T} = W^{-1}T_C + \langle T_L \rangle + \begin{bmatrix} v\tau & 0 & 0 & v\tau & 0 & 0 \end{bmatrix}' \tag{6.11}$$

6.7. Discrete Cosine Transform variable reduction

Reducing the dimensionality of the variables is an important step for the auto-focus to be able to efficiently estimate the correct transducer positions. When dealing with very large synthetic apertures (long tracks), the number variables (at least one for each sampling position) is still very high. The surface craft moves slowly and tries to follow a predefined path, so there should be a high degree redundancy in the along-track position data. The problem here is to obtain a set of functions that can describe the data obtained with the PCA method, but using a smaller number of variables. We can consider this data ($x[n]$) as a sum of sinusoidal functions ($X[k]\cos(\omega_k)$):

$$x[n] = \frac{X[0]}{N} + \frac{1}{N}\sum_{n=1}^{N-1} X[k]\cos\left[\frac{\pi}{N}k\left(n+\frac{1}{2}\right)\right] \tag{6.12}$$

Doing so, we can calculate the discrete cosine transform (DCT, [Khayam, S. A. (2003)]) of the data and use only the coefficients ($X[k]$) of this result that have amplitude above a suitable threshold (one that leaves enough coefficients to represent the data with an error bellow the target estimation error):

$$X[k] = \sum_{n=0}^{N-1} x[n]\cos\left[\frac{\pi}{N}k\left(n+\frac{1}{2}\right)\right] \tag{6.13}$$

This greatly reduces the number of variables the auto-focus routine has to search providing a faster and surer convergence to the correct solution ([Silva, S. et al (2009 c)]).

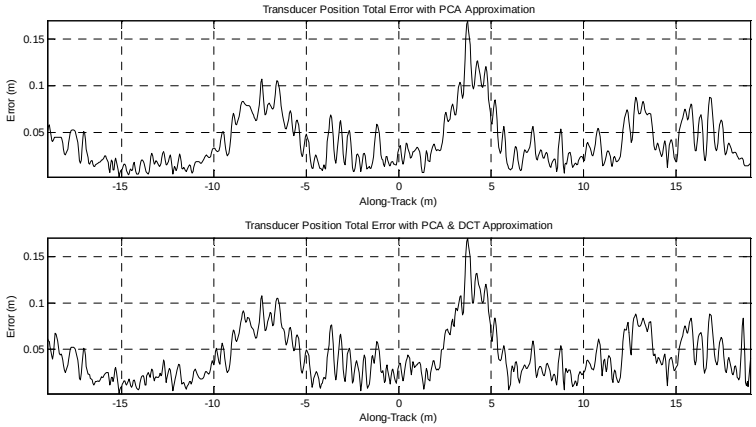

Figure 62: PCA and DCT transducer position approximation error.

Figure 62 represents the total error of the position approximation, calculated as the square root of the sum of the squares between the original position and the approximated position at each along-track sampling instance. In this case the data dimensionality was reduced from 12 (position and attitude for each transducer) to 1 and 16 DCT coefficient were used. As can be seen the approximation is suitable for synthetic aperture image formation with wavelength higher than 20 cm. Other reduction factors can lead to smaller errors.

6.8. Bias estimation

The problem of navigation data fusion can be regarded as a problem of estimating the low frequency inertial errors given some known positions.

Common algorithms use Kaman filter to produce these estimations. Kalman filter is a good solution for real-time systems but for off-line data processing other solutions exist.

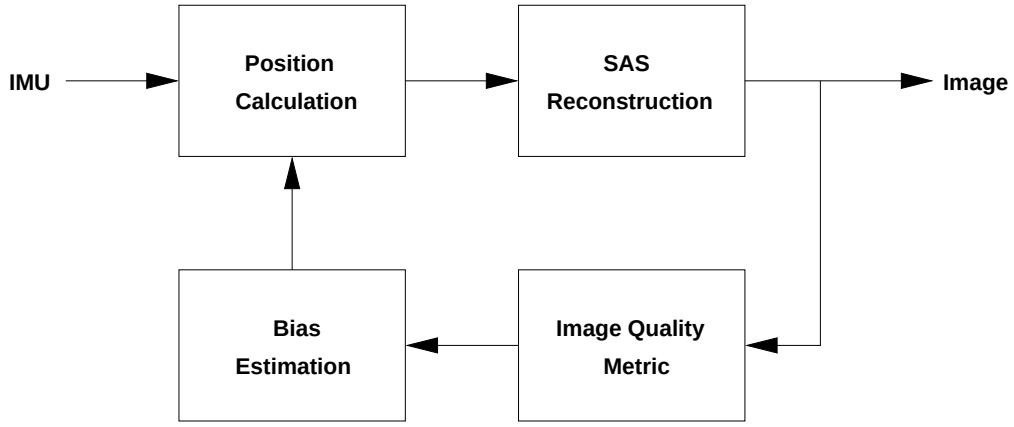

Figure 63: Schematic for IMU bias estimation through AF.

Bias estimation can also be made using the sonar data (Figure 63). Here the auto-focus algorithm is used to find a set of bias parameters that maximize and image quality metric. First an initial estimation of the bias parameters is found by traditional techniques such as the Kalman filter or fitting by minimum square error fitting. The bias estimation is then iterated in the auto-focus loop until a suitable measure of the image quality is achieved.

6.9. Sub-band processing

As see in the previous section, motion errors are critical to synthetic aperture image formation. Path deviations can be effectively corrected through the use of a back-projection algorithm if each along-track sampling position is known through the use of precise navigation systems such as a RTK-DGPS/IMU or a DPCA/IMU. Nevertheless, the use of high-frequency signals

imposes very tight restrictions to the uncertainty of the along-track positions given by these systems.

The navigation errors are hardly ever smaller than the original signal wavelength, and so create a solution surface that is difficult to search for the optimum set of parameters. However, if we have access to the raw data, by dividing the received signal bandwidth in several smaller bands and conjugate complex multiplying the pulse compressed signals obtained in each band one by the other, a new resulting signal is obtained with an effectively longer wavelength corresponding to the frequency difference between the two sub-bands. This longer wavelength effectively reduces the impact of phase fluctuation from the medium and platform motion uncertainties ([Silva, S. (2008 b)]).

Since the wavelength is longer, the aliasing effects that occur when the swath is under-sampled in the along-track direction are attenuated or even disappear if the new wavelength meats the new sampling criteria.

The goal of sub-band processing ([Silva, S. (2008 b)]) is to reduce the effects phase errors due to motion or other sources in very susceptible signals such as the ones with very short wavelength.

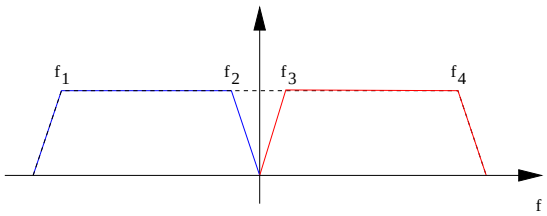

Figure 64: Chirp signal spectrum division schematic.

Considering a chirp signal with a spectrum as depicted in Figure 64:

$$p_m = rect\left(\frac{t}{T}\right)\cdot e^{j2\pi\left(\frac{f_4-f_1}{T}\right)t^2} \tag{6.14}$$

It is possible to divide the signal in several parts and thus its bandwidth into several sub-bands. Without loss of generality the case of the division by 2 is considered here:

$$\begin{cases} p_{m_1} = rect\left(\dfrac{t+T/4}{T/2}\right)\cdot e^{j2\pi\left(\frac{f_2-f_1}{T/2}\right)t^2} \\ p_{m_2} = rect\left(\dfrac{t-T/4}{T/2}\right)\cdot e^{j2\pi\left(\frac{f_4-f_3}{T/2}\right)t^2} \end{cases} \tag{6.15}$$

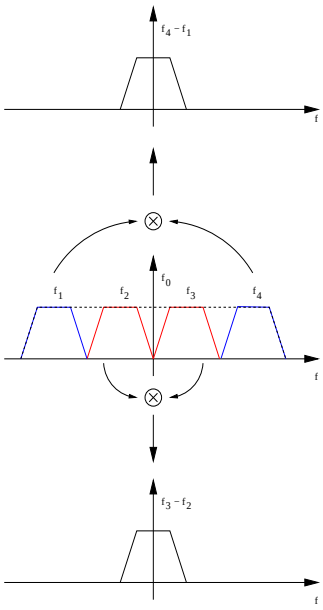

Figure 65: Grouping of the several sub-bands to form new signals with longer wavelength.

Each of these signals will now have its own center frequency. Using these signals to match filter the echo signal and after frequency re-centering, a point target will have an echo response given by:

$$\begin{cases} ss_{b_1} = \left\| ss_{b_1} \right\| e^{-j2\pi \langle f_2, f_1 \rangle t_f(\tau, x, y_s)} \\ ss_{b_2} = \left\| ss_{b_2} \right\| e^{-j2\pi \langle f_4, f_3 \rangle t_f(\tau, x, y_s)} \end{cases} \tag{6.16}$$

This new signal will necessarily have lower cross-track resolution which is given by its lower bandwidth ($\Delta f'$):

$$\delta_{XT} = \frac{c}{2\Delta f'} \tag{6.17}$$

By conjugate multiplying one by the other, one obtains a signal such that its wavelength corresponds to the beat frequency of the central frequencies of the two sub-bands:

$$\begin{aligned} & ss_{b_e} = ss_{b_2} \cdot ss_{b_1}{}^{*} = \\ & = \left\| ss_{b_2} \right\| \cdot \left\| ss_{b_1} \right\| \cdot e^{-j2\pi(\langle f_2, f_1 \rangle - \langle f_4, f_3 \rangle) t_f(\tau, x, y_s)} = \\ & = \left\| ss_{b_e} \right\| \cdot e^{-j2\pi f_e t_f(\tau, x, y_s)} \end{aligned} \tag{6.18}$$

This new equivalent (and longer) wavelength (f_e, λ_e) enables higher robustness to phase errors.

As can be seen in the schematic (Figure 65), the original echo data is divided into several bands (four in this example). Each sub-band originates a new set of echo data which is frequency reentered and pulse compressed with the corresponding chirp signal. Combining the each set of echoes one obtains two equivalent signals with longer wavelengths than the original one ($f_4 - f_1$ and $f_3 - f_2$). With different divisions of the original bandwidth it is possible to obtain different wavelengths.

Besides the obvious cross-track resolution loss, the along-track resolution also decreases. This is because the longer wavelength is not compensated by a longer along-track integration path width is maintained. As can be seen in Figure 66, the original synthetic aperture length is not changed (the target is still visible through [τ_1, τ_2]) and the new Doppler frequency is:

$$f_{D_e} = \frac{2v_r}{\lambda_e} = \frac{2v\theta(\tau)}{\lambda_e} =, \tau \in [\tau_1, \tau_2] \Rightarrow \frac{2v\theta_{3dB}}{\lambda_e} \leq \frac{2v\theta_{3dB}}{\lambda} \tag{6.19}$$

The equivalent center frequency is lower, which gives rise to a lower Doppler shift (f_{D_e}) and thus a lower resolution.

The new along-track resolution (δ_{AT_e}) is inferior by a factor equal to ratio between the original wavelength and the actual effective wavelength:

$$\delta_{AT_e} = r_0 \frac{\lambda_e}{D_{SA}} \tag{6.20}$$

Figure 67 shows the result of a simulation where a single target was processed using the original data and the data after sub-band division in two bands. The original data center frequency and bandwidth is respectively 200kHz and 30kHz. The cross-track resolution is lower by a factor of 2 and the along-track resolution lower by a factor of, approximately, 13.

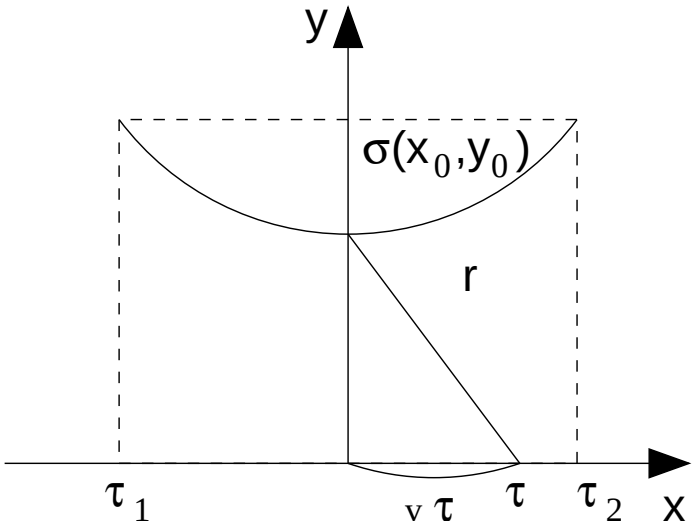

Figure 66: Synthetic aperture length.

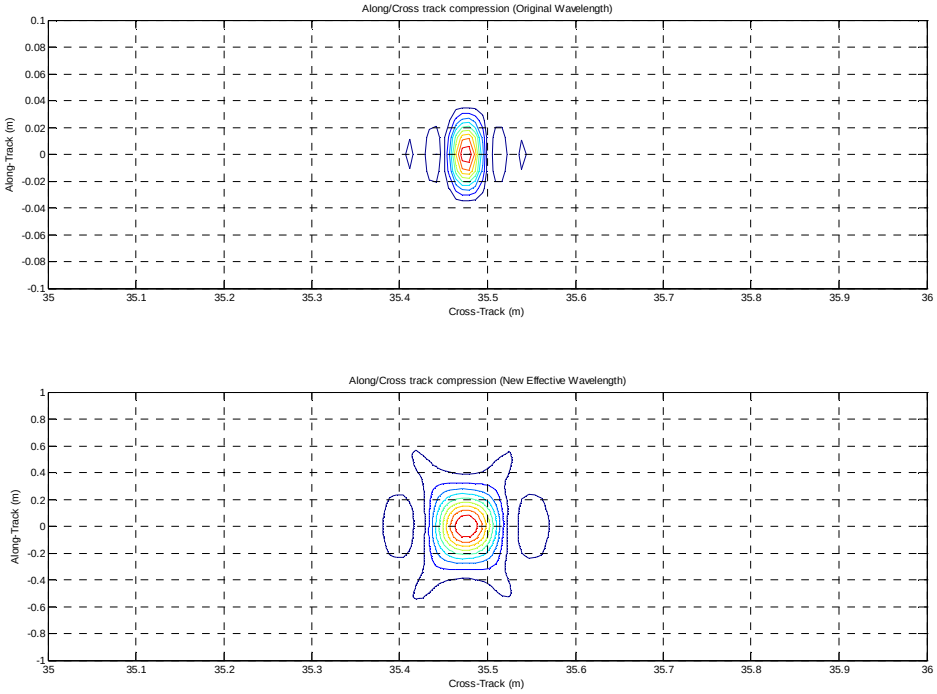

Figure 67: Resolution decrease in along/cross track when using sub-band processing.

Since the wavelength is longer, the aliasing effects that occur when the swath is under-sampled in the along-track direction are attenuated or even disappear if the new wavelength meats the new sampling criteria. This can be seen in the simulation results present in Figure 68.

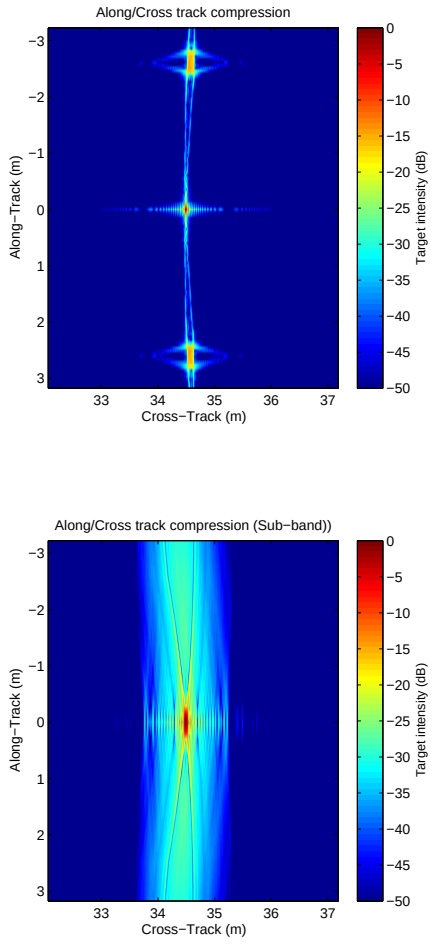

Figure 68: Single target under-sampled in the along-track direction: top, no sub-band processing and bottom, with sub-band processing.

Comparing a simulation of a set of targets processed with sub-band method, it can also be seen that there are no significant differences if the data at the original wavelength is sampled in the along-track dimension correctly or not as long as the sub-band processed data meets the sampling criterion (Figure 69). This is again a result of the fact that the integration time is not modified but the Doppler frequency is decreased. Reconstructing the image in the original wavelength would create grating lobes due to aliasing.

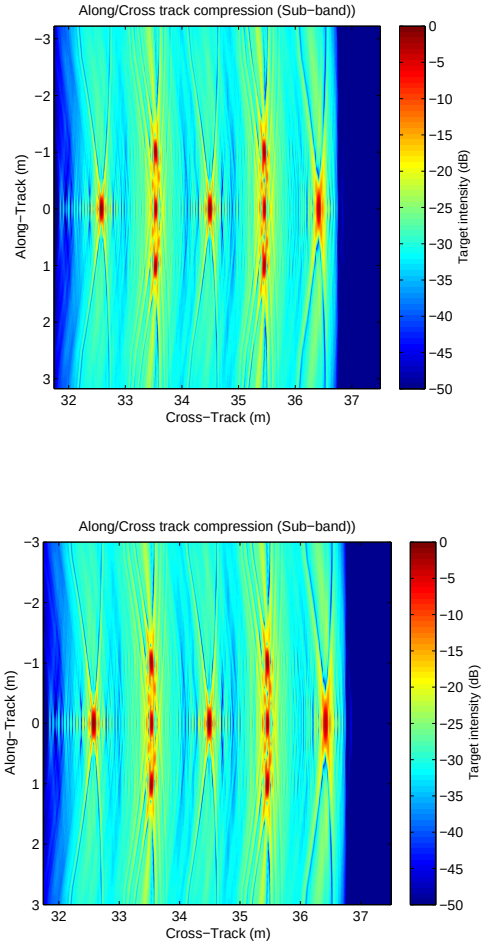

Figure 69: Set of targets: top under sampled and processed with sub-band; bottom correctly sampled but also processed with sub-band technique.

In a scene with multiple targets, combining the signals of the two sub-bands will create phase relations other than the desired ones. Considering for ease of analysis the case of sub-band processing a scene with two point targets, we get the following expression for the cross-track compressed sub-band signals:

$$\begin{cases} ss_{b_1} = e^{-j2\pi\langle f_2, f_1\rangle t_{f1}(\tau,x,y_s)} + e^{-j2\pi\langle f_2, f_1\rangle t_{f2}(\tau,x,y_s)} \\ ss_{b_2} = e^{-j2\pi\langle f_4, f_3\rangle t_{f1}(\tau,x,y_s)} + e^{-j2\pi\langle f_4, f_3\rangle t_{f2}(\tau,x,y_s)} \end{cases} \tag{6.20}$$

For notation simplicity the amplitude was dropped and the signal combination of these signals yields:

$$ss_{b_e} = ss_{b_2} \cdot ss_{b_1}{}^{*} = e^{-j2\pi\left[\langle f_4, f_3\rangle - \langle f_2, f_1\rangle\right]t_{f1}(\tau,x,y_s)} + e^{-j2\pi\left[\langle f_4, f_3\rangle - \langle f_2, f_1\rangle\right]t_{f2}(\tau,x,y_s)} + \\ + e^{-j2\pi\left[\langle f_4, f_3\rangle t_{f1}(\tau,x,y_s) - \langle f_2, f_1\rangle t_{f2}(\tau,x,y_s)\right]} + e^{-j2\pi\left[\langle f_4, f_3\rangle t_{f2}(\tau,x,y_s) - \langle f_2, f_1\rangle t_{f1}(\tau,x,y_s)\right]} \tag{6.21}$$

As it can be seen, this creates cross-products with a phase relations not given by $2\pi\left[\left\langle f_4, f_3\right\rangle - \left\langle f_2, f_1\right\rangle\right]t_f(\tau, x, y_s)$ which will degrade the synthetic image formation. But because their phase does not correlate to the matching filter hyperbole at the image coordinates, these cross-products will also be strongly attenuated. In the final image, the consequence of this will be that the image clutter level will be higher than if not using sub-band processing (smeared point target tails in Figure 69).

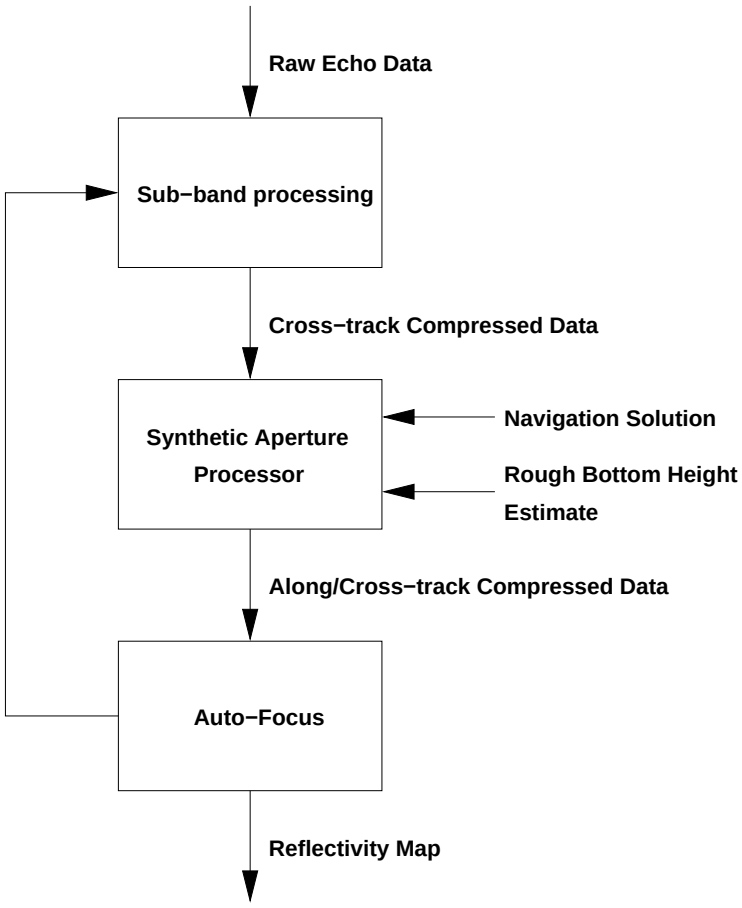

Figure 70: Use of the sub-band algorithm within a synthetic aperture processor.

As sub-band processing is a step that leads to the pulse-compressed images, it is possible to use this method with any synthetic aperture processor that can use pulse compressed

data. This includes the family of back-projection algorithms, the wave-number algorithm and the inverse scaled Fourier transform algorithm.

Figure 70 shows a signal flow diagram that situates the sub-band algorithm within a synthetic aperture processor.

6.10. Auto-Focus with Sub-Band Processing

The use of global auto-focus algorithm presents several advantages for synthetic aperture sonar image enhancing. Common auto-focusing algorithms require restrict along-track sample rates equal or higher than the Nyquist sample rate.

Global auto-focus algorithms correct not only phase errors due to navigation uncertainties, but also phase errors that are due to medium fluctuations. Nevertheless, these errors are seldom smaller than the original signal wavelength, and so create a solution surface that is difficult to search for the optimum set of parameters.

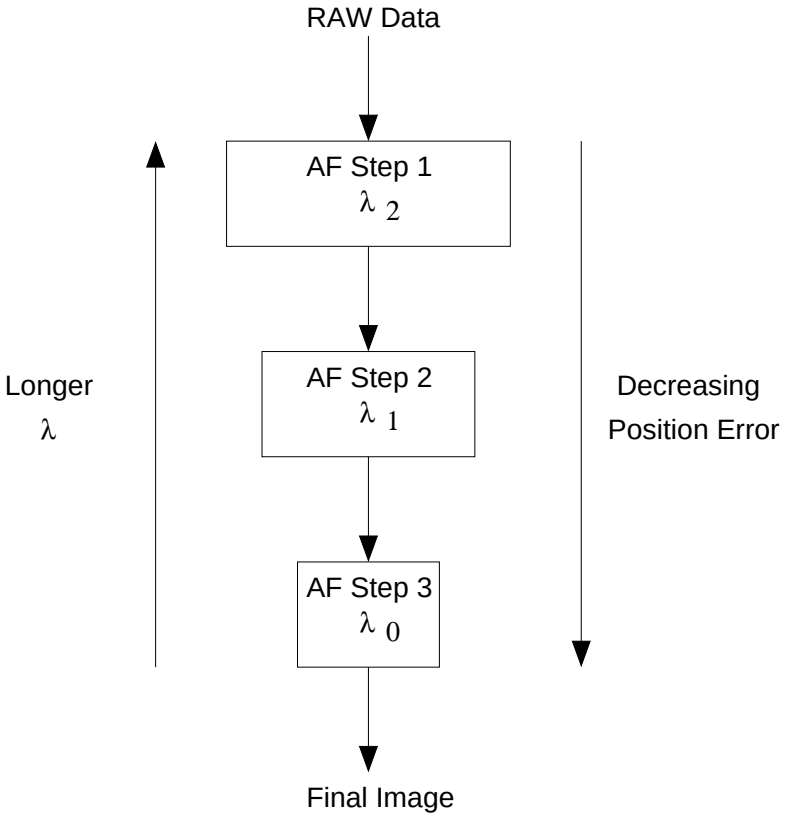

Figure 71: Auto-focus procedure with sub-band processing.

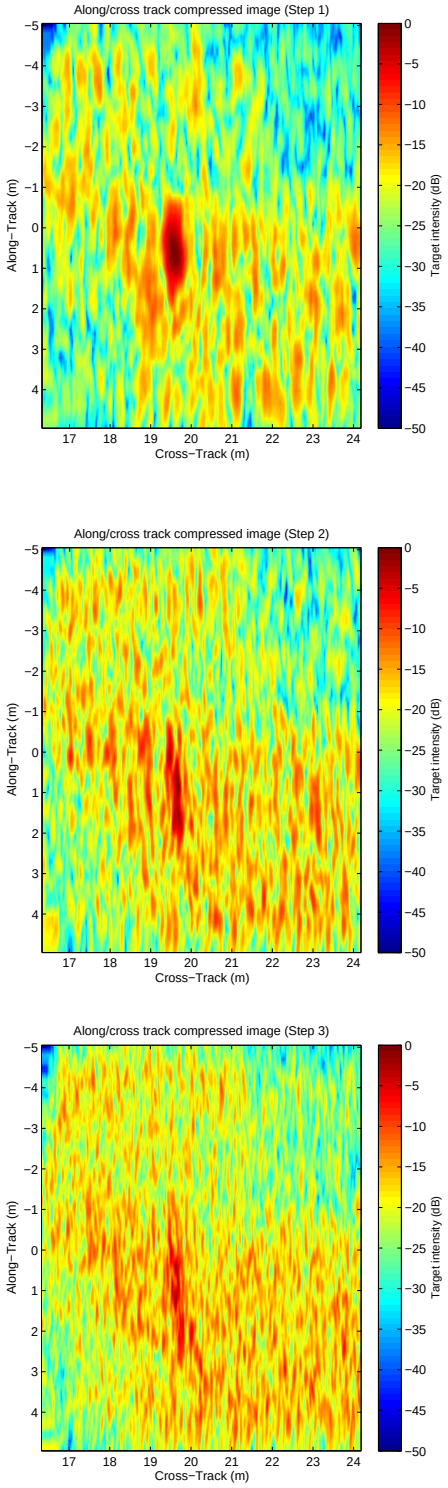

Figure 72: Sonar image of the artificial target through the various auto-focus steps: no autofocus, autofocus step 1, autofocus step 2, autofocus step 3.

Using sub-band processing, it is possible to divide the signal bandwidth into several sub-bands and combine them in to signals with different wavelengths.

At the first step, a large wavelength is used since the expected phase errors are also large. After achieving a predefined level of image quality, the auto-focus algorithm then proceeds by using a smaller wavelength and the previous estimated position parameters.

This step is repeated with decreasingly smaller wavelength and position error, until the original wavelength is used (Figure 71). The result is a faster progression through the solution surface, with lower probabilities of falling into local minima.

For image quality measurement, quadratic entropy is used, which provides a robust metric and enables faster convergence properties.

This auto-focus method is suitable for system working with an under-sampled swath and few transducers. No special image features are necessary for the algorithm to converge.

Figure 72 illustrates the result of the application of the described auto-focus procedure to a sample collected data by the sonar system described here, featuring a strong point-target. 3 successive auto-focus steps are shown. The algorithm starts with a longer wavelength thus producing a low resolution image. As it progresses through the process, the target gets a sharper appearance.

6.11. Interferometry with Sub-Band Processing

One of the major issues with height mapping estimation using interferometry is the correct phase unwrapping under low signal to noise ratio, especially in cases where there are many wraps.

This problem is more complex with smaller wavelengths, since they will produce more wraps for the same height variation.

As it will be explained in the next chapter and given by Figure 73, the height relates to the depression angle (ψ) and then to the measured range difference (ΔR) by:

$$(R_1 + \Delta R)^2 = R_2^{\,2} + B^2 - 2BR_2 \cos(\psi + \beta) \tag{6.22}$$

The range difference that enables the height computation is given by:

$$\Delta R = \frac{(\phi + 2\pi n)}{2\pi}\lambda \tag{6.23}$$

Here ϕ is the measured phase difference and n the number of integer wraps. So we see that for a ΔR resulting from the same height variation, the longer the wavelength, the smaller the phase variation and thus fewer wraps around 2π.

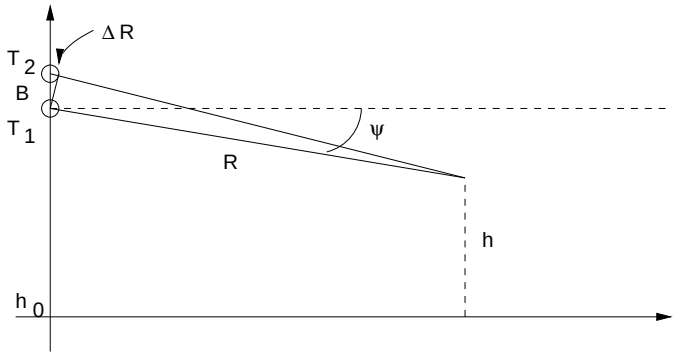

Figure 73: Schematic of the interferometric height mapping model.

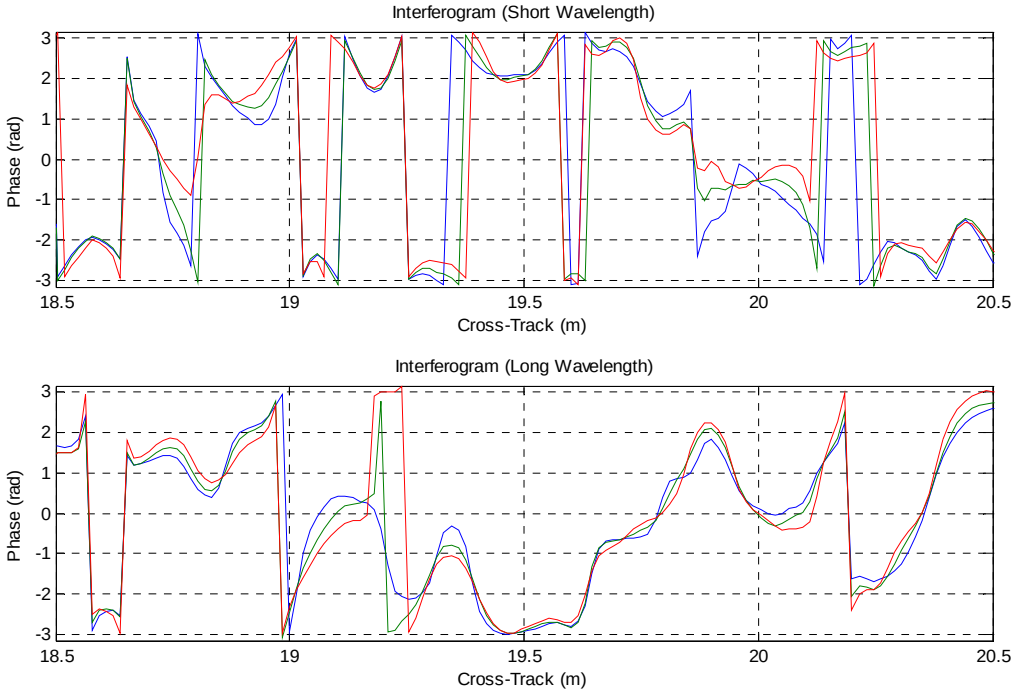

Figure 74: Two interferograms: short wavelength signal and long wavelength signal.

Figure 74 is presents two interferograms, in each graphic adjacent consecutive along-track lines are displayed in blue, red and green respectively. Both were produced using synthetic aperture images obtained in a test mission and derived from sub-band processing but with different wavelengths. It is evident here that with a longer wavelength the number of total

wraps is smaller for the same extension and height variation, thus easing the subsequent unwrapping step.

6.12. Speckle-reduction

In SAS images, the brightness distribution is not smooth and continuous, as careful inspection will show. Instead it is composed by complicated granular patterns of bright and dark spots that are called speckle. This is phenomenon that occurs in any form of coherent imaging where objects illuminated by coherent radiation have surface features that are rough on the scale of the illuminating wavelength.

Speckle results from the combination of many scattering centers within a given resolution cell, for this reason it bears little resemblance to any large-scale scattering properties of the ensonified terrain. As it consists of a superposition of randomly phased, diffraction-limited impulse-response functions, the nominal speckle size is a good measure of the achievable diffraction-limited resolution of a particular image. It is interesting that this phenomenon exists whether or not the coherent image is focused. Since speckle is formed from random phase scatterers, focusing only affects the phase of coherent scatters over the aperture and cannot affect the statistical properties of waves scattered by a rough surface. Therefore, the speckle size in a SAS image will indicate the achievable resolution. Qualitative comparison of the smallest speckle with the sharpness of a point reflector will indicate whether the diffraction-limited resolution has been achieved.

The probability density function of the speckle intensity, observed in the image, obeys negative exponential statistics ([Ghiglia, D. C., Pritt, M. D. 1998]). Therefore, fluctuations of the intensity about the mean are quite large. This can make visual interpretation difficult.

Speckle is an undesirable phenomenon in most applications, for it masks subtle image brightness transitions, destroys apparent brightness connectivity within common scattering features, alters visual resolvability of non-point-like features, and reduces image interpretability in general.

The mean intensity distribution in a coherent image is identical to the intensity that is observed if the object were illuminated with spatial incoherent waves with the same statistical properties. Incoherent illumination can be thought of as a rapidly varying random sequence of spatially coherent waveforms with highly complex phase structure. The structures of the wave-fronts at any fixed spatial location are independent from one time to the next.

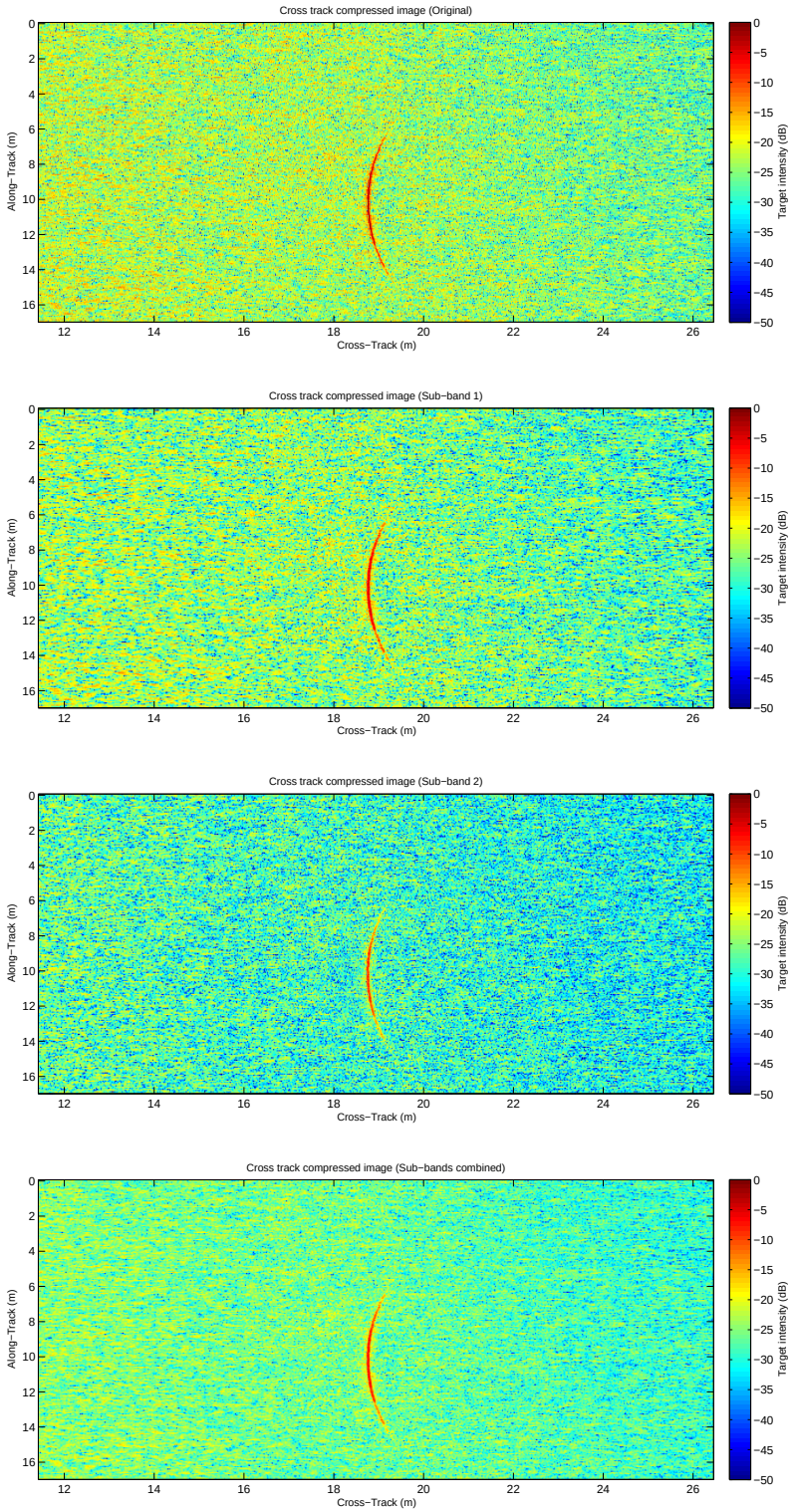

Figure 75: Effects of speckle in sub-band processing. From top to bottom: cross-track compressed image at the original wavelength, lower sub-band, higher sub-band and combined sub-bands.

Thus, (assuming equal bandwidths are involved) the time-integrated intensity is identical to the ensemble average intensity.

Typically, it is not possible to ensemble independent SAS images of a common surface over which to average the intensities. However, it is possible to partition the phase history into non-overlapping (with common aperture dimensions) and sun non-coherently the image intensities formed from each phase-history piece. Therefore, rather than taking several different looks at the submerged area, the similar effect is obtained by taking multiple looks from within the original phase-history domain.

Of course, the non-coherently integrated image formed in this fashion suffers a loss of resolution, because each image in the sum comes from a physically smaller portion of the phase history and thus a smaller synthetic aperture.

Mean and median filter are simple but effective ways to reduce speckle. More complex filters like the Lee filter or the Frost filter can also be used with improved results ([Baraldi, A.; Parmiggiani, F. (1995) and Mansourpour, M. et al (2006)]).

The sub-band processing technique describe here also has the property of reducing the impact of speckle. Speckle decorrelates from on band to the other, so by multiplying the sub-band signals the decorrelatated speckle portion will be attenuated. But it should be noted that, since the sub-bands are close in frequency the correlated speckle survives the process, making it possible to use the images for interferometry. Figure 77 shows a set of images obtained from the data collected in one test mission. In these images it is possible to see how the speckle varies between the original cross-track compressed images to the one processed with the sub-band technique.

Chapter 7: Bathymetry

7.1. The Bottom Height Estimation Problem

The bathymetry problem consists in obtaining a height estimation of the submerged bottom. Since by obtaining the reflectivity image the depth information is lost in the process of echo recording (echoes are obtained as a function of total distance to target), it is necessary to obtain more information about the reflectivity map as to unambiguously determine the depth coordinate of the submerged scene.

To obtain height estimation is necessary to take samples of the scene in such a way that the vertical dimension geometry is excited, in a similar a manner as it is necessary to have several samples in the horizontal dimension as to correctly obtain the along-track and slant-range reflectivity image. For this reason receiving transducers are placed along the vertical coordinate or the scene imaged from different parallel paths.

7.2. Volume Back-projection

It is possible to use back-projection in the vertical dimension. This method is usually referred to as volume back-projection since it consists in spreading the echoes in all the coordinates that define the image volume. The volume thus obtained has vertical resolution limited by the number of transducer typically used with a geometrical placement that enables the extraction of depth information.

With few transducers in the vertical direction it is also possible to obtain height estimation by calculating the most likely height in an area of the image. This is possible because the images will align when the correct height is used for their synthesis. So by generating several images at different heights, it is possible to create a surface of probable bottom depths [Barclay, P. et al (2003)]. Figure 76 shows two images (left and right figures) obtained from two vertically displaced synthetic apertures. The point target in the image is an artificial target placed in the river bottom at a depth of about 11m. Using this method it is possible to see that for the cross-track position of the target, the difference between the images (center figure) is minimum at that depth, which results in a rough estimate of the bottom height for that point.

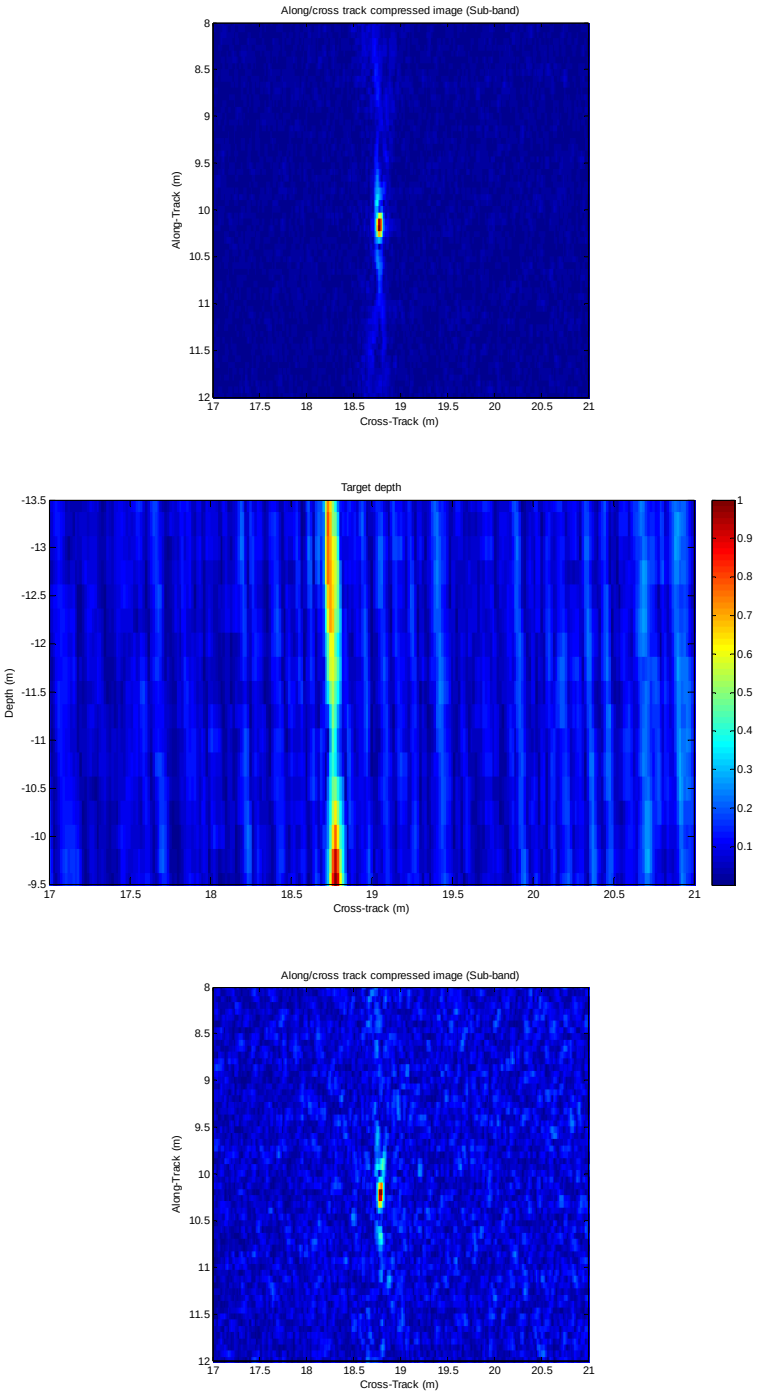

Figure 76: Height estimation through image alignment.

Of course, for this method to work correctly, it is necessary to have good bottom reflectivity characteristics across the images (no shadow or low reflectivity areas).

7.3. Interferometric Height Mapping

If the area observed is imaged by two synthetic apertures at displaced positions, the information included in the phase differences between the same target in the image can be used to obtain the third dimension and thus supply height estimation ([Richards, M. A. (2007)], [Silva, S. et al (2008 c)]).

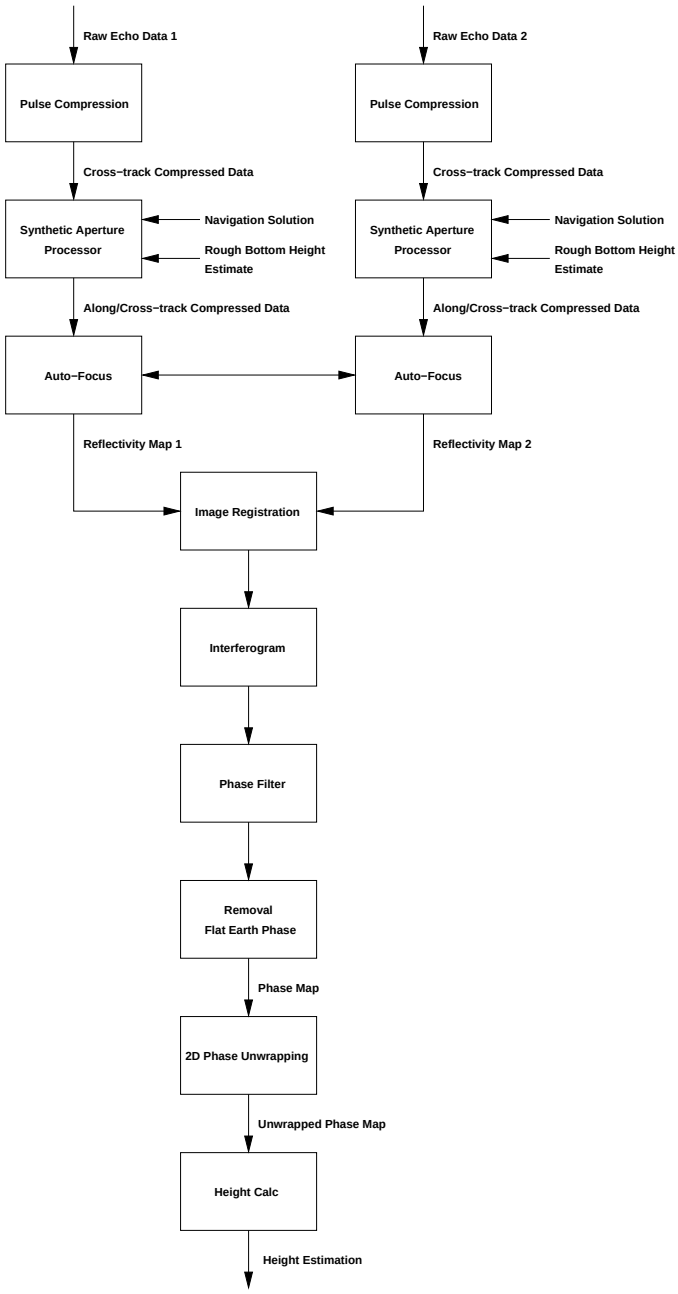

Figure 77: Interferometric processing flow diagram.

This is a known technique in synthetic aperture radar systems, but still a relatively unexplored concept in synthetic aperture sonar systems. The reason for this is linked to difficulties in obtaining good quality interferograms due to motion and medium induced phase errors. Nevertheless there have been studies that prove that medium phase stability is adequate ([Gough, P. T. (1989)]) and some very successful approaches to interferometric synthetic aperture sonar, as described in [Griffiths, H.D. (1997)], [Saebo, T. O. (2007)], [Silva, S. et al (2009 b)], and [Wang, L. (2001)].

Interferometric height mapping is not restricted to be used with synthetic aperture ([Pryor, D.E. (1998)]). It is possible to create interferograms from real aperture sonar images. Using synthetic aperture has several advantages inherent to the resolution abilities of the synthetic aperture sonar, especially when there is the need to cover large areas as the resolution is independent of the range.

Consider two synthetic aperture images and a swath formed from two separate apertures. These images are obtained from two different transducers mounted in the same system at different positions (single pass system) or from two different passages through the same swath (dual pass system).

Each synthetic aperture image is processed independently following the normal steps for image generation (Figure 77): pulse compression; synthetic aperture processing using known navigation data. Auto-focus being the only step that has to be linked has to not produce images with phase differences with other relation than height.

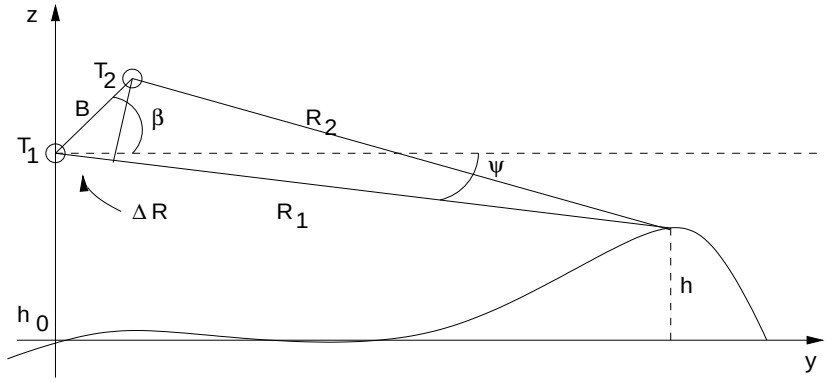

Figure 78: Interferometric height mapping geometry.

Prior to the creation of the interferogram, the images have to be carefully aligned through an image co-registration step. This step identifies several control points in both images, computes the geometric displacement relationship between each pair of control points and finally warps through interpolation one image so it overlays the other.

The geometric model of the interferometric height estimation can be seen in Figure 78 ([Griffiths, H. D. et al (1994)]). The relation in the ranges to the same point given by the two images is obtained by:

$$(R_1 + \Delta R)^2 = R_2^{\,2} + B^2 - 2BR_2 \cos(\psi + \beta) \tag{7.1}$$

Where B is the separation arm between the receiving transducers, β is the angle this arm makes with the horizontal line that intersects the transducer T_1 position and ψ is the angle of R_1 with also this horizontal line and is also called depression angle. ΔR can be given by the distance difference in number of phase cycles multiplied by the wavelength (λ):

$$\Delta R = \frac{\varphi}{2\pi}\lambda = \frac{(\phi + 2\pi n)}{2\pi}\lambda \tag{7.2}$$

Here the phase difference is decomposed in its fractional ϕ part, and the integer number of complete phase cycles n.

The height estimation can be given considering a small height variation in relation to a flat surface (mean height) in relation to a variation in the depression angle:

$$\delta h = h_0 \operatorname{cotan}\psi \delta\psi \tag{7.3}$$

The obtained images can be modeled by the following equations (where $ss_b(\tau, t)$ is the pulse compressed echo data and ff is the image obtained by synthetic aperture processing):

$$\begin{cases} ff_1(x, y_s) = \iint ss_{b_1}\left(\tau, t - \frac{2}{c}R_1\right) e^{-j\frac{4\pi}{\lambda}R_1}\, d\tau dt \\ ff_2(x, y_s) = \iint ss_{b_2}\left(\tau, t - \frac{2}{c}R_2\right) e^{-j\frac{4\pi}{\lambda}R_2}\, d\tau dt \end{cases} \tag{7.4}$$

The interferogram which gives rise to the phase differences matrix is given by the angular difference between each pixel in the images. For robustness against noise, this should be calculated as the angle of the conjugate multiply of the each pixel in the two complex images:

$$\phi = \angle\left(ff_2 \cdot ff_1^*\right) \tag{7.5}$$

7.4. Single Pass

When the two vertically displaced images are taken in the same pass, the interferometric height estimation is called single-pass. This is the more common case in sonar systems.

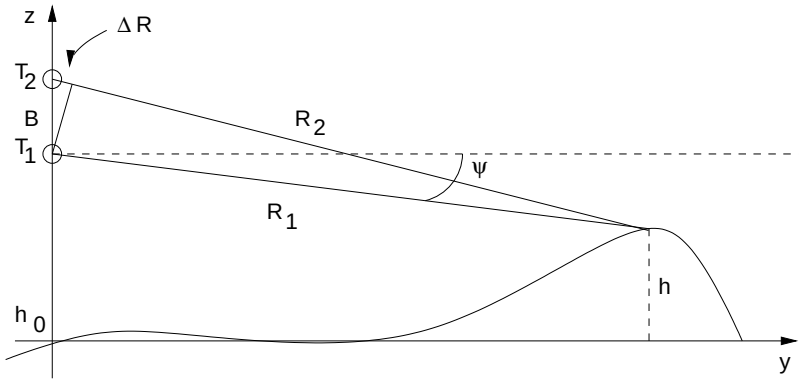

Figure 79: Interferometric height mapping geometry (1-pass).

In the particular case of this sonar system we can simplify the height estimation equation given the transducers are arranged in a vertical manner and so $\beta = 90^{o}$ (Figure 79). The measured phase is in this case:

$$\varphi = \frac{2\pi}{\lambda}\Delta R \tag{7.6}$$

And the difference in ranges due to the target height:

$$\Delta R = \sqrt{R_2^2 + B^2 + 2BR_2 \sin\psi} - R_1 \approx B\sin\psi \tag{7.7}$$

This approximation is valid for small values of ΔR when compared to R_1. The change in depression angle due to a change in the measured phase can be calculated using the above expression and taking the derivative in relation to the depression angle:

$$\delta\psi = \frac{\lambda}{2\pi B\cos\psi}\delta\varphi \tag{7.8}$$

It is also useful to filter the phase data due to noise and specular nature of the reflectivity maps. The phase is filtered by first converting the phase data to complex numbers and then applying separately a spatial 2-D linear filter (typically a low pass filter, $g[i, j]$) to the real and imaginary components or alternatively applying the filter to the pixel-by-pixel

conjugate multiply of the two complex images. The filtered phase data is then taken from the resultant complex number ([Ghiglia, D. C. et al (1998)]):

$$\hat{\phi}[k,l] = \arg\left\{\sum_{i=1}^{P}\sum_{j=1}^{P} g[i,j]\, ff_2[k-i,l-j]\, ff_1^*[k-i,l-j]\right\} \tag{7.9}$$

The phase variation due to a flat surface should be removed so that the interferogram includes only changes in phase that are due to height variation in the neighborhood of this constant height. The phase variation due to a constant height (flat-earth) is given by:

$$\varphi_{FE} = \frac{2\pi}{\lambda}\left[\sqrt{R_1^2 + B^2 + 2BR_1\sin(\psi_{FE})} - R_1\right] \tag{7.10}$$

Where ψ_{FE} is the depression angle considering the constant height h_0. This equation does not use the parallel beams approximation regarding the geometry, because this is not suitable for short ranges (shallow water) geometries.

The coherence of the interferogram (γ) should be estimated and is a indication of its quality:

$$\gamma = \frac{\sum_{i=1}^{P}\sum_{j=1}^{P} h[i,j]\, ff_1[k-i,l-j]\, ff_2^*[k-i,l-j]}{\sqrt{\sum_{i=1}^{P}\sum_{j=1}^{P} h[i,j]\, ff_1[k-i,l-j] ff_1^*[k-i,l-j] \sum_{i=1}^{P}\sum_{j=1}^{P} h[i,j]\, ff_2^*[k-i,l-j] ff_2^*[k-i,l-j]}} \tag{7.11}$$

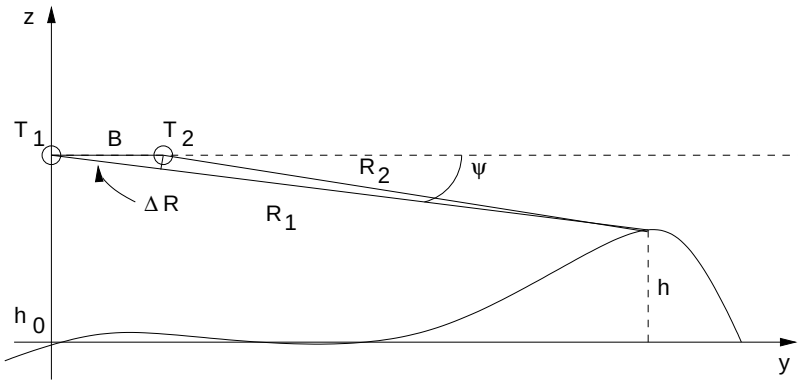

Figure 80: Interferometric height mapping geometry (2-pass).

Using the variation in phase due to a variation in height, it is possible to obtain from the flat-earth suppressed phase data the height estimation by using:

$$h = h_0 + H \cot\psi_{FE}\delta\psi \tag{7.12}$$

7.5. Dual Pass

Dual pass interferometric sonar images are difficult to obtain due to low medium coherence time, precision in vehicle positioning and motion errors. Since the presented system enables, in theory, dual pass interferometry, the problem will be briefly exposed here. For the dual passage case, the measure phase is:

$$\varphi = \frac{4\pi}{\lambda}\Delta R \tag{7.13}$$

The factor of 2 in this case in relation to the single-pass case is due to the fact that the height will introduce a difference in range in the transmit and received path of the second aperture. Since $\beta = 0^o$ (Figure 80), then:

$$\Delta R = \sqrt{R_2^2 + B^2 + 2BR_2\cos\psi} - R_1 \approx B\cos(\psi) \tag{7.14}$$

The depression angle change with the phase is given in this case by:

$$\delta\psi = \frac{\lambda}{2\pi B\sin\psi}\delta\varphi \tag{7.15}$$

7.6. Image Co-registration

The interferogram is obtained with the phase differences between the pixels of the same target within the different images. For this reason, the images must overlay perfectly with a sub-pixel accuracy.

Different geometries in the two apertures and changes in the terrain height result in image distortion that difficult image alignment.

Image registration is normally used to obtain a warping function that enables alignment of one image over the other ([Liu, Q.; Vesecky, J. F. (1990)], [(Wen Yang (2005)]).

A simple but effective method is to divide the image into several sub-images. A correlation matrix between all pairs of sub-images is obtained. The coordinates (tie points) of

the maximum of correlation of one image with the others is taken and used to form an image warping grid. This will actually be nothing less than a rough estimate of the bottom height variation.

Since registration to within a fraction of a pixel is necessary for good unwrapping results, the correlation peak must be found with good precision. For this several techniques can be applied such as image oversampling or quadrature interpolation of the correlation peak.

The quality of the sub-images correlations can be impaired by low reflectivity, shadowing or simply flat featureless bottom areas. In this case the warping information of these sub-images should be disregarded. The correlation quality (ξ) can be calculated by:

$$\xi = \frac{\max\left\{s^2_{ff_1 ff_2}\right\}}{\frac{1}{NM}\sum_k \sum_l s^2_{ff_1 ff_2}[k,l]} \tag{7.16}$$

Finally the warping is carried out using a two-dimensional interpolation method such as a bilinear or higher order interpolator depending on the target maximum phase error.

7.7. Phase Unwrapping

In equation (1.13), the phase difference ϕ is only known modulo 2π. This means that the absolute phase φ is related to ϕ through a wrapping operator:

$$\phi = W(\varphi) = \mathrm{mod}(\varphi + \pi, 2\pi) - \pi \tag{7.17}$$

Therefore to obtain the height map estimate it is necessary to know the integer value of cycles (since the beginning of the scene) and retrieve the absolute phase.

In the absence of noise, this is simply done by integrating the phase difference and adding 2π if the difference of value at a discontinuity is smaller than $-\pi$ or subtracting 2π if the value is larger than π.

However, in the presence of noise this phase integration will easily fail. Several algorithms have been derived to effectively solve this problem, such as path following methods or minimum norm methods ([Zebker H. A.; Lu, Y. (1998)], [Richards, M.A. (2007)]).

7.7.1. Path Following Techniques

Path following techniques enable phase unwrapping in the two dimensional domain by integrating along a path in the discrete plane of the interferogram. When integrating around a

close path, the difference between two adjacent pixels should not depend on the path taken. In reality, due to noise, aliasing (as caused by abrupt terrain changes) or other corruption sources, there can be a difference which is called a residue.

Residues can be detected by testing the integration path over every set of 2×2 pixels in the interferogram. To maintain phase unwrapping consistency in the data, integration through residues must be avoid. To do this, residues of opposite polarity are connected (branch cuts) and the integration is then be forbidden for paths that cross these branch cuts. Because in real data there is the probability of having a large number of branch cuts, entire areas can be left unwrapped, creating undefined zones in the output phase estimation. The path following algorithm must therefore be able to select the best way to connect the branch cuts as to minimize the unwrapped image areas. Discussion and details about this algorithms can be found in [Ghiglia, D. C. (1998)] and [Wilkinson, A.J.; Griffiths, H.D. (1994)].

7.7.2. Minimum Norm Techniques

A particular minimum norm method known as the least square phase unwrapping is commonly known because of its efficient implementation through frequency domain algorithms and capability of obtaining a phase unwrap estimate everywhere in the image. This is the algorithm used in this system. Given the wrapped phase values defined in a matrix with size $M\times N$ we wish to calculate the unwrapped phase values at the same pixel locations (i, j). The unwrapped phase differences in the x and y direction must agree with the wrapped phase differences in the same directions. Knowing this, the correct unwrapped phase is the solution that minimizes the error ε in the least-square sense ([Ghiglia, D. C. (1998)]):

$$\sum_{i=0}^{M-2}\sum_{j=0}^{N-1}\left(\varphi_{i+1,j}-\varphi_{i,j}-W\left\{\phi_{i+1,j}-\phi_{i,j}\right\}\right)^2+\sum_{i=0}^{M-1}\sum_{j=0}^{N-2}\left(\varphi_{i,j+1}-\varphi_{i,j}-W\left\{\phi_{i,j+1}-\phi_{i,j}\right\}\right)^2=\varepsilon^2 \tag{7.18}$$

After manipulation of this equation using the discrete variational approach, it is possible to arrive at the following results ([Ghiglia, D. C. (1998)]):

$$\left(\varphi_{i+1,j}-2\varphi_{i,j}+\varphi_{i-1,j}\right)+\left(\varphi_{i,j+1}-2\varphi_{i,j}+\varphi_{i,j-1}\right)=\left(\nabla\phi^{x}{}_{i,j}-\nabla\phi^{x}{}_{i-1,j}\right)+\left(\nabla\phi^{y}{}_{i,j}-\nabla\phi^{y}{}_{i,j-1}\right)=\rho_{i,j} \tag{7.19}$$

Where $\nabla\phi^{x}{}_{i,j}=W\left\{\phi_{i+1,j}-\phi_{i,j}\right\}$ and $\nabla\phi^{y}{}_{i,j}=W\left\{\phi_{i,j+1}-\phi_{i,j}\right\}$. This equation can be identified as the discrete Poisson's equation and it is possible to show ([Ghiglia, D. C. (1998)]) that its solution can be efficiently calculated through the discrete cosine transform (DCT):

$$\varphi = DCT^{-1}\left\{\frac{DCT\left\{\rho_{i,j}\right\}}{2\left(\cos\left(\pi i / M\right)+\cos\left(\pi j / N\right)-2\right)}\right\} \tag{7.20}$$

Note that the indexes are defined between 0 and $M-1$ for the x direction and 0 and $N-1$ in the y direction. 0 is assumed for values outside this region.

7.7.3. Network flow methods

In the network flow method, the phase unwrapping is done by writing the gradient of the interferogram as a constrained network with a value in each node that is an integer multiple of 2π ([Chen, C. W.; Zebker, H. A. (2000)]). The unwrapping problem is then posed as a global minimization problem with integer variables. The network structure of the problem allows the application of efficient solution algorithms.

Starting with the 2×2 integral of the phase gradient, the residues are calculated forming a grid of nodes with 0,-1 or 1 residue values. This grid can be seen as a transport network in which the flow on an arc is the difference in cycles between the unwrapped and wrapped phase gradients of neighboring pixels. To maintain the agreement with the data the total flow out of a node must be equal to the residue at that node which is equivalent to saying that the unwrapped phase field must be residue free.

The integer value of cycles in each flow arc is obtained solving the network flow problem using an algorithm such as the minimum cost flow. The integer values are then added to the pixels associated with that residue node, making possible to integrate the phase correctly.

Obviously the obtained solution is not unique since it is possible to add a cycle to the flow chart and still obtain a valid solution. Therefore some criterion must be used to choose one particular solution.

7.8. Error Sensitivity

To get an estimate of the height error (σ_h) as a function of phase noise (σ_φ) we can differentiate the height function with regard to phase:

$$\frac{\sigma_h}{\sigma_\varphi} = \frac{dh}{d\varphi} = \frac{dh}{d\psi}\cdot\frac{1}{\frac{d\Delta R}{d\psi}}\cdot\frac{d\Delta R}{d\varphi} \tag{7.21}$$

And get:

$$\sigma_h = -\frac{R_0}{B}\frac{\cos(\psi)}{\sin(\psi+\beta)}\frac{\lambda}{2\pi}\sigma_\varphi \tag{7.22}$$

Here, σ_φ is the phase error which can be derived from the signal-to-noise ratio and R_0 is the medium range to the scene.

For the arm variation (which is an important factor in dual-passage operation), the height estimation error can be given by:

$$\frac{dh}{d\varphi} = \frac{dh}{d\psi}\cdot\frac{1}{\frac{d\Delta R}{d\psi}}\cdot\frac{d\Delta R}{dB} \tag{7.23}$$

This in turn gives:

$$\sigma_h = \frac{R_0}{B}\frac{\cos(\psi)}{\cot(\psi+\beta)}\sigma_B \tag{7.24}$$

We can see that the longer the arm, the more insensible is the height error to the variation of the arm and phase measurement error. Figure 81 shows the height error estimate for increasing errors of phase and arm measurement error. It is possible to see that a longer arm produces lower height error for the same phase or arm error.

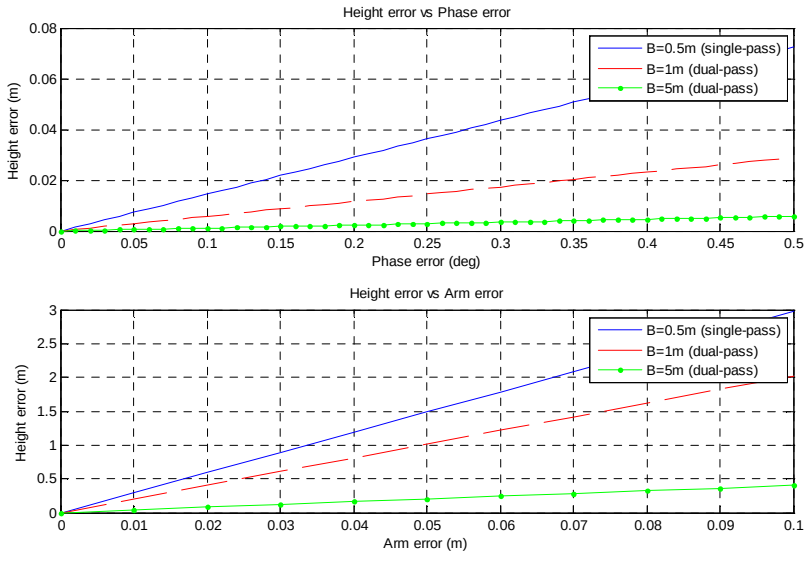

Figure 81: Height error sensibility.

7.9. Correlation between Interferometric Images

From the height error equation we see that the longer the arm between the transducers, the lower the height estimation error. The arm can't be arbitrarily long also due to decorrelation between the pair of synthetic aperture images. The maximum arm after which the images have an inappropriate degree of correlation is given by the following equation in the case of single-pass ([Richards, M.A. (2007)]):

$$B_{cv} = \frac{\lambda R_0}{\delta_{XT} \sin\psi} \tag{7.25}$$

And for the case of dual-pass:

$$B_{ch} = \frac{\lambda R_0 \cos\psi}{2\delta_{XT} \sin^2\psi} \tag{7.26}$$

δ_{XT} is the cross-track resolution. For the parameters of the sonar in consideration this results in 27m and 38m, respectively which is a comfortable large margin considering that the typical transducer separation ranges between 0.25 to 1.5m and parallel path separations between 1 and 5m. On the other hand, a longer transducer separation will also produce a greater phase variation, which is problematic since the phase is a wrapped measurement.

7.10. Fringe Spacing

The fringe spacing is the length of the unwrapped phase strips and is given by:

$$\frac{2\pi}{\frac{\delta\varphi}{\delta y}} = \frac{2\pi}{\frac{4\pi}{\lambda}\frac{\delta\Delta R}{\delta y}} = \frac{\lambda}{2}\frac{\delta y}{\delta\Delta R} \tag{7.27}$$

A factor of two in the phase value must be considered for the dual-pass case. $\delta\Delta R/\delta y$ is proportional to the arm length B which means that a longer arm will cause more phase wrappings in the interferogram and thus will make the phase unwrapping process more difficult.

7.11. Applications of Interferometry

The interferometric phase difference between two SAS images can be used to other purposes besides height estimation. InSAS presumes that there is no change in the image scene between the two images data collections, so phase difference are due only to height variations viewed from slightly different aspects angles. This assumption is true in one-pass systems (pair of interferometric images obtained in one pass run), but may not be true in two-pass systems (pair of interferometric images obtained in two different pass runs), a problem that is designated as temporal decorrelation. The time between passes in two-passes might be between minutes, but also might be weeks and so changes in the detected terrain or structure translate in phase changes. If we know before hand the height, these phase changes can be used to infer characteristics change over long period of time.

Therefore, a change in scattered height at a fixed location on the ground between two different times can be used to estimate sandbank growth or changes in submerged structures. As with InSAS static height mapping, it is assumed that the reflectivity of each pixel has negligible change between images. Because only a single receive aperture is used, the height change can be estimated at each pixel between imaging passes assuming that the terrain motion between passes is in the vertical dimension only. Any change in slant range between a scatterand the sonar will result in a detectable change in the interferogram, and a suitable method can be developed to extract this information.

These methods require two-pass operation. The time interval could be on the order of days or weeks to study the effects of such phenomena. Unlike radar system however, it might not be always possible to guarantee coherence between passes in sonar imagery due to medium instability (currents, sound propagation changes). Nevertheless, with man made structures, very strong echoes are recuperated and so high quality phase measurements are possible.

Ideally, exactly the same path should be followed on the two passes, so that the baseline between the two images is zero. In practice, this is very difficult, and a small non-zero baseline will be reflected in the data. This means the interferogram will have components due to both the temporal change in height, and the static height profile. One approach to removing the static component is to use an existing digital terrain model (for example from previous runs) to estimate it and then subtract its contribution to the phase.

Another application of interest is coherent change detection which is a two-pass application that compares two images taken from the same trajectory at different times. The time intervals are typically shorter, from a few minutes apart to many hours or days apart.

However, we now assume the height profile unchanged, but the reflectivity function does change. This could occur due to several disturbances like plant growth or sediment movements. In general, both the amplitude of the reflectivity and the phase will change. If there is no change in the reflectivity of a given pixel between passes, computing a normalized correlation coefficient of the two measurements should produce a value equal to unity. If the reflectivity has changed, a lesser value of the correlation coefficient should result.

Chapter 8: Results

8.1. Experiments in the atmosphere

To validate some hypothesis on the behavior of the synthetic sonar system and to validate the developed algorithms, an experiment was conducted in the atmosphere using an audio frequency sonar on a moving vehicle equipped with a low c6st GPS and inertial measurement unit. The test target was the face of the INESC building. This sonar was built using low cost audio devices, for the solely purpose of this experiments, but its value as a test bed for the more advanced under-water sonar system rose with its results. A 25cm wavelength was used.

The cross-track pulse compressed image is presented in Figure 82 it is possible to see several hyperboles described in the along-track by the strong reflections caused by corners and columns of the building. In Figure 83 a synthetic aperture image formation method was applied to the data, making possible to reconstruct objects (like the building column, corners and windows) to a better spatial resolution. The strong line like reflection in the center or the image is the curb of side-walk in front of the building. It is not straight as it should be due to the irregular motion of the vehicle.

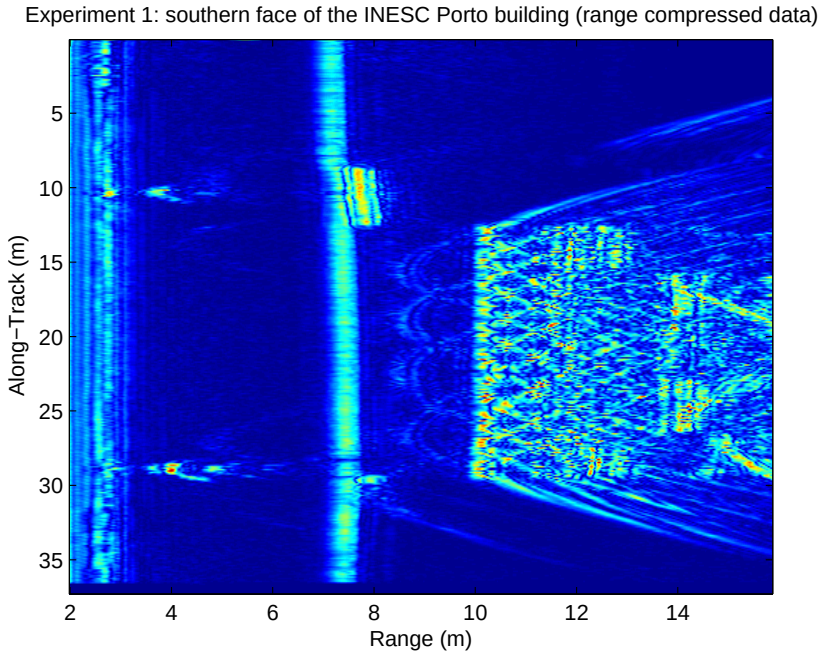

Figure 82: Face of the INESC Porto building: cross-track compression.

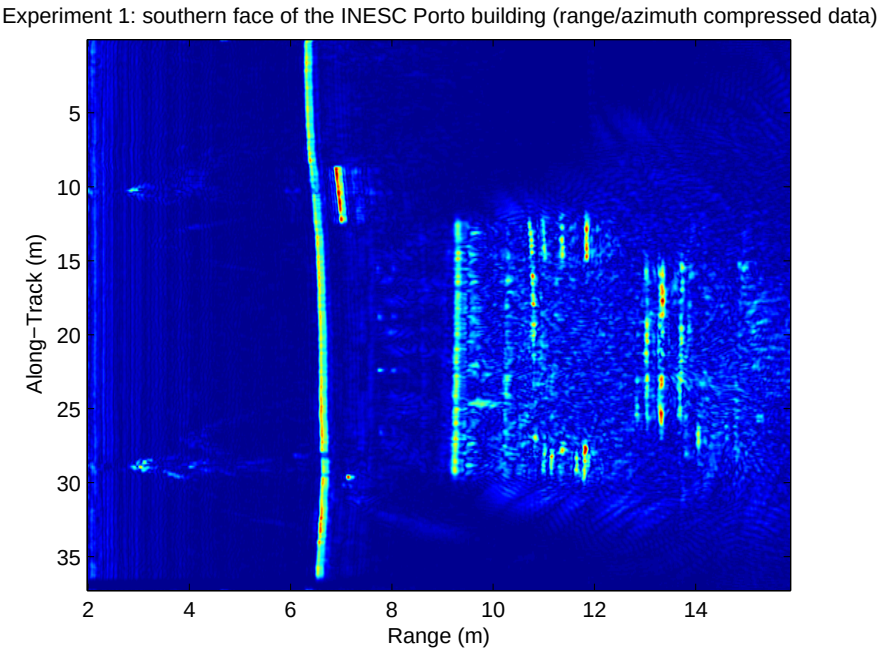

Figure 83: Face of the INESC Porto building: along and cross-track compression.

Some of the building columns (after the strong reflection line), that are describe as hyperboles in the cross-track compressed image, are correctly represented as a point while other are smeared. This was due to residual uncompensated velocity fluctuations of the vehicle.

8.2. Trials in the Douro River

The developed high frequency sonar system was tested in the Douro River, Portugal. During these tests it was possible to obtain some experimental data.

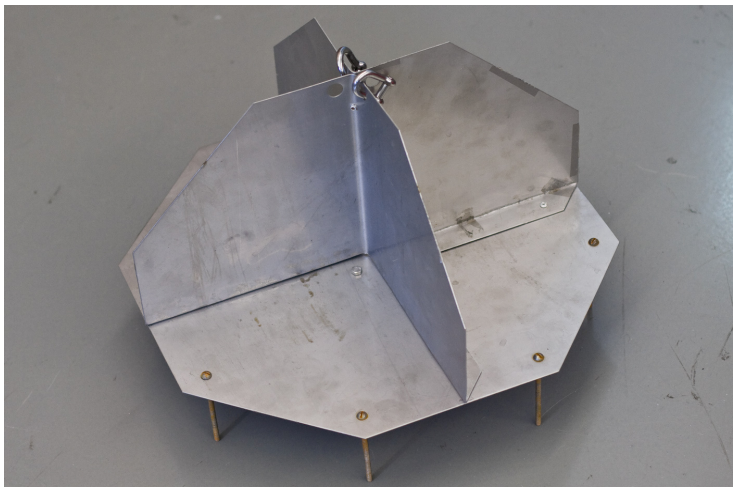

Figure 84: Artificial target used for resolution tests.

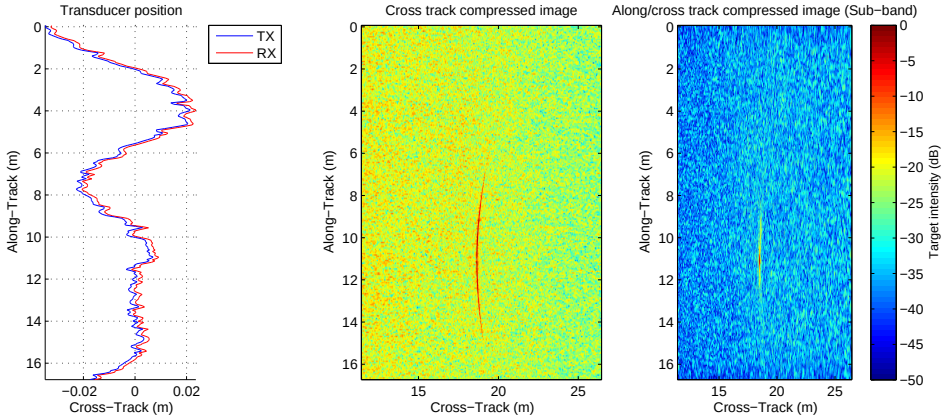

Figure 85: Transducer position, cross-track only and along/cross-track compressed image.

The test site has a muddy river bed with smooth height variations caused by accumulation of sand and mud. To test the synthetic aperture algorithm, an artificial target was placed on the river bottom that served for test site. The target is a half octahedral shape of $20\times20\times20$ cm (Figure 84) and the reflected signal is seen as a point after correct image synthesis. The results are shown in Figure 85. The autonomous vehicle followed a straight path in the vicinity of the target at a constant speed.

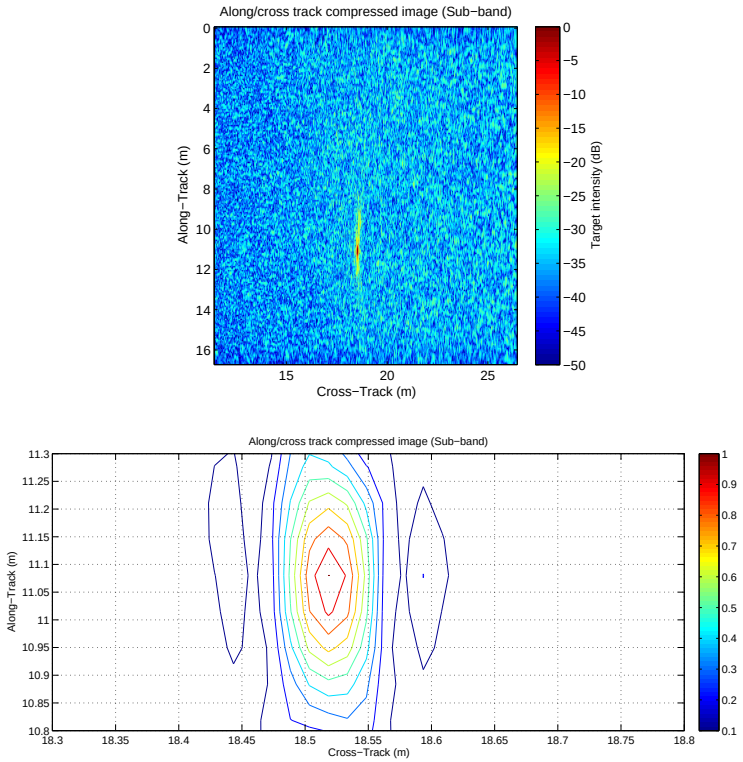

Figure 86: Reconstructed image of the artificial target with sub-band-processing (top) and counter plot of the center of the image (bottom).

The deviations from the predefined parameters are negligible and accounted for in the synthetic aperture processing step. The hyperbolic like target caused by the large aperture of the transducers is reduced, after synthetic aperture processing, to a small point.

This shows that the system can correctly form synthetic aperture images correctly combining the echoes using the navigation data. This image was processed using the sub-band method explained in the previously. A resolution of 25cm along-track and 4cm cross-track was obtained (Figure 86).

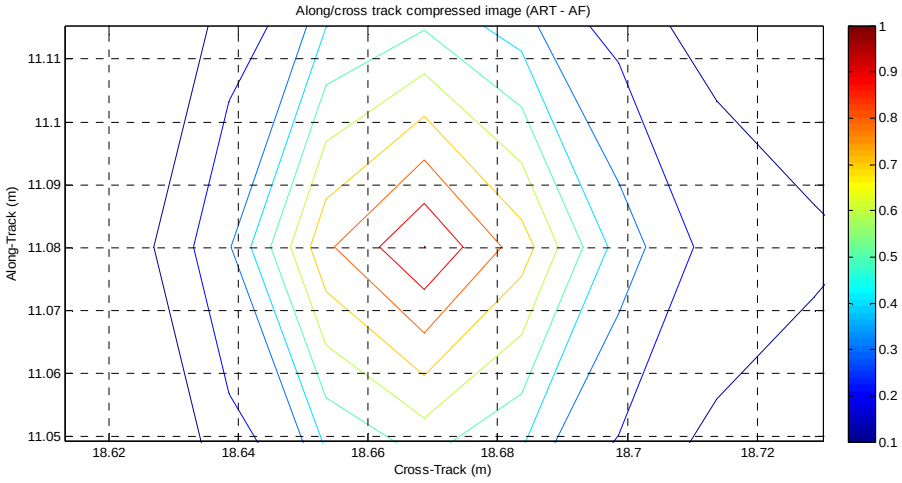

Figure 87: Contour plot of the target after auto-focus.

After auto-focus and using the original wavelength, the maximum theoretical resolution of 2.5 cm along and cross-track was achieved in this particular case (Figure 87). It must be noted though that reconstructing the image at the original wavelength inevitably produces grating lobes due to the along-track under-sampling.

In the case shown in Figure 88, the unmanned surface vehicle was programmed to make passages near the shore line. Rocks on the bottom of the river contrast with the typical mud bottom and are presented as scattered points in the synthetic aperture sonar image.

Due to strong under-sampling the image presented here was obtained with the described sub-band processing scheme and is thus of reduced quality. Even though it is still possible to observe the gain obtained by using the global optimization auto-focus algorithm described in the previous chapter. More trial runs are necessary to understand the performance of the autofocus algorithm and convergence characteristics in different cases such as distributed targets and specular targets.

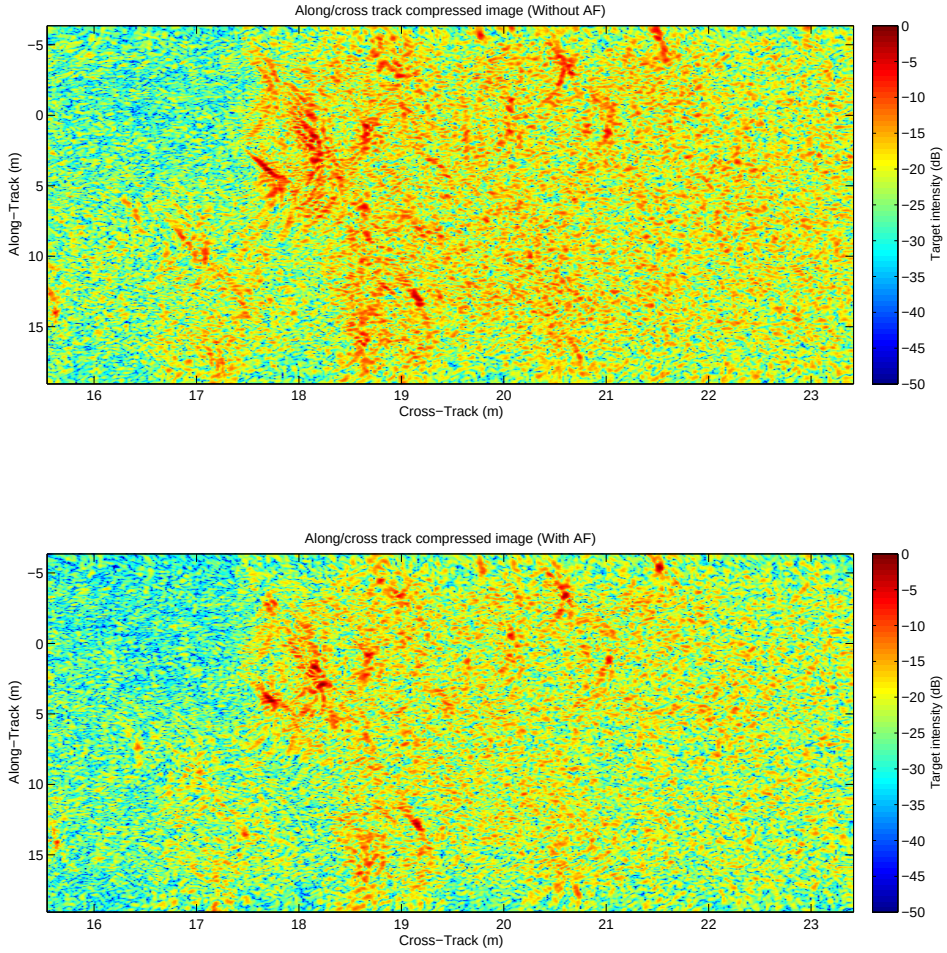

Figure 88: Image of the rocky river bottom before (top) and after auto-focus (bottom).

Figure 89 and Figure 90 show another sonar run near the shore line and the position of the sonar image in the map. Small rocks are shown as bright points and it is possible to see a set of closely spaced but distinguishable rocks in the upper part of the image.

A height map of an area near the shore and in close proximity to a harbor entrance is shown in Figure 91. At the entrance of the harbor there is a visible mound formed by accumulation of river sand. The height estimate was obtained using the interferometric process described in the previous section. High values for coherence were obtained for the interferogram that originated the height estimate (Figure 92). The unwrapping method used was the weighted least squares version of the minimum-norm algorithm. The weight map was calculated using the interferogram coherence values.

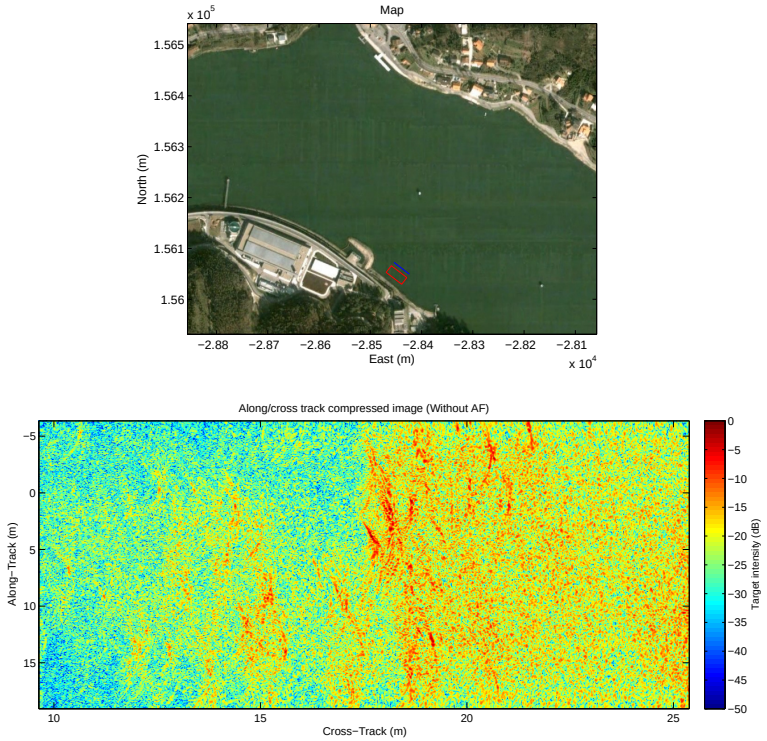

Figure 89: Position of imaged area on the map (top) and actual sonar image (bottom).

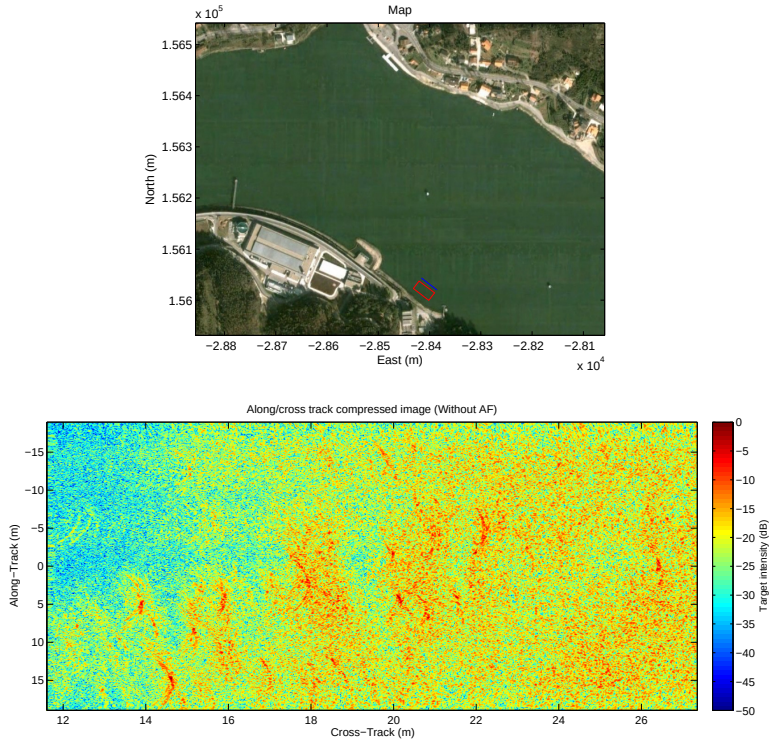

Figure 90: Position of imaged area on the map (top) and actual sonar image (bottom).

No other data was available for comparison of the obtained results for the height estimate and only a depth rudimentary depth measure was made that agreed with the average depth obtained.

The co-registration of swath reflectivity imagery with bathymetric data, enabled by interferometric synthetic aperture sonar, renders the results much more readable.

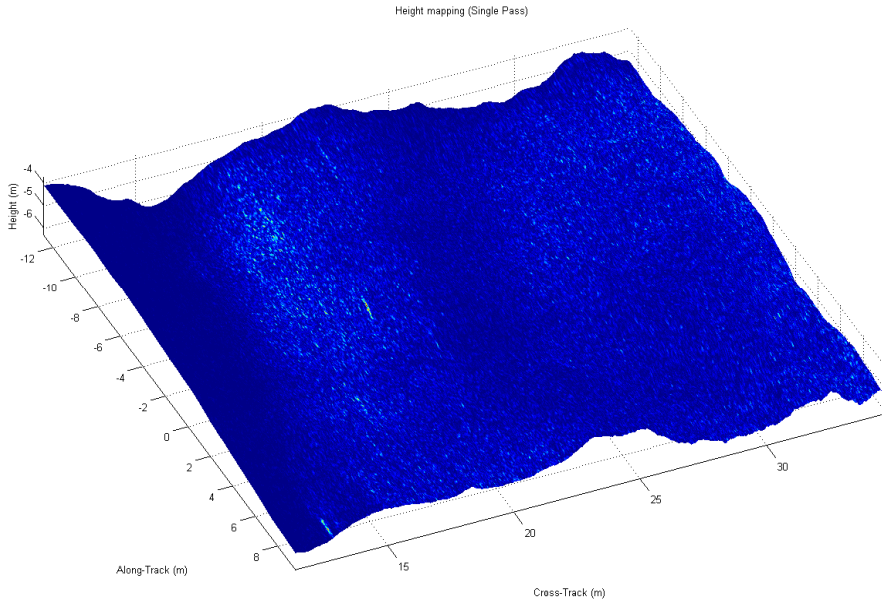

Figure 91: Height map of a harbor entrance in the Douro River.

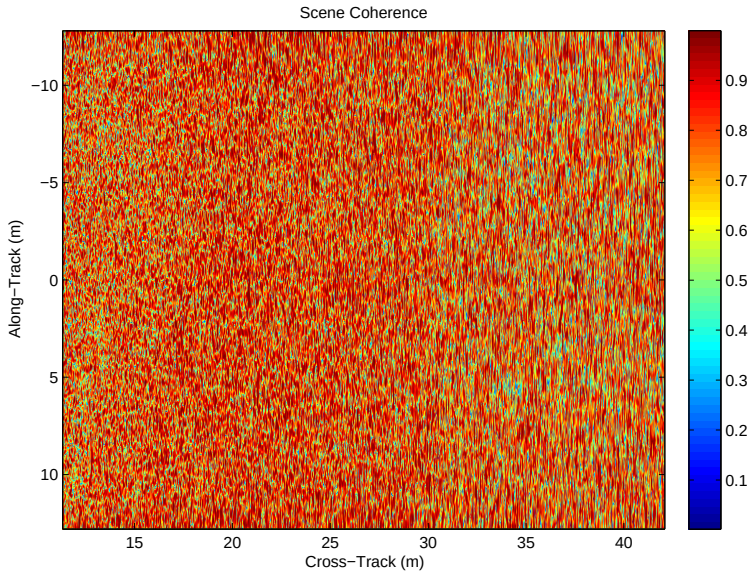

Figure 92: Coherency map.

8.3. Geo-radar

Geo-radar is a ground penetrating radar system that is used to study buried structures, typically in civil engineering applications like buried pipes, electric cables, etc ([Topczewski, L. et al (2007)], [Vasconcelos, G. et al (2008)]).

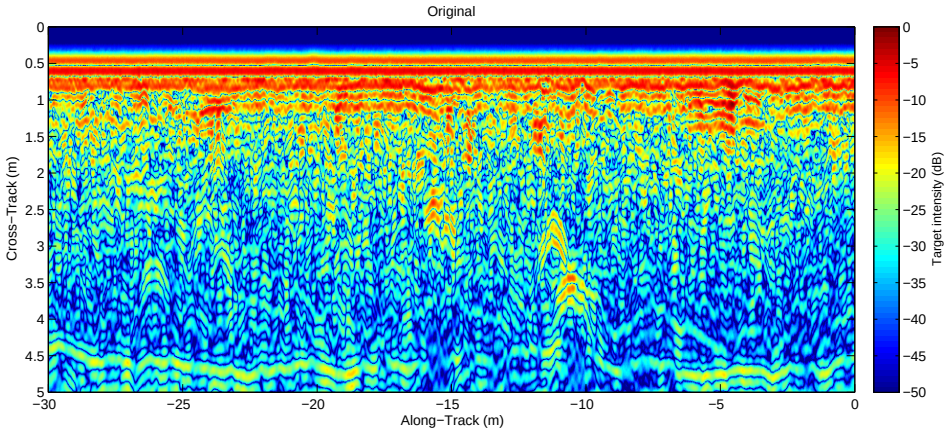

Figure 93: Geo-radar unprocessed image.

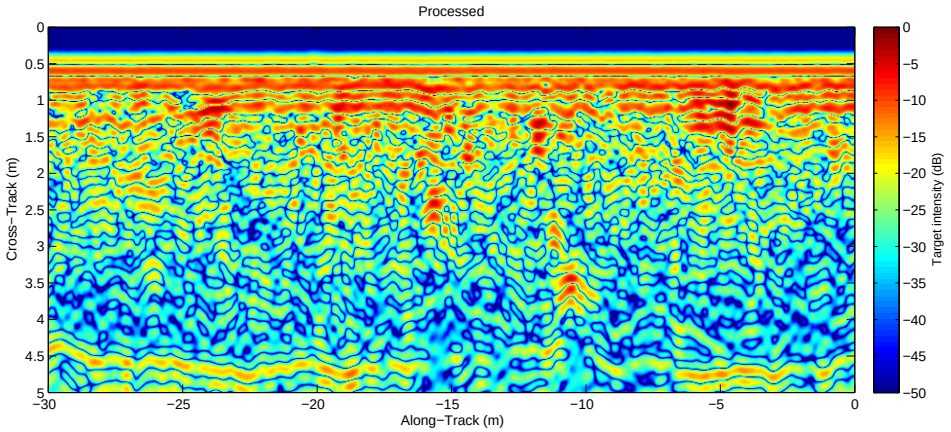

Figure 94: Geo-radar processed image.

Figure 93 shows a typical reflectivity image obtained with one of these systems. The data used to obtain the images presented here was gently provided by the geo-radar research group from the Civil Engineering Department of Minho University. Looking at the figure, it is possible to observe a hyperbole caused by some structures at a depth of about 3m.

Unfortunately, these systems usually do not provide phase information and so it is not possible to apply typical synthetic aperture algorithms directly. Nevertheless, if a sub-optimal

image is acceptable, the algebraic reconstruction algorithm can be used with the modulus data to reconstruct the image. Figure 94 shows the result of using this algorithm with the data from Figure 93.

It is possible to see that although the resolution gain is not very significant, as it was expected by using only the modulus data, the structures become more distinguishable from the clutter and the image more readable since also a range attenuation correction is applied.

Chapter 9: Conclusions and Future Work

A simple interferometric synthetic aperture sonar system that operates from small surface vehicles was presented. This system uses a simple vertically arranged dual element array to operate in shallow waters and retrieve bottom reflectivity and height maps. In this work it was found that a synthetic aperture sonar that uses a GPS/IMU system as navigation source can be successfully operated as long as the wavelength of the signal used is an order of magnitude higher than the navigation error. For the use of shorter wavelengths, the GPS/IMU navigation system can only be used to obtain a first solution to an iterative process where auto-focus and micronavigation algorithms are essential for effective image reconstruction.

An unmanned surface vehicle is a low cost solution that allows the use of satellite navigation for its control and motion compensation in the sonar image formation algorithms. It presents ease of use and deployment in areas suitable for its operation like rivers, dams and harbors. Because the sonar is fixed to the boat structure, the swath length is dependent on the depth and so is the coverage rate. This is, however, not an important shortcoming in the target environments as the depth is typically in the order of only a few tens of meters.

Not commonly considered motion errors were also studied such as surge, heave, roll, pitch and yaw. These sources of errors become more important because of the sonar platform used. Although the boat has a control system that tries to maintain a constant velocity, there are always velocity fluctuations (due to currents for example) that translate into surge errors. Again, because of the chosen platform, surface waves produce uncontrollable heave errors that must be corrected in the image reconstruction process.

The study of motion error impairments in synthetic aperture image synthesis lead to a better knowledge of the image reconstruction algorithms and their impact on image quality. Having a precise measurement of the boat position and attitude is of the most relevance. High frequency sonar system works with very short wavelengths, and so have quite high restrictions to motion compensation algorithms and navigation errors.

A frequency domain algorithm that enables motion compensation was developed, but it was shown that it is not suitable for high path distortions relative to a hypothetical sampling linear path. Using the back-projection algorithms one is no longer restricted to linear paths. Deviations from this path are not treated as errors, but simply as sampling positions. Images obtained in this way can be easily integrated in geographical information systems. Furthermore, the developed algebraic image reconstruction algorithm effectively integrates the available

sonar position and attitude information while also enabling fine image reconstruction adjustments through auto-focus procedures. This was one of the achievements of this work.

Sub-band processing eases the problems related to under-sampling and motion uncertainties if lower resolution images are acceptable. In the other hand, it can be used to aid the auto-focusing process, reducing the initial image formation requirements and progression through the auto-focus parameter optimization space. Since sub-band processing technique changes the received signals phase, more extensive studying and testing is needed as to understand its complete impact in interferometric and synthetic sonar imagery. The development of this technique was another accomplishment of this work.

The test trials in the Douro River were revealing. Not only it was possible to test the sonar platform and hardware developed, it was also possible to see in firsthand the problems that synthetic aperture image synthesis faces. Unwanted motion errors, medium fluctuation, and suitable echo signal level are among the most important. The developed hardware for the sonar system performed well, but the obtained results, however, put in evidence the simplicity of the sonar hardware. Improvements are still needed to enable higher power output, better signal-to-noise ratio, and more simultaneous acquisition channels to extend the system to an array of multiple along-track elements. Plain hardware can nonetheless produce good results through the use of smarter (and higher complexity) signal processing solutions such as the auto-focus procedures described here.

Working with high frequencies in water is challenging because of the increased attenuation and susceptibility to phase errors. Broad-beam transducers provide a potentially higher resolution, but also make the synthetic aperture process more difficult since shorter along-track sampling spacing is required to avoid aliasing. The quality of the images obtained by this system is mainly limited by under-sampling and the ability to obtain navigation solutions with a precision better than the used wavelength. The under-sampling results from the use a single element in the along-track direction and forcing the sampling of the swath at a rate lower than the imposed by the vehicle minimum velocity.

For future operational systems a multi-element array presents a better solution to the sampling problem. The use of multiple receiver systems can enable a higher area coverage ratio. Furthermore, a multi-element system eases the navigation system precision requirements through the use of micro-navigation techniques, avoiding or easing the burdensome auto-focus procedures.

Bottom height mapping is possible through the use of a double array of transducers and also by exploring the possibility of dual-pass interferometry. In this case the combination of

images of the same scene obtained from different positions of the platform allows the construction of three dimensional maps of the analyzed surfaces. In the future an effort should be made to correctly calibrate and assess the sonar height estimation capabilities using reference targets or by comparison against other available data sources. As this system enables gathering of both reflectivity and height maps, they can be combined together, thus enabling a better knowledge of the underwater territory.

Synthetic aperture sonar operated from the surface can be a powerful tool for shallow water remote sensing. Additionally enabling real-time operation will further extend the application possibilities of synthetic aperture sonar systems and with the maturing of image formation algorithms, the problems and difficulties inherent to synthetic aperture sonar will be solved, making its use widespread.

With this sonar and moving platform, the path was laid to develop state of the art sonar imaging systems that will hopefully help shed a light in the unknown submerged areas of our national territory and our planet.

Appendix A: System Parameters

The parameters of the synthetic aperture sonar system described in this thesis are summarized in the table bellow.

Table 1: Sonar system parameters.

Parameter	**Value**
Center frequency	200kHz
Band-width	40kHz
TX power	50W
Transducers beam-width	18º
Transducers diameter	0.05
Displacement between vertical receivers	0.25m
N.º of transmitter channels	1
N.º of receiver channels	2
Recording sampling frequency	100kHz
Pulse repetition frequency	10Hz
TX signal length	10ms
Maximum recording time	90ms

Bibliography

Asada, A.; Yabuki, T. (2001), “Synthetic aperture technique applied to a multi-beam echo sounder”, Journal of Earth, Planets Space, vol. 53, no.4, pp 321-326, Japan, 2001.

Banks, S. M.; Sutton, T. J.; Griffiths, H. D. (2001), “A technique for interferometric synthetic aperture sonar image processing”, Proceedings of the London Communication Symposium, UCL, London, UK, 2001.

Banks, S. M.; Griffiths, H. D. (2002), “Imaging and Motion Estimation for Synthetic Aperture Sonar Based on Fast Factorised Back-Projection”, Proceedings of the Sixth European Conference on Underwater Acoustics ECUA 2002, Gdansk, June, 2002.

Baraldi, A.; Parmiggiani, F. (1995), “An alternative form of the Lee filter for speckle suppression in SAR images”, Graphical models and image processing, vol. 57, no. 1, pp. 75-78, 1995.

Barker, R. H. (1953), "Group Synchronizing of Binary Digital Sequences", Communication Theory, London: Butterworth, pp. 273-287, 1953.

Bellettini, A.; Pinto, M. A. (2002), "Theoretical Accuracy of Synthetic Aperture Sonar Micronavigation Using a Displaced Phase Centre antenna", IEEE Journal of Oceanic Engineering, vol. 27, no. 4, pp. 780-789, October, 2002.

Berdjag, D.; Pomorski, D. (2004), “DGPS/INS data fusion for land navigation”, Proceedings of the 7th International Conference on Information Fusion, Stockholm, Sweden, 2004.

Bonifant, W. W. (1999), “Interferometric Synthetic Aperture Sonar Processing”, Master’s Thesis, Georgia Institute of Technology, July, 1999.

Brissette, M. B. (2006), “Mine Detection Using Swath Bathymetric Sonars: Tools and Techniques”, Proceedings of the UDT-Pacific, 2006.

Callow, H. J.; Hayes, M. P.; Gough, P. T. (2001), “Noncoherent Autofocus of Single Receiver, Broad-Band Synthetic Aperture Sonar Imagery”, Proceeding of the OCEANS, 2001. MTS/IEEE Conference and Exhibition , vol. 1, pp. 157-162, 2001.

Callow, H. J.; Hayes, M. P.; Gough, P. T. (2002), “A Wavefront Sensing Method for Synthetic Aperture Sonar Autofocus”, Proceedings of the sixth European Conference on Underwater Acoustics (ECUA), pp. 231-234, Gdansk, Poland, 2002.

Callow, H. J. (2003), “Signal processing for synthetic aperture sonar image enhancement”, PhD Thesis, Electrical and Electronic Engineering at the University of Canterbury, New Zealand, 2003.

Callow, H.J.; Hansen, R.E.; Sæbø, T.O. (2006), "Effect of Approximations in Fast Factorized Backprojection in Synthetic Aperture Imaging of Spot Regions", Proceedings of the Oceans 2006 Conference, pp. 1-6, 18-21, September, 2006.

Chang, E.; Tinkle, M. D. (2001), “Coherence of Pulsed Signals and Implications to Synthetic Aperture Aperture Processing”, Proceedings of the OCEANS, 2001. MTS/IEEE Conference and Exhibition , vol. 1, pp.193-201, 2001.

Chanussot, J.; Maussang, F.; Hetet, A. (2002), "Scalar image processing filters for speckle reduction on synthetic aperture sonar images", Proceedings of the Oceans '02 MTS/IEEE conference, vol. 4, pp. 2294-2301, October, 2002.

Chatillon, J.; Bouhier, M. E.; Zakharia, M.E. (1992), "Synthetic aperture sonar for seabed imaging: relative merits of narrow-band and wide-band approached", Journal of Oceanic Engineering, IEEE, vol. 17, no. 1, pp. 95-105, 1992.

Chatillon, J.; Adams, A. E.; Lawlor, M. A.; Zakharia, M. E. (1999), “SAMI: A Low-Frequency Prototype for Mapping and Imaging of the Seabed by Means of Synthetic Aperture”, IEEE Journal of Oceanic Engineering, vol. 24, no. 1, pp. 4-15, January, 1999.

Chatillon, J. (2000), “Absence of Nyquist criterion for spatial sampling of a strip-map synthetic aperture sonar”, Proceedings of the fifth European Conference on Underwater Acoustics (ECUA), vol. 1, pp 401-406, 2000.

Chen, C. W.; Zebker, H. A. (2000), “Network approaches to two-dimensional phase unwrapping: intractability and two new algorithms”, JOSA A, vol. 17, Issue 3, pp. 401-414, March, 2000.

Coffey and Hollingsworth (1972), “Report on Investigation of Sand Deposits, Middle Banks, Juno Point, Moreton Bay, Queensland”, Coffey and Hollingsworth: Brisbane, 1972.

Collins, T.; Atkins, P. (1998), "Doppler-sensitive active sonar pulse designs for reverberation processing", IEE Proceedings of Radar, Sonar and Navigation, vol. 145, no. 6, pp.347-353, 1998.

Collins, T.; Atkins, P. (1999), "Nonlinear frequency modulation chirps for active sonar," IEE Proceedings of Radar, Sonar and Navigation, vol. 146, no. 6, pp. 312-316, 1999.

Clay, C. S.; Medwin, H. (1977), “Acoustic Oceanography: Principles and Applications”, John Wiley & Sons, NY, USA, 1977.

Cruz, N.; Matos, A.; Cunha, S.; Silva, S. (2007), "ZARCO - An Autonomous Craft for Underwater Surveys", Proceedings of the 7th Geomatic Week, Barcelona, Spain, February, 2007.

Cumming, I. G.; Wong, F. H. (2005), “Digital Processing Of Synthetic Aperture Radar Data: Algorithms And Implementation”, Artech House Publishers, USA, January 2005.

Curlander, J. C.; McDonough, R. N. (1996), “Synthetic Aperture Radar: Systems and signal processing”, John Wiley & Sons, Inc., 6005 Third Avenue, New York, 1996.

Cutrona, L. J. (1975), “Comparison of sonar system performance achievable using synthetic-aperture techniques with the performance achievable by more conventional means”, The Journal of the Acoustical Society of America, vol. 58, Issue 2, pp. 336-348, August, 1975.

de Groot, S. J. (1996), “The physical impact of marine aggregate extraction in the North Sea”, ICES Journal of Marine Science, 53:1051-1053, 1996.

Denny, J. F.; O'Brien, T. F.; Bergeron, E.; Twichell, D.; Worley, C. R.; Danforth, W. W.; Andrews, B. A.; Irwin, B. (2006), “Advances in Shallow-Water, High-Resolution Seafloor Mapping: Integrating an Autonomous Surface Vessel (ASV) Into Nearshore Geophysical Studies”, AGU Fall Meeting Abstracts, pp. B1574+, 2006.

Farmer, D. M; Lemon, D. (1984), “The Influence of Bubbles on the Ambient Noise in the Ocean at high Wind Speeds”, Journal of Physical Oceanography, vol. 14, issue 11, pp. 1762-1778, 1984.

Farmer, D. M; Deane, G. B., Vagle, S. (2001), “The influence of bubble clouds on acoustic propagation in the surf zone”, IEEE J. Oceanic Engineering, vol. 26, pp. 113-124, 2001.

Farrell, J. L., Jr. (1976), “Integrated Aircraft Navigation”, Academic Press, USA, 1976.

Foo, K.Y.; Atkins, P.R.; Collins, T. (2003), "Robust underwater imaging with fast broadband incoherent synthetic aperture sonar", Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal 2003 (ICASSP '03), vol. 5, pp. V-17-20, April, 2003.

Fornaro, G. (1999), "Trajectory deviations in airborne SAR: analysis and compensation", IEEE Transactions on Aerospace and Electronic Systems, vol. 35, no. 3, pp. 997-1009, July, 1999.

Fortune, S. A.; Hayes, M. P.; Gough, P. T. (2001), “Statistical Autofocus of Synthetic Aperture Sonar Images using Image Contrast Optimization”, Proceedings of the OCEANS 2001 Conference, 2001.

Frerking, M. (1994), “Digital Signal Processing In Communications”, Springer, 1st edition, USA, January, 1994.

Ghiglia, D. C.; Pritt, M. D. (1998), “Two-Dimensional Phase Unwrapping: Theory, Algorithms, and Software”, ISBN: 978-0-471-24935-1, May, 1998.

Gimeno, E.; Lopez-Sanchez, J. M. (2001), “Near-field 2-D and 3-D Radar Imaging Using a Chirp Scaling Algorithm”, Proceedings of the International Geoscience and Remote Sensing Symposium 2001, vol. 1, pp. 354-356, July, 2001.

Gold, R. (1996), “Characteristic linear sequences and their coset functions”, SIAM J. Appl. Math, vol.14, Issue 5, pp. 980-985, 1996.

Gough, P. T.; Hayes, M. P. (1989), “Measurements of acoustic phase stability in Loch Linnhe, Scotland”, The Journal of the Acoustical Society of America, vol. 86, Issue 2, pp.837-839, August, 1989.

Gough, P. T.; Hawkins, D. (1997), “Imaging Algorithms for a Strip-Map Synthetic Aperture Sonar: Minimizing the Effects of Aperture Errors and Aperture Undersampling”, IEEE Journal of Oceanic Engineering, vol. 22, no. 1, pp. 27-39, January, 1997.

Gough, P. T.; Hawkins, D. (1998), “Unified Framework for Modern Synthetic Aperture Imaging Algorithms”, The International Journal of Imaging Systems and Technology, vol. 8, pp. 343-358, 1998.

Griffiths, H.D.; Rafik, T.A.; Meng, Z.; Cowan, C.F.N.; Shafeeu, H.; Anthony, D.K. (1997), "Interferometric synthetic aperture sonar for high resolution 3-D mapping of the seabed", IEE Proceedings - Radar, Sonar and Navigation, vol. 144, no. 2, pp.96-103, April, 1997.

Griffiths, H.D.; Baker, C.J.; Currie, A.; Voles, R.; Bullock, R.; Brennan, P.V. (1994), "Aircraft-borne interferometric SAR for 3-D high-resolution radar imaging", IEE Colloquium on Radar and Microwave Imaging, pp.3/1-3/4, November, 1994.

Groen, J.; Sabel, J. C. (2002), “Interaction Between Autofocusing and Synthetic Aperture Processing: a Study Based on Simulated and Experimental Rail Data”, Proceedings of the sixth European Conference on Underwater Acoustics (ECUA), pp. 235-240, Gdansk, Poland, 2002.

Hansen, R. E.; Sæbø, T.O.; Callow, H. J.; Hagen, P.E.; Hammerstad, E. (2005), "Synthetic aperture sonar processing for the HUGIN AUV", Proceeding of the Oceans 2005 – Europe Conference, vol. 2, pp. 1090-1094, 2005.

Hansen, R. E., Callow, H. J.; Sæbø, T. O. (2007), “The Effect of Sound Velocity Variations on Synthetic Aperture Sonar”, Proceedings of the 2nd International Conference on Underwater Acoustic Measurements, Greece, June, 2007.

Hawkins, D. W. (1996), “Synthetic Aperture Imaging Algorithms: with Application to wide bandwidth sonar”, PhD thesis, Department of Electrical and Electronic Engineering, University of Canterbury, October, 1996.

Hawkins, D. W.; Gough, P. T. (1997), “Synthetic Aperture Imaging Using the Accelerated Chirp Scalling Algorithms”, Proceedings of the International Geoscience and Remote Sensing Symposium, vol. 1, pp. 471-473, August, 1997.

Hawkins, D. W.; Gough, P. T. (2004), “Temporal Doppler Effects in SAS”, Sonar Signal Processing, vol. 26, p. 5, 2004.

Hayes, M. P. (1989), “A CTFM Synthetic Aperture Sonar”, PhD Thesis, Department of Electrical and Electronic Engineering, University of Canterbury, September, 1989.

Hayes, M.P.; Gough, P.T. (1992), "Broad-band synthetic aperture sonar", IEEE Journal of Oceanic Engineering, vol. 17, no. 1, pp. 80-94, January, 1992.

Haykin, S. (2000), “Communication Systems”, Wiley, 4th edition, USA, May 15, 2000.

Heering, P. (1982), “A Synthetic Aperture Sonar Study”, Technical report, Huntech-Lapp Systems Limited, Scarborough, Ontario, MIR 3A6, August, 1982.

Heering, P. (1984), “Alternate Schemes in Synthetic Aperture Sonar Processing”, IEEE Journal of Oceanic Engineering, vol. OE-9, no.4, 1984.

Hein, A. (2003), “Processing of SAR Data: Fundamentals, Signal Processing, Interferometry”, Springer, 1st edition, 2003.

Hogenauer, E. B. (1981), “An economical class of digital filters for decimation and interpolation”, IEEE Trans. on Acc., Speech, and Signal Processing, vol. ASSP-29, 155-162, 1981.

Hungerbühler, N. (1998), “Singular Filters for the Radon Backprojection”, Journal of Applied Analysis, vol. 5, no. 1, pp. 17–33, 1998.

Hunter, A. J. ; Hayes, M. P. ; Gough, P. T. (2003), “A Comparison of Fast Factorised Back-Projection and Wavenumber Algorithms For SAS Image Reconstruction”, Proceedings of the World Congress on Ultrasonics, Paris, France, September, 2003.

Hunter, A. J. (2006), “Underwater Acoustic Modelling for Synthetic Aperture Sonar”, PhD thesis, Electrical and Electronic Engineering at the University of Canterbury, New Zealand, 2006.

Jensen, J.; Kuperman, W.; Porter, M.; Schmidt, H. (2000), “Computational Ocean Acoustics”, Springer, NY, USA, 2000.

Johnson, K. A. (1992), “Sonar Motion Detection Using a Common Phase Shift Technique”, Master’s thesis, Department of Electrical and Electronic Engineering, University of Canterbury, 1992.

Jolliffe, I. T. (2002), “Principal Component Analysis”, Springer; 2nd Edition, October, 2002.

Khayam, S. A. (2003), “The Discrete Cosine Transform (DCT): Theory and Application”, Department of Electrical & Computer Engineering Michigan State University, March, 2003.

Kinsler, L.; Frey, A.; Coppens, A.; Sanders, S. (1982), “Fundamentals of Acoustics”, Addison Wesley Longman Ltd, USA, 1982.

Kuipers, J. B. (1999), “Quaternions and rotation Sequences: a Primer with Applications to Orbits, Aerospace, and Virtual Reality”, Princeton University Press, USA, 1999.

Legris, M.; Amate, M.; Hetet, A. (2004), “Navigation data Fusion for Low Frequency Synthetic Aperture Sonar Focusing”, Proceedings of the 7th European Conference on Underwater Acoustics, ECUA 2004 Delft, The Netherlands, 2004.

Legris, M.; Jean, F. (2007), “Comparison between DPCA Algorithm and Inertial Navigation on the Ixsea Shadows SAS”, Proceddings of the Oceans 2007 – Europe Conference, Aberdeen, UK, 2007.

Liu, Q.; Vesecky, J. F. (1990), “Registration of Interferometric SAR images”, Proceedings of the Geoscience and Remote Sensing Symposium IGARSS, 1990.

Liu, W.; Pokharel, P. P.; Principe, J. C.. (2006), “Correntropy: A Localized Similarity Measure”, International Joint Conference on Neural Networks, pp. 4919 – 4924, 2006.

Madsen, S. N. (2001), “Motion Compensation for Ultra Wide Band SAR”, Proceedings of the IEEE nternational Geoscience and Remote Sensing Symposium 2001, pp. 1436-1438, July, 2001.

Mackenzie, K.V. (1981), “Nine-term equation for the sound speed in the oceans”, J. Acoust. Soc. Am. 70(3), pp. 807-812, 1981.

Mansourpour, M.; Rajabi, M.; Blais, J. (2006), “Effects and Performance of Speckle Noise Reduction Filters on Active Radar and SAR Images”, Proceeding of the ISPRS, Ankara, Turkey, 2006.

McHugh, R.; Shaw, S.; Taylor, N. (1998), "Spatial sampling ambiguities in synthetic aperture sonar for broadside and squint look operations", Proceedings of the OCEANS 1998 Conference, vol. 2, pp. 960-964, October, 1998.

Mitola, J. (1997), “Software Radio Technology Challenges and Opportunities”, Software Radio Workshop, Brussels, Belgium, 1997.

Nelson, M. A. (1998), “DARPA Synthetic Aperture Sonar”, Proceedings of the Adaptive Sensor Array Processing (ASAP) Workshop, vol. 1, pp. 141-155, 15 May, 1998.

North, D. O. (1943), "An analysis of the factors which determine signal/noise discrimination in pulsed carrier systems", RCA Labs, Princeton, NJ, USA, Rep. PTR-6C, 1943.

Oppenheim A. V.; Schafer, R. W.; Buck, J. R. (1999), “Discrete-Time Signal Processing”, Prentice Hall, 2nd edition, USA, January 10, 1999.

Pinto, M. A.; Hollet, R. D.; Bellettini, A.; Chapman, S. (2002), “Bathymetric Imaging with Wideband Interferometric Synthetic Aperture Sonar”, IEEE Journal of Oceanic Engineering, 2002.

Porter, M. B., and the KauaiEx Group (2004), “Results from the KauaiEx experiments,” in Proceedings of the High-Frequency Ocean Acoustics Conference, AIP, 2004.

Pratson, L. F.; Edwards, M. H. (1996), “Introduction to advances in seafloor mapping using sidescan sonar and multibeam bathymetry data”, Journal of Marine Geophysical Researches, vol. 18, no. 6, December, 1996.

Preisig, J. (2005), “Performance Analysis of Adaptive Equalization for Coherent Acoustic Communications in the Time-Varying Ocean Environment”, J. Acoust. Soc. Am., vol. 118, pp. 263-278, 2005.

Pryor, D.E. (1998), “Theory and test of bathymetric side scan sonar”, Proceedings of the OCEANS 1988, 'A Partnership of Marine Interests', vol. 2, pp. 379-384, October, 1998.

Putney, A.; Anderson, R.H. (2005), "Reconstruction of undersampled SAS data using the WIPE algorithm", Proceedings of the OCEANS 2005 MTS/IEEE Conference, vol. 1, pp. 111-118, 2005.

Rihaczek, A.W. (1966), "Doppler-tolerant signal waveforms", Proceedings of the IEEE, vol. 54, no. 6, pp. 849-857, 1966.

Raney, R.K.; Runge, H.; Bamler, R.; Cumming, I.G.; Wong, F.H. (1994), "Precision SAR processing using chirp scaling", IEEE Transactions on Geoscience and Remote Sensing, , vol. 32, no. 4, pp. 786-799, July, 1994.

Rayleigh, J. W. S.; (1945), “The Theory of Sound: Two Volumes”, Dover Publications, Second Revised Edition edition, UK, January 1, 1945.

Richards, M.A. (2007), "A Beginner's Guide to Interferometric SAR Concepts and Signal Processing [AESS Tutorial IV]", IEEE Aerospace and Electronic Systems Magazine, vol. 22, no. 9, pp. 5-29, September, 2007.

Ruderman, Daniel L. (1994), “The statistics of natural images”, Network: Computation in Neural Systems, no. 5, pp. 517-548, UK, 1994.

Sæbø, T.O.; Callow, H.J.; Hansen, R.E.; Langli, B.; Hammerstad, E.O. (2007), "Bathymetric Capabilities of the HISAS Interferometric Synthetic Aperture Sonar", Proceedings of the OCEANS 2007 MTS/IEEE Conference, pp. 1-10, September, 2007.

Shippey, G.; Banks, S.; Phil, J. (2005), "SAS image reconstruction using Fast Polar Back Projection: comparisons with Fast Factored Back Projection and Fourier-domain imaging", Proceedings of the OCEANS 2005 MTS/IEEE-Europe Conference, vol. 1, pp. 96-101, June, 2005.

Silkaitis, J.M.; Douglas, B.L.; Hua Lee. (1995), "Synthetic-aperture sonar imaging: system analysis, image formation, and motion compensation", Conference Record of the Twenty-Ninth Asilomar Conference on Signals, Systems and Computers 1995, vol. 1, pp. 423-427, October, 1995.

Silva, S.; Chilro, R.; Cunha, S. (2001), "Building an Inexpensive Time Reference System using an OEM GPS Receiver and RT-Linux", Proceeding of the ION GPS Conference, Salt Lake City, USA, 2001.

Silva, S.; Cunha, S.; Matos, A.; Cruz, N. (2007 a). "An In-SAS System For Shallow Water Surveying", Proceedings of the 7th Geomatic Week, Barcelona, Spain, February, 2007.

Silva, S.; Cunha, S.; Matos, A.; Cruz, N. (2007 b), "An Autonomous Boat Based Synthetic Aperture Sonar," Proceedings of the OCEANS 2007 MTS/IEEE Conference, pp. 1-7, September, 2007.

Silva, S.; Cunha, S.; Matos, A.; Cruz, N. (2008 a), "An Algebraic Approach to Synthetic Aperture Sonar Image Reconstruction", Proceedings of the OCEANS 2008 MTS/IEEE Conference, September, 2008.

Silva, S.; Cunha, S.; Matos, A.; Cruz, N. (2008 b), "Sub-Band Processing of Synthetic Aperture Sonar Data", Proceedings of the OCEANS 2008 MTS/IEEE Conference, September, 2008.

Silva, S.; Cunha, S.; Matos, A.; Cruz, N. (2008 c), "Shallow Water Height Mapping With Interferometric Synthetic Aperture Sonar", Proceedings of the OCEANS 2008 MTS/IEEE Conference, September, 2008.

Silva, S.; Coiras, E.; Groen, J.; Cunha, S.; Matos, A.; Cruz, N.; Heremans, R.; Dupont, Y.; Acheroy, M.; Øyehaug, L.; Skartlien, R.; Moura, N, Simas, E.; Seixas, J.; Maussang, F.; Chanussot, J.; Rombaut, M.; Amate, M.; Thurman, E.; Riordan, J.; Toal, D.; Malthus, T.; Karpouzli, E.; Franco, F.; Burguera, A.; González, Y.; Oliver, G. (2009 a), "Advances in Sonar Technology", In-Tech, ISBN 978-3-902613-48-6, February, 2009.

Silva, S.; Cunha, S.; Matos, A.; Cruz, N. (2009 b), "Shallow Water Surveying Using Experimental Interferometric Synthetic Aperture Sonar", Marine Technology Society Journal, vol. 43, no. 1, pp. 50-63, 2009.

Silva, S.; Cunha, S.; Matos, A.; Cruz, N. (2009 c), "An Autonomous Surface Vehicle InSAS for Maintenance and Surveillance Applications", Proceedings of the SMART 2009 Conference, Porto, Portugal, July, 2009.

Tomiyasu, K. (1978), "Tutorial review of synthetic-aperture radar (SAR) with applications to imaging of the ocean surface", Proceedings of the IEEE, vol. 66, no. 5, pp. 563-583, May, 1978.

Topczewski, L.; Fernandes, F. M.; Cruz, P. J. S.; Lourenço, P. B. (2007), "Practical implications of GPR investigation using 3D data reconstruction and transmission tomography", Journal of Building Appraisal, 3(1), 59-76, 2007.

Ulander, L.; Hellsten, H.; Stenstrom, G. (2000), "Synthetic Aperture Radar Processing using Fast Factorized Back-Projection", Proceedings of the EUSAR 2000, 3rd European Conference on Synthetic Aperture Radar, pp. 753-756, Munich, Germany, 2000.

Ulander, L.; Hellsten, H.; Stenstrom, G. (2003), "Synthetic Aperture Radar Processing using Fast Factorized Back-Projection", IEEE Transactions on Aerospace and Electronic Systems , vol. 39, no. 3, pp. 760-776, July, 2003.

Urkowitz, H.; Bucci, N.J. (1992), "Using Complementary Sequences with Doppler Tolerance for Radar Sidelobe Suppression in Meteorological Radar", Geoscience and Remote Sensing Symposium, 1992. IGARSS '92. International , vol. 2, pp.1737-1738, 1992.

Vasconcelos, G.; Fernandes, F. M.; Lourenço, P. B.; Duarte, F. (2008), "Non-destructive evaluation of reinforced concrete masonry walls", 8th International Seminar on Structural Masonry, Istanbul Technical University, Turkey, 2008.

Waite, A. D. (2005), "Sonar for Practicing Engineers 3rd Edition", John Wiley & Sons, West Sussex, England, 2005.

Wang, L.; Belletini, A.; Fioravanti, S.; Chapman, S. ; Bugler, D. R.; Perrot, Y.; Hétet, A. (2001), "InSAS'00: Interferometric SAS and INS aided SAS imaging", Proceeding of the IEEE Oceans 2001 Conference, vol. 1, pp. 179-187, MTS/IEEE, 2001.

Wen Yang; Chuanzhao Han; Hong Sun; Yongfeng Cao, (2005), "Registration of high resolution SAR and optical images based on multiple features", Proceedings of the IEEE International Geoscience and Remote Sensing Symposium - 2005, vol. 5, Issue 25-29, pp. 3542-3544, July, 2005.

Welch, G.; Bishop, G. (2006), "An Introduction to the Kalman Filter", University of North Carolina at Chapel Hill, Department of Computer Science, USA, 2006.

Wettergren, T.A.; Traweek, C.M. (2005), "Optimization of conventional beamformer shading weights for conformal velocity sonar", IEEE Journal of Oceanic Engineering, vol. 30, no.1, pp. 213-220, 2005.

Whalen, J.E.; Wallace, G. (2005), "Multi-platform sonar-based systems for submerged threats operating in littorals", OCEANS, 2005. Proceedings of MTS/IEEE , vol. 1, pp. 684-689, 2005.

Wilkinson, A.J.; Griffiths, H.D. (1994), "Novel techniques for 3-D target reconstruction in SAR interferometry", IEE Colloquium on Radar and Microwave Imaging, pp. 1/1-1/6, November, 1994.

Wilkinson, D. R. (2001), “Efficient Image Reconstruction Techniques for a Multiple-Receiver Synthetic Aperture Sonar”, Master’s thesis, Department of Electrical and Electronic Engineering, University of Canterbury, April, 2001.

Wong, G. S. K.; Zhu, Z. (1995), “Speed of sound in seawater as a function of salinity, temperature and pressure”, J. Acoust. Soc. Am. 97(3), pp. 1732-1736, 1995.

Yang, C.; Duraiswami, R.; Gumerov, N. A.; Davis, L. (2003), “Improved Fast Gauss Transform and Efficient Kernel Density Estimation”, Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV’03), vol. 1, pp. 664 - 671, 2003.

Zebker H. A.; Lu, Y. (1998), “Phase unwrapping algorithms for radar interferometry: residue-cut, least-squares, and synthesis algorithms”, JOSA A, vol. 15, Issue 3, pp. 586-598, March, 1998.