

OPTI-ACOUSTIC STEREO IMAGING

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED
SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

HAKAN SAÇ

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
ELECTRICAL AND ELECTRONICS ENGINEERING

SEPTEMBER, 2012

Approval of the thesis:

OPTI-ACOUSTIC STEREO IMAGING

Submitted by **HAKAN SAÇ** in partial fulfillment of the requirements for the degree of **Master of Science in Electrical and Electronics Engineering Department, Middle East Technical University** by,

Prof. Dr. Canan Özgen
Dean, Graduate School of **Natural and Applied Sciences**

Prof. Dr. İsmet Erkmen
Head of Department, **Electrical and Electronics Engineering**

Prof. Dr. Kemal Leblebicioğlu
Supervisor, **Electrical and Electronics Engineering, METU**

Assoc. Prof. Dr. Özgül Salor Durna
Co-Supervisor **Electrical and Electronics Engineering, Gazi Univ.**

Examining Committee Members:

Prof. Dr. Uğur Halıcı
Electrical and Electronics Engineering, METU

Prof. Dr. Kemal Leblebicioğlu
Electrical and Electronics Engineering, METU

Prof. Dr. Gözde Bozdağı Akar
Electrical and Electronics Engineering, METU

Prof. Dr. Tolga Çiloğlu
Electrical and Electronics Engineering, METU

Assoc. Prof. Dr. Özgül Salor Durna
Electrical and Electronics Engineering, Gazi Univ.

Date: 13.09.2012

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name : Hakan SAÇ

Signature :

ABSTRACT

OPTI-ACOUSTIC STEREO IMAGING

Saç, Hakan

M. Sc., Department of Electrical and Electronics Engineering

Supervisor: Prof. Dr. Kemal Leblebicioğlu

Co-Supervisor: Assoc. Prof. Dr. Özgül Salor Durna

September 2012, 91 Pages

In this thesis, opti-acoustic stereo imaging, which is the deployment of two-dimensional (2D) high frequency imaging sonar with the electro-optical camera in calibrated stereo configuration, is studied.

Optical cameras give detailed images in clear waters. However, in dark or turbid waters, information coming from electro-optical sensor is insufficient for accurate scene perception. Imaging sonars, also known as acoustic cameras, can provide enhanced target details under these scenarios. To illustrate these visibility conditions, a 2D high frequency imaging sonar simulator as well as an underwater optical image simulator is developed. A computationally efficient algorithm is also proposed for the post-processing of the returned sonar signals.

Where optical visibility allows, integration of the sonar and optical images effectively provides binocular stereo vision capability and enables the recovery of three-dimensional (3D) structural information. This requires solving the feature

correspondence problem for these completely different sensing modalities. Geometrical interpretation of this problem is examined on the simulated optical and sonar images. Matching the features manually, 3D reconstruction performance of opti-acoustic system is also investigated. In addition, motion estimation from opti-acoustic image sequences is studied.

Finally, a method is proposed to improve the degraded optical images with the help of sonar images. First, a nonlinear mapping is found to match local the features in opti-acoustical images. Next, features in the sonar image is mapped to the optical image using the transformation. Performance of the mapping is evaluated for different scene geometries.

Keywords: Sonar simulator, underwater optical image simulator, epipolar geometry, 3D reconstruction, structure from motion, image enhancement.

ÖZ

OPTİ-AKUSTİK STEREO GÖRÜNTÜLEME

Saç, Hakan

Yüksek Lisans, Elektrik ve Elektronik Mühendisliği Anabilim Dalı

Tez Yöneticisi: Prof. Dr. Kemal Leblebicioğlu

Ortak Tez Yöneticisi: Doç. Dr. Özgül Salor Durna

Eylül 2012, 91 Sayfa

Bu tezde, iki boyutlu (2B) yüksek frekanslı bir sonarın, bir elektro-optik kamerayla kalibre edilerek birlikte kullanılması anlamına gelen opti-akustik stereo görüntüleme üzerinde çalışılmıştır.

Optik kameralar temiz sularda detaylı görüntüler vermektedir. Bununla birlikte, karanlık ve bulanık sularda, elektro-optik sensörlerden gelen bilgi sahneyi hassas bir biçimde algılamak için yetersiz kalmaktadır. Akustik kamera olarak da bilinen görüntüleme sonarları bu koşullar altında dahi sahne yerleşimiyle ilgili yeterli bilgiyi sağlayabilmektedir. Bahsedilen ortamlardaki görüşü modelleyebilmek için, bir yüksek frekanslı 2B görüntüleme sonarı simülatörü ve bir sualtı optik görüntü simülatörü geliştirilmiştir. Geri dönen sonar sinyallerinin işlenebilmesi için de sayısal olarak verimli bir algoritma önerilmiştir.

Optik görüntüleme açısından uygun olan koşullarda sonar ve optik görüntülerin birleştirilmesi stereo görüş kabiliyeti sağlamanın yanı sıra 3B sahne bilgisinin

iyileştirilmesine yardımcı olmaktadır. Bunu gerçekleştirebilmek içinse, tamamen farklı olan bu iki görüntüleme yönteminde eşleşen özellikler bulunmalıdır. Bu problemin geometrik olarak yorumlanması, simülasyonla elde edilen sonar ve optik görüntüler üzerinde yapılmıştır. Özelliklerin elle eşleştirilmesiyle, opti-akustik stereo sistemin 3B geriçatım performansı da incelenmiştir. Buna ek olarak, opti-akustik görüntülemeyle hareket kestirimi üzerine çalışılmıştır.

Son olarak, bozuk optik görüntülerin sonar görüntüsü yardımıyla iyileştirilmesi için bir yöntem önerilmiştir. Öncelikle, opti-akustik görüntülerde yerel olan bazı özellikleri eşleştirerek lineer olmayan bir haritalama bulunmuştur. Daha sonra, bu dönüşüm yardımıyla, sonar görüntüsündeki özellikler optik görüntüye haritalanmıştır. Bulunan haritalamanın performansı, farklı sahne geometrileri için karşılaştırılmıştır.

Anahtar Kelimeler: Sonar simülatörü, sualtı optik görüntü simülatörü, eşkutupsal geometri, 3B geriçatım, hareketten-yapı, görüntü iyileştirme.

ACKNOWLEDGEMENTS

I would like to thank my supervisor Prof. Dr. Kemal Leblebiciođlu for his guidance, advice, criticism, patience, and encouragement throughout the completion of thesis.

I would like to thank my co-supervisor Assoc. Prof. Dr. Özgöl Salor Durna for her guidance, encouragement and support during my studies.

I wish to express my special thanks to Prof. Dr. Gözde Bozdađı Akar for her guidance, advice and patience throughout the studies.

I also would like to thank Prof. Dr. Uđur Halıcı for her advice in my studies.

Thanks to each one of my officemates. It is their companion that made my time in Computer Vision Laboratory so enjoyable and valuable.

This work is funded by TÜBİTAK with in the scope of the project 111E267.

TABLE OF CONTENTS

ABSTRACT	iv
ÖZ	vi
ACKNOWLEDGEMENTS	viii
TABLE OF CONTENTS	ix
LIST OF TABLES	xii
LIST OF FIGURES	xiii
CHAPTERS	1
1 INTRODUCTION	1
1.1 Literature Review	2
1.2 Scope of the Thesis	4
1.3 Outline of the Thesis	5
2 2D IMAGING SONAR SIMULATOR	6
2.1 Basics of 2D Sonar Imaging	6
2.2 The Acoustic Wave Equation.....	7
2.3 Sonar Simulation Models	8
2.3.1 Frequency Domain Models	8
2.3.2 Finite Difference Models	8
2.3.3 Ray Tracing.....	8
2.4 SIGMAS Simulator.....	9
2.5 Model Simplification	10
2.6 Image Formation	12
2.6.1 Scanning the Scene	12
2.6.2 Processing of Data.....	13
2.6.2.1 Multiple Returns and Bright Areas	15
2.6.2.2 Acoustic Shadows	19

2.6.2.3	Decimation	20
2.6.2.4	Transformation	21
2.7	Performance of the Sonar	21
2.7.1	Noise	22
2.7.2	Reverberation	22
2.8	Results	23
3	UNDERWATER OPTICAL IMAGE SIMULATOR	28
3.1	Underwater Vision	28
3.2	Image Components.....	29
3.2.1	Absorbtion.....	29
3.2.2	Scattering.....	30
3.2.3	Jaffe-McGlamery Model.....	31
3.2.3.1	Signal	32
3.2.3.2	Forward Scatter	32
3.2.3.3	Backscatter	33
3.3	Image Formation	34
3.3.1	Pinhole Camera Model.....	34
3.3.2	Simulation Steps	35
3.4	Results	36
4	EPIPOLAR GEOMETRY OF OPTI-ACOUSTIC STEREO IMAGING	39
4.1	Introduction	39
4.2	Epipolar Geometry	40
4.2.1	Sonar to Optical Mapping	41
4.2.2	Optical to Sonar Mapping	44
4.3	Results	45
5	3D RECONSTRUCTION FROM OPTI-ACOUSTIC STEREO IMAGING	48
5.1	3D Reconstruction in Stereo Optical System.....	48
5.2	3D Reconstruction in Opti-Acoustic System	49
5.2.1	Indirect Method	51
5.2.2	Direct Method	52
5.2.3	First-Order Error Analysis	53
5.3	Results	53
6	3D MOTION ESTIMATION FROM OPTI-ACOUSTIC STEREO IMAGING ..	59

6.1	Introduction	59
6.2	Preliminaries	61
6.2.1	Motion Transformation and Plane Representation	61
6.2.2	Image Measurements and Transformations	62
6.3	Up-to-Scale Estimation	63
6.4	Resolving the Scale Factor Ambiguity	64
6.5	MLE Formulation	64
6.6	Results	65
7	UNDERWATER OPTICAL IMAGE ENHANCEMENT	71
7.1	Relationship between Matching Optical and Sonar Image Points	71
7.2	RBF Neural Networks	73
7.3	Mapping Performance	74
7.3.1	Planar Scene	74
7.3.2	Non-Planar Scene	75
7.4	Results	79
7.4.1	Planar Scene	79
8	CONCLUSIONS	82
	APPENDIX	86
	BIBLIOGRAPHY	88

LIST OF TABLES

TABLES

Table 2.1: Distribution functions.	22
Table 2.2: Estimated parameters for 40°–60° grazing angle	23
Table 2.3: Results for selected sites and for 40°–60° grazing angle.....	23
Table 3.1: The colors of the visible light spectrum.....	29
Table 3.2: Volume absorption and scattering coefficients for pure water.	30
Table 4.1: Variation in the shape of the epipolar contours	44

LIST OF FIGURES

FIGURES

Figure 1.1: Optical image of an outboard gear in turbid water	1
Figure 1.2: Sonar image of a pipeline	2
Figure 2.1: Beam emitted from transducer.	6
Figure 2.2: Lambert illumination model.	11
Figure 2.3: 3D model of a barrel and its simplified sonar image.....	11
Figure 2.4: Rays transmitted from point source in specified directions.	12
Figure 2.5: Intersection of the emitted ray with a scene point	13
Figure 2.6: Ray sampling techniques. a) Angular sampling, b) spatial sampling.....	14
Figure 2.7: a) Received raw data, b) data after sorting.	15
Figure 2.8: Multiple returns to sonar receiver.....	15
Figure 2.9: Rays and their intersections with a) cuboid, b) sphere and c) ellipsoid. .	16
Figure 2.10: Scatter profile of a) plane, b) sphere and c) ellipsoid.....	17
Figure 2.11: a) Received signal, b) resulting sonar signal after summation.....	18
Figure 2.12: Region that can not be illuminated by sound pulses.	19
Figure 2.13: Acoustic shadows in processed intensity signal.	20
Figure 2.14: a) A rectangular image and b) its polar form.....	21
Figure 2.15: Scene 1 which consists various geometrical figures.	24
Figure 2.16: Image of scene 1 after decimation step	24
Figure 2.17: a) Polar image, b) effect of noise and reverberation.....	25
Figure 2.18: a) ROV approaching the bridge footers, b) sonar image of footers taken with Blueview P900, c) simulated image.	26
Figure 3.1: An underwater scene about 5 meters deep.	28
Figure 3.2: Volume scattering function.	31
Figure 3.3: Rays for the signal and backscatter	32
Figure 3.4: Light scattered towards the LOS by nearby particles.....	33
Figure 3.5: Pinhole camera	34
Figure 3.6: Points where the optical rays are directed	35
Figure 3.7: a) True illumination of scene in Figure 2.15, b) the effect of attenuation	36
Figure 3.8: Effect of forward scatter in a) and backscatter in b).....	37
Figure 4.1: Relationship between various parameters of a typical configuration.....	40
Figure 4.2: Sonar coordinate system.....	41
Figure 4.3: Epipolar contours in optical image are hyperbolas for nearly parallel configuration, a) and become ellipses for larger vergence angles b).....	42

Figure 4.4: Feature points in sonar image a) and corresponding epipolar curves in optical image b).....	45
Figure 4.5: Feature points in optical image a) and corresponding epipolar curves in sonar image b).	46
Figure 5.1: Ideal 3D reconstruction	48
Figure 5.2: Minimization of geometric error.	49
Figure 5.3: Geometric interpretation of range solution a) and azimuth solution b)...	50
Figure 5.4: Klein bottle	53
Figure 5.5: Reconstruction of Klein bottle points from noisy data for baseline of 25 cm and average distance of 1.25 m.	54
Figure 5.6: Reconstruction of Klein bottle from noisy points for baseline of 50 cm and average distance of 1.25 m	55
Figure 5.7: Reconstruction of Klein bottle from noisy points for baseline of 25 cm and average distance of 3 m.	56
Figure 5.8: Reconstruction of Klein bottle from noisy points for baseline of 50 cm and average distance of 3 m.	57
Figure 6.1: Rotation and translation in monocular system a) and induced motions in monocular and stereo systems b).	60
Figure 6.2: Position of opti-acoustic system relative to the scene.	66
Figure 6.3: Images of the scene in consecutive views.	66
Figure 6.4: 3D reconstruction based on initial estimation of motion.....	67
Figure 6.5: Planar estimator.	68
Figure 6.6: MLE with clustering.....	69
Figure 7.1: RBF network architecture.....	73
Figure 7.2: Training points in optical a) and sonar b) image of the rectangular grid. 74	
Figure 7.3: Test points on sonar image a), result of the sonar to optical mapping b).75	
Figure 7.4: Training points in optical a) and sonar b) image of cuboids, result of the sonar to optical mapping of training points.....	76
Figure 7.5: Test points on sonar image a), result of the sonar to optical mapping b).77	
Figure 7.6: Matching feature points of a cuboid.	77
Figure 7.7: Transformation of training a) and test b) points to optical image with a modified mapping.	78
Figure 7.8: Rectangular grid a), its optical b) and sonar image c).	79
Figure 7.9: Training points on a) optical image, b) sonar image and c) selected test points on sonar image.....	80
Figure 7.10: Result of the sonar to optical mapping for various mappings.	81

CHAPTER 1

INTRODUCTION

Visual inspection, repair, maintenance, security, search and survey are among critical underwater applications in marine sciences. Autonomous underwater vehicles (AUVs) and remotely operated vehicles (ROVs), also together known as unmanned underwater vehicles (UUVs), equipped with specialized imaging systems are effective platforms for these operations. Automation of these operations reduces the costs as well as eliminates the risk on human divers in underwater surveys [1].

Optical cameras give high-detailed images of underwater environment under good visibility conditions. However, in deep and turbid waters, these sensors can become totally ineffective where the visibility range decreases to several meters as shown in Figure 1.1. This deficiency can be overwhelmed with the integration of a different sensing modality such as two dimensional (2D) imaging sonar.



Figure 1.1: Optical image of an outboard gear in turbid water.

In recent years, high frequency 2D sonar cameras have emerged (e.g., Blueview [2] and Dual Identification Sonar (DIDSON) [3]). Despite the lower resolution and higher noise levels compared to optical cameras, these systems provide high enough visual details (Figure 1.2) to human operator in underwater surveys even in muddy waters where the information obtained from optical camera is inadequate.

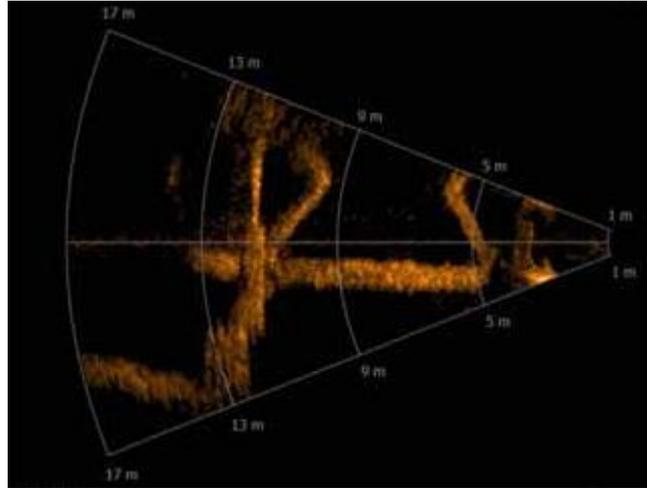


Figure 1.2: Sonar image of a pipeline [2].

Under sufficient visibility conditions for optical camera, fusion of optical and acoustical information can improve the performance of these systems compared to processing of images from each sensor, alone. Hence, exploiting the unique strengths and properties of these complementary sensors, limitations and shortcomings of each sensing modality can be overcome. This has motivated an investigation in deploying these multimodal cameras in stereo configuration namely “*opti-acoustic stereo imaging*” as a novel strategy for enhanced scene perception of underwater environments which requires the development of computer vision algorithms that provide three dimensional (3D) interpretation of sonar imagery.

1.1 Literature Review

Opti-acoustic stereo imaging is a relatively new discipline in computer vision area. Hence, limited number of research has been presented up-to now. These studies on opti-acoustical stereo imaging and image simulators (sonar and optical) can be summarized as follows:

In [4], Bell presents an application of optical ray tracing technique to the simulation underwater side-scan sonar imagery where the images are generated by the use of acoustic signals. Basic methodology of the sonar simulation model and the techniques used to circumvent the time based display problems is illustrated.

In [5], Coiras et al. proposed a method for the simulation of synthetic aperture sonar images which is based on the frequency domain approach. In this implementation, the sonar image is generated by expressing the Fourier transform of the received signal in terms of transmitted signal.

In [6], Lyons et al. presents the results for the statistical characterization of shallow-water seafloor reverberation. Several appropriate models of backscatter probability density function (PDF) is examined in order to better describe the measured amplitude distributions.

In [7], Metari et al. proposed a polychromatic method for light dispersion based on multiple scattering. This model involves two basic concepts, namely attenuation and ambient illumination. The resulting model can be used to restore the degraded images as well.

In [8], Sedlazeck et al. proposed a method to simulate underwater images using physical models for light attenuation and scattering. The Jaffe-McGlamery model [9, 10] for effects on colors is used as a base for simplification. Extension of the basic model to incorporate color images, shadows and several light sources is investigated.

In [11], Negahdaripour presents the epipolar geometry of opti-acoustical stereo imaging using a similar approach in traditional stereo optical vision. The constraint equations for the epipolar geometry is derived by using the asymmetrical projection models of these two sensing modalities.

In [1], Pirsiavash et al. proposed a method for the calibration of opti-acoustic stereo system. Matching the features manually, rotation and translation of optical camera relative to sonar is computed by minimizing the Mahalanobis distance between measured and reprojected points where the features are obtained from planar grid.

In [12], Sekkati et. al presents 3D reconstruction using opti-acoustic stereo imaging. Initial noisy closed form solutions is obtained using the epipolar constraints in [11]. Improvements on initial solutions is obtained with direct and indirect optimization methods.

In [13], Negahdaripour et al. proposed a method to estimate 3D motion by using the visual cues in sonar and optical image sequences for vision based navigation. This approach improves the estimates and resolves the scale factor ambiguity of motion based on solely monocular vision sequences.

1.2 Scope of the Thesis

In this thesis, advantages of using 2D imaging sonar with an optical camera in calibrated stereo configuration is examined. Not having yet an opportunity to use these completely different imaging systems in a controlled real environment, a 2D high frequency imaging sonar and an underwater optical image simulator is developed for the examination of various opti-acoustic imaging algorithms.

In the next step, theoretical background of epipolar geometry of the opti-acoustic system is presented and derived results are tested on the simulated sonar and optical images of a 3D scene with known geometry.

In addition to studying the correspondence problem, 3D reconstruction from opti-acoustic images is examined. Closed form solutions are obtained from geometrical constraints and improvements on these results via direct and indirect optimization methods is evaluated. These results are compared with the results of traditional stereo optical systems for the same configurations.

Motion estimation from opti-acoustic stereo sequences relative to planar surfaces is also investigated. In this part, up-to-scale estimates of the translation and plane normal is obtained from monocular optical image sequences, then the scale factor ambiguity is resolved from a minimum of one sonar correspondence. This result is further improved using all available optical and sonar correspondences via optimization.

Finally, exploiting the advantages of 2D sonar in dark and turbid waters over optical cameras, a method is proposed for the fusion of the opti-acoustic images. Using a nonlinear mapping which is based on Radial Basis Function (RBF) neural network, features points in sonar image are transformed to optical image to enhance the scene perception where the parameters of mapping is obtained via optimization. The performance of the mapping is evaluated for different scene geometries in terms of the sonar imagery.

1.3 Outline of the Thesis

The organization of the thesis is as follows: In Chapter 2, a 2D high frequency forward-looking sonar simulator is developed. Steps and requirements of image formation process is examined in detail. The chapter ends with the example sonar images from the developed simulator.

In Chapter 3, an underwater optical image simulator is developed. Steps of image formation process is explained in detail. The chapter ends with an example of underwater optical image from the developed simulator.

In Chapter 4, epipolar geometry of the opti-acoustic stereo imaging is investigated. Theoretical results are examined on the simulated sonar and optical images.

In Chapter 5, 3D reconstruction techniques from the opti-acoustic images is examined. Theoretical results are examined on a 3D structure with known geometry.

In Chapter 6, motion estimation from the opti-acoustic image sequences is investigated. Theoretical results are examined on a 3D structure with known geometry.

In Chapter 7, a method is proposed to improve the optical images by means of sonar images. Performance of the method is evaluated for different scene geometries.

The thesis is concluded with the final discussions and comments in Chapter 8.

CHAPTER 2

2D IMAGING SONAR SIMULATOR

In this chapter, simulation of 2D high frequency sonar images is studied. Steps and requirements of image formation process is explained in detailed.

2.1 Basics of 2D Sonar Imaging

Sonar uses propagation of sound to detect and identify objects mostly under the surface of water. There are two types of sonar. First one is the passive sonar which listens the sounds radiating from the environment. Second one is the active sonar which emits sound pulses into the area to be observed then listens for the echoes coming from objects that might be present in that area [14]. 2D imaging sonar is of active type.

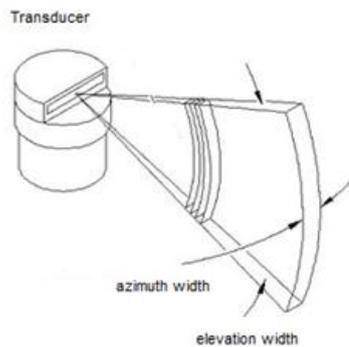


Figure 2.1: Beam emitted from transducer.

A single beam emitted from sonar transducer is shown in Figure 2.1. A 2D imaging sonar consists of an array of such transducers arranged in an order that yields the desired illumination of the scene by sound pulses. In Figure 2.1 azimuth width (θ_B)

and elevation width (ϕ_B) show the horizontal and vertical beamwidth of the emitted sound pulse respectively.

2D imaging sonars, sometimes referred to as acoustic cameras operate at high frequencies to provide the desired resolution. Blueview P900 sonar [2] operates at 900 KHz which utilizes 512 beams with $\theta_B = 1^\circ$ azimuth and $\phi_B = 20^\circ$ elevation widths produced by acoustic lens and sonar curvature. Transmitted sound pulses are spaced 0.18° apart horizontally, covering a total of 90° in horizontal axis.

A typical sonar image is obtained by plotting the intensity recorded at the receiver versus time for each beam in a particular azimuth direction. The geometry of the narrow sound pulses provides good localization accuracy in azimuth direction. However, it is considered as a 2D imaging system because there is a $\pm(\phi_B/2)^\circ$ uncertainty in the elevation of an imaged 3D scene point. This is a trade off because the narrower the elevation width of the pulse, the less uncertainty in elevation of the scene points at the cost of the narrower angle of view.

2.2 The Acoustic Wave Equation

Assuming the sound speed (c) is constant in the water column, the propagation of the acoustic wave can be written as [15]

$$\frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial y^2} + \frac{\partial^2 p}{\partial z^2} - \frac{1}{c} \frac{\partial^2 p}{\partial t^2} = -\delta(x - x_s, y - y_s, z - z_s) s(t) \quad (2.1)$$

where p is the acoustic pressure, $s(t)$ is the wave signal produced by point source located at the 3D point (x_s, y_s, z_s) and δ is the Dirac's delta function.

If sound propagates as in (2.1), the acoustic pressure which is a function of time and space is given by [5]

$$p(\mathbf{x}; t) = \frac{s(t - \frac{r}{c})}{4\pi r} \quad (2.2)$$

where $\mathbf{x} = [x; y; z]$ and r is the distance between \mathbf{x} and acoustic source.

$$r = \sqrt{(x - x_s)^2 + (y - y_s)^2 + (z - z_s)^2}. \quad (2.3)$$

Eq. (2.2) shows that the acoustic pressure is inversely proportional to r which is a consequence of spherical spreading law. Since 2D imaging sonar is active sonar, this attenuation in acoustical pressure applies twice because the sound travels first sonar to object and then object to sonar. Hence the sonar receives shifted and reduced sound signal. On the other hand Eq. 2.2 is valid only if the acoustic wave is transmitted by monopole point source.

2.3 Sonar Simulation Models

In practice, several models are used frequently for simulating the sonar image [5]. Some of these methods are as follows:

2.3.1 Frequency Domain Models

This approach uses the relationship between Fourier transform of the transmitted and received sound pulses. Filtering operations and including the point spread function of sonar to image formation process are easier in this approach.

2.3.2 Finite Difference Models

In this approach, the acoustic wave equation in (2.1) is solved numerically. The derivative of the acoustic pressure $p(\mathbf{x},t)$ w.r.t. time and space can be found using finite differences. Given initial conditions, i.e., $p(\mathbf{x},0)$, the acoustical pressure is calculated in space at any time. However, using finite difference methods for sonar which operates at high frequencies are impractical because of the high computational requirements.

2.3.3 Ray Tracing

Ray tracing is a technique used in computer graphics for generating images by tracing the path of light through the scene. With this technique, high quality sonar images can be obtained [16]. The main drawback of ray tracing is its high computational load.

2.4 SIGMAS Simulator

SIGMAS is a sonar simulator which was developed at Nato Underwater Research Center (NURC) [17]. It uses the frequency domain model in section 2.3.1. In sonar simulations there is trade off between the time required to form the image and the quality of the image. Hence, a sonar simulator should be sufficiently fast but at the same time should produce the realistic images. To assure these conditions, some assumptions should be made.

In view of 2D forward looking sonar imaging, since the objects of interest are in generally in the range of several meters, the sound of the speed in water is assumed to be constant which simplifies the tracing the paths of the sound pulses. Furthermore, to apply the Kirchoff scattered field approximation, imaged objects are assumed to made of discrete patches [5].

The acoustic pressure at the sonar receiver \mathbf{x}_r is a function of frequency. Using Green's theorem, the pressure at the surface patch A can be written as [18, 19]

$$P(\mathbf{x}_r; f) = \iint_{x \in A} [G(\mathbf{x}_r, \mathbf{x}; f) \nabla P(\mathbf{x}; f) - P(\mathbf{x}; f) \nabla G(\mathbf{x}_r, \mathbf{x}; f)] \mathbf{n}(\mathbf{x}) dA \quad (2.4)$$

where \mathbf{n} is the surface normal and G is Green's function. For homogenous media, G is described as [20]

$$G(\mathbf{x}_r, \mathbf{x}; f) = \frac{e^{ik|\mathbf{x}-\mathbf{x}_r|}}{4\pi|\mathbf{x}-\mathbf{x}_r|} \quad (2.5)$$

where $k = 2\pi f/c$ is the wave number.

Part of the pressure of the transmitted sound pulse transferred to the hitting surface patch. The surface reflectivity coefficient R gives proportion of reflected pressure (P_i) to transmitted wave pressure P . P_i and P are related as [5]

$$P(\mathbf{x}; f) = (1 + R(\mathbf{x}; f))P_i(\mathbf{x}; f) \quad (2.6)$$

$$\frac{\partial P(\mathbf{x}; f)}{\partial n} = 1 - R(\mathbf{x}; f) \quad (2.7)$$

where n shows the changes in the direction of the normal.

Assuming an acoustically rigid object, i.e., $R = 1$, further simplifies the calculations. Substituting 1 for R in (2.6) and (2.7) and using (2.5) twice for two way spherical spreading loss yields the equation:

$$P(\mathbf{x}_r; f) = \iint_{x \in A} -\frac{i e^{ik|\mathbf{x}-\mathbf{x}_r|} e^{ik|\mathbf{x}-\mathbf{x}_s|}}{\lambda |\mathbf{x} - \mathbf{x}_r| |\mathbf{x} - \mathbf{x}_s|} P(\mathbf{x}_s; f) \frac{\mathbf{x} - \mathbf{x}_r}{|\mathbf{x} - \mathbf{x}_r|} \mathbf{n}(\mathbf{x}) dA \quad (2.8)$$

where \mathbf{x}_s is the location of the acoustic source.

If the surface of the object assumed to made of small patches with an equal area of a^2 , the integral in (2.8) can be written as a summation [5]

$$P(f) = \frac{a^2}{\lambda} \sum_k \frac{\hat{\mathbf{n}}_k \cdot \hat{\mathbf{r}}_k}{\|\mathbf{r}_k\|^2} S(f) e^{i\frac{2\pi f}{c} \|\mathbf{r}_k\|} \quad (2.9)$$

where r_k is the vector to k_{th} surface patch, $S(f)$ is the Fourier transform of the $s(t)$ and caps indicate the unit vectors.

2.5 Model Simplification

Images produced using (2.9) is very realistic but this model discretizes the simulated objects facets smaller than the wavelength of the transmitted signal, which requires high computational load.

If the acoustic rigidity of objects ignored, Eq. (2.9) can be further simplified. On the other hand, if surfaces are rough considering the wavelength of the transmitted signal, complex exponentials can be omitted and same reflectivity values can be obtained using bigger surface patches [5]. To compensate the effects of the complex exponentials, experimentally determined noise model should be used [16]. Reduction in the pressure level at the received signal due to transmission loss can be improved by Time-Varying-Gain (TVG) which removes the need for the inversely squared term in (2.9). If all the calculations are made in image space then there is no need to the Fourier transform which results the intensity I of a pixel at point \mathbf{r}

$$I(\mathbf{r}) = K(\mathbf{r}) \hat{\mathbf{n}}_k \cdot \hat{\mathbf{r}}_k R_k(\mathbf{r}) \chi_k(\mathbf{r}) \quad (2.10)$$

where the sonar is at origin, K includes the all scaling factors, R_k is the acoustical reflectivity of the k_{th} surface patch and χ_k is the characteristic function of the patch. (2.10) is fundamentally Lambertian illumination model for diffuse reflectance [21] which is also known as ideal diffuse reflection. In this model the observed intensity of a scene point only depends on the direction of the illumination.

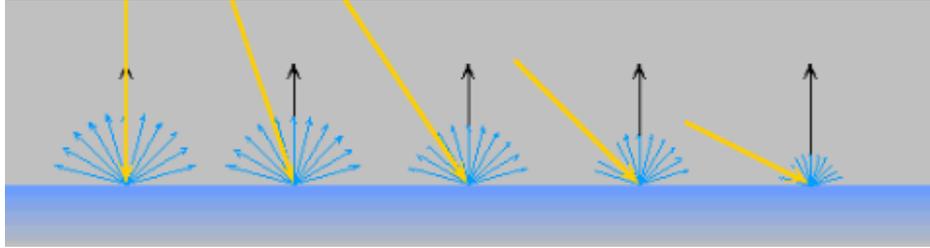


Figure 2.2: Lambert illumination model.

Figure 2.2 shows the ideal diffuse reflectance model. The perceived brightness of a scene point is proportional to the cosine of the angle between incidence beam and surface normal.

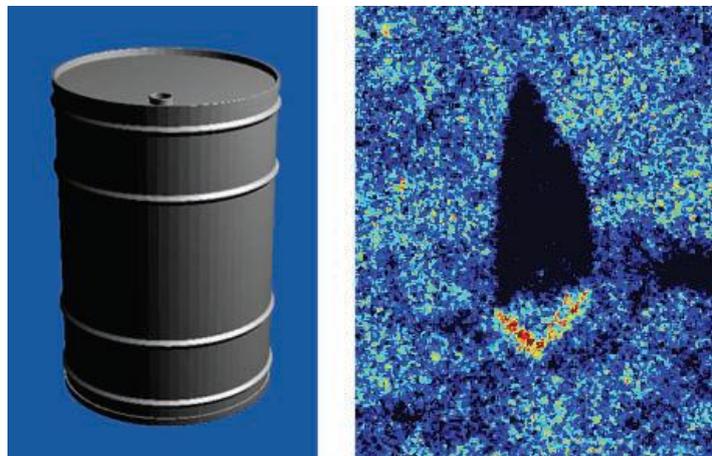


Figure 2.3: 3D model of a barrel and its simplified sonar image [5].

The assumptions made on the frequency based model simplifies the calculations of image formation process. After this point any standard computer graphics software can be used to obtain the sonar image. Figure 2.3 shows the 3D model of the barrel

sitting on the seafloor and its sonar image produced by the simplified frequency domain model.

2.6 Image Formation

Sonar image of given 3D model of a scene is produced by optical ray tracing technique in section 2.3.3 based on Eq. (2.10). By means of ray tracing, required distance, relative angle and intensity information of the targets is obtained to form the image. The steps of the image formation process are given detailed in following sections.

2.6.1 Scanning the Scene

Given a 3D model of the scene and location and orientation of the sonar, the scene can be scanned by optical rays if the azimuth (θ_B) and elevation (ϕ_B) beam of the sonar is specified. For the Blueview P900 [2] sonar, $\theta_B = 1^\circ$ and $\phi_B = 20^\circ$ and there are 512 beams with a total coverage of $\theta_T = 90^\circ$ in azimuth direction. Therefore horizontal beam spacing is $\Delta\theta = 90/512 \approx 0.18^\circ$. On the other hand, referring to Figure 2.1, the sound wave is continuous vertically. It is impractical to create continuous beam so the wave is sampled at $\Delta\phi = 20/512 \approx 0.04^\circ$ intervals. Using the specifications of the Blueview P900 sonar, assuming a point acoustic source and spherical spreading of the beams, the scene is scanned at horizontally $\Delta\theta$ and vertically $\Delta\phi$ degrees intervals with a total coverage of 90° and 20° , respectively. Figure 2.4 shows the directions of these rays transmitted from sonar.

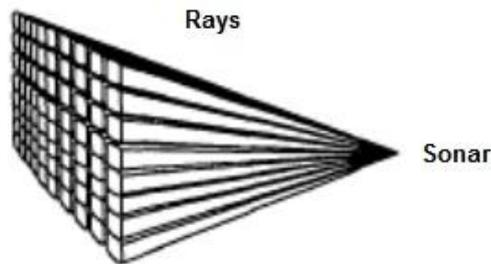


Figure 2.4: Rays transmitted from point source in specified directions.

After the rays are emitted from the acoustic source, their trajectories traced until they intersect an object in 3D space found at point \mathbf{P} . For each ray, 3 values are calculated and stored. The first one is the Euclidean distance between source and \mathbf{P} , second one is intensity of the returned signal and finally the third one is the surface number which will be used in uniforming the non-uniformly distributed sonar signal. These informations for each ray is stored in the 512x750 matrices (for each data type), where columns of the matrices corresponds to one sound beam in Figure 2.1 and rows corresponds to the angular sampling of the continuous beam.

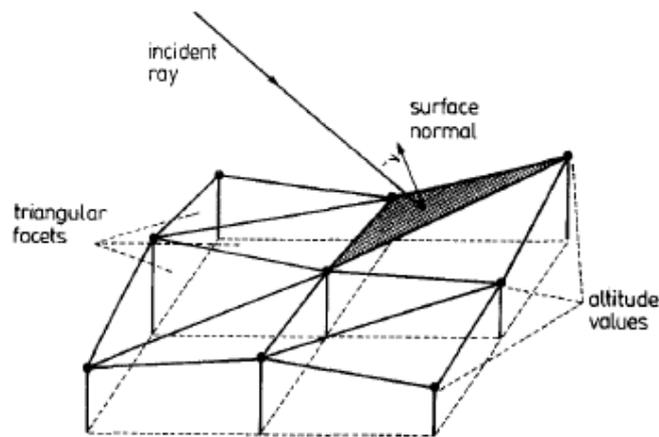


Figure 2.5: Intersection of the emitted ray with a scene point [4].

Figure 2.5 shows one of these rays transmitted from point source. The intensity value returned from point \mathbf{P} is calculated with (2.10) which corresponds to the Lambert illumination model. As a side product of this process, depth map of the scene is obtained.

2.6.2 Processing of Data

Modelling the continuous sound wave as angularly sampled discrete rays causes some problems in displaying the received data. Hence the stored data in matrices must be processed before displaying as an image.

Sonar forms an image using the time-of-flight of the sound pulses. After sampling and quantizing the received signal, it can be plotted directly as a function of time. Each column of the image corresponds to one continuous sound wave and rows of

the image corresponds to the range of the objects. However, the intensity values of the rows in sonar image must show the uniformly increment time instants because the pixels of a digital image is considered as a equal sized and equal distant pixels. This requires that the values stored in rows of the data matrix represent the constant range increments.

The first way to obtain constant time increments is adjusting the to pulse emitting angles so that they correspond to the spatially sampled rays, otherwise the resulting intensity values leads to unacceptable results [4]. On the other hand, even the exact settling of the scene is known, spatial sampling of the pulses is difficult and requires careful calculations.

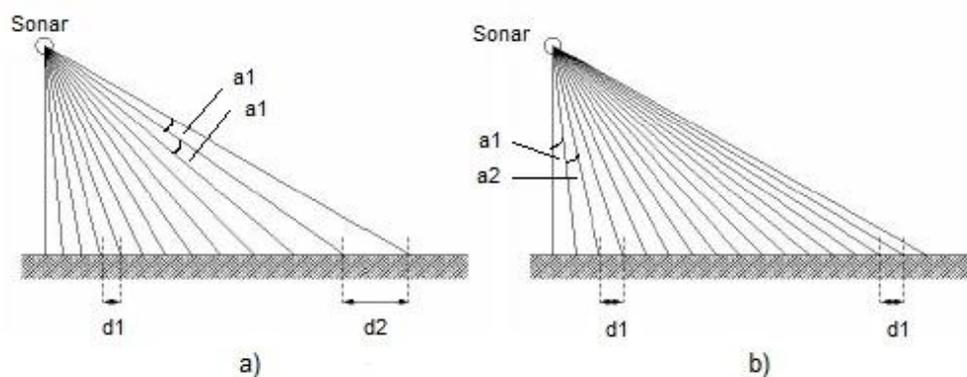


Figure 2.6: Ray sampling techniques; a) angular sampling, b) spatial sampling.

The second way to produce equally spaced range samples is angular sampling of the rays then process the intensity information, which is a function time, without violating the sampling theory. Figure 2.6 shows the ray sampling techniques where the angular sampling technique is used in our simulations.

The first step in angular sampling method is to sort the data in ascending range order. Although the rays emitted to provide ascending time returns, the received signal is not necessarily ensure this condition since the scene probably does not ensure the isovelocity conditions. Figure 2.7 a) shows the received raw data for one sound pulse, and Figure 2.7 b) shows sorted data in ascending range.

To complete the image formation process, multiple returns, acoustic shadows, decimation and transformation stages should be considered.

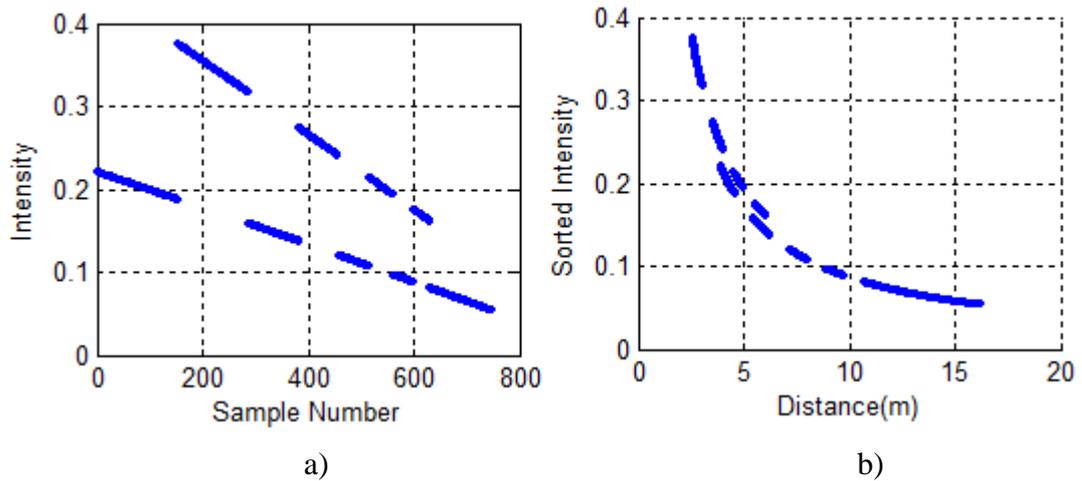


Figure 2.7: a) Received raw data, b) data after sorting.

2.6.2.1 Multiple Returns and Bright Areas

Considering (2.10), the brightness of point \mathbf{P} in sonar image is proportional to cosine of the angle between normal of the surface at \mathbf{P} and intersecting ray. As a result, if the direction of the sound ray inclined in a surface normal, these points will produce brighter areas in sonar image. Furthermore, since sound is a continuous wave, for most of the geometrical shapes, returns that reach the receiver at the same time is inevitable. This case is basically shown in Figure 2.8. Ray 1 and ray 2 are the angular samples of the continuous beam and have the equal distance to the points $\mathbf{P1}$ and $\mathbf{P2}$ found on the seabed respectively. Hence, reflections from these points will reach the receiver at the same time. In real sonar image, the pixel belong to these points will be brighter relative to other pixels.

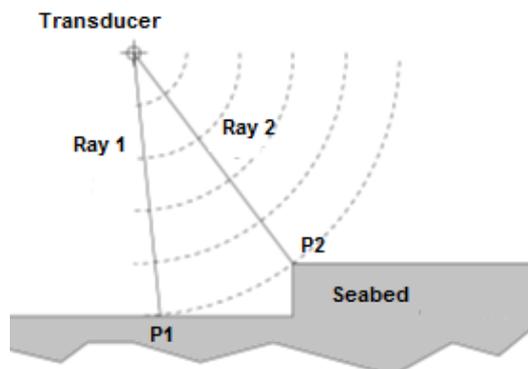


Figure 2.8: Multiple returns to sonar receiver [4].

To evaluate the multiple returns, intensity values received at the same instants must be added. This process is not straightforward because the sound waves are not sampled spatially which creates nonuniform increments of received data in time. The proposed method for this operation is based on the examining the scatter profiles of returning signals for different geometrical figures such as planes, spheres and ellipsoids. Intersection of transmitted rays with these geometric figures is shown in Figure 2.9 in cross sections. In Figure 2.9 a) P1 and P3, in Figure 2.9 b) N and F1 (also F2) and in Figure 2.9 c) M and M1 are, the points where maximum and minimum signal returns are received, respectively. Corresponding signal returns are shown in Figure 2.10 where the characteristic of different types of surfaces can be

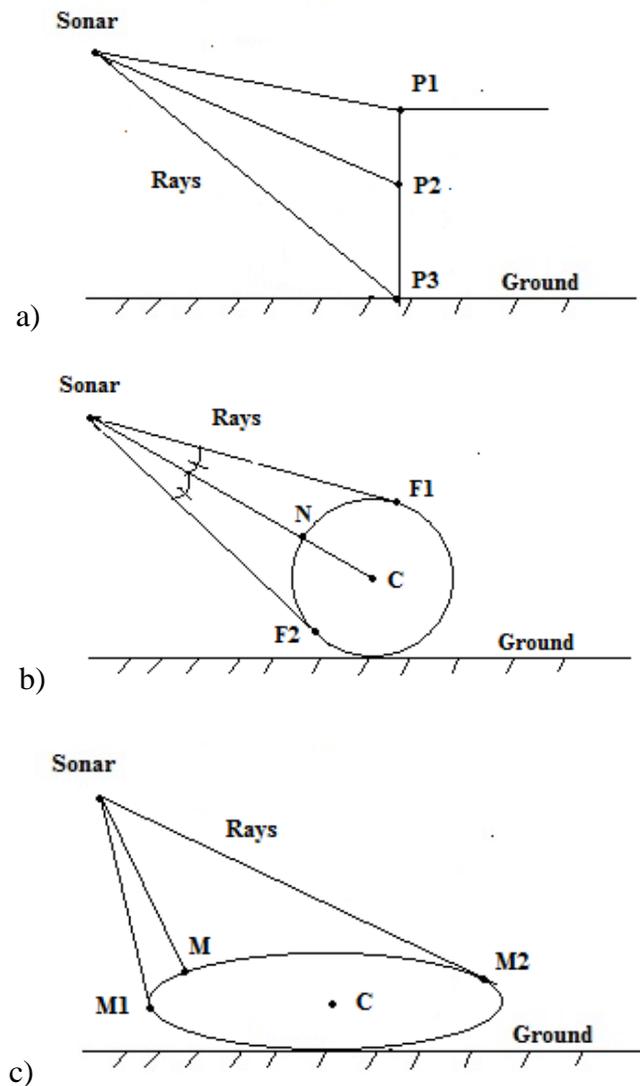


Figure 2.9: Rays and their intersections with a) cuboid, b) sphere and c) ellipsoid.

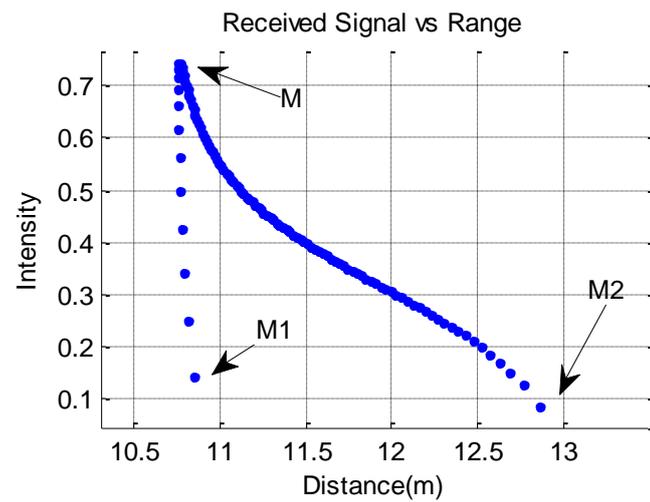
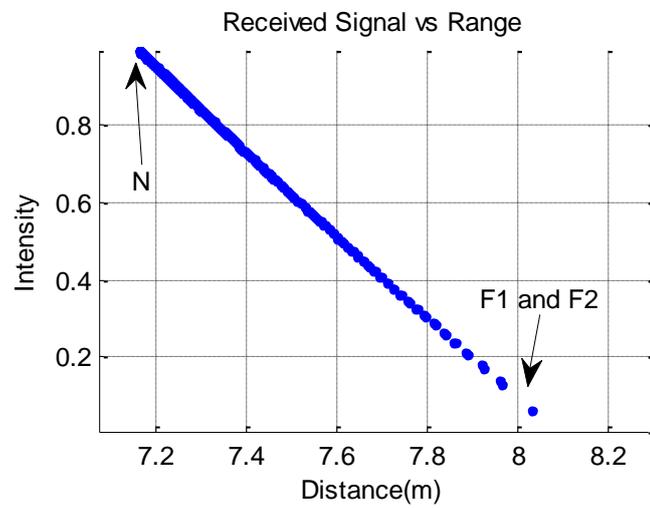
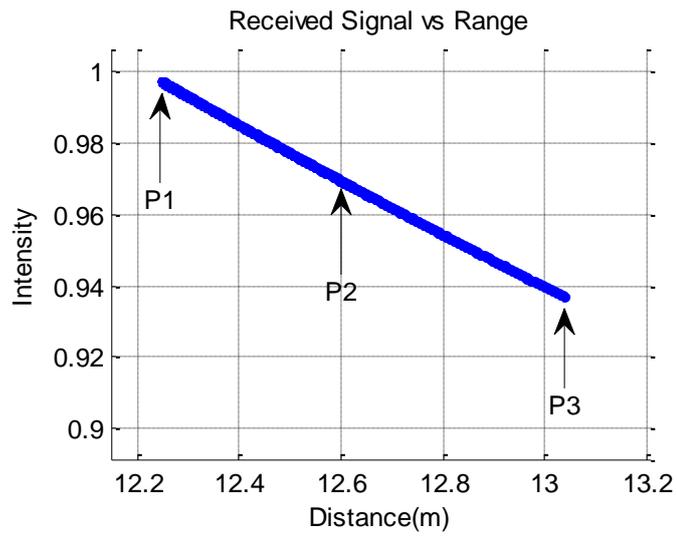


Figure 2.10: Scatter profile of a) plane, b) sphere and c) ellipsoid.

seen. Among these surfaces, the profile of the sphere might be tricky. Referring to Figure 2.9 b), N is the closest point of circle to the sonar transmitter which is, at the same time, the point where the maximum intensity value is obtained. This intensity value decreases as the sound waves intersect the furthest accessible points F1 and F2. Since the cross section of the sphere is symmetrical around N relative to sonar, discriminating the returning signals from the surface patches N-F1 and N-F2 is impossible. So intensity values in Figure 2.10 b) must be multiplied by 2 before summation.

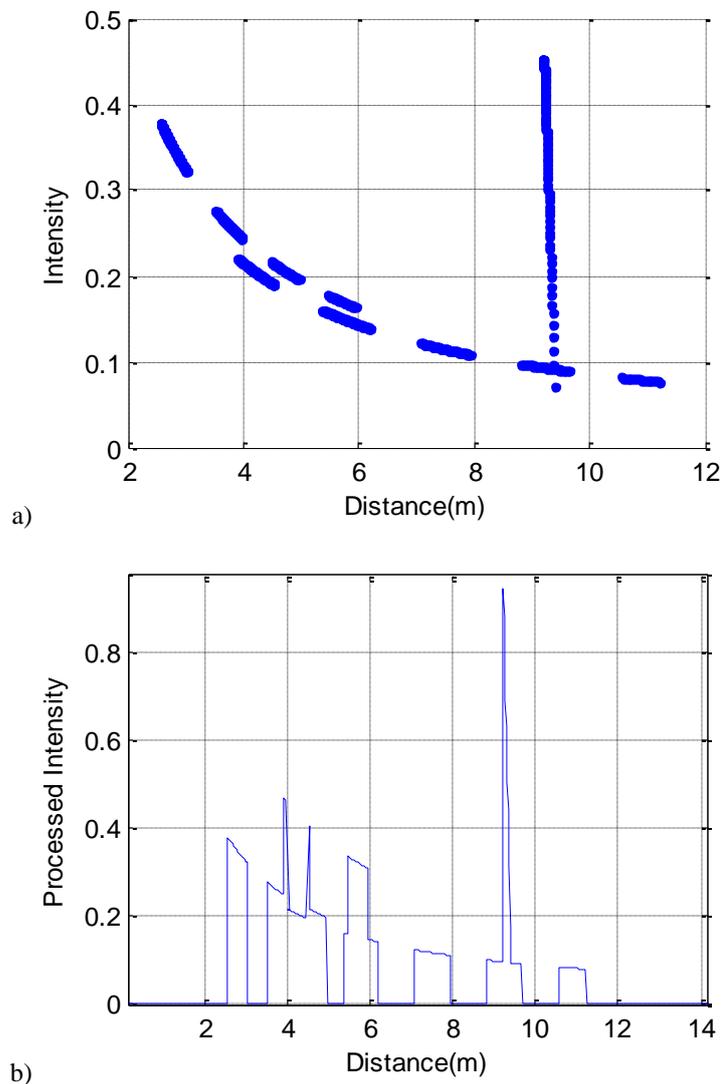


Figure 2.11: a) Received signal, b) resulting sonar signal after summation.

To perform the summation, surface numbers stored in matrix should be used. Considering the scatter profiles of different surface types, the nonuniform data samples in time should be transformed to uniform samples by interpolation for every surface in the matrix so that the summation is achieved without distorting the data. Figure 2.11 a) shows the received intensity signal from a scene which consists of planes and ellipsoids and Figure 2.11 b) shows the result of summation.

2.6.2.2 Acoustic Shadows

In an optical image shadow zones are areas where the light can not illuminate or partly illuminate. Similarly, while sonar scans the scene, some part of the objects and ground can not be ensonified by acoustic pulses. This will result in acoustic shadow zones in sonar image where any return can not be received. This case is illustrated in Figure 2.12. It is impossible to illuminate the lower step of the floor with rays.

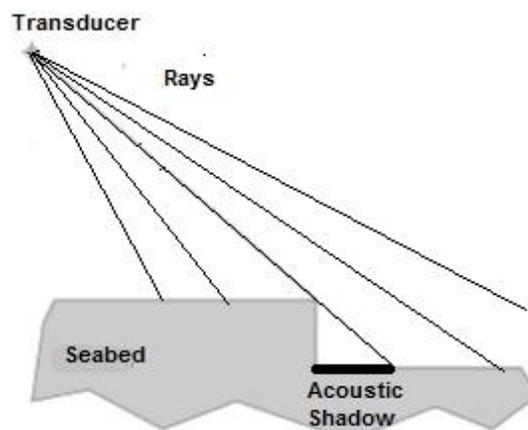


Figure 2.12: Region that can not be illuminated by sound pulses.

Shadows are one of the most important features in sonar image. In 2D optical images getting the depth information which enhances understanding the 3D scene structure is important. Sonar also produces 2D image but shadow zones in the sonar image provide 3D information about the scene structure. Sometimes this additional information is crucial for the accurate localization of the objects where the optical cameras can not provide enough information due to the turbidity of water and insufficient lighting.

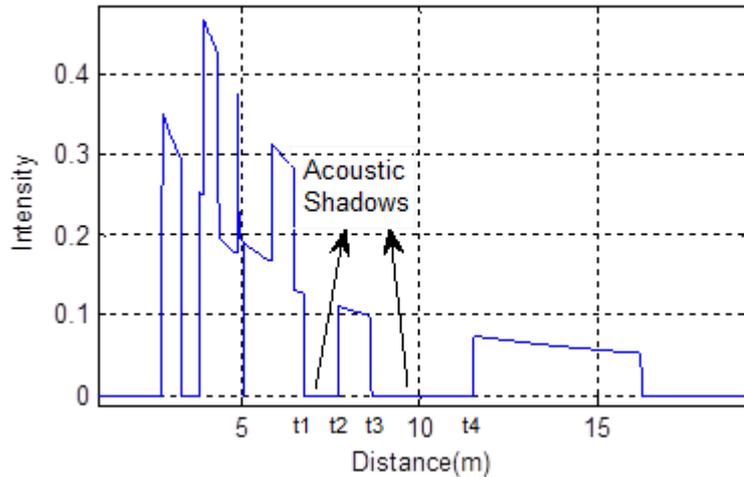


Figure 2.13: Acoustic shadows in processed intensity signal.

A shadow region might be created an object sitting on the floor or preventing the illumination of an object with sound waves by another object. This situation can be seen Figure 2.13 where the intervals t_1 - t_2 and t_3 - t_4 are corresponds to shadow regions on seabed. On the other hand, transition to Figure 2.11 a) to Figure 2.13 is not straightforward. If the scattered data in Figure 2.13 is interpolated with a sinc function or linearly, the lacking intensity values are completed with unrealistic ones. To overcome this problem, a threshold (s_t) should be specified for the shadow zones which can be determined using the detection range of sonar. If the time difference between any consecutive samples in scattered data is greater than s_t , then 0 should be added between these samples at sampling instants (t_s) which is also used in interpolation of scattered data. Furthermore, this step is required to guarantee constant time increments of the sonar data.

2.6.2.3 Decimation

The signal processing operations in previous sections generate more samples than number of pixels in rows of the displayed image. Hence, it is necessary to reduce number of samples without distorting it. This operation is achieved by decimation which is a two-step process. First the signal is filtered with anti-aliasing filter then downsampled to specified sample size.

2.6.2.4 Transformation

Forward looking sonar images differ from optical images geometrically. Since the sensor from which the light projected onto is rectangular, so are the optical images. On the other hand, the projection model of the sonar is different where a pixel on image shows the relative angle w.r.t the looking direction and range of corresponding area. Taking into account the ray directions emitted from sonar, it is not surprising that forward looking sonar produces non-rectangular images. Hence, after the decimation step, the image which is directly produced from decimated data, must be transformed to polar form using the detection interval and θ_T . Otherwise, geometrical distortions in the image are inevitable. Rectangular to polar conversion of the Matlab's peaks function is shown in Figure 2.14 .

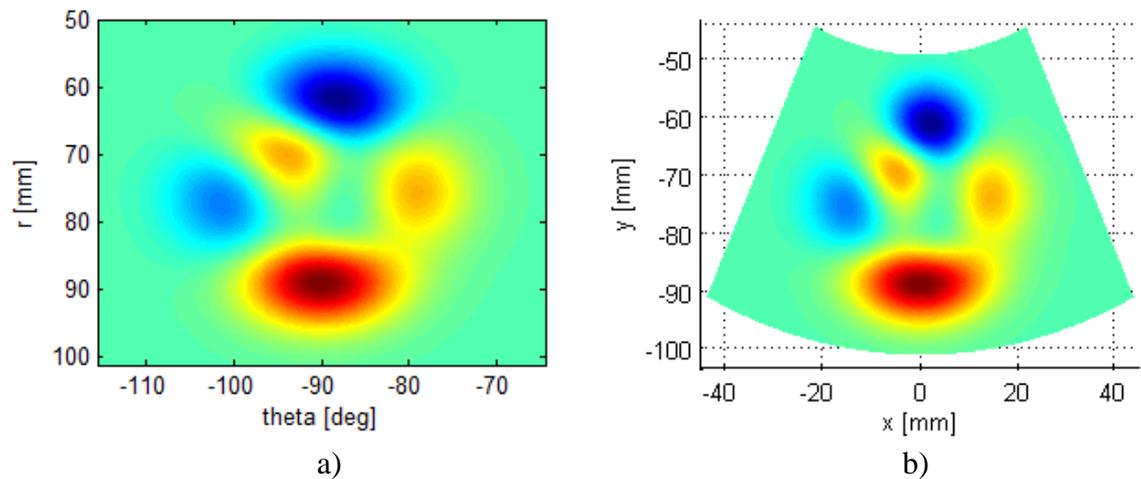


Figure 2.14: a) A rectangular image and b) its polar form.

2.7 Performance of the Sonar

Two factors effect the operation capability and detection performance of the sonar systems. These are noise and reverberation. In general, detection performance of the sonar is limited either noise or reverberation. As to high frequency 2D imaging sonar, reverberation is the main factor that limits the operation performance [22]. Explanation and effects of these factors are given in following sections.

2.7.1 Noise

Sonar must detect signals received from targets against background called noise. Ambient noise originates from the environment where the sonar is used. The sources of this noise are generally beyond our control. Shipping, waves, thermal agitation, rainfall and animals are the potential sources of ambient noise. The frequency of the shipping noise is from 10 Hz to mid-hundreds of Hz while the frequency of the wave noise is from mid-hundreds of Hz to 50 KHz. However, thermal noise, which is the thermal attack of electrons to receiver, is dominating noise above 50 KHz [23].

2.7.2 Reverberation

When the sonar emits a pulse, it receives echoes from sources apart from the interested targets. Returns coming from non-target objects are called reverberation. Surface, bottom and volume of the water are main reverberation sources. However, if the sonar is not close to surface, bottom reverberation dominates provided that the sonar operates at low volume reverberation levels.

Table 2.1: Distribution functions.

Function	PDF
Rayleigh	$(2x/\lambda)e^{-x^2/\lambda}$
Weibull	$\alpha\beta x^{\beta-1}e^{-\alpha x^\beta}$
K	$\sqrt{\alpha}\Gamma(v)(x/\sqrt{\alpha})^v K_{v-1}(2x/\sqrt{\alpha})$
Rayleigh mixture	$\sum_{i=1}^m \epsilon_i (2x/\lambda_i) e^{-x^2/\lambda_i}$

Bottom reverberation is the sum of many elementary echoes scattered by points in the seabed. It is a complex quantity and converges to Gaussian distribution in the limit. Assuming independent and uncorrelated real and imaginary parts, the amplitude and phase of the reverberation are Rayleigh and uniformly distributed respectively [16]. However, K, Weibull and mixture Rayleigh models generally fits the bottom reverberation statistics better than Rayleigh distribution [6]. Probability density functions (pdf) of these models are given in Table 2.1.

Table 2.2: Estimated parameters for 40°–60° grazing angle [6].

Bottom Type	Weibull (α, β)	Mixture ($\epsilon_i; \lambda_i$)	K (α, ν)
sand/shell	1.009, 1.955	0.369, 0.262, 0.369; 1.101, 0.717, 1.101	0.033, 30.20
coarse sand	1.031, 1.862	0.353, 0.375, 0.272; 1.172, 1.172, 0.542	0.108, 9.224
posidonia	1.092, 1.656	0.228, 0.588, 0.185; 2.049, 0.823, 0.269	0.377, 2.654
mud	1.018, 1.918	0.316, 0.338, 0.346; 0.754, 1.112, 1.115	0.042, 24.03
medium sand	1.087, 1.658	0.546, 0.399, 0.055; 0.652, 1.606, 0.040	0.337, 2.970
shells	1.060, 1.754	0.286, 0.372, 0.342; 0.348, 1.300, 1.218	0.242, 4.128

Estimated parameters of the distributions in Table 2.1 is shown in Table 2.2 for different surface types and 40°–60° grazing angle. As shown in Table 2.3, K and mixture Rayleigh distributions show the best fit for most surface types. On the other hand, mixture Rayleigh distribution is used in our simulations because of the simplicity of the model compared to K distribution.

Table 2.3: Results for selected sites and for 40°–60° grazing angle [6].

Bottom Type	Rayleigh	Weibull	Mixture	K
sand/shell	0.652	0.929	0.961	0.962
coarse sand	0.864	0.905	0.864	0.797
posidonia	1.28310×10^{-7}	0.367	0.989	0.996
mud	0.487	0.691	0.784	0.874
medium sand	4.25310×10^{-4}	0.716	0.881	0.615
shells	0.002	0.508	0.963	0.866

2.8 Results

In this section, sonar images which is simulated by the method mentioned in previous sections will be presented step by step. Simulated scenes consist of geometrical shapes such as cuboids, spheres, ellipsoids and cylinders as shown in Figure 2.15.

In this simulation, the location of the sonar is $L_s = [0, -6, 2]^T$ and it is directed to the origin of the coordinate system.

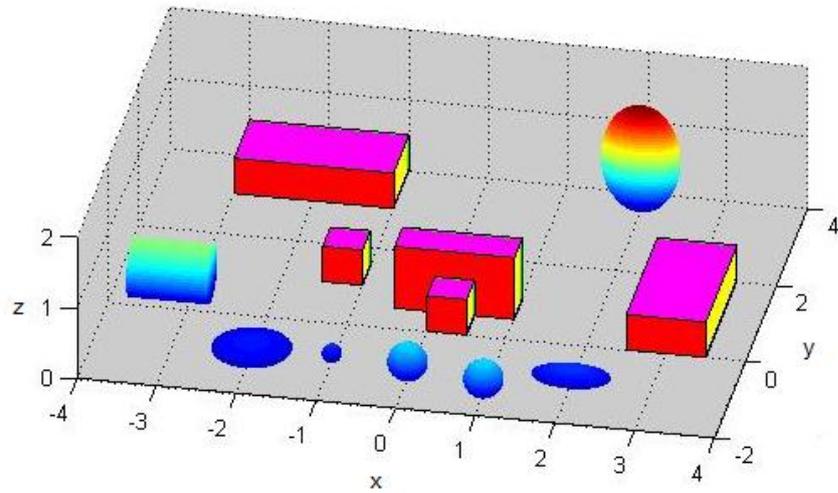


Figure 2.15: Scene 1 which consists various geometrical figures.

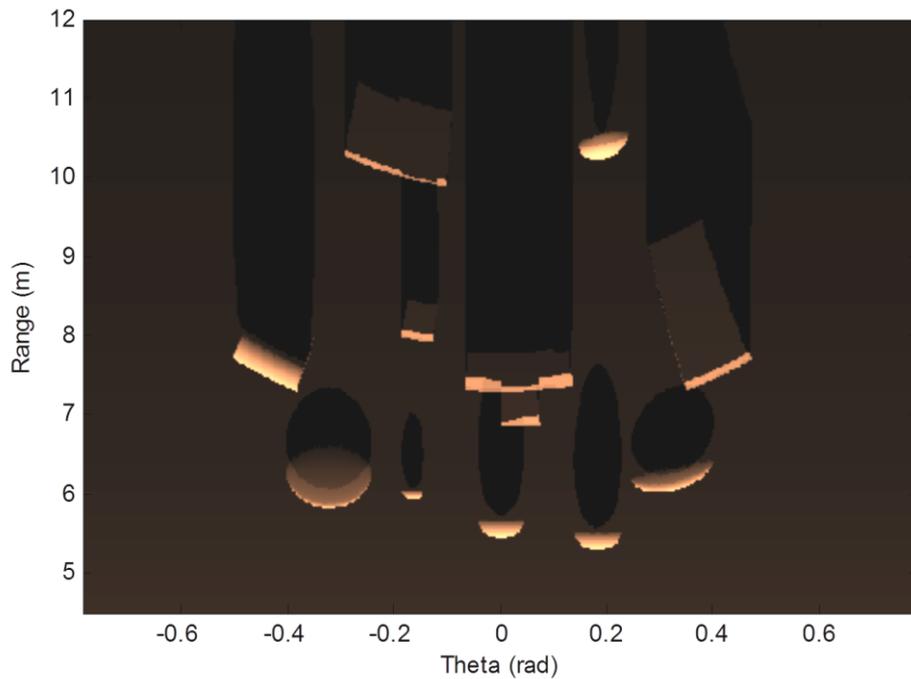


Figure 2.16: Image of scene 1 after decimation step.

Image directly formed after the decimation step is shown in Figure 2.16. To prevent geometrical distortions it should be transformed to the polar form using the

specifications of the sonar. The result of the transformation is shown in Figure 2.17 a) and finally the effect of noise and reverberation is shown in Figure 2.17 b).

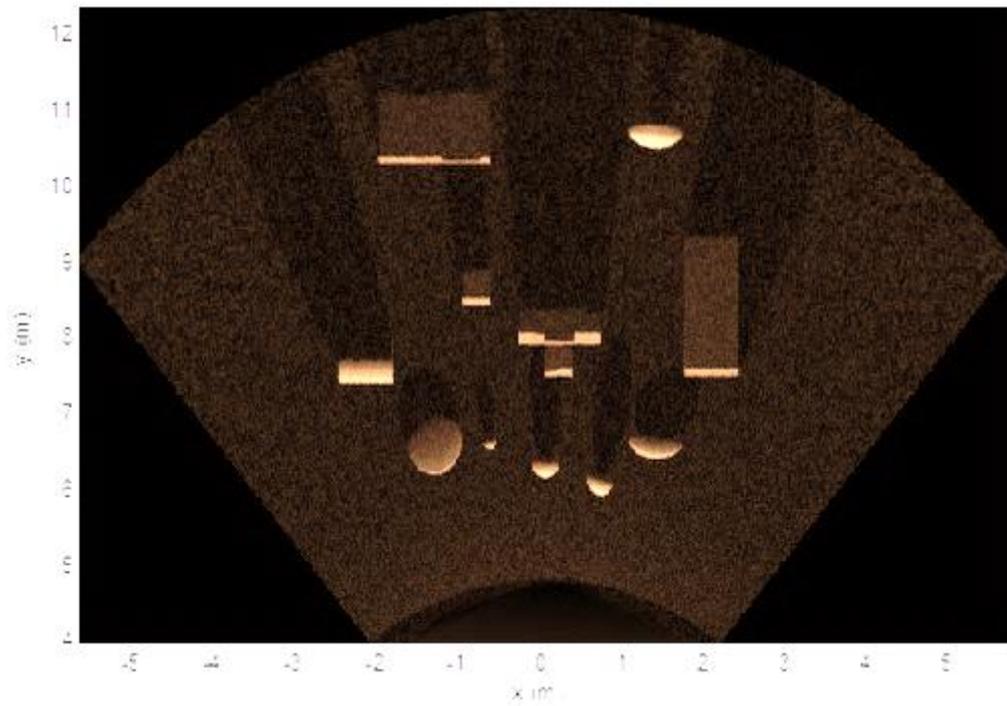
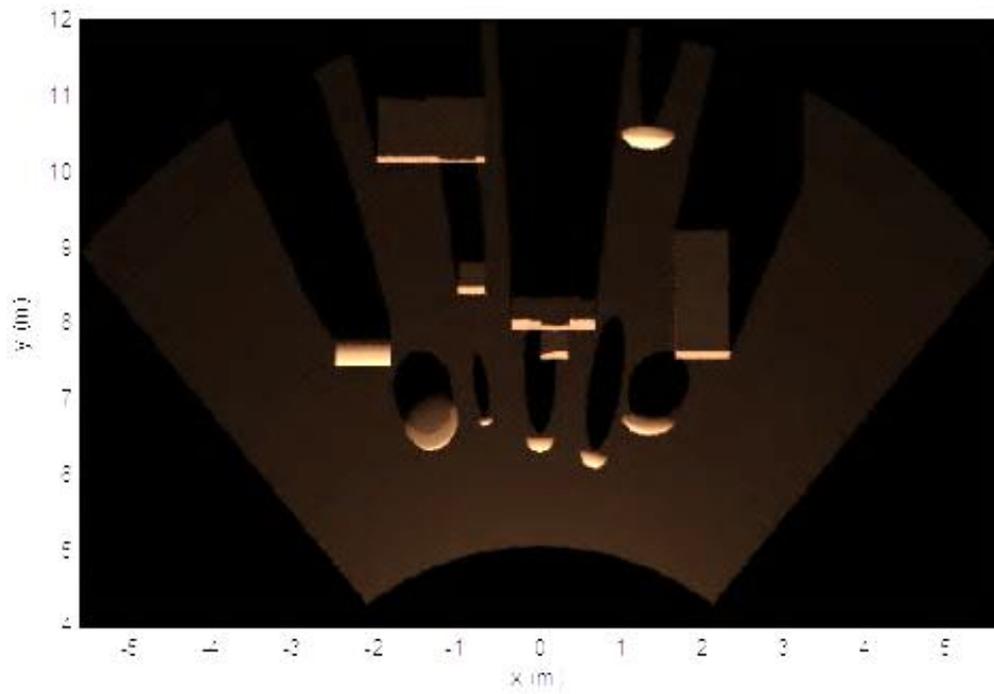


Figure 2.17: a) Polar image, b) effect of noise and reverberation.

Acoustic shadows can be clearly seen in Figure 2.17 a). However, noise and reverberation somewhat mask shadow areas as in Figure 2.17 b). Comparison of a real sonar image with simulated image is also made and shown in Figure 2.18.

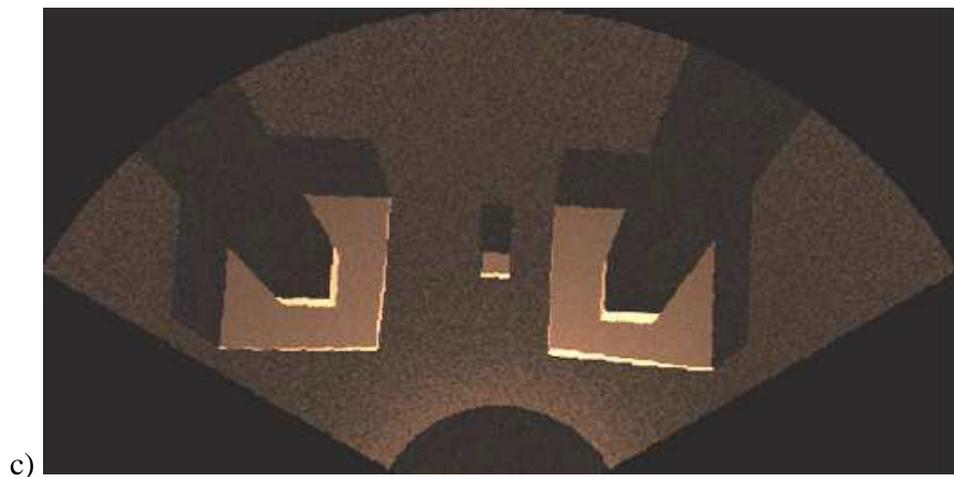
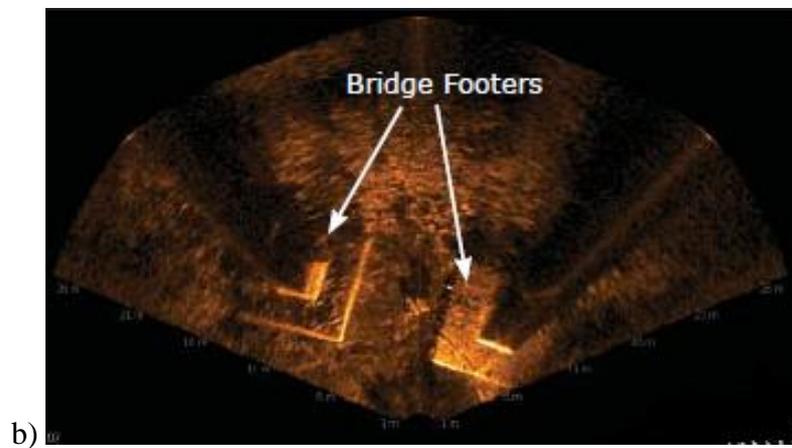
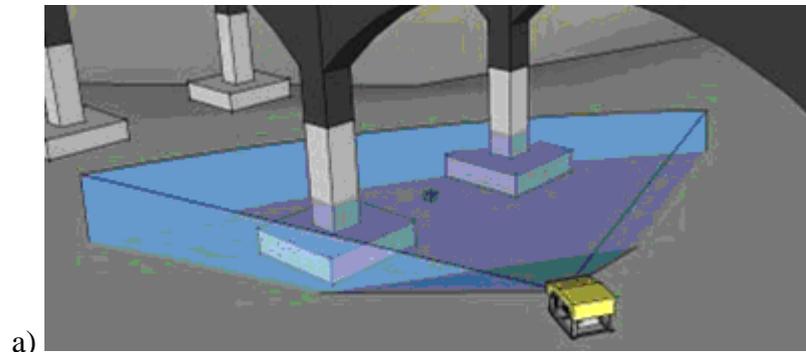


Figure 2.18: a) ROV approaching the bridge footers, b) sonar image of footers taken with Blueview P900 [2], c) simulated image.

The calculations are done on a personal computer (Core 2 Duo, 2.53GHz). All algorithms are implemented using MATLAB running on Windows 7. The average simulation time for a scene as in Figure 2.15 is 3.5 minutes which is a relatively short time for the simulation of high frequency sonar. This shows the computational efficiency of the method when compared with the method given in [4]. However, this computation time can be reduced much more by using a faster compiling environment such as C and a faster processor. The code can also be optimized for efficient computation.

CHAPTER 3

UNDERWATER OPTICAL IMAGE SIMULATOR

In this chapter, simulation of underwater optical images is studied. Steps and requirements of image formation process is explained in detailed.

3.1 Underwater Vision

Underwater vision is limited due to the challenging environmental conditions. Visibility decreases quickly as a function distance as shown in Figure 3.1. The objects are scattered in the scene so the effect of weakening conditions are spatially varying [24] like in open-air vision in fog. However, simulating underwater conditions is more complicated than open-air counterpart because of the interaction of light with water. Image components (color model) and effects of this interaction to image formation process are described in the following sections.



Figure 3.1: An underwater scene about 5 meters deep.

3.2 Image Components

Light rays which consist RGB color channels are absorbed and scattered while travelling in water. Absorption of light in water is proportional to its wavelength. Red light has the wavelength bigger than green and blue light hence more absorbed than others. This is the reason why underwater images are bluish [8]. Wavelengths of lights in the visible spectrum are given in Table 3.1.

Table 3.1: The colors of the visible light spectrum.

Color	Wavelength Interval
Red	~ 700–635 nm
Orange	~ 635–590 nm
Yellow	~ 590–560 nm
Green	~ 560–490 nm
Blue	~ 490–450 nm
Violet	~ 450–400 nm

The basic model which explains the absorption and scattering events is the Jaffe-McGlamery [9, 10] model.

3.2.1 Absorption

While travelling in water, photon, which is the discrete bundle of light energy, is absorbed by molecules and particles found in the water. Degree of this absorption depends on the wavelength of the light and amount of the particles. The loss of photons is parameterized by distance travelled by photon and the effect of different factors is combined in a single absorption coefficient a [m^{-1}] which is a function of wavelength λ [7],

$$E(z, \lambda) = E(0, \lambda)e^{-za(\lambda)} \quad [\text{Wm}^{-2}] \quad (3.1)$$

where $E(0, \lambda)$ is the initial irradiance and z is the distance travelled by light in water. (3.1) shows the exponential decline of irradiance with distance.

3.2.2 Scattering

Photons are scattered when colliding with small particles in water. Scattering of photons occur in certain angle $\psi \in [0, \pi]$ by volume spread function (VSF) $\beta(\psi, \lambda)$. To obtain the scattering coefficient, VSF is integrated in all angles [25],

$$b(\lambda) = 2\pi \int_0^\pi \beta(\psi, \lambda) \sin(\psi) d\psi \quad [m^{-1}] \quad (3.2)$$

where $b(\lambda)$ expresses the total scattering coefficient of water. The reduction in irradiance due to scattering is

$$E(z, \lambda) = E(0, \lambda) e^{-zb(\lambda)} \quad [Wm^{-2}] \quad (3.3)$$

and the overall attenuation coefficient is [8],

$$c(\lambda) = a(\lambda) + b(\lambda) \quad [m^{-1}]. \quad (3.4)$$

The values of $a(\lambda)$ and $b(\lambda)$ is given in Table 3.2 for some wavelengths.

Table 3.2: Volume absorption and scattering coefficients for pure water [26].

λ	a	b
700nm	0.6229	0.0005
670nm	0.4346	0.0006
620nm	0.3047	0.0009
580nm	0.1016	0.0011
530nm	0.0505	0.0017
470nm	0.0130	0.0027
420nm	0.0054	0.0043

Scattering not only causes reduction in irradiance. Due to multiple scattering, photons are scattered towards light beam leads to increase in irradiance and can be modeled as the sum of forward and backward scattering events which occur in the intervals $\psi \in [0, \pi/2]$ and $\psi \in [\pi/2, \pi]$ respectively.

Forward scattering describes light scattered forward at small angles, causes blur in image and can be modeled with distance dependent point spread function [27].

To model backward scattering, volume spread function is used. The Einstein-Smoluchowski parameterization for the VSF is

$$\beta_w(\psi, \lambda) = \beta_w(90, \lambda_0) \left(\frac{\lambda_0}{\lambda}\right)^{4.32} \quad (3.5)$$

where λ_0 and $\beta_w(90, \lambda_0)$ is found in [28]. (3.5) does not take into account particles hence Kopelevich's formulation for particle scattering should be used which is given as

$$\beta(\psi, \lambda) = \beta_w(\psi, \lambda) + v_s \beta_s(\psi) \left(\frac{\lambda_0}{\lambda}\right)^{1.7} + v_l \beta_l(\psi) \left(\frac{\lambda_0}{\lambda}\right)^{0.3} \quad (3.6)$$

where v_s and v_l are the amounts and β_s and β_l are the VSF's of small and large particles in water respectively. β_s , β_l , β_w and β versus scattering angle is shown in Figure 3.2.

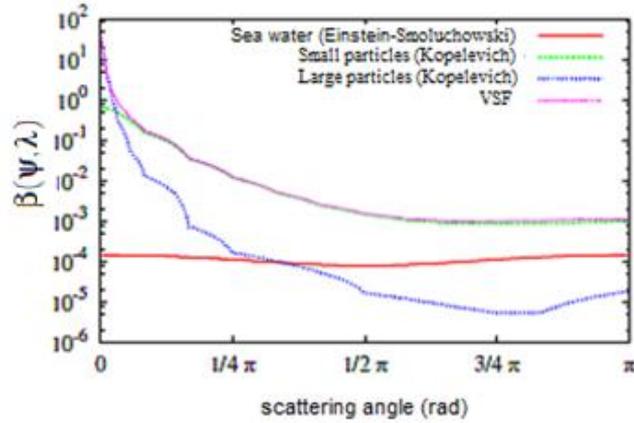


Figure 3.2: Volume scattering function [8].

3.2.3 Jaffe-McGlamery Model

In this model, irradiance sensed by the optical sensor is the sum of direct light, forward and backward scatter and written as

$$E_T = E_D + E_{fs} + E_{bs} \quad (3.7)$$

where T, D, fs and bs stand for total, direct, forward and backward scatter respectively. Following subsections describes these components in detailed.

3.2.3.1 Signal

Direct light is the main component of the sensed irradiance which is attenuated while travelling towards scene structures. The degree of reduction is related with attenuation coefficient c which is described in sections 3.2.1 and 3.2.2. In calculations, point light sources are used and these sources are described by their 3D positions and powers. In the following context $L(\lambda)$ symbolizes the point source where λ emphasizes the dependence to the wavelength.

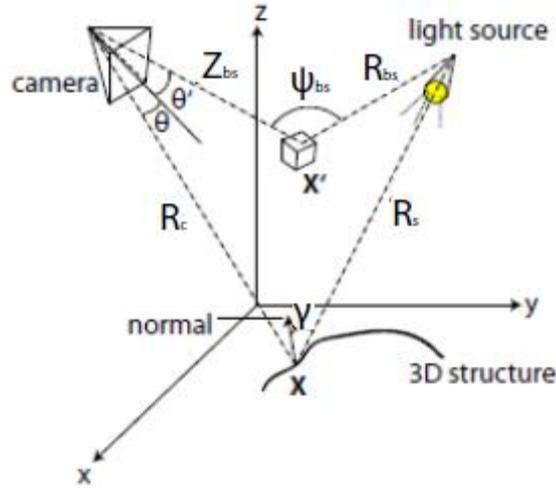


Figure 3.3: Rays for the signal and backscatter [9].

For geometric computations, rays originating from light source is intersected with scene point \mathbf{x} as in Figure 3.3. The irradiance sensed at this point by direct transmission is [8]:

$$E_D(\mathbf{x}, \lambda) = L(\lambda) \cos(\gamma) \frac{e^{-c(\lambda)R_s}}{R_s^2} \quad (3.8)$$

where R_s is the distance between this point and light source and γ is the angle between normal at point \mathbf{x} and intersecting light ray.

3.2.3.2 Forward Scatter

Light scattered at small angles by particles along line of sight (LOS) is called forward scatter [24] as shown in Figure 3.4. Forward scatter increases the signal irradiance but causes image blur given by the convolution:

$$E_{fs}(\mathbf{x}, \lambda) = E_D(\mathbf{x}, \lambda) * g(\mathbf{x}|R_c, G, c(\lambda), B) \quad (3.9)$$

where g is the filter that changes with distance R_c and empirical constants G and B . g is a distance dependent filter because forward scatter cumulates as the distance travelled by the light increases which results in spatially increasing blur. This filter can be approximated as a Gaussian kernel based low pass filter which is given in [27].

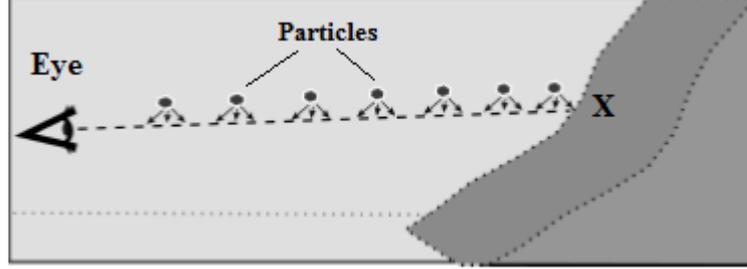


Figure 3.4: Light scattered towards the LOS by nearby particles [24].

3.2.3.3 Backscatter

During multiple scattering of photons while travelling in water, photons can also be scattered into other light beams causes backscatter which is also known as veiling light. Veiling light is the dominant component that degrades visibility.

For geometric computations 3D space is sliced into N parallel planes to image plane each of which have thickness Δz_i [10]. The backscatter component of light is the sum of veiling light obtained from those slices.

Veiling light accumulates with increasing distance and is computed similar to forward scatter and given as [8],

$$E_s(\mathbf{x}') = E_{s,d}(\mathbf{x}') + E_{s,fs}(\mathbf{x}') \quad (3.10)$$

where

$$E_{s,d}(\mathbf{x}') = L(\lambda) \frac{e^{-c(\lambda)R_{bs}}}{R_{bs}^2} \beta(\psi, \lambda) \quad (3.11)$$

and

$$E_{fs}(\mathbf{x}') = E_{s,d}(\mathbf{x}') * g(\mathbf{x}' | R_{bs}, G, c(\lambda), B) \quad (3.12)$$

where $\beta(\psi, \lambda)$ is the volume scattering function and R_{bs} is the distance between light source and scattering volume \mathbf{x}' . Using (3.10), the overall backscatter component is

$$E_{bs}(\mathbf{x}, \lambda) = \sum_{i=1}^N \frac{e^{-c(\lambda)Z_{bs_i}} E_s(\mathbf{x}', i) \Delta z_i}{\cos\theta(\mathbf{x}')} \quad (3.13)$$

where i denotes the index of backscatter slice and Z_{bs_i} is distance between the camera and center of the slice i .

3.3 Image Formation

Underwater image of the given 3D model of a scene is produced by a pinhole camera which is based on optical ray tracing technique similar to used in section 2.6.

3.3.1 Pinhole Camera Model

A camera that doesn't use lens to focus light and has a point aperture is known as pinhole camera. Pinhole camera model explains the projection of 3D scene point onto the image plane of an ideal pinhole camera.

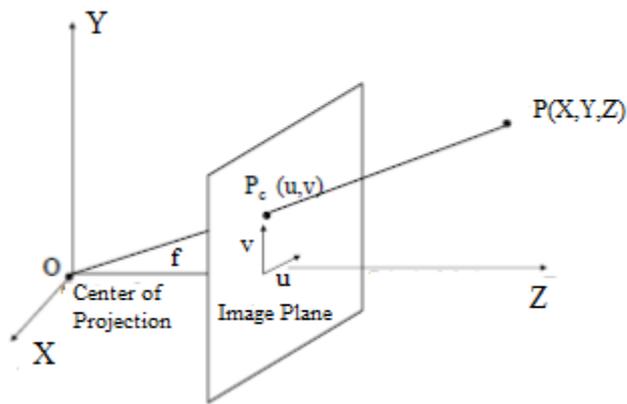


Figure 3.5: Pinhole camera.

In Figure 3.5, a point in 3D space and its projection onto image plane, which is parallel to XY plane, is shown. Using the projection model of pinhole camera, location $P_c(u,v)$ of 3D point $P(X,Y,Z)$ on image plane is calculated using the similarity of triangles. Assuming Z is the principle axis

$$u = f \frac{X}{Z} \quad \text{and} \quad v = f \frac{Y}{Z} \quad (3.14)$$

where f (focal length) is the coordinates of image plane in principle axis.

3.3.2 Simulation Steps

Given a 3D model of the scene and location and orientation of the camera, the scene can be scanned by optical rays if the specifications of the sensor (size and pixel) and focal length of the camera is known. In simulations, the sensor is taken as 36x24 mm (full frame) and 600x400 pixel. That is to say, 600x400 grid is formed from 36x24mm sensor with equal grid heights and widths. Then, optical rays are sent to the centers of those grids from the center of projection (camera location) as in Figure 3.6. Thanks to the pre-specified ray directions, this system acts as an ideal pinhole camera without the need to calculate the locations of scene points on the sensor as in (3.14).

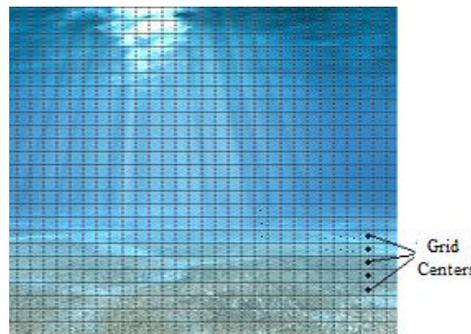


Figure 3.6: Points where the optical rays are directed.

After the rays are emitted to specified directions, their trajectories traced until they intersect an object in 3D space found at point \mathbf{P} . For each ray, 3 values are calculated and stored. The first one is the Euclidean distance between the camera and \mathbf{P} , second one is perceived brightness (intensity) of the scene point \mathbf{P} and finally the third one is

the color of the intersected point where distance and color information will be used to model attenuation, forward scatter and backscatter events later.

3.4 Results

In this section, underwater optical images which is simulated by the method mentioned in section 3.2 will be presented step by step. Simulated scenes consist of geometrical shapes such as cuboids, spheres, ellipsoids and cylinders as shown in Figure 2.15. In this simulation, the location of the camera is $L_c = [-0.25, -6, 2]^T$, focal length is 50mm and it is directed the origin of the coordinate system. In the first

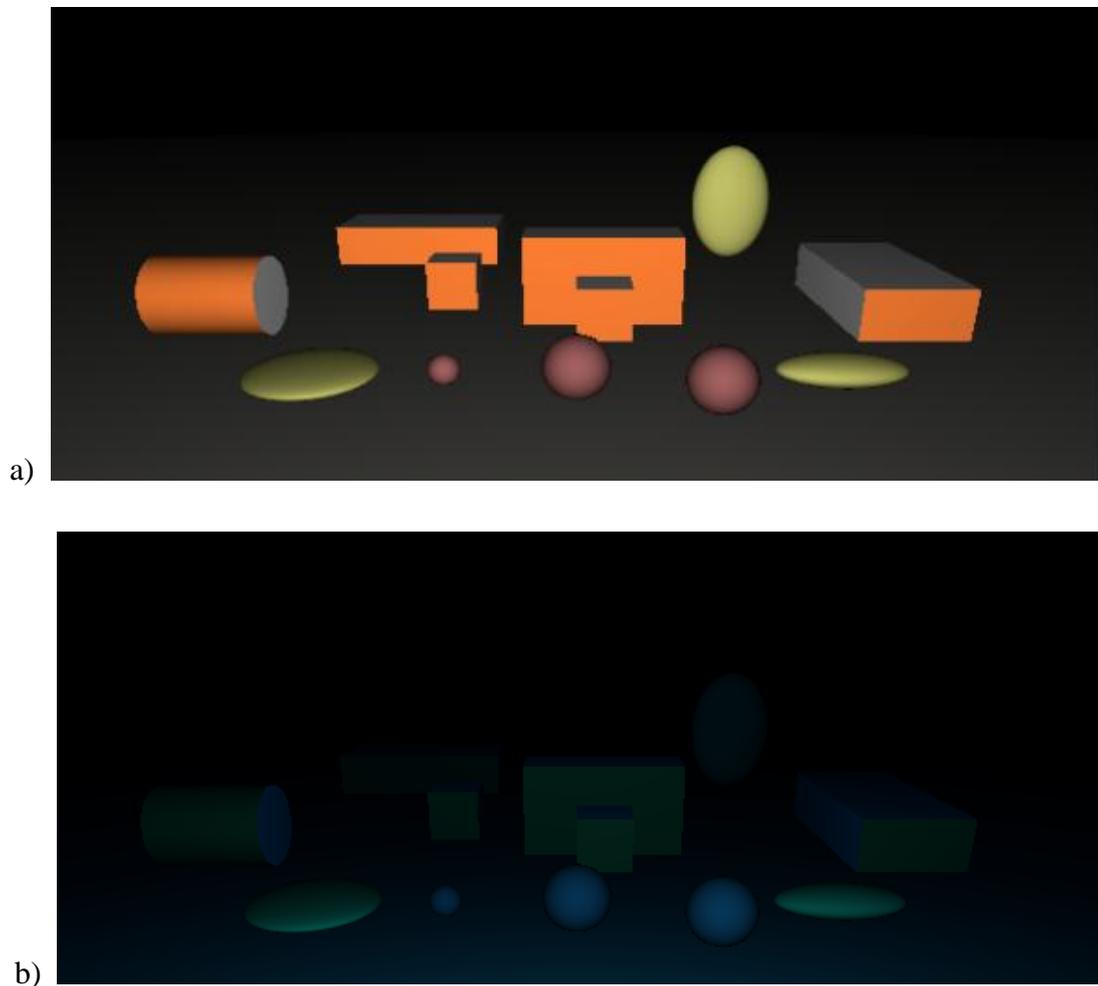


Figure 3.7: a) True illumination of the scene in Figure 2.15, b) the effect of attenuation.

step, various colors are assigned to the surfaces of objects in Figure 2.15 and the scene is scanned with light rays. Again, Lambert illumination model is used for intensity calculations. The true colors of the illuminated scene is shown in Figure 3.7 a). In the second step the effect of attenuation is considered. Since attenuation is wavelength dependent, the channels of the colors of scene points are processed separately. Total attenuation factor (c) for R, G, B channels are taken 0.41, 0.066 and 0.018, respectively. Using the stored distance and color information, attenuation is calculated as in (3.1) and (3.3). The effect of attenuation is shown in Figure 3.7 b). where R and G channel is attenuated more than B channel is clearly seen. In the third step, forward scattering of light is considered. Forward scattering is calculated with (3.9) where the parameters of the Gaussian filter are adjusted using the distance

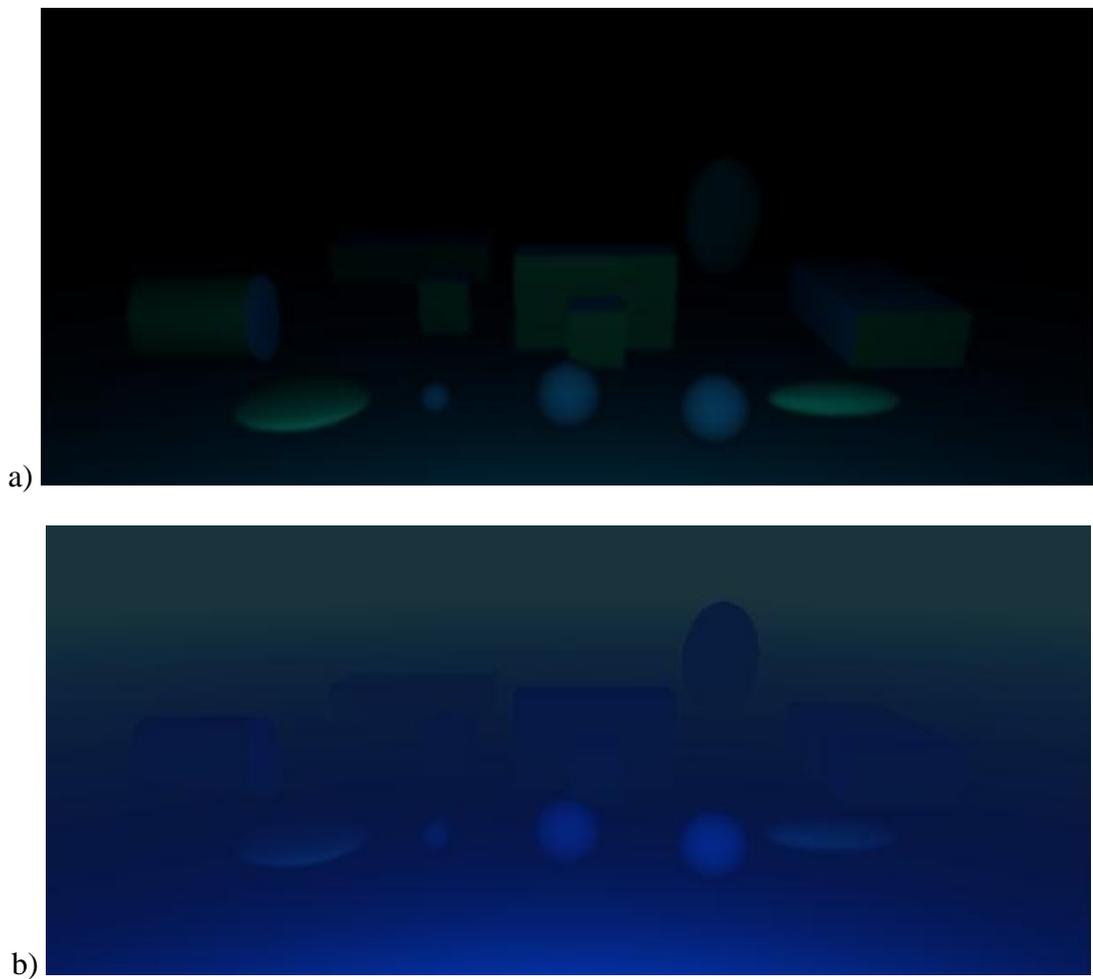


Figure 3.8: Effect of forward scatter in a) and backscatter in b).

information. In Figure 3.8 a) the effect of forward scatter in small angles is shown which mainly adds blur to image. In the final step, backscatter is calculated using (3.13). Backscatter (veiling blue light) is the dominant component which limits the visibility in underwater and its effect is clearly seen in Figure 3.8 b) where contrast of the scene elements decreases considerably.

All algorithms are implemented using MATLAB as in the sonar simulations. The average simulation time for a scene as in Figure 2.15 is 3 minutes and this computation time can be reduced by using a faster compiling environment.

CHAPTER 4

EPIPOLAR GEOMETRY OF OPTI-ACOUSTIC STEREO IMAGING

In this chapter, feature correspondence problem in opti-acoustic images is investigated. Geometrical interpretation of this problem is examined on the simulated optical and sonar images given in Chapter 2 and 3.

4.1 Introduction

Opti-acoustic stereo imaging is the usage of traditional optical camera with 2D sonar in stereo configuration. The epipolar geometry deals with the feature correspondences of these sensing modalities. Although an effective general method isn't proposed for automatic matching of opti-acoustic features, establishing the epipolar geometry helps the operators in locating correct matches thanks to constricting the feature search to 1D region. It also provides a unified framework for automatic detection of underwater structures [11].

In the following sections, matrices and vectors are denoted with bold characters. \mathbf{P}_o and \mathbf{P}_s represents the coordinates of 3D point $\mathbf{P} = [X, Y, Z]$ in optical and sonar coordinate systems, respectively. Points in the optical image have coordinates (x, y) . Range and azimuth in sonar image are denoted by polar coordinates (\mathfrak{R}, θ) and ϕ is the elevation angle of point \mathbf{P} . Corresponding rectangular coordinates are expressed as $(x_s, y_s) = \mathfrak{R}(\sin\theta, \cos\theta)$. \mathbf{R} and $\mathbf{t} = [t_x, t_y, t_z]$ represent the rotation and translation of optical camera in sonar coordinate frame. r_i , \tilde{r}_i and r_{ij} denotes the rows, columns and elements of \mathbf{R} , respectively, $(i, j = 1, 2, 3)$. 3x3 matrix \mathbf{U} with

elements u_{ij} describes the epipolar contours in the optical image. Eigenvalues of \mathbf{U} are λ_i ($i = 1, 2, 3$). Finally θ_y (vergence angle) denotes the angle between the baseline and X-axis of the camera coordinate frame. A typical configuration of opti-acoustic system is shown in Figure 4.1.

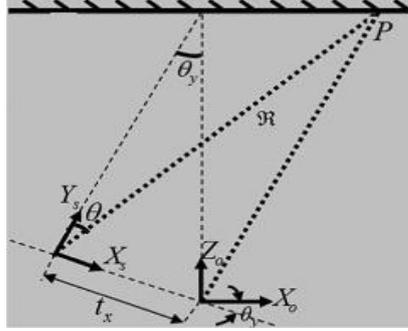


Figure 4.1: Relationship between various parameters of a typical configuration [11].

A 3D point \mathbf{P} can be expressed by rectangular and spherical coordinates $[X, Y, Z]^T$ and $[\theta, \phi, \mathfrak{R}]^T$, respectively. Referring to Figure 4.2, these coordinate systems are related by

$$\mathbf{P} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \mathfrak{R} \begin{bmatrix} \cos\phi \sin\theta \\ \cos\phi \cos\theta \\ \sin\phi \end{bmatrix} \quad (4.1)$$

and the inverse of this transformation is given by

$$\theta = \tan^{-1}\left(\frac{X}{Y}\right), \quad \phi = \tan^{-1}\left(\frac{Z}{\sqrt{X^2 + Y^2}}\right) \text{ and} \quad (4.2)$$

$$\mathfrak{R} = \sqrt{X^2 + Y^2 + Z^2}.$$

4.2 Epipolar Geometry

Assuming the ideal projection model in section 3.3.1, the optical image coordinates $\mathbf{p} = (x, y, f)$ of a 3D point $\mathbf{P}_o = [X_o, Y_o, Z_o]^T$ are given by

$$x = f \frac{X_o}{Z_o} \text{ and } y = f \frac{Y_o}{Z_o} \quad (4.3)$$

where f is the focal length of the ideal pinhole camera. With $\mathbf{P}_s = [X_s, Y_s, Z_s]^T$ denotes the 3D coordinates of same point in sonar coordinate system, coordinates of the matching point \mathbf{p} in the sonar image is given by range and azimuth measurements $\mathbf{s} = (\theta, \mathfrak{R})$ where

$$\theta = \tan^{-1}(X_s/Y_s) \text{ and } \mathfrak{R} = \sqrt{X_s^2 + Y_s^2 + Z_s^2}. \quad (4.4)$$

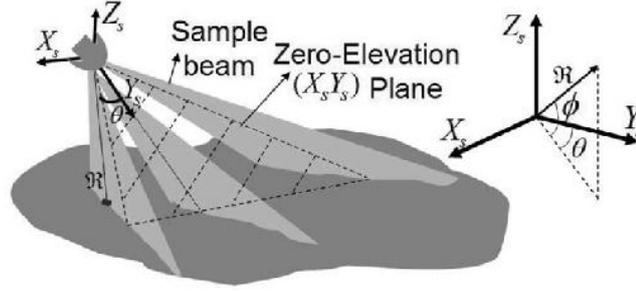


Figure 4.2: Sonar coordinate system [11].

The relationship or constraint is found between pairs \mathbf{p} and \mathbf{s} to establish epipolar geometry. In other words, given a point in sonar or optical image, the locus points, i.e. curve, should be determined in other image. These matching points corresponds to epipolar lines in stereo optical configuration [29]. On the other hand, opti-acoustic system has an asymmetrical projection model. Hence, epipolar curves in the optical image aren't lines, but conic sections. The exact shape of these conic sections depend on the relative orientation of the cameras and can be visualized by the intersection of the optical image plane with a cone comprising the locus of points with θ at distance \mathfrak{R} from the sonar [11] as shown in Figure 4.3.

4.2.1 Sonar to Optical Mapping

This mapping is looking for the locus of the matching point $\mathbf{p} = (x, y, 1)$ in optical image, given the $\mathbf{s} = (\theta, \mathfrak{R})$ in sonar image. Using the relationship between the optical and sonar coordinates of 3D point \mathbf{P}

$$\mathbf{P}_s = \mathbf{R}\mathbf{P}_o + \mathbf{t} \quad (4.5)$$

where \mathbf{R} is the 3×3 rotation matrix and $\mathbf{t} = [t_x, t_y, t_z]$ is the translation vector defining the rigid body transformation. Determination of \mathbf{R} and \mathbf{t} is an extrinsic calibration problem and the details can be found in Appendix A.

Assuming \mathbf{R} and \mathbf{t} is known, we can write

$$\mathfrak{R} = \|\mathbf{P}_s\| = \|\mathbf{R}\mathbf{P}_o + \mathbf{t}\| = \sqrt{\|\mathbf{P}_o\|^2 + 2\mathbf{t}^T \mathbf{R}\mathbf{P}_o + \|\mathbf{t}\|^2}. \quad (4.6)$$

Arranging (4.6) yields

$$\|\mathbf{P}_o\|^2 + 2(\mathbf{t}^T \mathbf{R})\mathbf{P}_o + (\|\mathbf{t}\|^2 - \mathfrak{R}^2) = 0. \quad (4.7)$$

By combining (4.3) with (4.7) we can write [11]

$$\|\mathbf{p}\|^2 Z_o^2 + 2(\mathbf{t}^T \mathbf{R}\mathbf{p})Z_o + (\|\mathbf{t}\|^2 - \mathfrak{R}^2) = 0. \quad (4.8)$$

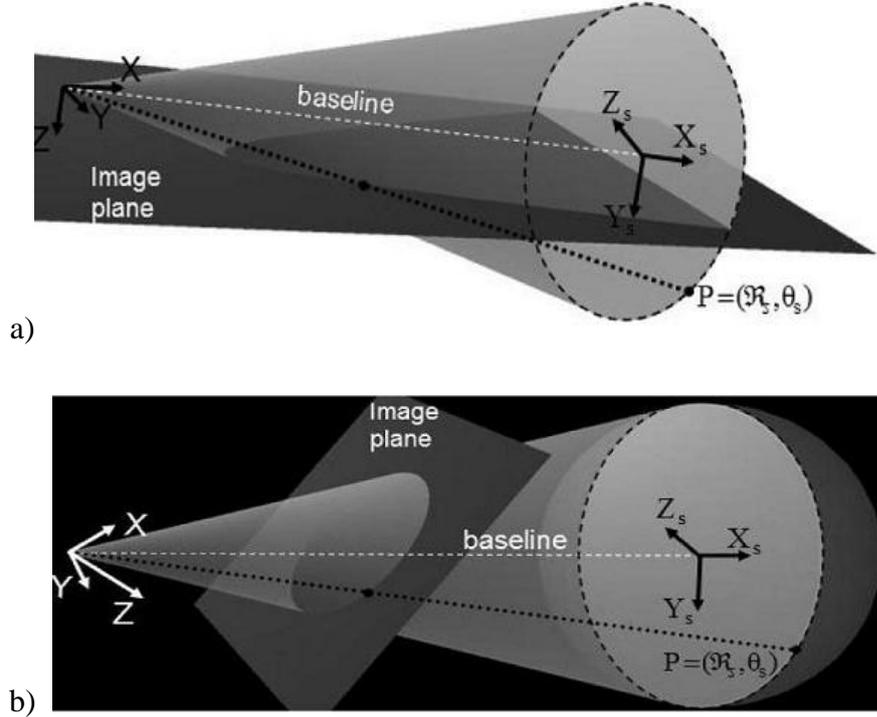


Figure 4.3: Epipolar contours in optical image are hyperbolas for nearly parallel configuration a) and become ellipses for larger vergence angles b) [11].

On the other hand, by expanding (4.6) we have

$$X_s = \mathbf{r}_1 \cdot \mathbf{P}_o + t_x \quad \text{and} \quad Y_s = \mathbf{r}_2 \cdot \mathbf{P}_o + t_y. \quad (4.9)$$

Substituting these into sonar azimuth equation yields

$$(\mathbf{r}_1 - \tan\theta \mathbf{r}_2) \cdot \mathbf{P}_o + (t_x - \tan\theta t_y) = 0 \quad (4.10)$$

Combining (4.3) with (4.10), we can write

$$Z_o = \frac{\tan\theta t_y - t_x}{(\mathbf{r}_1 - \tan\theta \mathbf{r}_2) \cdot \mathbf{p}} \quad (4.11)$$

Using (4.11) in (4.8) we obtain the epipolar constraint [11]

$$\mathbf{p}^T \mathbf{U} \mathbf{p} = 0 \quad (4.12)$$

where

$$\mathbf{U} = (\tan\theta t_y - t_x)^2 \mathbf{I} + (\|\mathbf{t}\|^2 - \mathfrak{R}^2)(\mathbf{r}_1 - \tan\theta \mathbf{r}_2)^T (\mathbf{r}_1 - \tan\theta \mathbf{r}_2) + (\tan\theta t_y - t_x)((\mathbf{r}_1 - \tan\theta \mathbf{r}_2)^T (\mathbf{t}^T \mathbf{R}) + (\mathbf{R}^T \mathbf{t})(\mathbf{r}_1 - \tan\theta \mathbf{r}_2)). \quad (4.13)$$

By expanding (4.13) in the form

$$u_{11}x^2 + 2u_{12}xy + u_{22}y^2 + 2u_{13}x + 2u_{23}y + u_{33} = 0 \quad (4.14)$$

contour shape of the epipolar constraint can be determined using the eigenvalues of the upper left 2x2 subblock of \mathbf{U} :

$$\Pi_\lambda = \lambda_1 \lambda_2 = u_{11}u_{22} - u_{12}^2. \quad (4.15)$$

Types of the epipolar curves for different values of Π_λ are given in Table 4.1.

Table 4.1: Variation in the shape of the epipolar contours.

Π_λ	Epipolar Curve
> 0	ellipse, circle, point, no curve
$= 0$	parabola, 2 parallel lines, 1 line, no curve
< 0	hyperbola, 2 intersecting lines

4.2.2 Optical to Sonar Mapping

This is the reverse mapping which is looking for the locus of the matching point $\mathbf{s} = (\theta, \mathfrak{R})$ in sonar image, given the optical image projection $\mathbf{p} = (x, y, 1)$ of 3D point \mathbf{P} . Starting with (4.5), we can write

$$\mathbf{P}_o = \mathbf{R}^{-1} \begin{bmatrix} x_s \cos\phi - t_x \\ y_s \cos\phi - t_y \\ \mathfrak{R} \sin\phi - t_z \end{bmatrix} \quad (4.15)$$

where $x_s = \mathfrak{R} \sin\theta$ and $y_s = \mathfrak{R} \cos\theta$ rectangular coordinates of the corresponding pixel. Combining (4.3) with (4.15) and solving for $\sin\phi$ and $\cos\phi$ yields [11]

$$\begin{aligned} & (q_{31}\sigma_2 - q_{32}\sigma_1)^2 + ((q_{12}\sigma_1 - q_{11}\sigma_2)\sin\theta + (q_{22}\sigma_1 - q_{21}\sigma_2)\cos\theta)^2 \\ & = \mathfrak{R}^2((q_{31}q_{12} - q_{32}q_{11})\sin\theta + (q_{31}q_{22} - q_{32}q_{21})\cos\theta)^2 \end{aligned} \quad (4.16)$$

$$\begin{aligned} q_{k1} &= yr_{k3} - r_{k2}, \quad q_{k2} = xr_{k3} - r_{k1} \quad (k=1, 2, 3) \\ \sigma_i &= t_x q_{1i} + t_y q_{2i} + t_z q_{3i} \end{aligned}$$

Arranging (4.16) we obtain the epipolar constraint

$$\mathfrak{R} = \sqrt{N_{\mathfrak{R}}(\theta)/D_{\mathfrak{R}}(\theta)} \quad (4.17)$$

where

$$\begin{aligned} N_{\mathfrak{R}}(\theta) &= (q_{31}\sigma_2 - q_{32}\sigma_1)^2 + \\ & \quad ((q_{12}\sigma_1 - q_{11}\sigma_2)\sin\theta + (q_{22}\sigma_1 - q_{21}\sigma_2)\cos\theta)^2 \\ D_{\mathfrak{R}}(\theta) &= ((q_{31}q_{12} - q_{32}q_{11})\sin\theta + (q_{31}q_{22} - q_{32}q_{21})\cos\theta)^2 \end{aligned} \quad (4.18)$$

4.3 Results

In this section, (4.12) and (4.17) is evaluated on simulated sonar and optical images in sections 2.8 and 3.4, respectively. In these simulations, the cameras are in nearly parallel configuration where rotation, translation and vergence angle is given by

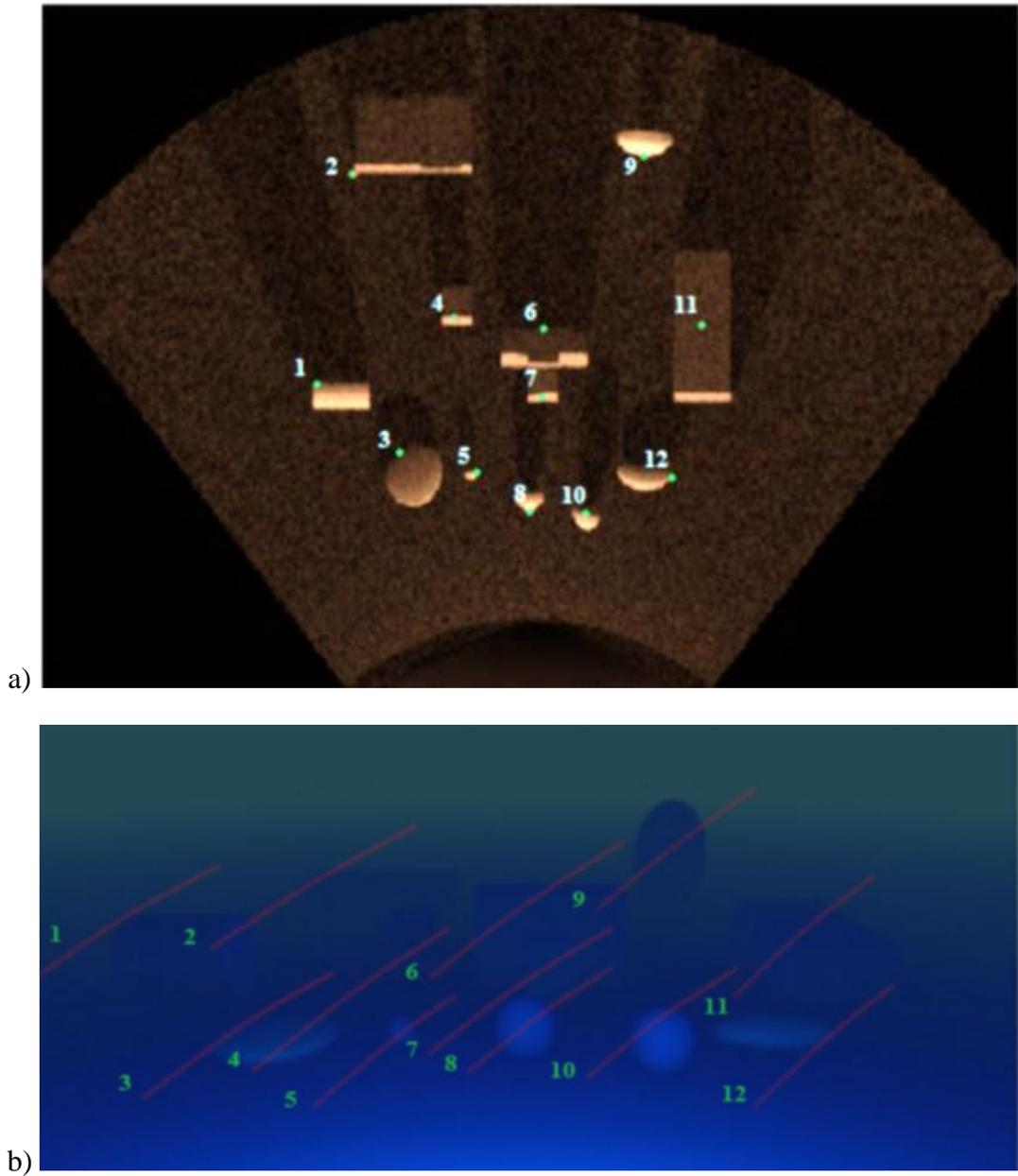


Figure 4.4: Feature points in sonar image a) and corresponding epipolar curves in optical image b).

$$\mathbf{R} = \begin{bmatrix} 0,99 & 0,04 & 0 \\ -0,04 & 0,99 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \mathbf{t} = \begin{bmatrix} -0,25 \\ 0 \\ 0 \end{bmatrix} \text{ and } \theta_y = 2.39^\circ.$$

Note that, Blueview P900 [2] beams have $\phi_B = 20^\circ$ width in elevation, that is to say, sonar view is bounded to $\pm(\phi_B/2)^\circ$ in elevation. This visibility constraint further reduces the search area for matching features [11].

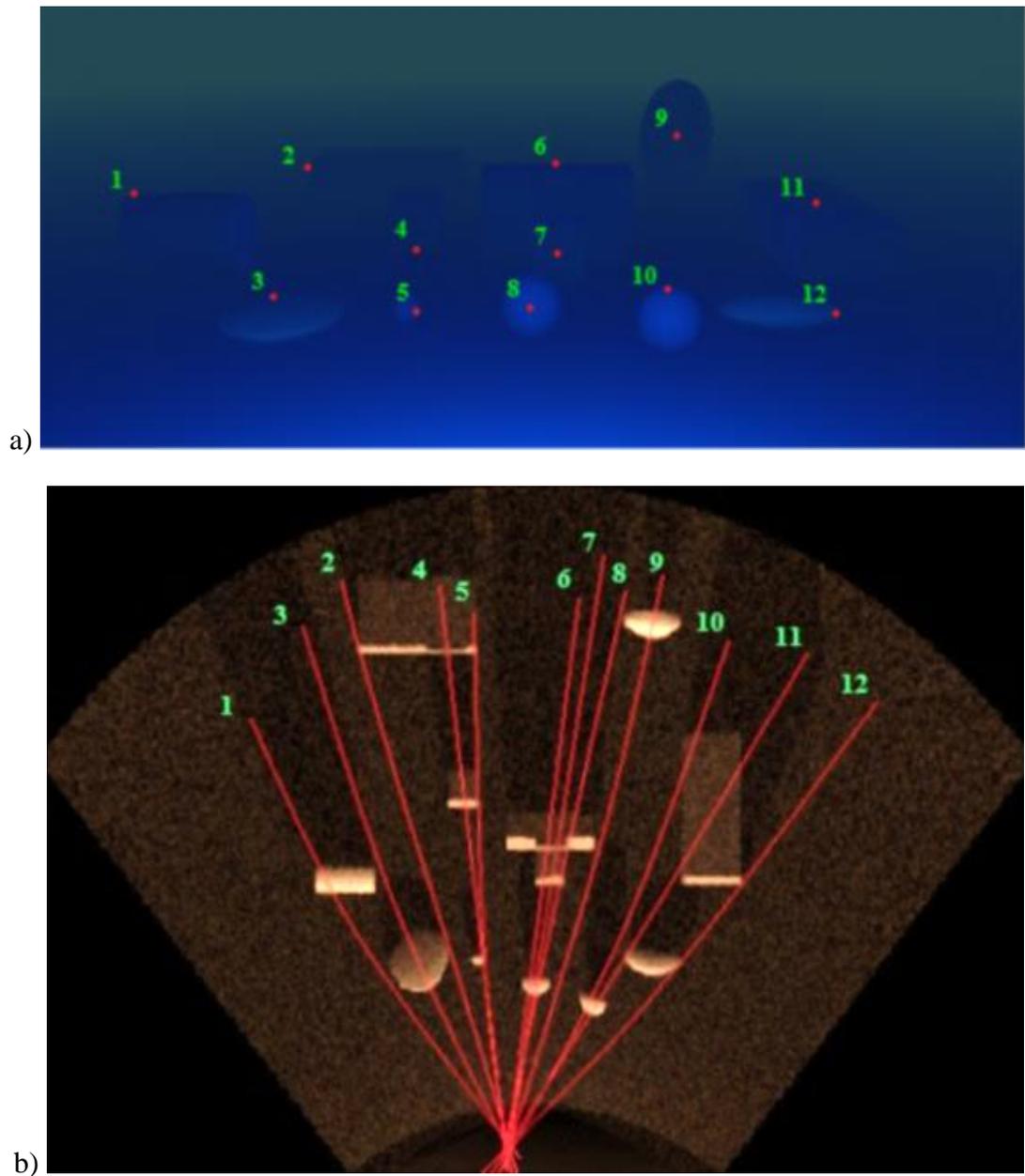


Figure 4.5: Feature points in optical image a) and corresponding epipolar curves in sonar image b).

In Figure 4.4 a), some feature points in sonar image are marked and numerated and in Figure 4.4 b) result of the sonar to optical mapping is shown. Since θ_y is relatively small ($< 30^\circ$), epipolar contours remain as hyperbolas.

In Figure 4.5 a), some feature points in optical are numerated and in Figure 4.5 b) result of the optical to sonar mapping is given where the epipole can be clearly identified.

CHAPTER 5

3D RECONSTRUCTION FROM OPTI-ACOUSTIC STEREO IMAGING

3D reconstruction from multiple views has been deeply studied for traditional electro-optical systems. In this chapter, 3D reconstruction from opti-acoustic system will be analyzed and compared with the reconstruction performance of stereo optical system.

5.1 3D Reconstruction in Stereo Optical System

Assuming ideal pinhole camera model, the 3D reconstruction problem in stereo optical system is finding the intersection of two lines passing through the camera projection centers and projection of unknown sought after scene point on image planes as shown in Figure 5.1. This process is called triangulation. In Figure 5.1, lines through the points \mathbf{x} and \mathbf{x}' are intersected in 3D scene point \mathbf{X} , because the image measurements are ideal. However, in practice, measurements \mathbf{x} and \mathbf{x}' are noisy so these lines do not intersect. Hence an optimal solution for 3D reconstruction must be found.

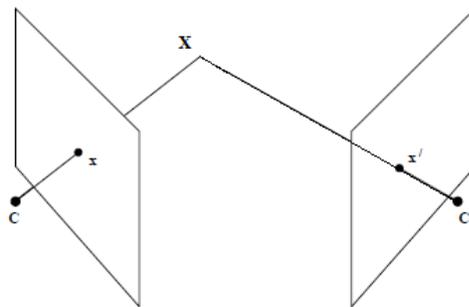


Figure 5.1: Ideal 3D reconstruction.

In general, noisy measurements \mathbf{x} and \mathbf{x}' do not satisfy the epipolar constraint. The correct locations of these noisy measurements $\hat{\mathbf{x}}$ and $\hat{\mathbf{x}}'$, which satisfy the epipolar constraint ($\hat{\mathbf{x}}'^T \mathbf{F} \hat{\mathbf{x}} = 0$), are in close neighborhood of the noisy measurements. Therefore, these points can be found by minimizing the reprojection error for a 3D point $\hat{\mathbf{X}}$, which can be expressed as the following optimization problem [29]:

$$\text{Minimize } \|\mathbf{x} - \hat{\mathbf{x}}\|^2 + \|\mathbf{x}' - \hat{\mathbf{x}}'\|^2 \quad (5.1)$$

subject to $\hat{\mathbf{x}}'^T \mathbf{F} \hat{\mathbf{x}} = 0$, where \mathbf{F} is the fundamental matrix.

Relationship between noisy and optimal points is shown in Figure 5.2 where d and d' are the Euclidean distances between measured and reprojected points.

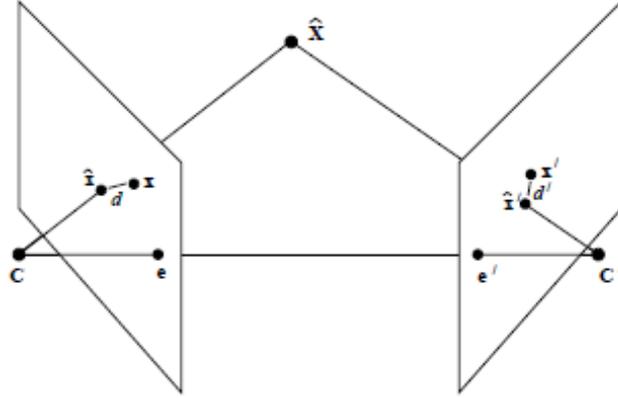


Figure 5.2: Minimization of geometric error.

5.2 3D Reconstruction in Opti-Acoustic System

In an opti-acoustic system, 3D reconstruction is determining the 3D point $\mathbf{P} = [X, Y, Z]$, given the matches $\mathbf{p} = (x, y, f)$ and $\mathbf{s} = (\theta, \mathfrak{R})$ which is generally more accurate than stereo optical system [12]. However, noisy opti-acoustic correspondences do not satisfy the epipolar geometry as in stereo optical systems.

Four constraints arise from projection equations in opti-acoustic system which is given as

$$x = f(X_o/Z_o) \text{ and } y = f(Y_o/Z_o) \quad (5.2)$$

$$\theta = \tan^{-1}(X_s/Y_s) \text{ and } \mathfrak{R} = \sqrt{X_s^2 + Y_s^2 + Z_s^2}. \quad (5.3)$$

Using three out of four constraints provides different estimates of 3D point for each opti-acoustic matches. Since the equations in (5.2) and (5.3) are asymmetrical because of the different projection models in sonar and optical camera, various closed form solutions can be derived [11]. On the other hand, these suboptimal solutions should be optimized by minimizing some geometrical error in maximum likelihood (ML) sense as in the stereo optical system.

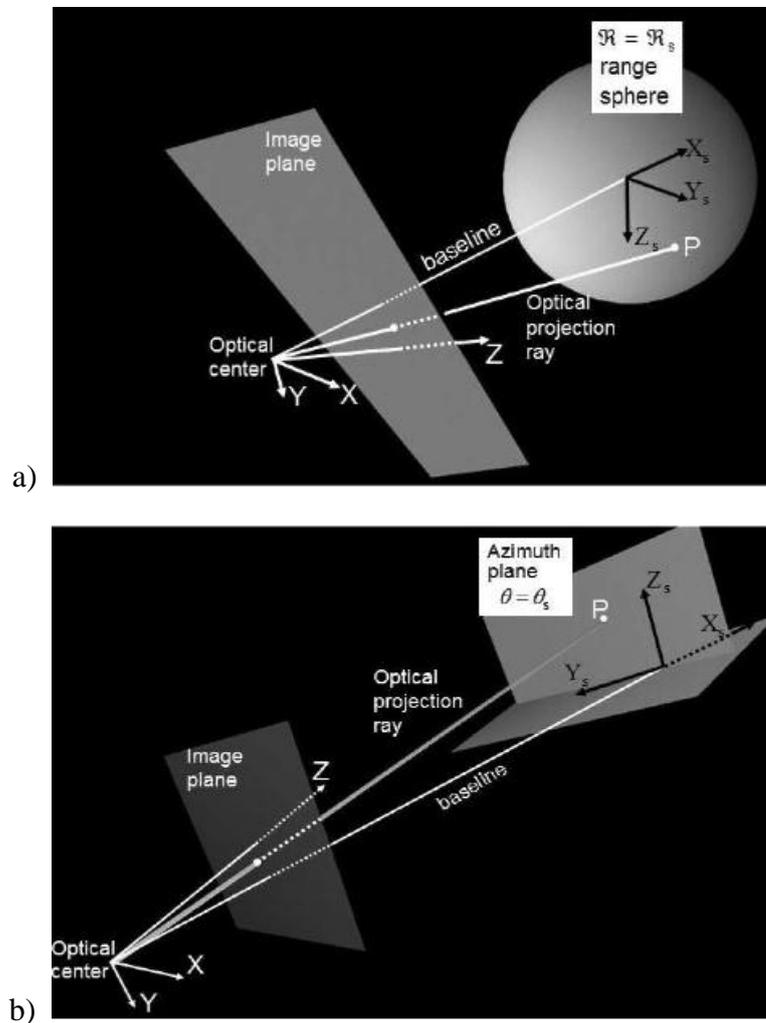


Figure 5.3: Geometric interpretation of range solution a) and azimuth solution b) [11].

Range solution is the intersection of the optical ray with the range sphere. Using (4.5), (5.2) and range equation in (5.3), range solution is obtained as

$$Z_{\mathfrak{R}} = f \frac{\mathfrak{R}\|\mathbf{p}\| - \mathbf{t}^T \mathbf{R}\mathbf{p}}{\|\mathbf{p}\|^2}. \quad (5.4)$$

Similarly, azimuth solution is the intersection of the optical ray with the azimuth plane. Using (4.5), (5.2) and azimuth equation in (5.3), azimuth solution is found as

$$Z_{\theta} = f \frac{\tan\theta t_y - t_x}{(\mathbf{r}_1 - \tan\theta \mathbf{r}_2) \cdot \mathbf{p}}. \quad (5.5)$$

Illustration of these closed form solutions is shown in Figure 5.3.

While (5.4) and (5.5) give the exact solution with ideal measurements, they produce incorrect estimates with noisy data since the corresponding points in sonar and optical image do not satisfy the epipolar constraint [1]. So, by applying an appropriate cost function, geometric error should be minimized to improve noisy estimates.

5.2.1 Indirect Method

Perfect matching optical and sonar features satisfy the epipolar constraint given in (4.12). However, this constraint can not be achieved in real images due to the noisy observations. Assuming, both images corrupted by additive Gaussian noise, imperfect observations are given as

$$\begin{cases} \hat{x} = x + n(0, \sigma_x) \\ \hat{y} = y + n(0, \sigma_y) \end{cases} \quad \begin{cases} \hat{\mathfrak{R}} = \mathfrak{R} + n(0, \sigma_{\mathfrak{R}}) \\ \hat{\theta} = \theta + n(0, \sigma_{\theta}) \end{cases} \quad (5.6)$$

where $n(0, \sigma_i)$ denotes the normal distribution with zero mean and variance σ_i . On the other hand, removal of the outliers with an appropriate robust estimation method (such as RANSAC) is also assumed which allows us to study with only the inliers for desired distribution. Under these assumptions, indirect 3D reconstruction method involves the following steps [12]:

1) Find Maximum Likelihood Estimate (MLE) of $\mathbf{p} = (x, y, f)$ and $\mathbf{s} = (\theta, \mathfrak{R})$ that satisfy (4.12) while minimizing the Mahalanobis distance between vectors $X = (x, y, \mathfrak{R}, \theta)$ and $\hat{X} = (\hat{x}, \hat{y}, \hat{\mathfrak{R}}, \hat{\theta})$;

2) Apply closed form solutions to MLE.

First step can be formulated as a constrained optimization problem:

$$\text{Minimize } (\mathbf{X} - \hat{\mathbf{X}})^T \boldsymbol{\Sigma}^{-1} (\mathbf{X} - \hat{\mathbf{X}}) \quad (5.7)$$

subject to $\mathbf{p}^T \mathbf{U} \mathbf{p} = 0$, where $\boldsymbol{\Sigma} = \text{E} \left[(\mathbf{X} - \hat{\mathbf{X}})(\mathbf{X} - \hat{\mathbf{X}})^T \right]$. Assuming measurement errors in optical and sonar images are independent and utilizing a penalty function instead of strictly enforcing the epipolar constraint, (5.7) can be expressed as the following unconstrained optimization problem:

$$\text{Minimize } \frac{(x - \hat{x})^2}{\sigma_x^2} + \frac{(y - \hat{y})^2}{\sigma_y^2} + \frac{(\mathfrak{R} - \hat{\mathfrak{R}})^2}{\sigma_{\mathfrak{R}}^2} + \frac{(\theta - \hat{\theta})^2}{\sigma_{\theta}^2} + \lambda (\mathbf{p}^T \mathbf{U} \mathbf{p})^2 \quad (5.8)$$

Solution of (5.8) will be close to satisfying epipolar constraint by suitable selection of λ .

5.2.2 Direct Method

Direct method also involves two steps [12]:

1) Compute closed form solution of 3D point $\hat{\mathbf{P}}$ using noisy measurements $\hat{\mathbf{p}} = (\hat{x}, \hat{y}, f)$ and $\hat{\mathbf{s}} = (\hat{\mathfrak{R}}, \hat{\theta})$;

2) Find Maximum Likelihood Estimate (MLE) of $\mathbf{p} = (x, y, f)$ and $\mathbf{s} = (\theta, \mathfrak{R})$ by minimizing the Mahalanobis distance between vectors X and \hat{X} based on (5.2) and (5.3).

In this method the function to be minimized is

$$\frac{(x - \hat{x})^2}{\sigma_x^2} + \frac{(y - \hat{y})^2}{\sigma_y^2} + \frac{(\mathfrak{R} - \hat{\mathfrak{R}})^2}{\sigma_{\mathfrak{R}}^2} + \frac{(\theta - \hat{\theta})^2}{\sigma_{\theta}^2}. \quad (5.9)$$

(5.8) and (5.9) are nonlinear optimization problems and can be solved by applying Levenberg-Marquardt algorithm [30].

5.2.3 First-Order Error Analysis

Reconstruction error can be quantified for exact and estimated depths Z and \hat{Z} based on mean square error (MSE), i.e., $E[(Z - \hat{Z})^2]$. For each of the closed form solutions given in (5.4) and (5.5), first-order analytical MSE can be obtained by deriving Jacobian and given as

$$MSE = \left(\frac{\partial Z}{\partial x}\right)^2 \sigma_x^2 + \left(\frac{\partial Z}{\partial y}\right)^2 \sigma_y^2 + \left(\frac{\partial Z}{\partial \mathfrak{R}}\right)^2 \sigma_{\mathfrak{R}}^2 + \left(\frac{\partial Z}{\partial \theta}\right)^2 \sigma_{\theta}^2 \quad (5.10)$$

which gives relatively accurate results for nondegenerate scene geometries.

5.3 Results

In this section, the performance of both stereo optical and opti-acoustic systems are compared for the reconstruction of 3D points on Klein bottle shown in Figure 5.4. First, closed form solutions $Z_{\mathfrak{R}}$ and Z_{θ} is computed. Second, improvements on range solution with direct and indirect method is evaluated. Finally, reconstruction from stereo optical system using linear solution based on singular value decomposition is given.

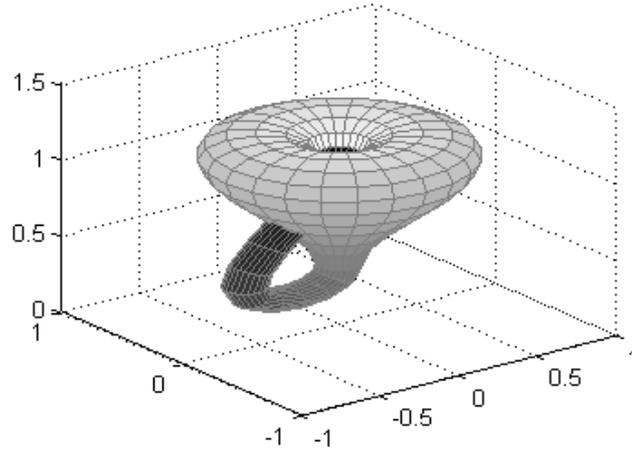


Figure 5.4: Klein bottle.

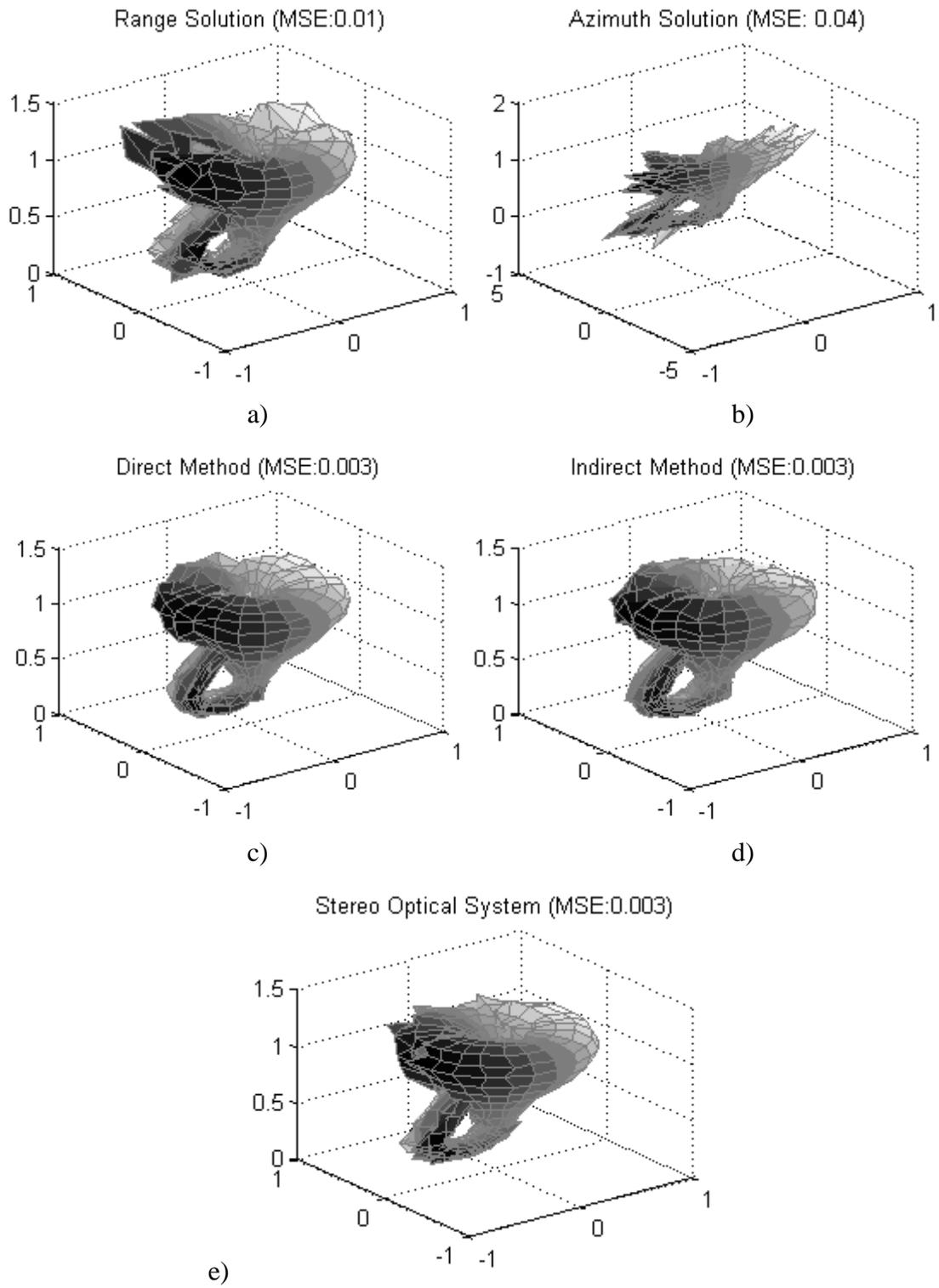


Figure 5.5: Reconstruction of Klein bottle points from noisy data for baseline of 25 cm and average distance of 1.25 m.

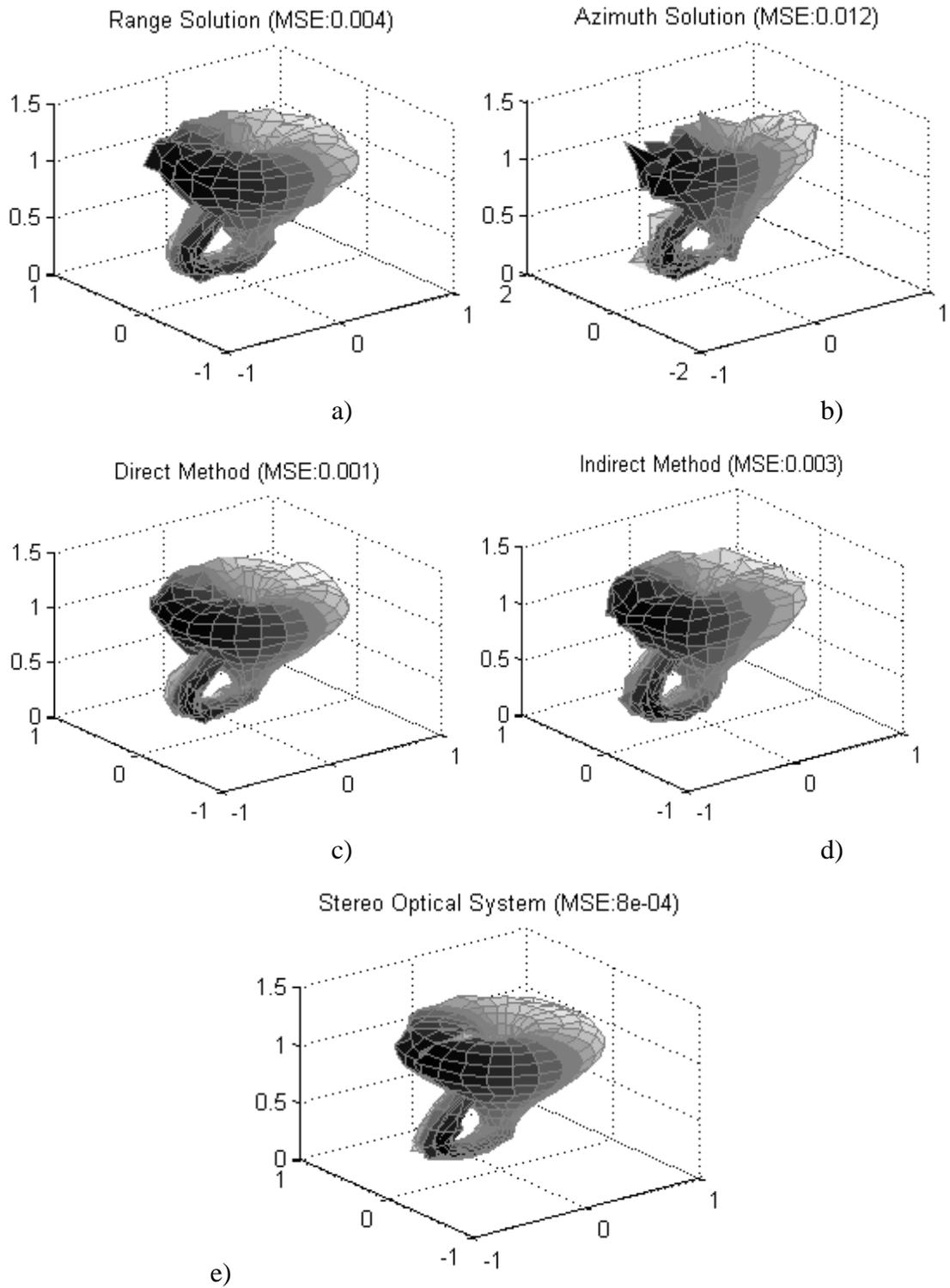


Figure 5.6: Reconstruction of Klein bottle from noisy points for baseline of 50 cm and average distance of 1.25 m.

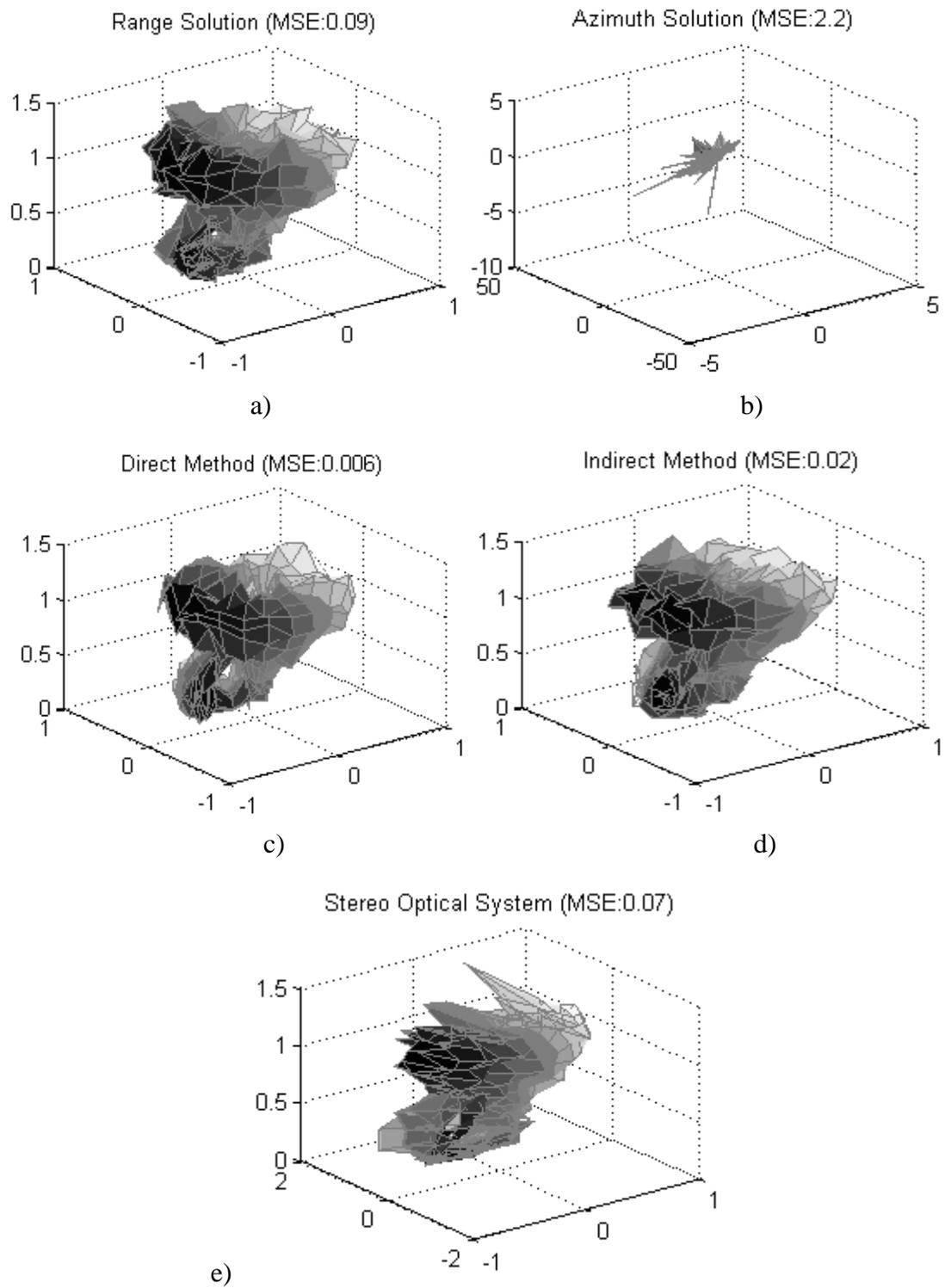


Figure 5.7: Reconstruction of Klein bottle from noisy points for baseline of 25 cm and average distance of 3 m.

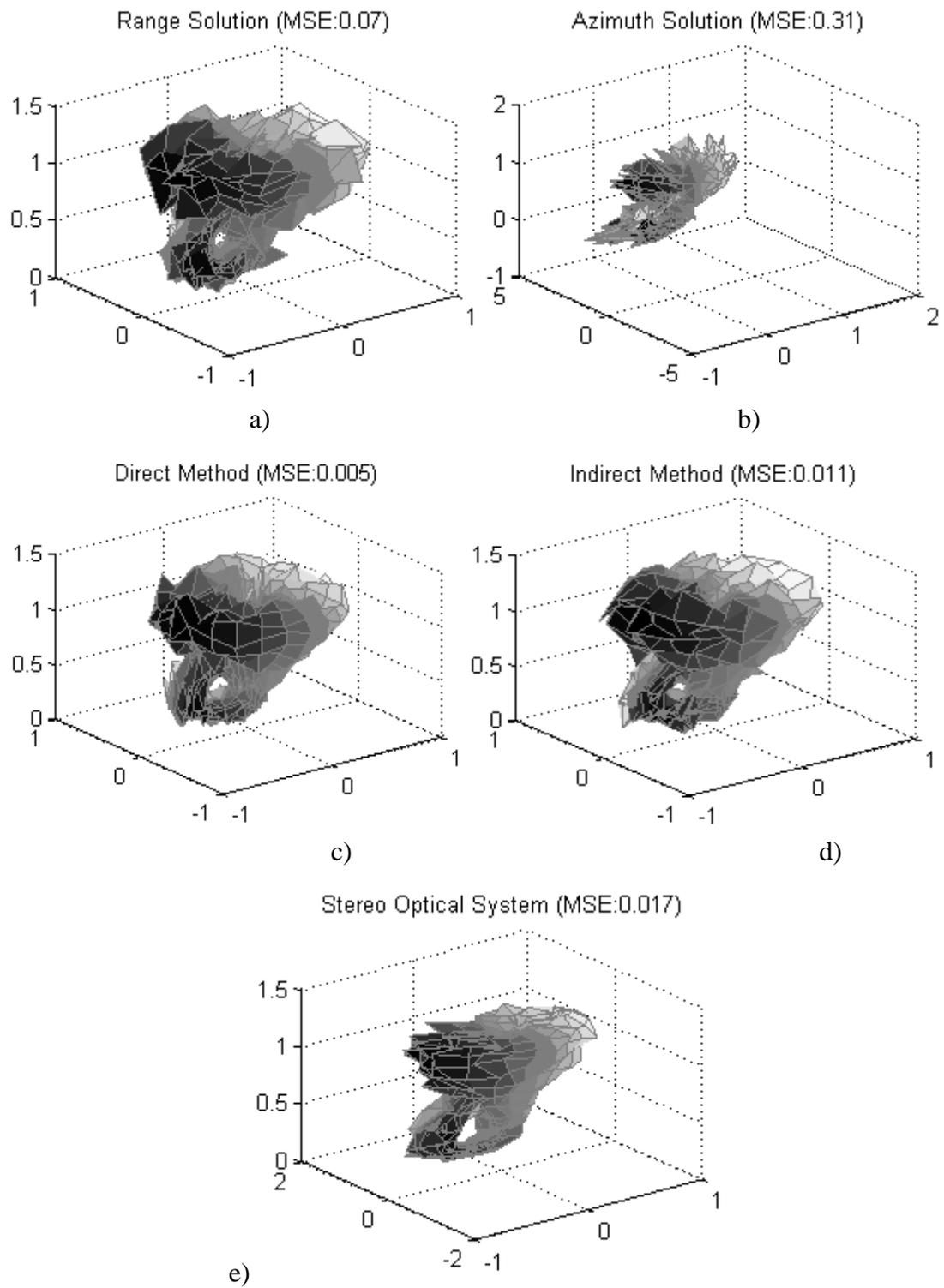


Figure 5.8: Reconstruction of Klein bottle from noisy points for baseline of 50 cm and average distance of 3 m.

In simulations, cameras are in nearly parallel configuration and 3D reconstruction performances are evaluated based on MSE for varying baseline, which is the distance between cameras, and average distance from target. Noise variances are assumed $\sigma_x = \sigma_y = 2$ pixels, $\sigma_{\text{sr}} = 2$ cm and $\sigma_\theta = 0.03$ radians.

In Figure 5.5, baseline is 0.25m and average distance from bottle is 1.25m. In this configuration stereo optical and opti-acoustical system show the same performance.

In Figure 5.6, baseline is increased to 0.5m while the average distance from bottle is kept same. Performance of both stereo systems increases (especially azimuth solution) compared to the system in Figure 5.5. In this configuration, stereo optical system and opti-acoustic system with direct method show similar performances though stereo optical system has a slight edge.

In Figure 5.7, baseline is again adjusted to 0.25m and average distance from bottle is increased to 3m. Performance of both stereo systems decreases (especially azimuth solution and stereo optical system) due to the increasing target distance. In this configuration, opti-acoustic system is superior to stereo optical system.

In Figure 5.8, baseline is increased to 0.5m while the average distance from bottle is kept same (3m). Performance of both stereo systems increases (especially stereo optical system) compared to the system in Figure 5.7, though opti-acoustic system shows better performance than stereo optical system.

CHAPTER 6

3D MOTION ESTIMATION FROM OPTI-ACOUSTIC STEREO IMAGING

In this chapter, 3D motion estimation from opti-acoustic image sequences is studied. Unlike 3D reconstruction problem, optical and sonar image correspondences are treated separately to obtain motion unknowns. Theoretical results are examined on a 3D structure with known geometry.

6.1 Introduction

It is essential to estimate 3D motion and scene structure for 3D scene recognition and positioning. This problem has been deeply studied for monocular motion and it is natural to integrate sonar and optical cameras for the solution of motion estimation problem.

The advantage of stereo system in motion estimation is undisputable. In monocular system, side-to-side translations and rotations induces similar displacements in some scene geometries as shown in Figure 6.1. On the other hand, a second camera located at sufficiently large distance from the first camera, induces different motions hence two motions can be discriminated.

Estimating the scene structure from a sonar and optical camera requires finding matching features for two views which is the most complex problem in opti-acoustical stereo imaging due to the uncorrelated image intensity characteristics of these modalities. By evaluating motion in sonar and optical image sequences separately, this problem can be prevented. Furthermore, motion estimation accuracy

and robustness can be improved by integrating visual cues and range information from optical and sonar sequences, respectively [13]. On the other hand, dual solution ambiguity for the motion relative to planar surfaces can be resolved using the sonar motion transformation.

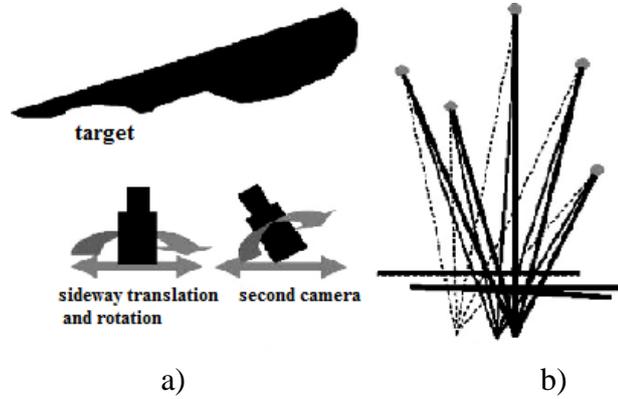


Figure 6.1: Rotation and translation in monocular system a) and induced motions in monocular and stereo systems b) [13].

Feature points on the scene surfaces are generally located on some dominant plane π_g . Using the projections of minimum four non-collinear coplanar points on π_g in two views, the rigid body motion $\{\mathbf{R}_o, \mathbf{t}_o\}$ of the optical camera and normal vector \mathbf{n}_o of π_g can be determined, up to well-known scale factor of monocular vision [29]. Alternatively, arbitrary five points can be used to compute essential matrix which then can be decomposed into $\{\mathbf{R}_o, \mathbf{t}_o\}$ up-to-scale.

Up-to-scale motion obtained from optical sequence can be expressed in sonar coordinate frame provided that the opti-acoustic system is calibrated. The scale factor ambiguity, which prevents the accurate scene reconstruction, can be solved with a single match in the consecutive sonar views [13]. This solution can be further improved by an appropriate maximum likelihood formulation which utilizes all available optical and sonar correspondences $\{\mathbf{p}_o, \mathbf{p}'_o\}$ and $\{\mathbf{p}_s, \mathbf{p}'_s\}$. Finally, 3D position of points on the π_g can be determined using the optimum rotation and translation.

6.2 Preliminaries

In the following sections, $\mathbf{P} = (X, Y, Z)^T$ and $\tilde{\mathbf{P}} = (\mathbf{P}^T, 1)^T$ denotes a 3D point P and its homogenous coordinates. 3D coordinates and corresponding homogenous coordinates of P in the first sonar and optical coordinate systems are denoted $\mathbf{P}_s, \tilde{\mathbf{P}}_s$ and $\mathbf{P}_o, \tilde{\mathbf{P}}_o$ respectively. 2D coordinates and homogeneous coordinates of P in the first optical image are $\mathbf{p}_o = (x, y)$ and $\tilde{\mathbf{p}}_o = \lambda(\tilde{\mathbf{p}}_o^T, 1)^T$. Spherical coordinates $[\theta, \phi, \mathfrak{R}]^T$ are used in the sonar measurements for the range, azimuth and elevation angle:

$$\mathbf{P}_s = \begin{bmatrix} X_s \\ Y_s \\ Z_s \end{bmatrix} = \mathfrak{R} \begin{bmatrix} \cos\phi \sin\theta \\ \cos\phi \cos\theta \\ \sin\phi \end{bmatrix} \quad (6.1)$$

and the inverse transformation is

$$\theta = \tan^{-1}\left(\frac{X_s}{Y_s}\right), \quad \phi = \tan^{-1}\left(\frac{Z_s}{\sqrt{X_s^2 + Y_s^2}}\right) \text{ and} \quad (6.2)$$

$$\mathfrak{R} = \sqrt{X_s^2 + Y_s^2 + Z_s^2}.$$

6.2.1 Motion Transformation and Plane Representation

Coordinates of point P in optical coordinate frame can be expressed in sonar coordinate frame as

$$\mathbf{P}_s = \mathbf{R}\mathbf{P}_o + \mathbf{t} \quad (6.3)$$

where \mathbf{R} is the 3x3 rotation matrix and $\mathbf{t} = [t_x, t_y, t_z]$ is translation vector defining the rigid body transformation between cameras.

Previous and current viewing positions of point P in optical and sonar cameras are denoted $\mathbf{P}_s, \mathbf{P}'_s$ and $\mathbf{P}_o, \mathbf{P}'_o$, respectively. These points are related as

$$\begin{aligned} \mathbf{P}'_s &= \mathbf{R}_s \mathbf{P}_s + \mathbf{t}_s \\ \mathbf{P}'_o &= \mathbf{R}_o \mathbf{P}_o + \mathbf{t}_o \end{aligned} \quad (6.4)$$

where $\{\mathbf{R}_s, \mathbf{t}_s\}$ and $\{\mathbf{R}_o, \mathbf{t}_o\}$ are the sought after rigid body motion unknowns.

Since rigid body transformation is valid for both previous and current viewing positions, using (6.3) and (6.4), we can write

$$\begin{aligned} \mathbf{R}_s &= \mathbf{R}\mathbf{R}_o\mathbf{R}^T \\ \mathbf{t}_s &= (\mathbf{I} - \mathbf{R}\mathbf{R}_o\mathbf{R}^T) + \mathbf{R}\mathbf{t}_o. \end{aligned} \quad (6.5)$$

A point P on plane π_g with normal \mathbf{n}_i ($i = \{s, o\}$) satisfies the equation

$$\mathbf{P}_i \cdot \mathbf{n}_i = -1 \quad (6.6)$$

Using (6.3) and (6.6), the relation between normal vector of π_g in sonar and optical coordinate frames can be written as

$$\mathbf{n}_s = \left(\frac{1}{1 - \mathbf{t}^T \mathbf{R} \mathbf{n}_o} \right) \mathbf{R} \mathbf{n}_o. \quad (6.7)$$

6.2.2 Image Measurements and Transformations

Using the perspective projection model in section 3.3.1, projection of 3D point P onto previous and current optical views are

$$\begin{aligned} \tilde{\mathbf{p}}_o &= \mathbf{C} \tilde{\mathbf{P}}_o \\ \tilde{\mathbf{p}}'_o &= \mathbf{C}' \tilde{\mathbf{P}}'_o \end{aligned} \quad (6.8)$$

where \mathbf{C} and \mathbf{C}' are 3x4 camera matrices of consecutive views.

The matching points \mathbf{p}_o and \mathbf{p}'_o in consecutive optical views, which are the projection of plane π_g , satisfies the up-to-scale projective transformation [29],

$$\mathbf{p}'_o \cong \mathbf{Q}_o \mathbf{p}_o \quad (6.9)$$

where \cong denotes an up-to-scale equality. The 3x3 projective transformation matrix \mathbf{Q}_o can be derived with the minimum of four non-collinear coplanar scene points [31]. Alternatively, 5 non-degenerate matching points can be used to estimate the essential matrix with more recent 5-point algorithm [32]. Projection of P onto sonar image is expressed with range and azimuth measurements and denoted by

$\mathbf{p}_s = (x_s, y_s) = \Re(\sin\theta, \cos\theta)$ and corresponding homogenous coordinates are $\tilde{\mathbf{p}}_s = (\tilde{\mathbf{p}}_s^T, 1)^T$.

The matching 3D points \mathbf{P}_s and \mathbf{P}'_s in consecutive sonar views satisfy [13]

$$\mathbf{P}'_s = \mathbf{Q}_s \mathbf{P}_s \quad (6.10)$$

where $\mathbf{Q}_s = \mathbf{R}_s - \mathbf{t}_s \mathbf{n}_s^T$. On the other hand, matching sonar image points satisfy

$$\tilde{\mathbf{p}}'_s = \mathbf{H}'_s \tilde{\mathbf{p}}_s \quad \mathbf{H}_s = \begin{bmatrix} \alpha q_{11} & \alpha q_{12} & \beta q_{11} \\ \alpha q_{21} & \alpha q_{22} & \beta q_{11} \\ 0 & 0 & 1 \end{bmatrix} \quad (6.11)$$

where q_{ij} denotes the elements of \mathbf{Q}_s , $\alpha = \cos\phi / \cos\phi'$ and $\beta = \Re \sin\phi / \cos\phi'$.

6.3 Up-to-Scale Estimation

Up-to-scale transformation \mathbf{Q}_o , also known as the homography of the consecutive views is fixed by the orientation of the scene plane and motion of the camera:

$$\mathbf{Q}_o \cong \mathbf{R}_o - \mathbf{t}_o \mathbf{n}_o^T \quad (6.12)$$

\mathbf{Q}_o decomposed into $\{\mathbf{R}, k^{-1} \mathbf{t}_o, k \mathbf{n}_o\}$, where k represents the unknown scale factor ambiguity. The normal \mathbf{n}_o of plane π_g may be normalized such that $\hat{\mathbf{n}}_o = k_o \mathbf{n}_o$ where $k_o = 1 / \|\mathbf{n}_o\|$. Hence the translation vector becomes $\hat{\mathbf{t}}_o = k_o^{-1} \mathbf{t}_o$.

By decomposing \mathbf{Q}_o , we obtain up to 8 interpretations. Using reference plane non-crossing and reference point visibility constraints, 6 incorrect interpretations can be eliminated [33]. Remaining solutions are, also known as true and dual solution, can be derived in terms of each other in closed form. If 5 point is used instead of 4 planar points, decomposition of essential matrix into $\{\mathbf{R}, k \mathbf{t}_o\}$, yields up to 10 interpretations. Alternatively, single up-to-scale interpretation can be derived with eight point algorithm [34]. To estimate homography, features in the optical sequences may be matched manually for improved accuracy and a RANSAC-based implementation [35] allows us to remove outliers.

6.4 Resolving the Scale Factor Ambiguity

The surface normal in the sonar coordinate is

$$\mathbf{n}_s = \|k_s\| \hat{\mathbf{n}}_s; \quad k_s = \frac{k_o^{-1}}{1 - k_o^{-1} \mathbf{t}^T \mathbf{R} \hat{\mathbf{n}}_o}, \quad \hat{\mathbf{n}}_s = \mathbf{R} \hat{\mathbf{n}}_o. \quad (6.13)$$

By determining the magnitude of the surface normal k_s , scale factor ambiguity can be resolved [13]. Using (6.6)

$$\mathbf{P}_s \cdot \|k_s\| \hat{\mathbf{n}}_s = -1, \quad \mathbf{n}_s = [\hat{n}_x \ \hat{n}_y \ \hat{n}_z] \quad (6.14)$$

which can be expanded as

$$k_s \left((\hat{n}_x \sin\theta + \hat{n}_y \cos\theta) \cos\phi + \hat{n}_z \sin\phi \right) = -1 / \mathfrak{R}. \quad (6.15)$$

It follows that

$$\begin{aligned} \phi &= \{\phi_s, \pi - \phi_s\} \\ \phi_s &= \tan^{-1} \left(\frac{\hat{n}_x \sin\theta + \hat{n}_y \cos\theta}{\hat{n}_z} \right) \\ &+ \sin^{-1} \left(\frac{-1}{k_s \mathfrak{R} \sqrt{(\hat{n}_x \sin\theta + \hat{n}_y \cos\theta)^2 + \hat{n}_z^2}} \right). \end{aligned} \quad (6.16)$$

A single match in the two sonar image corresponding to a point on π_g , gives two equations to determine the elevation angle of that point. Using these equations in (6.11), the unknown $\cos\phi'$ can be eliminated. The correct solution for the elevation angle can be chosen based on the visibility constraint of the sonar and the scale factor is recovered.

6.5 MLE Formulation

The solution for k_s obtained from a single sonar match is sub-optimal in ML sense since we have not use all available motion cues [36]. This solution can be used as

initial guess in a nonlinear optimization method that uses all available optical and sonar matches $\{\mathbf{p}_o, \mathbf{p}'_o\}$ and $\{\mathbf{p}_s, \mathbf{p}'_s\}$, respectively.

To model uncertainty in the measurements \mathbf{p}_o and \mathbf{p}_s , Gaussian noise is added. This assumption allows to minimize the Euclidean distance between observations and projections of the estimated 3D points in ML sense. Hence, we minimize

$$f(\mathbf{R}, \mathbf{t}_o, \mathbf{n}) = \sum_N \frac{\|\mathbf{p}_o - \widehat{\mathbf{p}}_o\|^2 + \|\mathbf{p}'_o - \widehat{\mathbf{p}}'_o\|^2}{\sigma_o^2} + \lambda \frac{\|\mathbf{p}_s - \widehat{\mathbf{p}}_s\|^2 + \|\mathbf{p}'_s - \widehat{\mathbf{p}}'_s\|^2}{\sigma_s^2} \quad (6.17)$$

where N is the number of matches, $(\widehat{\cdot})$ denotes any reprojection and λ is a regularization parameter which accounts for the relative noise levels of optical and sonar measurements.

To obtain reprojections, 3D points corresponding to the observations $\{\mathbf{p}_o, \mathbf{p}'_o\}$ and $\{\mathbf{p}_s, \mathbf{p}'_s\}$ is determined by triangulation (see Appendix B for details) using the initial estimates of \mathbf{R} and \mathbf{t}_o and the. Then, (6.17), which is a nonlinear function, is minimized by applying the Levenberg-Marquardt algorithm [29]. To increase the estimation accuracy, points that do not belong to the plane π_g are discarded by RANSAC [35, 37]. However, each match is reconsidered for labeling after each iteration.

6.6 Results

In this section, motion estimation accuracy of the opti-acoustic system is examined using two consecutive image pairs of perpendicular planes. The scene and position of stereo rig relative to the planes are given in Figure 6.2 where O1, O2 and S1, S2 denotes the first and second positions of optical camera and sonar, respectively. The average distance between the system and scene is approximately 1.5m. The image of the scene for consecutive views is shown in Figure 6.3.

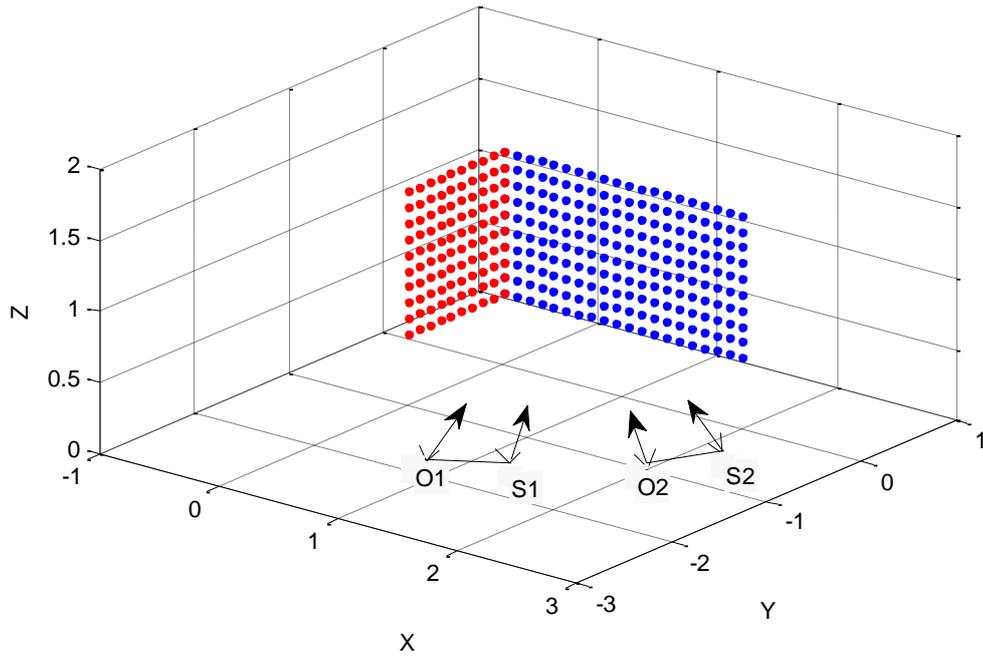


Figure 6.2: Position of opti-acoustic system relative to the scene.

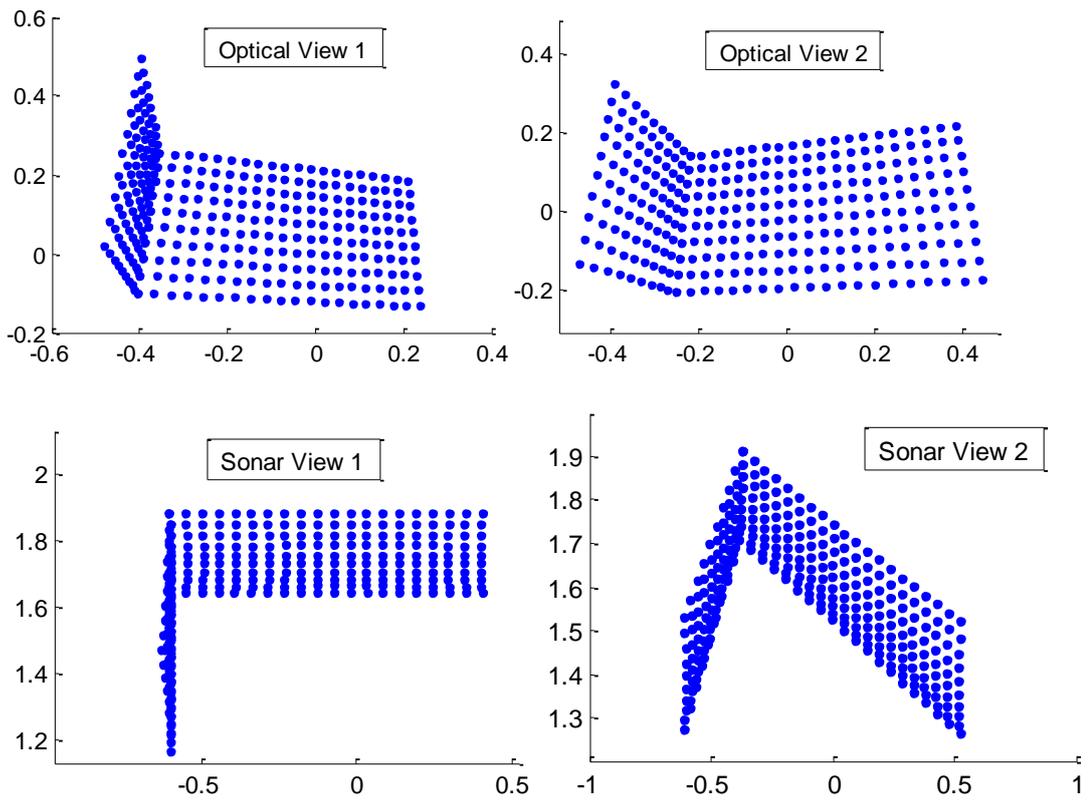


Figure 6.3: Images of the scene in consecutive views.

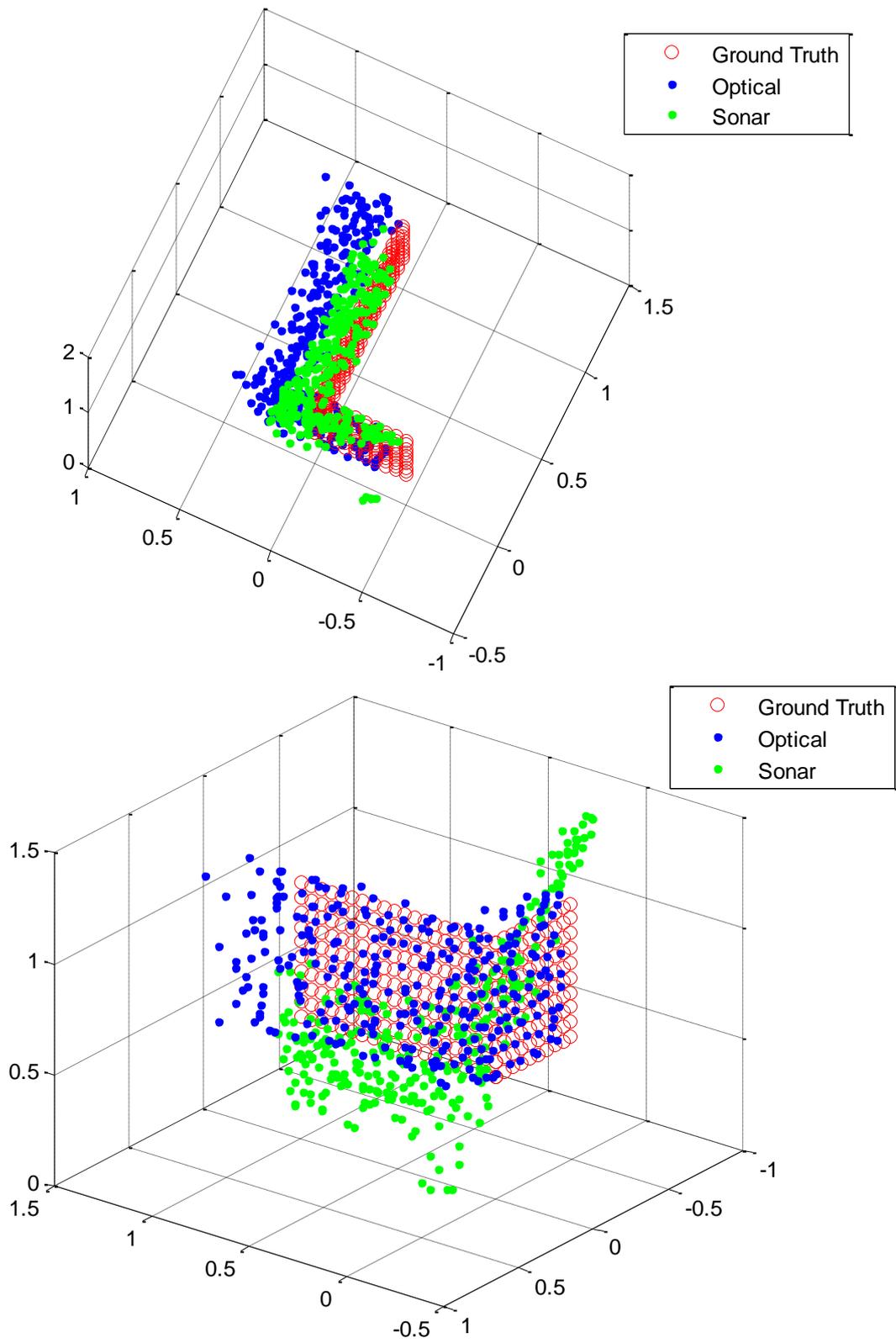


Figure 6.4: 3D reconstruction based on initial estimation of motion.

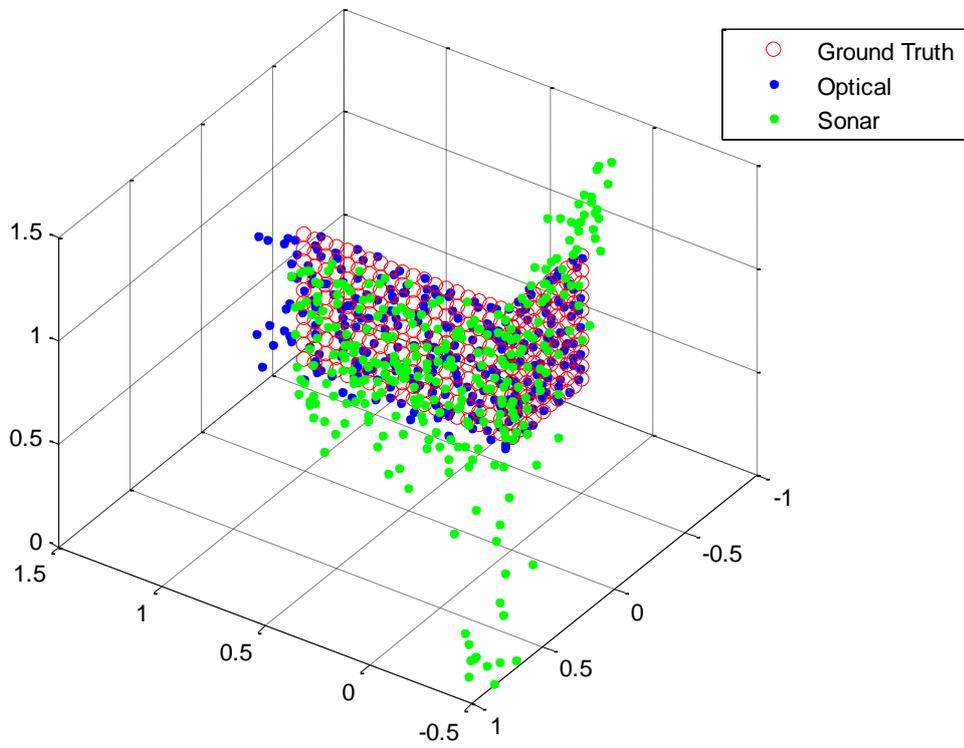
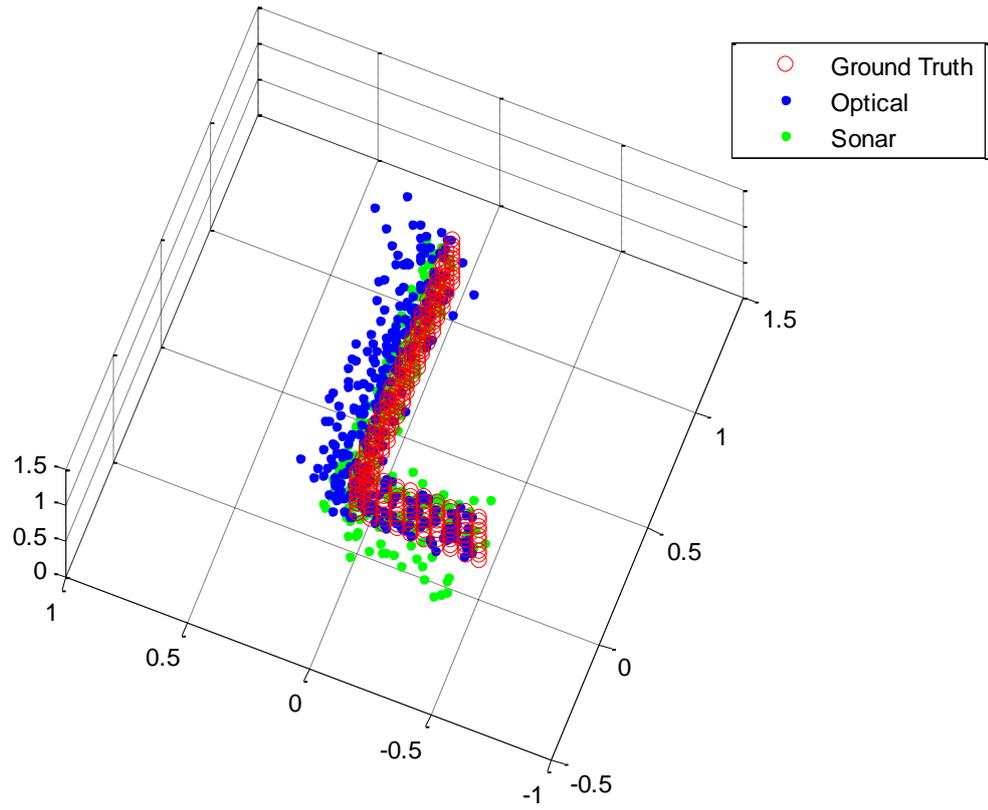


Figure 6.5: Planar estimator.

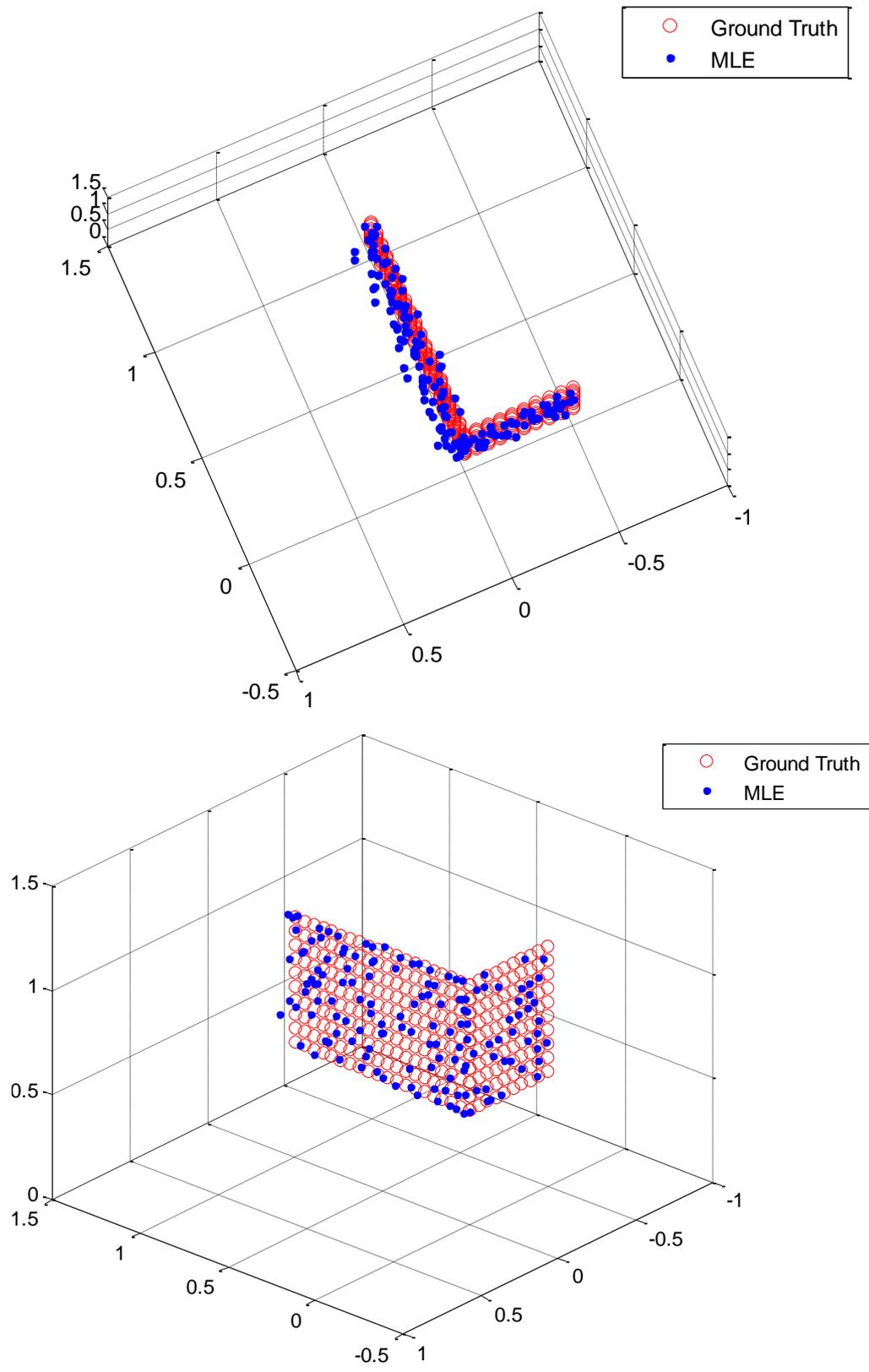


Figure 6.6: MLE with clustering.

Measurement noise standard deviations of the optical camera and sonar are set to 1cm. Reconstructed structure using the initial estimates of \mathbf{R} and \mathbf{t}_o is shown in Figure 6.4. Since the initial estimate is based on a single sonar match and outliers are also used in computations, the reconstructed structure is noisy. On the other hand, it is clear from Figure 6.4 that reconstruction accuracy of the stereo optical system is better than that of stereo sonar system for the same noise level.

Estimation of the motion parameters assuming the scene as a single planar surface still yields relatively good result as shown in Figure 6.5. While not valid for this scene, planar scene assumption removes some of the outliers (20 percentage). As a result, the scale factor approaches to its optimal value. After obtaining the improved scale factor, the structure is reconstructed using all available samples to compare with initial estimation.

Finally, noisy 3D data is clustered into two sets to determine the dominant plane in the scene. Using these clusters, the perpendicular planes in the scene can be represented in a more realistic way. Hence, RANSAC performs better in outlier removal. Approximately 50 percentage of the all available matches was eliminated by RANSAC in the final (8th) iteration and accuracy of the motion estimation is considerably increased as shown in Figure 6.6.

CHAPTER 7

UNDERWATER OPTICAL IMAGE ENHANCEMENT

Visibility in underwater is limited due to the nature of light propagation in water. Causes of the image degradation varies spatially as mentioned in Chapter 3. To improve visibility in underwater, methods such as histogram equalization and high pass filtering are not useful since these are spatially invariant [24].

2D sonars gives more clear images even in turbid waters. Hence, degraded optical images can be improved via integration of information coming from acoustic sensor. In this chapter, a method is proposed to realize and verify this idea which is based on mapping of features in the sonar image to the optical image via a nonlinear transformation. The transformation is obtained using radial basis function (RBF) neural network. Before giving the details, underlying reasons to select such a method will be presented.

7.1 Relationship between Matching Optical and Sonar Image Points

Let $\mathbf{P}_o = [X_o, Y_o, Z_o]^T$ and $\mathbf{P}_s = [X_s, Y_s, Z_s]^T$ denote a 3D point $\mathbf{P} = [X, Y, Z]$ in optical and sonar coordinate frames, respectively. Taking the optical frame as the world frame, coordinates of the point \mathbf{P} in the optical image is

$$x_o = f \frac{X}{Z} \text{ and } y_o = f \frac{Y}{Z}. \quad (7.1)$$

and these coordinate systems are related as

$$\mathbf{P}_s = \mathbf{R}\mathbf{P}_o + \mathbf{t}. \quad (7.2)$$

Coordinates of point \mathbf{P} in sonar image is

$$\begin{aligned} \theta &= \tan^{-1}(X_s/Y_s) \text{ and } \mathfrak{R} = \sqrt{X_s^2 + Y_s^2 + Z_s^2} \\ x_s &= \mathfrak{R}\sin\theta \text{ and } y_s = \mathfrak{R}\cos\theta. \end{aligned} \quad (7.3)$$

Using (7.1), (7.2) and (7.3)

$$\begin{aligned} x_s &= \sqrt{\sum_{i=1}^3 \mathbf{r}_i \cdot \mathbf{P} + t_i} \sin\left(\tan^{-1}\left(\frac{\mathbf{r}_1 \cdot \mathbf{P} + t_x}{\mathbf{r}_2 \cdot \mathbf{P} + t_y}\right)\right) \\ y_s &= \sqrt{\sum_{i=1}^3 \mathbf{r}_i \cdot \mathbf{P} + t_i} \cos\left(\tan^{-1}\left(\frac{\mathbf{r}_1 \cdot \mathbf{P} + t_x}{\mathbf{r}_2 \cdot \mathbf{P} + t_y}\right)\right) \end{aligned} \quad (7.4)$$

where \mathbf{r}_i and t_i denotes the i th rows of rotation and translation matrix, respectively and

$$\mathbf{P} = \begin{bmatrix} x_o Z/f \\ y_o Z/f \\ Z \end{bmatrix}. \quad (7.5)$$

From (7.4), it is clear that the relation between pairs (x_o, y_o) and (x_s, y_s) is nonlinear. Furthermore, due to implicit dependency to unknown depth Z , the mapping varies from point to point. Hence, finding a global mapping as in the optical stereo pairs, also known as homography, is not possible.

It was shown in Chapter 5 that given a match (x_o, y_o) and (x_s, y_s) , the unknown depth Z can be recovered by finding and optimizing the closed form solution. On the other hand, there are a limited number of matches in degraded optical and sonar images. However, a nonlinear transformation can be found using all available matches. Mapping the features from sonar image to optical image where the value of transformation is unknown can be done by interpolation. Using RBF neural network is a clear way of finding such a transformation.

7.2 RBF Neural Networks

RBF network is an artificial neural network that uses RBFs as activation functions. These networks have typically three layers as shown in Figure 7.1. Hidden and output layers implement a set of nonlinear basis functions and linear summation, respectively [38].

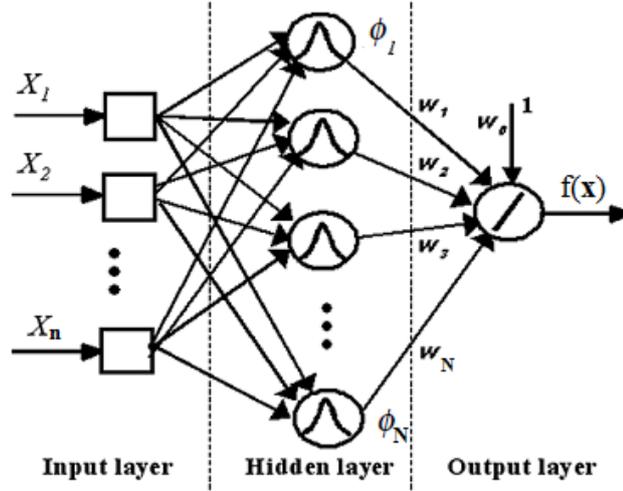


Figure 7.1: RBF network architecture.

The output $f(\mathbf{x}): \mathbb{R}^n \rightarrow \mathbb{R}$ of the network is

$$f(\mathbf{x}) = \sum_{i=1}^N w_i \phi(\|\mathbf{x} - \mathbf{c}_i\|) \quad (7.6)$$

where N is the number of neurons in hidden layer, \mathbf{c}_i is center of i th neuron, w_i are the linear weights and $\phi(\cdot)$ are the basis functions. In general, $\phi(\cdot)$ taken to be Gaussian

$$\phi(\|\mathbf{x} - \mathbf{c}_i\|) = e^{-\frac{\|\mathbf{x} - \mathbf{c}_i\|^2}{2\sigma_i^2}} \quad (7.7)$$

where σ_i is the variance of the i th neuron. The unknowns w_i, \mathbf{c}_i and σ_i are determined so that the fit between input and output data is optimum. In our problem, these unknowns are found using the available optical and sonar matches where the input and output of transformation are image coordinates of sonar and optical correspondences, respectively.

7.3 Mapping Performance

Multiple returns to sonar receiver and its effect on the produced image was mentioned in section 2.6.2.1. Hence, it is not surprising that the performance of the mapping depends on the scene geometry. In the following sub-sections, applicability of proposed method is tested for different scene geometries in terms of 2D sonar imaging.

7.3.1 Planar Scene

First, the mapping is evaluated on a scene with a rectangular grid lying on the ground. Optical image and sonar image of the grid and the selected training points (corners and midpoints of the edges) for the mapping is given in Figure 7.2. Using these training points, the mapping is determined with RBF network so that the transformation is perfect in the matching of sonar and optical pairs.

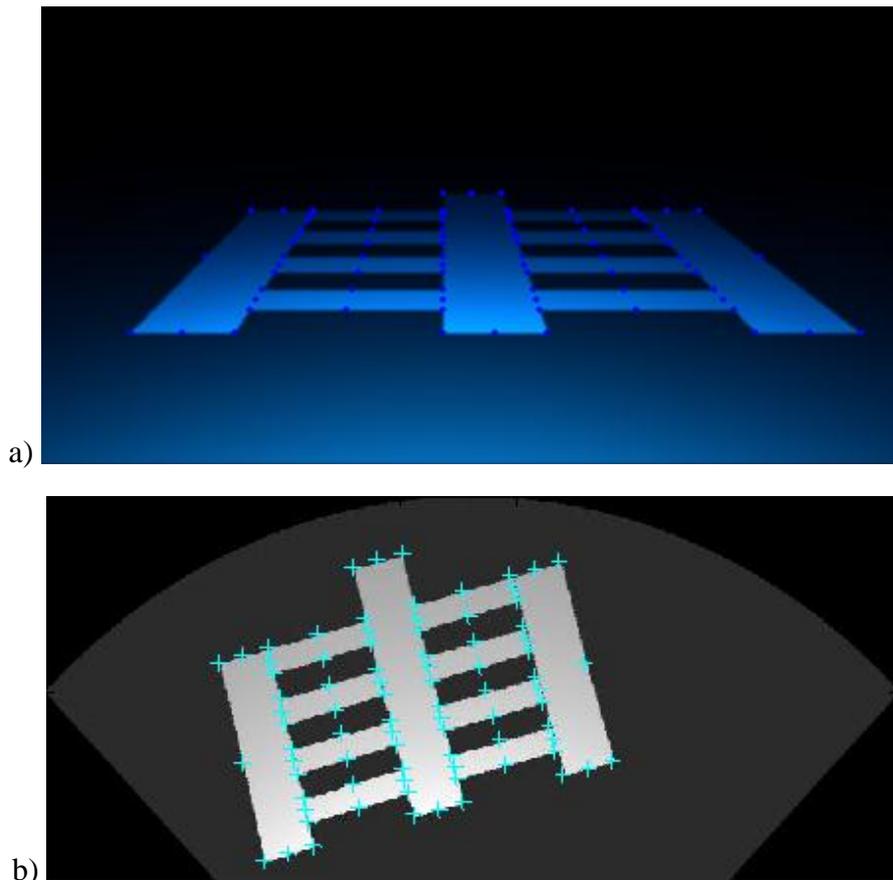


Figure 7.2: Training points in optical a) and sonar b) image of the rectangular grid.

For testing, points with a distance $1/3$ of the edges to corners are used. Result of the mapping is shown in Figure 7.3 where perfect fit of the transformation is clearly seen.

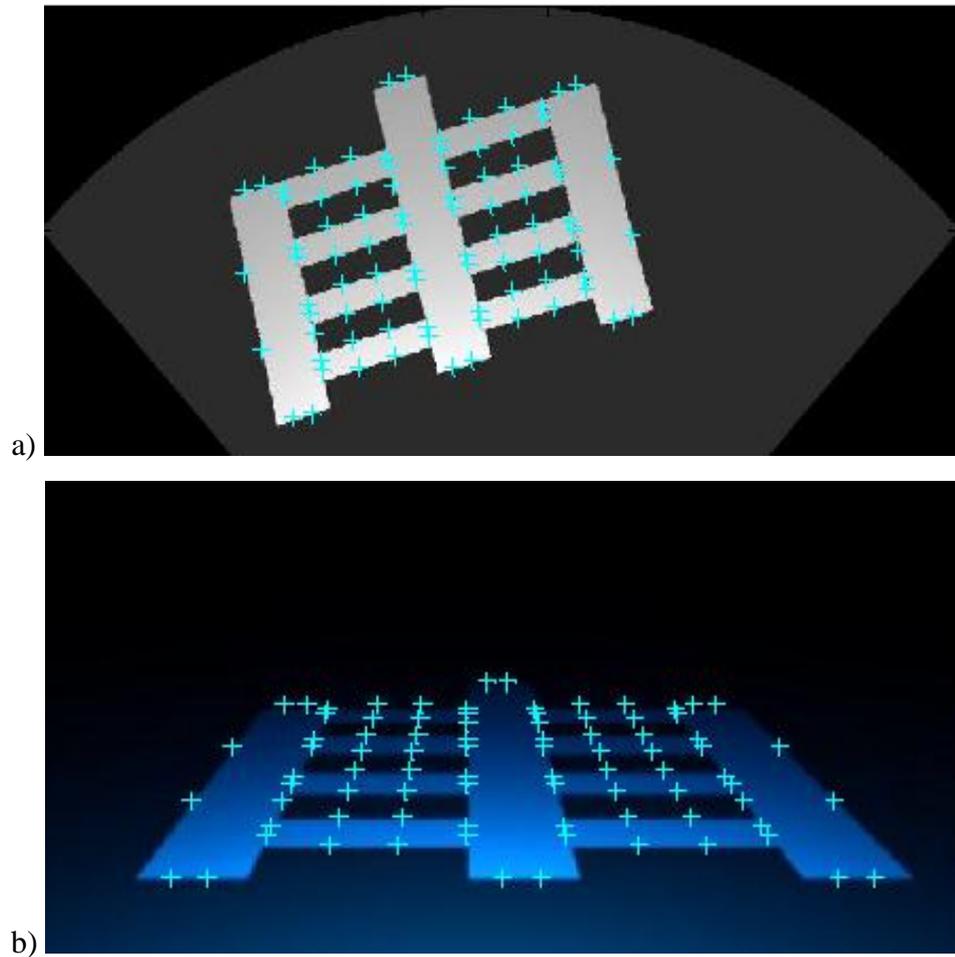


Figure 7.3: Test points on sonar image a), result of the sonar to optical mapping b).

7.3.2 Non-Planar Scene

The performance of the mapping is evaluated on a scene with cuboids. Optical and sonar image of the scene as well as the selected training points (corners and midpoints of the edges) for the mapping is given in Figure 7.4 a) and b). First, using these training points, a mapping is determined so that that the transformation perfectly fits in the process of matching sonar and optical pairs as in the grid example. Transformation of the sonar *training points* to optical image is shown in Figure 7.4 c).

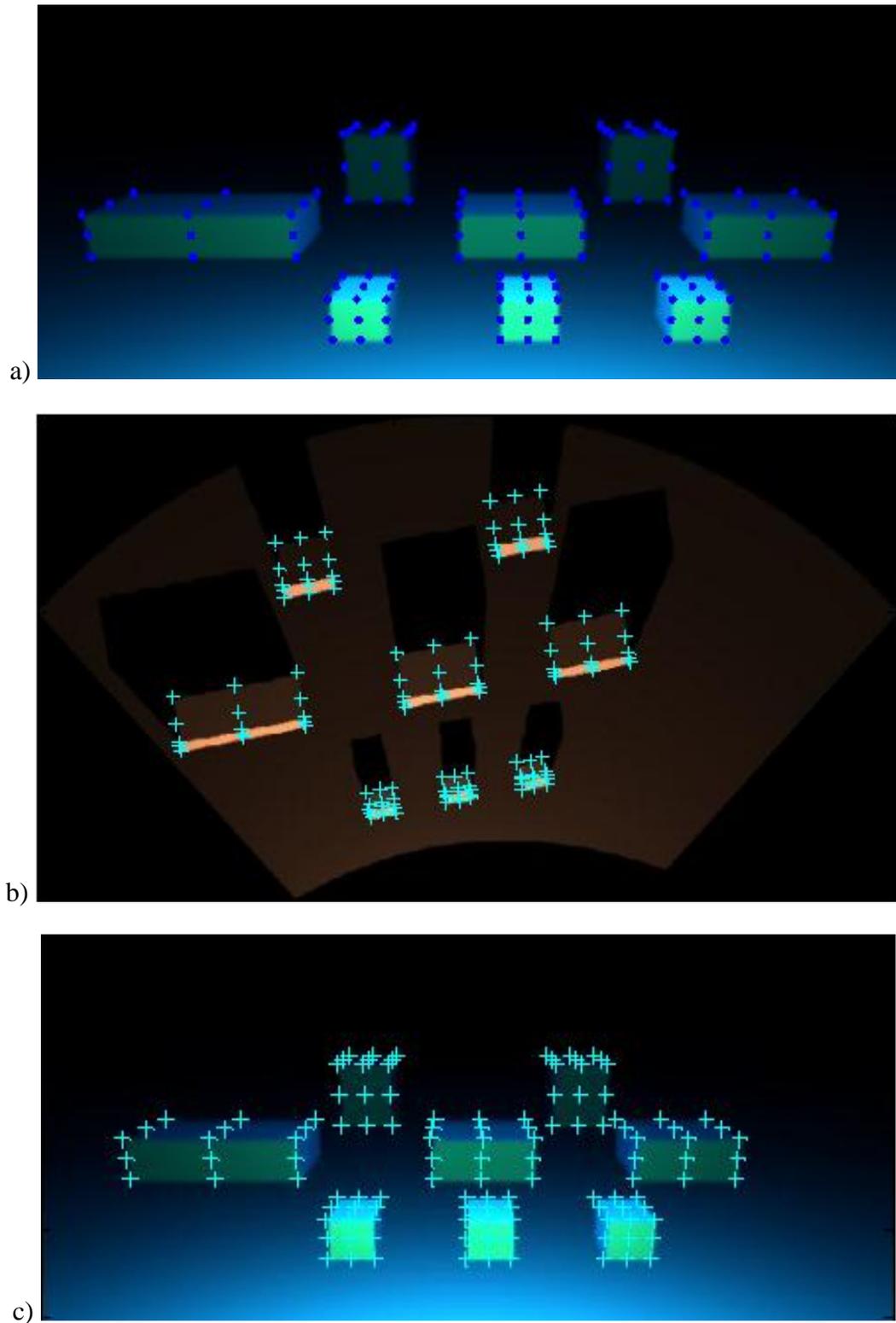


Figure 7.4: Training points in optical a) and sonar b) image of cuboids, result of the sonar to optical mapping of training points.

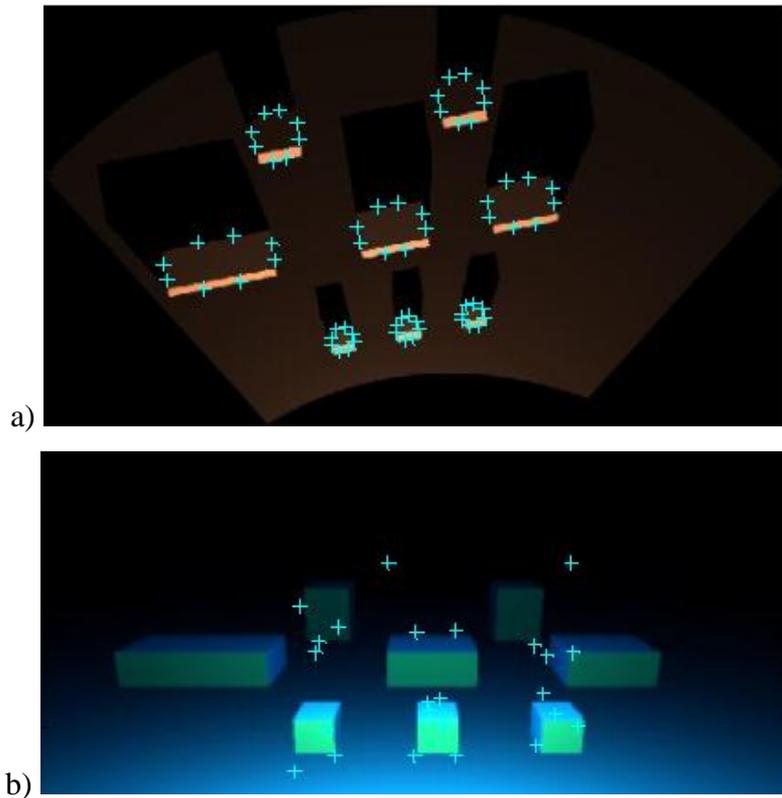


Figure 7.5: Test points on sonar image a), result of the sonar to optical mapping b).

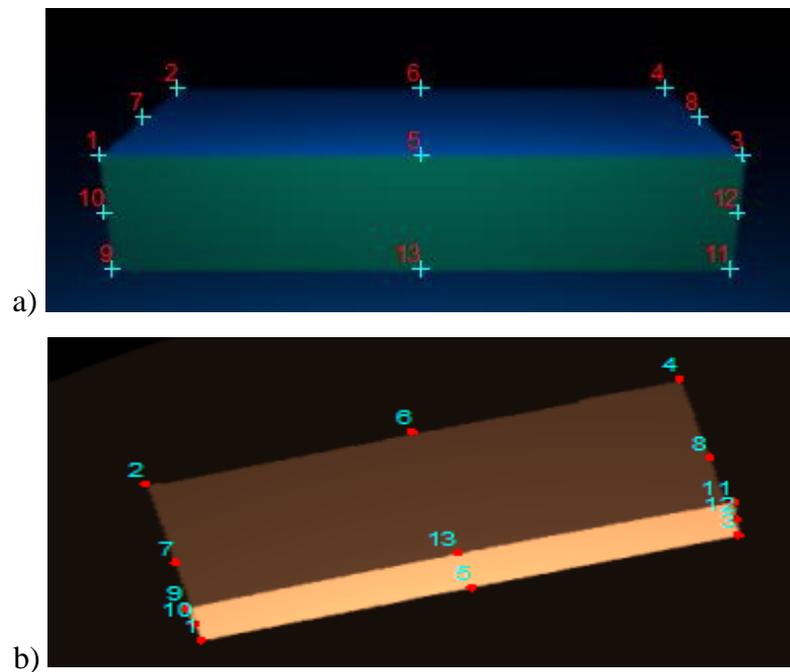


Figure 7.6: Matching feature points of a cuboid.

For testing, points on top face of the cuboids with a distance $1/3$ of the edges to corners are used. Result of the mapping is shown in Figure 7.5 where none of the test points in sonar image transformed to their correct locations in the optical image. Considering the similar optical and sonar image of the cuboid in Figure 7.4, this result may look erroneous since the mapping interpolates the unknown points. To comprehend these results, please look at the Figure 7.6 where the matching features in sonar and optical images are given. Many of the matching points are located as expected but feature points 1, 3, 5, 9, 11, 13 are not. Due to the degenerate scene geometry in terms of 2D sonar imaging, obtaining the missing points by interpolation is not possible. However, using a mapping where the unknown w_i , c_i and σ_i are determined so that the mapping does not perfectly fits the matching pairs, i.e., MSE of the mapping is not zero, these results may be improved. Result of using such a transformation is shown in Figure 7.7. As seen in Figure 7.7 a), modified transformation tends to map the training points to a dominant plane. As a result, test points, approximately located on the same dominant plane, are mapped to more realistic locations as shown in Figure 7.7 b).

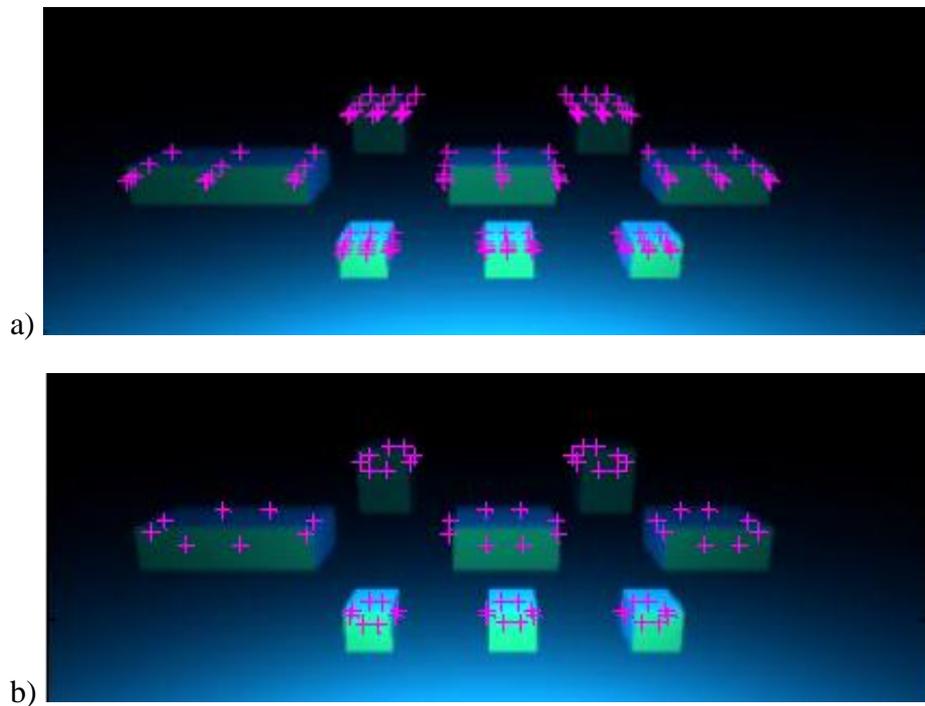


Figure 7.7: Transformation of training a) and test b) points to optical image with a modified mapping.

7.4 Results

In this section, performance of the proposed method in the recovery of scene points in degraded optical images is evaluated.

7.4.1 Planar Scene

A rectangular grid is used for planar scene test as in section 7.3.1. Settlement of the grid in the scene, its sonar image and degraded optical image are given in Figure 7.8. In the optical image, only the close points of the grid to view point can be perceived whereas the grid can be completely identified in the sonar image.

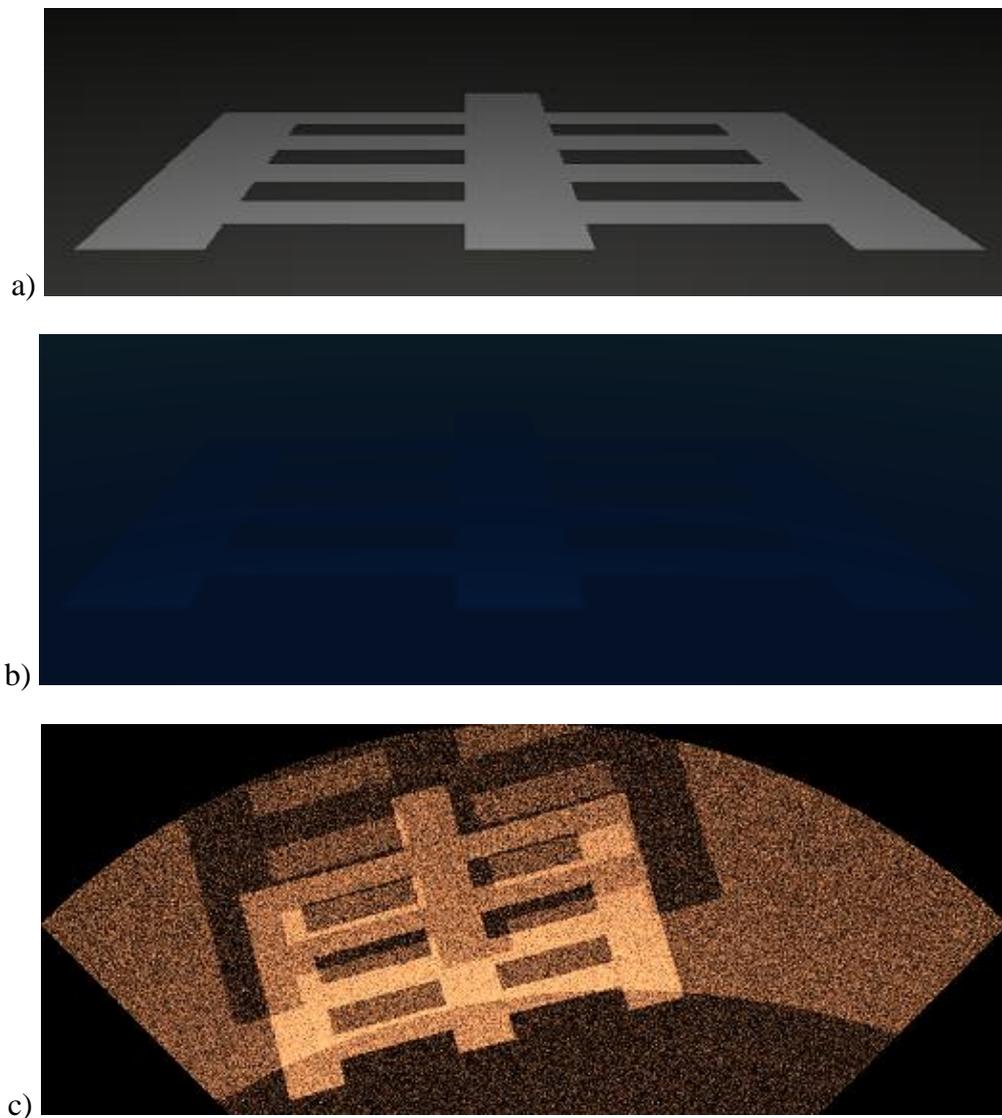


Figure 7.8: Rectangular grid a), its optical b) and sonar image c).

To determine mapping, visible feature points in the optical image and corresponding points in the sonar image are used as shown in Figure 7.9 a) and b). Selected test points, scattered to the entire grid, are shown in Figure 7.9 c).

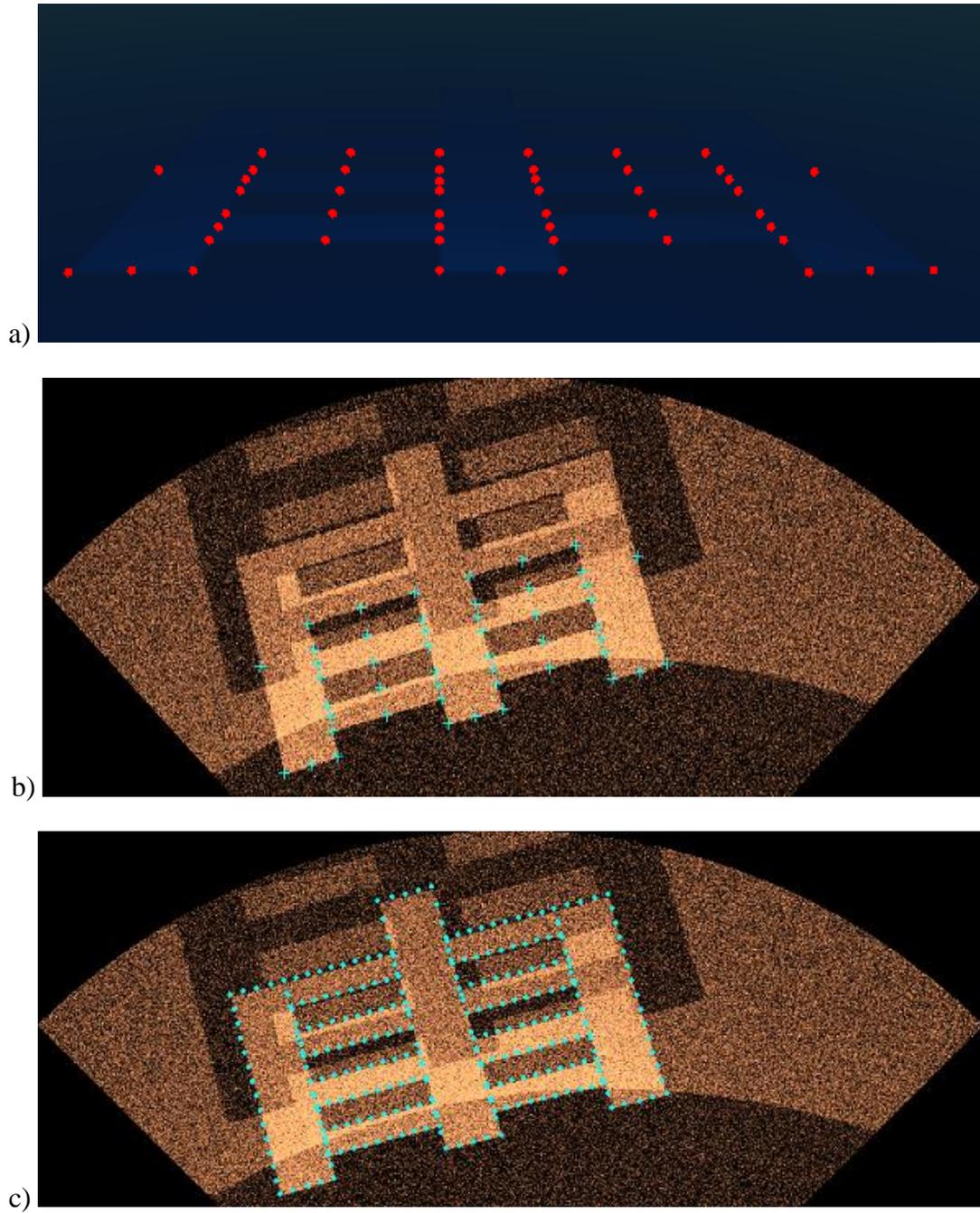


Figure 7.9: Training points on a) optical image, b) sonar image and c) selected test points on sonar image

The result of the mapping is shown for variances $\sigma_i = 1000$, 500 and 200 in Figure 7.10 a), b) and c), respectively. Even though MSE of all the mappings are zero, different results are obtained. Still, all the mappings considerably improves the degraded optical image.

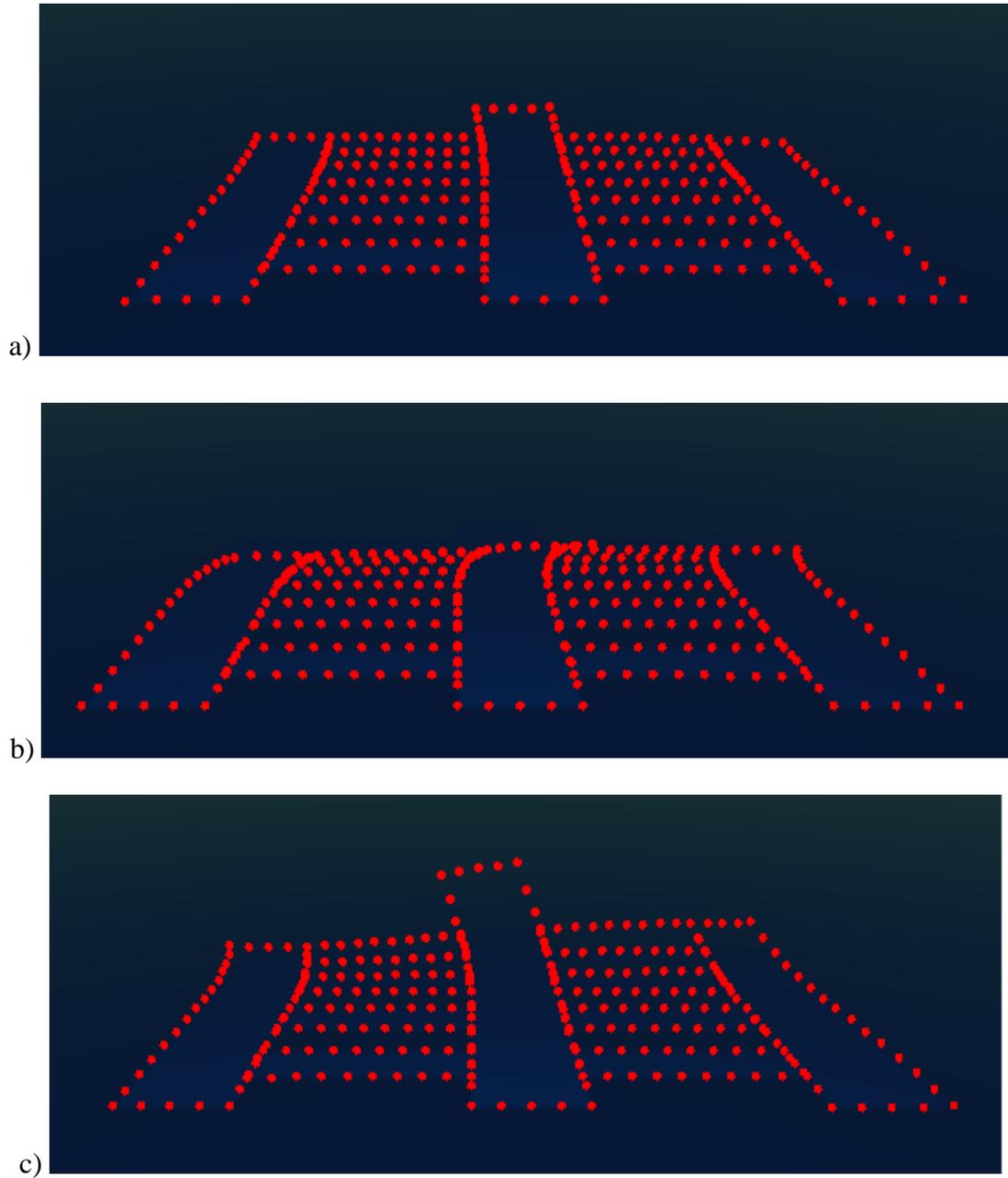


Figure 7.10: Result of the sonar to optical mapping for various mappings.

CHAPTER 8

CONCLUSIONS

Opti-acoustic stereo imaging is quite a new paradigm in computer vision area. In this thesis, imaging capabilities of opti-acoustic systems is investigated.

First of all, a 2D high frequency forward-looking sonar simulator is developed which permits the visualization of the received acoustic data in image form. The method is based on the optical ray tracing [4] combined with the frequency domain approach [5], but several fundamental modifications were required to produce the 2D sonar image. These modifications are generally associated with time-based display problems and solutions to these problems have been clarified in detail.

Determination of the bright and shadow areas were explained and a computationally efficient algorithm is proposed to process the multiple returns to sonar receiver. The method is based on storing the assigned number of the intersected surface and processing the returned signals for different surfaces separately considering their acoustic scatter profiles. This operation removes the need of using too many optical rays and reduces the computation time. Images produced with this method are sufficiently realistic and the simulator can be used in real time applications by preprocessing the scene in a different compiling environment.

Secondly, an underwater optical image simulator using physical models for light, attenuation and scattering is developed. In order to achieve more realistic images, the change in colors is simulated by the Jaffe-McGlamery model [9, 10] where the total attenuation is calculated considering the wavelength of the colors. Forward scatter,

namely image blur, was computed by the convolution of the attenuated signal with a Gaussian kernel based distance dependent low-pass filter. Finally, backscatter, the dominant degradation component, was computed by slicing the 3D space into parallel planes. Then, the underwater image was formed by the summation of degraded signal, forward scatter and backscatter which gives satisfactory results. However, the illumination model can be extended for materials other than having diffuse reflection and floating particles can also be added for more realistic simulations.

An effective general method isn't proposed for automatic matching of opti-acoustic features. Hence, assuming a calibrated system, epipolar geometry of the opti-acoustical stereo imaging was investigated on simulated sonar and optical images to see the geometrical relationship between the corresponding points in these images. Given a point in one of the images, the region where the matching point in other image was constricted to a 1D curve thanks to establishing epipolar geometry. Also, it was observed in opti-acoustical system that, epipolar curves in the optical image aren't lines, but conic sections unlike a traditional stereo optical system which is the result of asymmetrical projection models in sonar and optical images [11]. Considering the vertical beam width of the sonar pulses, epipolar curves can be further constrained which is an advantage of opti-acoustic system over stereo optical system.

After the geometrical background was examined, 3D reconstruction from opti-acoustic images was investigated by manual feature matching. Given an opti-acoustic match, in the first stage, corresponding 3D point was computed from closed form (range and azimuth) solutions. However, an optimal solution in the maximum likelihood sense was derived from a nonlinear method which requires a good initial estimate to improve the convergence rate [12]. Hence, range solution was used in the optimization procedure which provides a better initial condition compared to the azimuth solution. In the next stage, range solution was optimized with direct and indirect methods where the performance of the direct method is better than the indirect method.

Using the same stereo configuration, these results were compared to that of the traditional stereo optical system. It was observed that stereo optical and opti-acoustic systems show similar performance at short target ranges and relatively large baseline. However, opti-acoustic system performed better than stereo optical system for increased target ranges and smaller baseline. Although the azimuth solution in opti-acoustic system suffers from the same disadvantage, range and azimuth solutions may be fused optimally according to target range and stereo baseline to improve the 3D reconstruction performance.

In the next step, motion estimation from opti-acoustic system is examined. Different from 3D reconstruction, the estimation is based on optical and sonar matches in two consecutive views which doesn't require the knowledge of opti-acoustic correspondences. First, the homography was determined from minimum of four optical matches of non-collinear coplanar points which then decomposed into the correct and dual estimates of rotation and up-to-scale translation and plane normal. Dual estimate of the motion was eliminated by means of the sonar motion transformation. Then, using a single match in sonar measurements, the scale factor was recovered. To obtain MLE estimates of motion parameters, this initial estimate was used in a nonlinear optimization method that utilizes all available optical and sonar matches [13]. The weight in MLE formulation can be selected based on the feature localization accuracy of the optical and acoustic systems to improve estimation accuracy in various environmental conditions. Furthermore, the integration of motion and epipolar geometry may be used to explore a geometric approach to solve the opti-acoustic stereo correspondence problem.

Finally, improving the degraded underwater optical image via sonar image was considered. First, using all available matches, a nonlinear transformation which maps the features in the sonar image to the optical image is found using an RBF neural network. Then, selected test points in the sonar image is transformed to the optical image and performance of the mapping is evaluated for different scene geometries in terms of sonar imaging.

The result of the mapping is satisfactory for planar scenes. Even though, results change slightly for different variances, degraded optical images can be improved more or less and the method can be extended for non planar scenes by transforming the features that are found on the locally planar surfaces.

Some of the topics that remain as future work are noted below:

- Calibration of the opti-acoustic system.
- Removal of the radial distortion in sonar images.
- Modification of the proposed image enhancement method for non-planar scenes.
- Improving the sonar-to-optical imagery so that the overall optical image enhancement system will be used on real time.

APPENDIX A

The points on a plane satisfy

$$f/Z_o = -(\mathbf{n}_o \cdot \mathbf{p}_o) \quad (\text{A.1})$$

where $\mathbf{n}_o = (n_{ox}, n_{oy}, n_{oz})^T$ is the inward surface normal in the optical camera system and Z_o is the depth of the point assuming Z axis is the principal axis. The surface normal of the same point $\mathbf{n}_s = (n_{sx}, n_{sy}, n_{sz})^T$ in the sonar coordinate frame is given by (6.7). To find the relative pose of the camera and sonar, orthogonality of the \mathbf{R} has to be enforced. Decomposing the rotation into three axes gives

$$\mathbf{R}(\alpha_x, \alpha_y, \alpha_z) = \mathbf{R}_z(\alpha_z)\mathbf{R}_y(\alpha_y)\mathbf{R}_x(\alpha_x) \quad (\text{A.2})$$

where $\mathbf{R}_u(\alpha_u)$ denotes a α_u degree rotation about axis u. Using (7.4), each match provides two constraints in terms of rotation, translation and normal \mathbf{n}_o of the calibration plane. Then, the unknown pose parameters can be found by minimizing the reprojection errors of reconstructed planar grid points [1].

Let $\hat{\mathbf{M}} = [\hat{\mathbf{R}}, \hat{\mathbf{t}}]$ and $\hat{\mathbf{n}}_o$ be the initial estimates of unknown parameters. Using the optical coordinates $\hat{\mathbf{p}}_o$ of a feature point, the depth of the corresponding point can be estimated using (A.1). Using (5.2), we have an initial estimate for \mathbf{P}_o . 3D coordinates of the corresponding point $\hat{\mathbf{P}}_s$ in sonar frame is obtained by using this initial estimate in (6.3). The elevation angle of the sonar match $\hat{\mathbf{p}}_s$ can be calculated with (6.16) and 3D coordinates of the corresponding point \mathbf{P}_s is obtained from (6.1). Transforming \mathbf{P}_s into optical frame yields $\hat{\mathbf{P}}_o$. The optimal pose parameters is found by minimizing

$$f(\mathbf{R}, \mathbf{t}, \mathbf{n}) = \sum_N (\mathbf{P}_o - \hat{\mathbf{P}}_o)^T \Sigma_{p_o}^{-1} (\mathbf{P}_o - \hat{\mathbf{P}}_o) + (\mathbf{P}_s - \hat{\mathbf{P}}_s)^T \Sigma_{p_s}^{-1} (\mathbf{P}_s - \hat{\mathbf{P}}_s) \quad (\text{A.3})$$

where the covariances can be estimated analytically based on first-order approximation. This nonlinear optimization problem is solved by Levenberg-Marquardt algorithm [30].

APPENDIX B

Let (\mathfrak{R}, θ) and (\mathfrak{R}', θ') be the measurements of a 3D point \mathbf{P} in two sonar views. Each of these 4 observations gives a constraint on the 3 unknowns in \mathbf{P} . Using three out of four constraints, various linear closed form solutions can be obtained. While $\|\mathbf{P}\| = \mathfrak{R}$ is a nonlinear equation, the 3 measurements $(\theta, \mathfrak{R}', \theta')$ provides [39]

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \mathbf{P} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \quad (\text{B.1})$$

where

$$\begin{aligned} a_1 &= [-1, \tan\theta, 0] & b_1 &= 0 \\ a_2 &= (\tan\theta')\mathbf{r}_2 - \mathbf{r}_1 & b_2 &= t_x - (\tan\theta')t_y \\ a_3 &= \mathbf{t}^T \mathbf{R} & b_3 &= (\mathfrak{R}'^2 - \mathfrak{R}^2 - \|\mathbf{t}\|^2/2). \end{aligned} \quad (\text{B.2})$$

Discarding the range measurements in first view yields,

$$\mathbf{P}_o = \mathbf{A}_o^{-1} \mathbf{b}_o; \quad \mathbf{A}_o = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \quad \mathbf{b}_o = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}. \quad (\text{B.3})$$

On the other hand, using the fourth constraint gives

$$\mathfrak{R}^2 = \mathbf{P}^T \mathbf{P} \approx (\mathbf{A}_o^{-1} \mathbf{b}_o)^T \mathbf{P}. \quad (\text{B.4})$$

Using this approximation we can write

$$\widetilde{\mathbf{P}}_o = \widetilde{\mathbf{A}}_o^{-1} \widetilde{\mathbf{b}}_o; \quad \widetilde{\mathbf{A}}_o = \begin{bmatrix} \mathbf{A}_o \\ (\mathbf{A}_o^{-1} \mathbf{b}_o)^T \end{bmatrix} \quad \widetilde{\mathbf{b}}_o = \begin{bmatrix} \mathbf{b}_o \\ \mathfrak{R}^2 \end{bmatrix}. \quad (\text{B.5})$$

Fixing length equal to the range yields

$$\mathfrak{R} = \widetilde{\mathbf{P}}_o^T \mathbf{P} \approx \frac{(\mathbf{A}_o^{-1} \mathbf{b}_o)^T}{\|(\mathbf{A}_o^{-1} \mathbf{b}_o)^T\|} \mathbf{P}. \quad (\text{B.6})$$

where $\widehat{\mathbf{P}}_o^T = \mathbf{P}_o / \|\mathbf{P}_o\|$. Hence, the improved estimate is obtained as [39]

$$\widetilde{\mathbf{P}}_o' = \widetilde{\mathbf{A}}_o'^{-1} \widetilde{\mathbf{b}}_o'; \quad \widetilde{\mathbf{A}}_o' = \begin{bmatrix} \mathbf{A}_o \\ \widehat{\mathbf{P}}_o^T \end{bmatrix} \quad \widetilde{\mathbf{b}}_o' = \begin{bmatrix} \mathbf{b}_o \\ \mathfrak{R} \end{bmatrix}. \quad (\text{B.7})$$

BIBLIOGRAPHY

- [1] S. Negahdaripour, H. Sekkati and H. Pirsiavash, "Opti-Acoustic Stereo Imaging: On System Calibration and 3-D Target Reconstruction," *IEEE Transactions on Image Processing*, vol. 18, no. 6, pp. 1203-1214, 2009.
- [2] BlueView, [Online]. Available: <http://www.blueview.com/P900-Series.html>. [Accessed 5 August 2012].
- [3] Sound Metrics, [Online]. Available: <http://www.soundmetrics.com/products/didson-sonars>. [Accessed 12 08 2012].
- [4] J. M. Bell, "Application of optical ray tracing techniques to the simulation of sonar images," *Optical Engineering*, vol. 36, no. 6, p. 1806–1813, 1997.
- [5] S. R. Silva, "Simulation and 3D Reconstruction of Side-looking Sonar Images," in *Advances in Sonar Technology*, In-teh, 2009, pp. 1-15.
- [6] A. P. Lyons and D. A. Abraham, "Statistical characterization of high-frequency shallow-water seafloor backscatter," *The Journal of the Acoustical Society of America*, vol. 106, no. 3, pp. 1307-1315, 1999.
- [7] S. Metari and F. Deschênes, "A novel polychromatic model for light dispersion," in *4th international conference on Image and signal processing*, Montreal, 2010.
- [8] A. Sedlazeck. and R. Koch, "Simulating Deep Sea Underwater Images Using Physical Models for Light Attenuation, Scattering, and Refraction," in *Vision, Modeling, and Visualization*, Berlin, 2011.
- [9] J. Jaffe, "Computer modeling and the design of optimal underwater imaging systems," *IEEE Journal of Oceanic Engineering*, vol. 15, no. 2, pp. 101-111, 1990.
- [10] B. L. McGlamery, "Computer Analysis and Simulation of Underwater Camera System Performance," San Diego, 1975.

- [11] S. Negahdaripour, "Epipolar Geometry of Opti-Acoustic Stereo Imaging," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 10, pp. 1776-1788, 2007.
- [12] H. Sekkati and S. Negahdaripour, "Direct and Indirect 3-D Reconstruction from Opti-Acoustic Stereo Imaging," in *3D Data Processing, Visualization, and Transmission*, Carolina, 2006.
- [13] S. Negahdaripour and A. Taatian, "3-D motion estimation by integrating visual cues in 2-D multi-modal opti-acoustic stereo sequences," *Computer Vision and Image Understanding*, vol. 114, no. 8, pp. 928-941, 2010.
- [14] A. D. Waite, *Sonar for Practising Engineers*, West Sussex: Wiley, 2002.
- [15] D. S. Drumheller, *Introduction to Wave Propagation in Nonlinear Fluids and Solids*, Cambridge University Press, 1998.
- [16] J. M. Bell and L. M. Linnet, "Simulation and Analysis of Synthetic Sidescan Sonar," *IEEE Proceedings on Radar, Sonar and Navigation*, vol. 144, no. 4, pp. 219-226, 1997.
- [17] NATO Underwater Research Center, [Online]. Available: <http://www.cmre.nato.int/index.php>. [Accessed 5 August 2012].
- [18] C. Clay and H. Medwin, *Acoustical Oceanography: Principles and Applications*, Wiley, 1977.
- [19] I. Karasalo and P. Skogqvist, "Identification of Object Parameters from Transient Bistatic Scattering Data," in *Proceedings of the Boundary Influences in High Frequency Shallow Water Acoustics Conference*, Bath, 2005.
- [20] W. Neubauer, *Acoustic Reflection from Surfaces and Shapes*, Washington, D.C: Naval Research Laboratory, 1986.
- [21] R. Zhang, P. Tsai, J. Cryer and M. Shah, "Shape from Shading: A Survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 8, pp. 690-706, 1999.
- [22] C. D. Loggins, "A comparison of forward-looking sonar design alternatives," in *IEEE Oceans*, Santa Barbara, 2001.

- [23] R. P. Hodges, *Underwater Acoustics*, West Sussex: Wiley, 2010.
- [24] Y. Y. Schechner and N. Karpel, "Recovery of Underwater Visibility and Structure by Polarization Analysis," *IEEE Journal of Oceanic Engineering*, vol. 30, no. 3, pp. 570-587, 2005.
- [25] T. Treibitz and Y. Y. Schechner, "Instant 3Descatter," in *IEEE Computer Vision and Pattern Recognition*, New York, 2006.
- [26] "Wet Lab," [Online]. Available: <http://www.wetlabs.com/iopdescript/Table%201.pdf>. [Accessed 10 August 2012].
- [27] S. Metari and F. Deschênes, "A New Convolution Kernel for Atmospheric Point Spread Function," in *11 th International Conference on Computer Vision*, Rio De Janerio, 2007.
- [28] D. Mobley, *Light and Water: Radiative Transfer in Natural Waters*, Academic Press, 1994.
- [29] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed., Cambridge: Cambridge University Press, 2004.
- [30] D. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," *Society for Industrial and Applied Mathematics*, vol. 11, no. 2, 1963.
- [31] H. Longuet-Higgins, "The visual ambiguity of a moving plane," *Proceedings of the Royal Society of London. Series B*, vol. 223, pp. 165-175, 1984.
- [32] D. Nister, "An efficient solution to the five-point relative pose problem," *Pattern Analysis and Machine Intelligence*, vol. 26, no. 6, pp. 756 - 770 , 2004.
- [33] E. Malis and M. Vargas, "Deeper understanding of the homography decomposition," INRIA, Sevilla, 2007.
- [34] R. Hartley, "In defence of the 8-point algorithm," in *ICCV*, 1995.
- [35] M. Fischler and R. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 1, pp. 220-226, 1981.

- [36] S. Negahdaripour and A. Taatian, "3-D motion and structure estimation for arbitrary scenes from 2-D optical and sonar video," in *IEEE Oceans*, Quebec, 2008.
- [37] P. D. Kovesi, "MATLAB and Octave Functions for Computer Vision and Image Processing," Centre for Exploration Targeting, [Online]. Available: <http://www.csse.uwa.edu.au/~pk/Research/MatlabFns/index.html>. [Accessed 22 08 2012].
- [38] C. Bishop, *Neural Networks for Pattern Recognition*, Oxford: Clarendon Press, 1995.
- [39] S. Negahdaripour, "On 3-D reconstruction from stereo FS sonar imaging," in *IEEE Oceans*, Seattle, 2010.