

OBJECT TRACKING WITH RANGE GATED CAMERA SYSTEM

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF INFORMATICS
OF
THE MIDDLE EAST TECHNICAL UNIVERSITY



BY

MUSTAFA YAĞCIOĞLU

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE
OF
DOCTOR OF PHILOSOPHY
IN
THE DEPARTMENT OF INFORMATION SYSTEMS

SEPTEMBER 2015

OBJECT TRACKING WITH RANGE GATED CAMERA SYSTEM

Submitted by **Mustafa Yağcıođlu** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy in Information Systems, Middle East Technical University** by,

Prof. Dr. Nazife Baykal
Director, **Informatics Institute**

Prof. Dr. Yasemin Yardımcı Çetin
Head of Department, **Information Systems**

Assoc. Prof. Dr. Alptekin Temizel
Supervisor, **Modeling and Simulation, METU**

Dr. Kuthan Yelen
Co-Supervisor, **Laser Systems Design Department, ASELSAN A.Ş.**

Examining Committee Members:

Prof. Dr. Yasemin Yardımcı Çetin
Information Systems, METU

Assoc. Prof. Dr. Alptekin Temizel
Modeling and Simulation, METU

Assist. Prof. Dr. P. Erhan Eren
Information Systems, METU

Assist. Prof. Dr. Sinan Kalkan
Computer Engineering, METU

Assist. Prof. Dr. Behçet Uđur Töreyin
Informatics Institute,
Istanbul Technical University

Date: 07.09.2015

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: Mustafa Yağcıođlu

Signature : _____

ABSTRACT

OBJECT TRACKING WITH RANGE GATED CAMERA SYSTEM

YAĞCIOĞLU, Mustafa

Ph.D., Department of Information Systems

Supervisor: Assoc. Prof. Alptekin TEMİZEL

Co-Supervisor: Dr. Kuthan YELEN

September 2015, 113 pages

Imaging systems can be categorized as active or passive. Passive systems act only as a receiver and do not send any energy to their environment. Frequently used night vision systems such as thermal imagers and image intensifier systems are examples of passive systems. Active imaging systems on the other hand, send energy to their environment and detect the reflected energy from the objects. While passive systems require some ambient light or object radiation, active imaging systems can work without any external light source. Range-gated imaging systems are active systems that use a high-power pulsed-light source and control the opening and closing times of the camera shutter in conjunction with the light source. By calculating the arrival time of the reflected light from the object, the camera shutter is opened for a short time period to form an image using the returned light. This allows generating high

contrast images of the objects in difficult lighting conditions. To track an object with a range-gated system, beside the horizontal and vertical position of the object, the distance of the object should also be tracked. The distance information is used to arrange shutter timing of the camera. In this thesis, a range-gated camera system test bed was constructed and a robust object tracking algorithm integrated with this system was developed. The distance of the object is calculated by the proposed algorithm with various image comparison methods and the performances of comparison methods are compared.

Keywords: Range Gated Imaging, Object Tracking

ÖZ

MESAFE KAPILI KAMERA SİSTEMİ İLE NESNE TAKİBİ

YAĞCIOĞLU, Mustafa

Ph.D. Bilişim Sistemleri Bölümü

Tez Yöneticisi: Doç. Dr. Alptekin TEMİZEL

Ortak Tez Yöneticisi: Dr. Kuthan YELEN

Eylül 2015, 113 sayfa

Görüntüleme sistemleri, “aktif” ve “pasif” sistemler olarak ikiye ayrılabilir. Pasif sistemler, sadece alıcı olarak çalışır ve çevrelerine enerji yaymazlar. Yaygın olarak kullanılan termal görüntüleme sistemleri ve görüntü kuvvetlendirici sistemler, pasif görüntüleme sistemlerinin örneklerindedir. Aktif görüntüleme sistemleri, etraflarına enerji yayıp, nesnelere yayılan enerjiyi görüntüleyen sistemlerdir. Pasif sistemler, az da olsa ortam ışığına veya ay ışığına ihtiyaç duyarken, aktif görüntüleme sistemleri ihtiyaç duymazlar. Mesafe kapılı görüntüleme sistemleri yüksek enerji atımlı ışık kaynağını kullanırlar ve kameranın açma-kapama sürelerini ışık kaynağı ile ilişkili olarak belirlerler. Görüntülenecek nesneden yansıyan ışığın kameraya ulaşma zamanı hesaplanır, kamera perdesi kısa bir süre açılarak, yansıyan ışık ile

görüntü elde edilir. Bu yöntemle, zor aydınlanma şartlarında bile yüksek çözünürlüklü görüntü elde edilir.

Mesafe Kapılı Sistemlerde nesne takibi yapabilmek için, nesnenin yatay ve dikey hareketinin yanında, kameraya uzaklığının da takip edilmesi gerekmektedir. Uzaklık bilgisi, kameranın açılıp kapanma zamanlamalarında kullanılmakta olup nesnenin görüntüsünün alınabilmesi için doğru olarak bilinmelidir. Bu tez çalışmasında, mesafe kapılı kamera sistemi kurularak, kurulan sistem ile entegre olarak çalışan gürbüz bir nesne takibi algoritması geliştirilmiştir. Nesnenin uzaklığı, önerilen algoritma ve çeşitli görüntü kıyaslama yöntemleri kullanılarak hesaplanmış ve görüntü kıyaslama yöntemlerinin performansları kıyaslanmıştır.

Anahtar Kelimeler: Mesafe Kapılı Sistem, Nesne Takibi

ACKNOWLEDGEMENT

I would like to state my special thanks and gratitude to Assoc. Prof.Dr. Alptekin Temizel for his supervision, encouragement, patience and support at all phases of my thesis. Under his guidance, I have chance to develop myself both academically and personally.

I would like to thank to Dr. Kuthan Yelen for his support and throughout the thesis. Without his support and experience, it has been difficult to progress on thesis work.

I would like to thank to Prof. Dr. Yasemin Yardımcı, Assist. Prof. Dr. Erhan Eren, Assist. Prof. Dr. Sinan Kalkan and Assist. Prof. Dr. Behçet Uğur Töreyin for accepting to read and review this thesis and for their valuable contributions.

This work was supported by the Laser Systems Design Department of Aselsan MGEO division. I would like to thank to Laser Systems Design Department Manager Birol Erentürk for his support for the system construction.

I am also thankful to my manager Alper Ünsoy for his endless support and motivation throughout the thesis and my friends Azmi Çağrı Erçakır, Gökhan Çapar, Emrah Yeşilyurt, İbrahim Büyükgüllü and Ahmetcan Uralcan for their help in capturing the test videos.

Finally, my biggest gratitude is to my wife Derya and my sons Mehmet and Ahmet for their endless patience, support, love and trust.



To My Wife and Sons

TABLE OF CONTENTS

ABSTRACT	iv
ÖZ	vi
TABLE OF CONTENTS	x
LIST OF TABLES	xii
LIST OF FIGURES	xiii
CHAPTER 1	1
1.1 Statement of the Problem	3
1.2 Purpose of the Study	3
1.3 Significance of the Study	4
1.4 Assumptions	4
1.5 Summary	5
CHAPTER 2	7
2.1 Range Gated Active Imaging Systems	7
2.2 Imaging with Range Gated System	9
2.3 Object Tracking	11
CHAPTER 3	15
3.1 Range Gated Active Imaging System	16
3.1.1 Working Principle	16
3.1.2 Range Gated System Construction	18
3.1.3 Algorithms	19
3.1.4 Research Questions	19
3.2 Methodology	20
3.2.1 System Components	20
3.2.2 Construction of the System	23
3.2.3 Gated Imaging	26
3.2.4 Tracking of Moving Objects	28
CHAPTER 4	37
4.1 X-Y Tracking Module	37
4.2 Z-Tracking Module	39
4.2.1 Ground-Truth for Object Location	39
4.2.2 Limitations	39
4.2.3 Reference Image	44
4.2.4 Algorithm	47

4.2.5	Testing Environment	61
CHAPTER 5	65
5.1	Method Evaluation	65
5.1.1	Experiment 1: Stationary Object Scenario	66
5.1.2	Experiment 2: Moving Object Scenario	70
5.1.3	SSIM STR Analysis	75
5.1.4	Reference Object Analysis	78
5.1.5	Brightness Invariance	79
5.1.6	Translation Invariance	82
5.1.7	Experiments with Artificial Images	84
5.1.8	Performance Analysis	85
5.2	Test Results and Conclusion	88
5.2.1	Test Scenarios	88
5.2.2	Performance Analysis of Test Scenarios	101
5.2.3	Discussions on Test Results	102
CHAPTER 6	107
6.1	Conclusion	107
6.2	Future Work	108
REFERENCES	109

LIST OF TABLES

Table 1 Error mean and STD values for stationary object scenario. Bold is used to highlight the best result.	70
Table 2 Error Mean and STD values for different methods. Object is located at 600 mm away from the previous location.	73
Table 3 RMS Error values for different methods. Object is located at 600 mm away from the previous location. Bold is used to highlight the best result.	74
Table 4 Error Mean and STD values for different methods. Object is located at the same previous location.	74
Table 5 RMS Error values for different methods. Object is located at the same previous location. Bold is used to highlight the best result.	74
Table 6 Strong and weak sides of methods under different conditions	87
Table 7 Object Distances.....	88
Table 8 Error Mean and STD values for all methods for Standing Man and Sitting Man Scenarios. Object is located 15 locations between -1000 mm and 1000 mm. Bold is used to highlight the best result.	93
Table 9 Error Mean and STD values for all methods for Three Antenna Scenarios. Object locations between -900 mm and 900 mm are used. Bold is used to highlight the best result.	99
Table 10 Success values for all methods and all test scenarios	102
Table 11 Success rates for Sitting Man scenario with two types of reference image. Bold is used to highlight the changes in SSIM STR and SSIM CNTRST methods.	104
Table 12 Success rates for Standing Man Scenario for 2, 3, 4 ns step sizes	106

LIST OF FIGURES

Figure 1 A thermal image acquired with a Long Wave Infrared Camera covering the spectral range 8-12 micrometer [1]	2
Figure 2 Camera and object model	3
Figure 3 Effect of the averaging 1, 3, 5, and 10 images on image quality taken from [7].	10
Figure 4 Active tracking system components	15
Figure 5 Working principle of range gated systems. (a) Laser pulse is emitted; camera shutter is closed, (b) Emitted laser is reflected, shutter is opened for a short time period	16
Figure 6 Range Gated Camera System Components and Their Relations.....	18
Figure 7 Laser Unit developed at Aselsan, (a) on board, (b) in its case	20
Figure 8 Components of the Intensified Camera Unit	21
Figure 9 Working principle of a gated intensifier. Low amount of light enters to image intensifier. Photocathode converts photons into electrons, MCP multiplies the electrons, and Phosphor screen converts electrons to light. Thickness of the arrows indicates the amount of light or electrons	22
Figure 10 Electron multiplication in one channel of MCP	23
Figure 11 Camera system with the camera and the control unit taken from [42].....	24
Figure 12 Spectral sensitivity of the intensifier taken from [42]	25
Figure 13 Distance-time model for gated viewing.....	27
Figure 14 Imaging functions for two different parameter (sensor delay time, laser pulse width, sensor gate width) set	28
Figure 15 Selection of the object that is being tracked	29
Figure 16 Histogram for the prior distribution.....	30
Figure 17 Result of the particle filter with Gaussian window around current state. Image order is from left to right, top to bottom	34
Figure 18 Optical flow (left) and Result of the particle filter with optical flow feedback (right)	35
Figure 19 Optical flow output (left) and particle filter output with optical flow feedback (right)	36
Figure 20 Summary of the X-Y Tracking Module. Optical flow determines the coarse movement and Particle filter makes the fine adjustment	37
Figure 21 Summary of the Robust and Fast Image Tracking in Range Gated System....	38
Figure 22 Average image intensities for each delay time are shown in black squares and the polynomial fitted to the black points is shown as line.	40
Figure 23 Four images taken with the same delay time, showing frame to frame brightness variation. For illustration purposes, dark pixels enhancement is applied by log transform.	40
Figure 24 Setup for noise calculation	41
Figure 25 Image taken from the Noise Setup.....	42
Figure 26 Illumination centroid distribution	43
Figure 27 Illumination distance distribution	43

Figure 28 Changes in reference images with new tracked images. Initial reference image (a), and the updated reference images (b), (c) and (d). For illustration purposes, dark pixels enhancement is applied by log transform.	45
Figure 29 Changes in reference images. Object moves and partial occlusion occurs. (a) is the initial reference image, (b)-(e) are the reference images updated at each tracking iteration and (f) is the final reference image. For illustration purposes, dark pixels enhancement is applied by log transform.	46
Figure 30 Image slices around the tracked object	47
Figure 31 Summary of Object Tracking Algorithm	48
Figure 32 Sample images (left) and their corresponding histograms (right). For illustration purposes, dark pixels enhancement is applied by log transform.	51
Figure 33 Diagram of structural similarity measurement system	53
Figure 34 (a)-(k) 11 images taken with 11 different delay times. For illustration purposes, dark pixels enhancement is applied by log transform.	56
Figure 35 Some of the selected region images. For illustration purposes, dark pixels enhancement is applied by log transform.	57
Figure 36 Reference Image. For illustration purposes, dark pixels enhancement is applied by log transform.	57
Figure 37 (a) Mutual Information, (b) Image Brightness, (c) Pixel-wise Comparison, (d) Histogram Comparison, (e) Peak Signal to Noise Ratio and (f) Structural Similarity metrics for 11 images	58
Figure 38 An example probability density function for 11 images in the case of object moving in one direction.	60
Figure 39 Test object to be tracked	62
Figure 40 Image showing the test object and the tower	63
Figure 41 Histogram Comparison (a), Pixel-wise Comparison (b), Image Brightness (c), Mutual Information (d), Peak Signal to Noise Ratio (e) and the Structural Similarity (f) metrics for 15 set of images (each image set consists of 11 images captured with different delay times).....	69
Figure 42 Error values (in mm) for Histogram Comparison, Pixel-wise Comparison, Image Brightness, Mutual Information and Structural Similarity methods for 15 set of images. (Interquartile range method with $k=0.02$ is applied) (Stationary object scenario)	69
Figure 43 Observed object locations (in mm) for Histogram Comparison, Pixel-wise Comparison, Image Brightness, Mutual Information, Peak Signal to Noise Ratio and Structural Similarity methods for 15 set of images. Object is located at 600 mm away from the previous location. Interquartile range method is not applied (a) and applied with $k = 0.02$ (b).	71
Figure 44 Observed object locations (in mm) for Histogram Comparison, Pixel-wise Comparison, Image Brightness, Mutual Information, Peak Signal to Noise Ratio and Structural Similarity methods for 15 set of images. Interquartile range method is applied with different k values. Object is located at 600 mm away from the previous location.	72
Figure 45 Error values (in mm) for Histogram Comparison, Pixel-wise Comparison, Image Brightness, Mutual Information, Peak Signal to Noise Ratio and Structural	

Similarity methods for 15 set of images. (Interquartile range method is applied with different k values) Object is located at the same previous location.	72
Figure 46 (a) is a typical first image and (b) is a typical last image in collected data set. For illustration purposes, dark pixels enhancement is applied by log transform.	75
Figure 47 Comparison results for the SSIM methods. Red line is the interquartile threshold value for Interquartile with $k = 0.07$	77
Figure 48 Comparison results taken from brightness analysis section. Images 1- 2 are the first images, 3-4 are the last images.	77
Figure 49 Comparison of the updated reference image and the selected object for each updated reference image	78
Figure 50 Images used for brightness invariance analysis. (a)-(b) are the images with no object, (c)-(h) are the images with object but different brightness levels. (f) is used as the reference image. For illustration purposes, dark pixels enhancement is applied by log transform	80
Figure 51 Similarity results between the reference image and the selected images for all methods. Results 1-4 are the not object images and the other results (5 to 18) are the images with different (increasing) brightness levels where 14th image having the most similar brightness value to the reference image. MI, IB, PwC, HC, and PSNR are shown in (a), other methods are shown in (b).	81
Figure 52 Images used for movement invariance analysis. (a) is an image with no object, (b)-(f) are the images with object but different locations. (e) is used as the reference image. For illustration purposes, dark pixels enhancement is applied by log transform.	82
Figure 53 Images not having the target object and images with the target object has gone through translation (i.e. present at different locations). Samples 1 and 2 are the results for images not having the target object; samples 3-11 having images of the object with different translation amounts and sample 12 is the result for the reference image. MI, IB, PwC, HC, and PSNR are shown in (a), other methods are shown in (b).	83
Figure 54 Images used for artificial image analysis. (a) is an image not having the target object, (b) and (c) are the two examples of the artificial images. (d) is used as the reference image. For illustration purposes, dark pixels enhancement is applied by log transform.	84
Figure 55 Comparison results for the images. Sample 1 corresponds to image not containing the object, Samples 2-9 for the artificial images and Sample 10 for the reference image.	85
Figure 56 PDF of Error while the object is located at previous location.....	86
Figure 57 PDF of Error while the object is located 600 mm away from previous location.....	86
Figure 58 Objects used for performance tests of the comparison methods. (a) is named as “Standing Man”, (b) is “Sitting Man” and (c) is “antenna”. (d) is the images of all three objects taken from the observation tower.....	89
Figure 59 Image taken from standing man object (a) and the object region selected manually (b). For illustration purposes, dark pixels enhancement is applied by log transform.	90

Figure 60 Observed locations (top) and the error values (bottom) for all methods for Standing Man Scenario. Black line is the ground-truth line.	91
Figure 61 Image taken from sitting man object (a) and the object region selected manually (b). For illustration purposes, dark pixels enhancement is applied by log transform.	92
Figure 62 Observed locations (top) and the error values (bottom) for all methods. (Sitting Man Scenario)	94
Figure 63 Image taken from Antenna (a) and the object region selected manually (b). For illustration purposes, dark pixels enhancement is applied by log transform.	95
Figure 64 Observed locations (top) and the error values (bottom) for all methods. (Antenna-3ns StepSize Scenario).....	97
Figure 65 Observed locations (top) and the error values (bottom) for all methods. (Antenna-2ns StepSize Scenario).....	98
Figure 66 Image taken from Antenna (a) and the object region selected manually (b). For illustration purposes, dark pixels enhancement is applied by log transform.	99
Figure 67 Observed locations (a) and the error values (b) for all methods. (Antenna-2ns StepSize Zoom Out Scenario)	100
Figure 68 Laser pulse characteristic used in test setup. Power is not absolute value but a scale to show the signal characteristic.....	101
Figure 69 Imaging Function.....	102
Figure 70 Observed locations for all methods. Sitting Man scenario is run but an observed image is used as reference object.....	104
Figure 71 Observed locations (top) and the error values (bottom) for SSIM LUM method for 2ns, 3ns and 4ns step sizes.....	105
Figure 72 Image taken from Antenna (a) and the object region selected manually (b). For illustration purposes, dark pixels enhancement is applied by log transform.	106

CHAPTER 1

INTRODUCTION

Surveillance systems pose a number of challenges. These systems are expected to work in low light levels and low visibility conditions. Using a thermal sensor (8-12 micrometer or 3-5 micrometer) is an effective way of identifying targets compared to other passive imaging systems since thermal sensors are sensitive to heat emission from the imaged object and they are independent of the environmental light conditions. However, measurement of the dissipated heat by the sensor may not be at sufficient level to construct an image because of low dissipated energy or distance to the object. An example of thermal image is shown in Figure 1. Since the thermal cameras detect the dissipated heat, the images acquired with these cameras do not contain details such as textures and object reflectance. Besides that, thermal imaging is an expensive and complicated technology compared to classical Charge Coupled Device (CCD) or Complementary Metal Oxide Semiconductor (CMOS) technologies. Another passive system, intensified cameras, which are based on the principle of increasing the number of incoming photons, suffers from high noise because such systems also increase the number of unwanted photons.



Figure 1 A thermal image acquired with a Long Wave Infrared Camera covering the spectral range 8-12 micrometer [1]

Contrary to passive imaging systems, active imaging systems send energy to their environment and detect the reflected energy from the objects. Although active systems do not require ambient light, they suffer from low signal-to-backscattering noise ratio. Fog, rain or snow which reflects the illuminator light in front of the scene might adversely affect the image quality and might even render the images useless for practical purposes. Besides, strong light source in the vicinity of the object might also degrade the image quality of the object in question.

Range-gated imaging systems improve the signal-to-backscattering noise ratio by reducing the amount of backscattered light captured by the sensor. This is achieved by synchronizing the arrival of pulsed light returning from the target with the gating of an intensified camera. Since the photon time of flight is known, short sensor exposure time is arranged and only information that arrives at the sensor within certain timing window contributes to the imaging process.

Distance of the object is a key parameter in range gated systems. For object tracking with a range gated system, one should also track the distance of the object beside the movement in X and Y directions to arrange suitable gating timing.

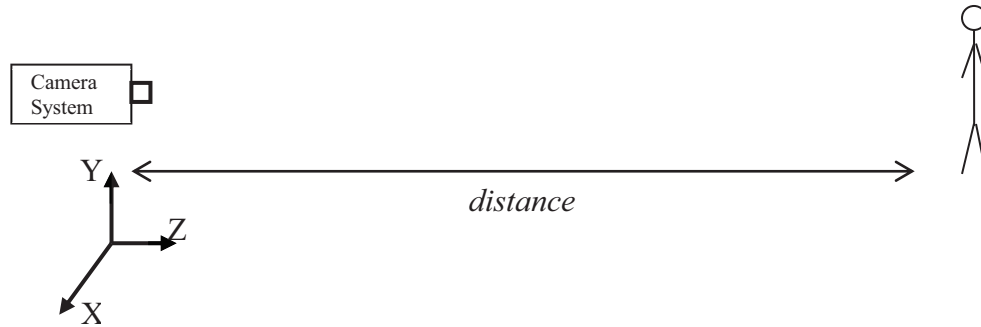


Figure 2 Camera and object model

1.1 Statement of the Problem

Active imaging systems with extremely short pulses on the order of nanoseconds only recently became feasible thanks to advances in laser and photodetecting technologies. Although there are many studies aiming to produce and increase the quality of reflectance images with range gated systems, none of them deals specifically with object tracking problem with range gated systems. Different to tracking using visible light or thermal cameras, besides tracking the object in X and Y dimensions, in a range gated system, distance of the object to camera should also be tracked. Calculated distance information is required to synchronize with the arrival of the pulsed target irradiance with the gating of an intensified camera to keep the object in sight.

1.2 Purpose of the Study

The purpose of this study is to develop hardware using actively illuminated range gated imaging approach and propose a robust real time object tracking algorithm that

works collaboratively with the system by providing object distance to the system continuously as input.

1.3 Significance of the Study

Current surveillance systems use the day camera or thermal sensors for object tracking. Day camera systems suffer from low light and might be blind if there is too much illumination. On the other hand, while thermal cameras do not require any illumination, they do not provide texture or reflectance information of the object and the object should dissipate heat to be visible in thermal images. Range gated imaging systems have the capability of getting images from long distances and from any object at the absence of light in the environment. In range gated systems, distance of the object should precisely be known. Although there many studies dealing with the quality of reflectance images and 3D reconstruction with range gated systems, none of them is processing images for tracking the movement of the object on the scene. This study proposes an algorithm for object tracking continuously without resorting to an additional laser or radar based range finder unit even the object distance changes during the tracking. Range estimation is solely based on image processing.

The present study on object tracking with range gated camera system will help to track an object with range gated camera system without an operator support.

1.4 Assumptions

For this study, it is assumed that the object does not move out of the scene partially or completely during tracking.

For this study, objects are located in the range 204 to 315 meters and all images are taken at summer night clear weather. Changes in atmospheric attenuation and scattering are ignored due to the limited distance to the object and clear weather. For longer range tracking, atmospheric attenuation and scattering must be considered.

1.5 Summary

In Chapter 2, a literature survey is presented. History of the range gated applications starting with first granted patent and proposed range gated systems are described. Previous imaging application based on 3D scene reconstruction and 2D imaging at low light conditions are summarized in this chapter. Lastly, 2D object detection and tracking methods are described.

Chapter 3 presents the working principles of the range gated system and the properties of the system used in this study. This chapter also introduce the basics of the methodologies used in the study.

Finally in Chapter 4, the proposed algorithm for object tracking in range gated camera system is described. A robust algorithm that tracks the movement of the object in all three directions is presented. Details of the proposed methods are given and the performances of these methods are analyzed in this chapter. Additionally the test scenarios and test results are described. Finally, concluding remarks are given in this chapter.



CHAPTER 2

BACKGROUND AND LITERATURE REVIEW

The range gated camera systems have been actively researched in order to achieve better performance in night vision applications. Since such systems require multidisciplinary work and applications are mainly on military area, there is limited amount of published work on this subject and most of the articles on range gated camera systems focus on the system construction. In the literature, there are two main application areas for range gated camera systems: 3D scene construction and 2D imaging at low light conditions. Automated object tracking on such a system is desirable as it would eliminate the need for a human operator to continuously set the gate timing for the changing object distance and allow the object to remain in sight all the time without user intervention. However, there is no object tracking work specifically designed for a range gated system to our knowledge.

2.1 Range Gated Active Imaging Systems

History of the range gated camera system concepts starts from 1960s. Neuman [2] was granted the first patent in 1968 from United States Patent Office. They proposed a system with Q-switched 60 nanoseconds gated laser pulse and 50 nanoseconds camera gate. Kerpchar [3] proposed an underwater range gated imaging system in 1972. System is integrated on a diver; gate and delay times of the system are configurable by the diver. At 1990, Ulich et al. [4] were granted a patent for using multiple cameras which are independently configurable with different delay and gating times. They claimed that by using such a system, sub-images of a scene can be

obtained by only one light pulse. One year later, Ulich et al. [5] were granted a patent for an imaging camera which includes pulsed light source, camera and timing electronics internally.

Most of the academic works on range gated camera systems are based on the system description. Bonnier and Larochelle [6] describes the properties of their range gated system that they named airborne laser based enhanced detection and observation system (ALBEDOS) for coast guard and maritime. ALBEDOS is based on a powerful laser diode array illuminator and a range-gated low-light-level TV camera.

Repasi et al. [7] locate laser and the camera at different locations (in classical approach, laser source and camera are located at the same location) and acquire images. They also aim to acquire images even when the object is not in the line of sight of the camera by using the walls as mirror. By this method, they can get the image of a man who is standing 20 meter away from the illuminated wall.

Fect and Rothe [8] discuss the development and compare the experimental results of two imaging laser radar systems based on range gating. One of the systems operates with 1574nm eye safe laser and the other operates with 532nm laser. Ofer et al. [9] developed an active night vision system using gated imaging principles. There are also some underwater applications. Weiqi et al. [10] describe an underwater imaging system and claim that they reach 30 meter visibility at turbid medium. Tan et al. [11] also work on the turbid water medium and compare the images obtained from different turbidity levels. The U.S. Army and the U.S. Air Force are aggressively pursuing programs that utilize laser range-gated, shortwave IR imaging systems to

perform long range target identification. Driggers et al. [12] investigated the impact of speckle on laser range-gated shortwave infrared imaging system on target identification performance.

Determining the properties of the system is a challenging task. Power of the laser unit, gating time of the laser pulse and the camera unit and performance of the intensifier are the main properties which need to be determined. Zhang et al. [13] give detailed information for the properties of the system. First question they asked is “how much light returns from the object?” Some parameters affecting the amount of light are the power of the laser and the distance of the object. They give a table of reflected light for different laser powers and the object distances. They also proposed atmospheric transmission values for the 808 nm light. Second question they asked is; “can the detector discriminate the returned light from noise? Spectral response of the intensifier is the main parameter for the detector responsivity”. In conclusion, they state that theoretically when the returned light is higher than the detector responsivity, the target can be detected. However, if the light is higher than 2 times detector responsivity should be more realistic.

2.2 Imaging with Range Gated System

Improvement of the quality of the obtained range-gated images is another field of study. Gilles [14] proposes some image processing algorithms to eliminate the artifacts on range gated images. He compares the effects of applying mean and median filters and evaluated the results. Repasi et al. [7] propose taking the average of images to reduce the speckle. Effects of the image averaging are shown in Figure 3. An approach to modelling the performance of active imaging systems has been

introduced by Richard et al. [16]. This approach leverages physically accurate models of speckle, scintillation, and psychophysically accurate models of target acquisition.

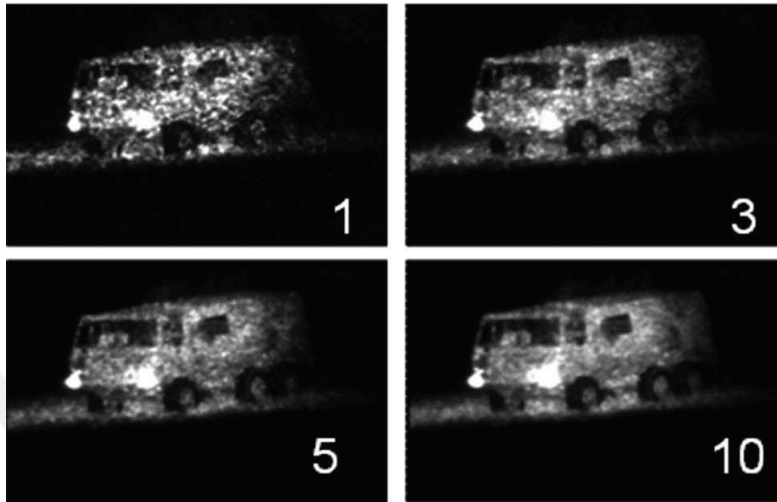


Figure 3 Effect of the averaging 1, 3, 5, and 10 images on image quality taken from [7].

Another line of work utilizing range gated camera systems is on 3D reconstruction. Busck and Heiselberg [17] give a methodology for 3D object reconstruction by using the images taken from a gated system. They obtain better than 1 mm range accuracy by using 200 picoseconds gate time. Laurenzis [18] claims that the whole depth information can be calculated from a minimum number of two range-gated images via the super-resolution depth mapping technique. Laurenzis and Bacher [19] discuss a method of image coding by multiple exposures of range gated images. Their claim is that they can enlarge the depth mapping range by a factor of 12 by using three range gated images.

One year later, Laurenzis et al. [15] proposed an improved algorithm that uses quasi-ambiguous coding sequences to enlarge the depth mapping by a factor of 13 instead of 12. Anderson [20] suggests two different methods and investigates algorithm performance through Monte Carlo simulations using a simplified system and

imaging model. Monnin et al. [21] state that by precisely synchronizing both the illuminator and the camera shutter, it is possible to acquire "slices" of the scene at specific known distances. They show that even with large laser pulses and without megahertz-capable electronics, the third dimension can be recovered for the whole range of the scene by processing only two images acquired in specific conditions. Steinvall et al. [22] review some of the work on range gated imaging undertaken at the Swedish Defence Research Agency (FOI). The basis for 3-D scene reconstruction is an image sequence acquired using a sliding gate delay time. Göhler and Lutzmann [48] investigated the influence of the number of averaged images on the resulting range accuracy. They located a well-defined plate as an object and used brightness of each pixel of the plate to calculate the range error for different number of averaged images.

2.3 Object Tracking

Object tracking with range gated camera system is a naive unexplored research field because of the lack of wide availability of range gated systems. However, there is a great deal of interest in object tracking algorithms using intensity images. Yilmaz et al. [23] provides an extensive survey on object tracking area. In object tracking, an object can be defined as anything that is of interest. Objects are represented by their shapes and appearances. Comaniciu et al. [24] introduced the kernel-based tracking technique which uses the basin of attraction of the similarity function. Yilmaz et al. [25] state that contour (boundary of an object) and the silhouette (the region inside the contour) representations are suitable for tracking complex nonrigid shapes.

Object skeleton is another object representation method and can be extracted by applying medial axis transform to object silhouette. Ali and Aggarwal [26] used skeleton model for shape representation. Beside the stated shape representations, there are also some appearance representations of the objects. Probability density of object appearance is a commonly used method in object representation. Zhu and Yuille [27] used Gaussian and Paragios and Deriche [28] used mixture of Gaussians for probability densities of the objects. The probability densities of appearance features (color, texture etc.) can be computed from the regions specified by object shape models (silhouette or inside of a contour) [23]. Fei Yin et al. [29] deals with the problem of performance evaluation of motion tracking methods. They propose a rich set of metrics to assess different aspects of performance of motion tracking. Guo-liang et al. [30] propose an algorithm by using adaptive mean-shift method. They claim that the algorithm can track fast moving objects of changing size successfully. Beauchemin and Barron [31] describe the basic terms and principles of optical flow techniques. Fleet and Weiss [32] describe the details of gradient based optical flow algorithms.

Beside determination of the object representation, selecting the right feature plays a critical role in object tracking [23]. Object features should be unique in order to easily distinguish the object from the background. Yilmaz et al. [23] states that feature selection is closely related with the object representation. For example, color is used as a feature for histogram-based appearance representations and object edges are usually used as features for contour-based representation. Combination of these features can also be used in tracking algorithms.

Most of the object tracking methods are based on the 2D intensity images. Kristan et al. [33] construct a tracker that combines optical flow and color based particle filter. Kodama et al. [34] propose the technique for tracking the velocity of objects that combines the particle filter with optical flow. Tung and Matsuyama [35] propose a solution to human motion tracking by using a color-based particle filter driven by optical flow. Beyan and Temizel [43] proposed a surveillance system for indoor environments which is capable of tracking multiple objects using both visible and thermal band images. Spies et al. [36] discuss the computation of the instantaneous 3D displacement vector fields of deformable surfaces from sequences of range data. Barron and Spies [37] combines the 2D optical flow method and 3D range flow methods to achieve better tracking performance.

Every tracking method requires an object detection mechanism for initialization or need to be manually initialized. Besides using the information on a single frame, temporal information can also be used for object detection. Background subtraction is a popular method using temporal information. Background of the scene is modeled and the blobs obtained through differencing of the new images and the background model are detected as objects. Wren et al. [38] propose a human tracker system that they called Pfinder (person finder). Since the camera is fixed, Pfinder models the background scene that is relatively static. Then, when a human enters the scene it begins to build up a model of that person. Beyan and Temizel [44] present a fully automatic multiple-object tracker based on mean-shift algorithm. They detect new objects entering to the field of view and objects that are leaving the scene by using foreground detection.

Speeding the object tracking algorithms by the help of Graphics Processing Unit (GPU) is another field of study. Guler et al. [45] described the implementation of a real time intelligent video surveillance system on GPU. The system is based on background subtraction and composed of motion detection, camera sabotage detection, abandoned object detection, and object-tracking algorithms. They claim that GPU makes the system up to 21.88 times faster than the central processing unit. Gurcan [46] analyzed the tracking framework called TLD (Tracking-Learning-Detection) to optimize it computationally on a CPU-GPU hybrid setting and developed a solution via using GP-GPU (General Purpose GPU) programming.

Although there are a number of studies on range gated imaging systems, most of the studies focus on the system construction, 3D reconstruction and enhancement of images obtained by range gated systems. While there are various studies on object tracking with 2D images or 3D depth data, real time object tracking by range gating technology is an uncharted territory.

CHAPTER 3

METHODOLOGY

The purpose of this study is to develop a range gated camera system and propose an active object tracking system which consists of a tracking algorithm working in association with the range gated camera system. In the scope of this study, a range gated system consisting of an intensified gated camera and a laser system, has been constructed. By using the output images of the range gated system, object tracking is done in all three dimensions. The distance of the object estimated by the tracking algorithm is used by the range gated system to obtain the subsequent image without resorting to an additional laser or radar based range finder unit. Range estimation is solely based on image processing. Components of the active tracking system are shown in Figure 4.

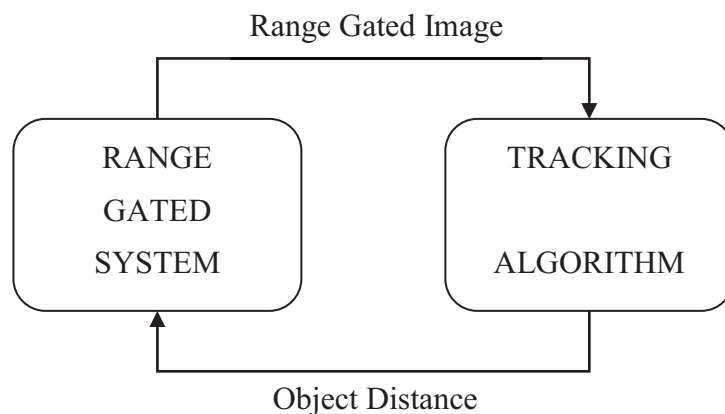


Figure 4 Active tracking system components

3.1 Range Gated Active Imaging System

Working principles and a brief summary of the system components are described in this section.

3.1.1 Working Principle

In a range gated imaging system, a camera with tightly controlled opening and closing shutter times is used in conjunction with a high power pulsed light source. Exposure time of the camera and the light source pulse width are arranged according to the return time of the emitted light pulse. Working principle of the gated imaging system is shown at Figure 5.

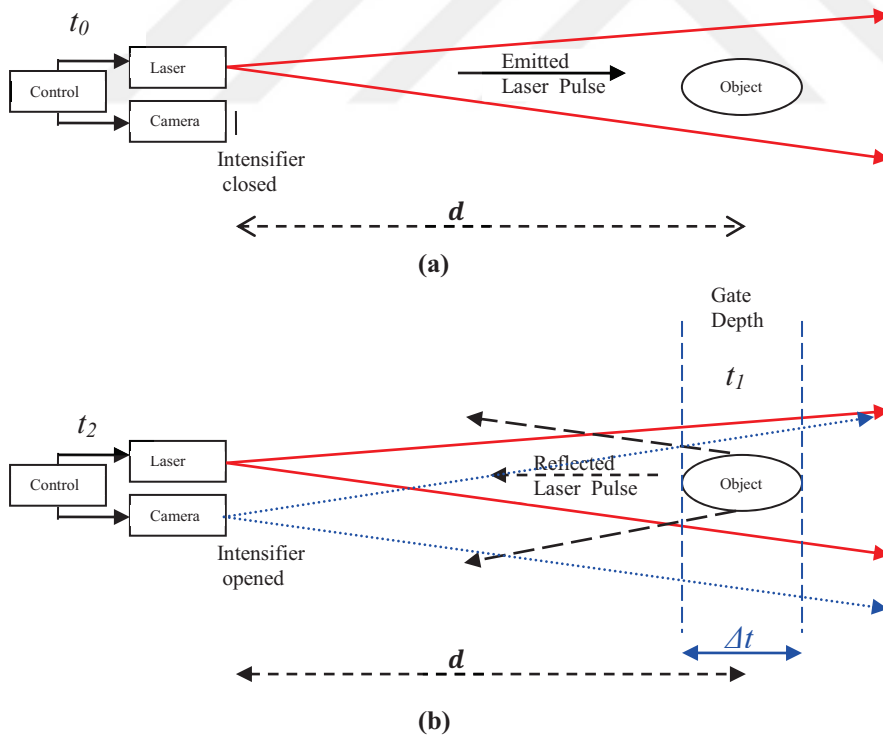


Figure 5 Working principle of range gated systems. (a) Laser pulse is emitted; camera shutter is closed, (b) Emitted laser is reflected, shutter is opened for a short time period

Steps of the Figure 5 can be summarized as:

- A laser pulse is emitted at t_0 while camera shutter is closed.
- At t_1 , emitted light pulse is reflected from the targeted object that is located at distance d .
- At t_2 , the camera is opened for a short period (Δt) corresponding to the desired depth of view.

Return time r that is defined as $t_2 - t_0$ is calculated as;

$$r = (2 \times d) / c, \quad (\text{Equation 3.1})$$

where c is the speed of the light and d is the distance of the object. It is the duration that the emitted light reaches the object and returns to the camera. Controller sets the opening time of the intensifier by using this r value.

‘Depth of view’ is the depth of the desired illumination area. It is used for the calculation of the ‘gate time’ of the camera.

To visualize the range gating concepts better, numerical values will be given. For an object located at 3 km distance ($d = 3000m$) with 9 meters depth (depth of view = 9 m), the return time $t_2 - t_0$ and gate time (Δt) can be calculated as follows:

$$t_2 - t_0 = 2 \times 3000 / 3 \times 10^8 = 2 \times 10^{-5} \text{ second.} \quad (\text{Equation 3.2})$$

$$\Delta t = 2 \times \text{desired depth of view} / c = 2 \times 9 / 3 \times 10^8 = 6 \times 10^{-8} \text{ second.} \quad (\text{Equation 3.3})$$

3.1.2 Range Gated System Construction

A range gated camera system has three main components: Laser unit, intensified camera unit and the Host Computer as shown in Figure 6. After a brief description of these units in this section, details will be given in Section 3.2.1.

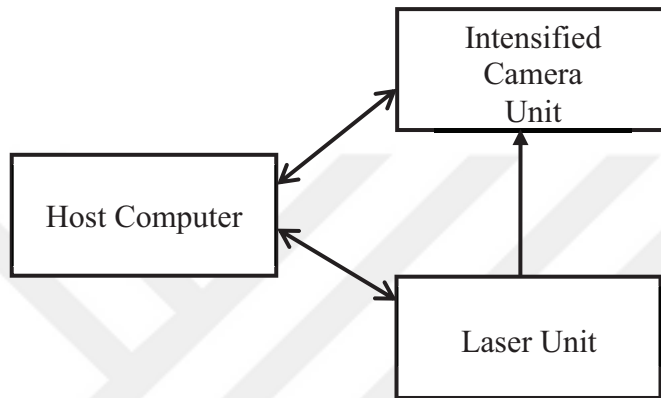


Figure 6 Range Gated Camera System Components and Their Relations

Laser Unit:

The Laser unit is the illumination source of the system. The laser beam can travel long distances with minimum expansion hence it can deliver bright light spot to the intended area. Its pulse duration can be precisely controlled and during this pulse duration it can achieve extremely high powers. Laser unit triggers the Camera Unit when it emits a laser pulse.

Intensified Camera Unit:

Since the amount of the emitted light reflected from the distant object is very low, sensitivity of the camera is an important parameter. Image intensifier is used with the

CCD to achieve high gain values. The camera can also provide small exposure times (in the level of nanoseconds).

Host Computer:

Host Computer is the master unit controlling the system. Pulse width of the laser and camera exposure timings are controlled by the Host Computer. Besides that, images are processed on the host computer and the system parameter “object distance” is calculated on the host computer.

3.1.3 Algorithms

The purpose of this study is to implement an active tracking system which consists of an object tracking algorithm working in association with a range gated camera system. In the scope of this thesis, application of some basic tracking algorithms such as optical flow and particle filter are investigated. Image Brightness, Pixel-wise Comparison, Histogram Comparison, PSNR and Structural Similarity approaches are used to determine the distance of the object to the camera. The object can move in all X, Y and Z directions.

3.1.4 Research Questions

The research question is “Can we implement a robust real time object tracking algorithm working in association with a range gated camera system?”. Besides that, “What is the effect of the fluctuations of system timings on the tracking algorithm?” and “What is the effect of fog on a near infrared wavelength based range gated camera system?” are the complementary questions of this thesis.

3.2 Methodology

There are two main steps in this study. The first one is constructing the system and the second is developing a robust object tracking algorithm on this system.

3.2.1 System Components

Figure 6 describes the Range gated camera system. In this section, details of each unit will be given.

3.2.1.1 Laser Unit

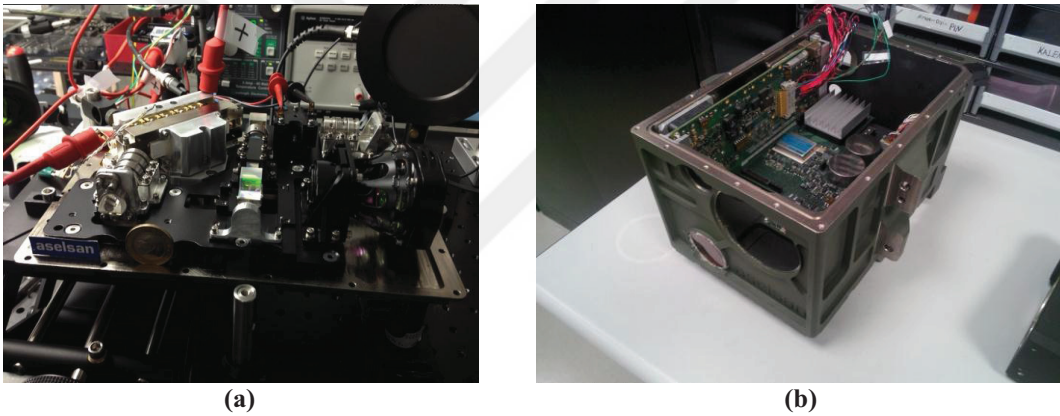


Figure 7 Laser Unit developed at Aselsan, (a) on board, (b) in its case

In the scope of this thesis, a laser unit shown at Figure 7 developed at ASELSAN was used. The Laser Unit is able to send 10 ns pulses in the near infrared region with a peak power of 200 kW up to 20 Hz.

3.2.1.2 Intensified Camera Unit

Intensified camera unit consists of Camera, Gated Intensifier and Related Optics and Mechanics as illustrated in Figure 8. Gated intensifier is used to increase the intensity of the incoming light and the camera is used for capturing the intensified image.

Related Optics and Mechanics are used for the mechanical and optical connections of these two components.

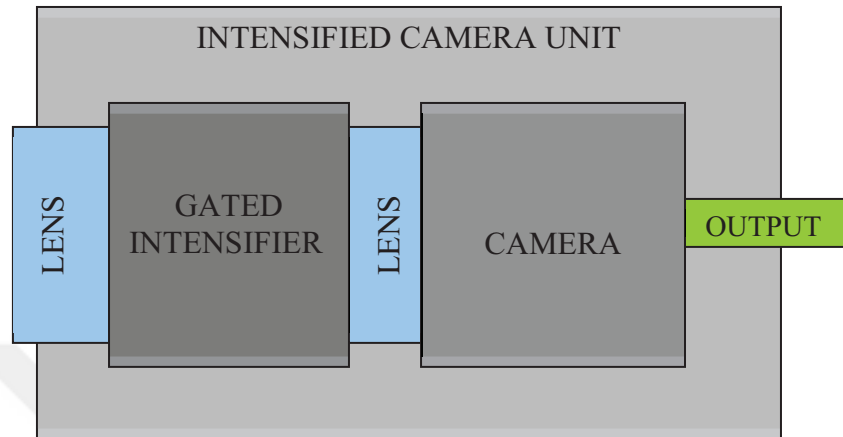


Figure 8 Components of the Intensified Camera Unit

3.2.1.2.1 Camera

Camera is a standard CCD camera with an analog video output. Its frame rate is 50 Hz and the resolution is 720x480 pixels. However, output of the camera is an interlaced video which means that the two fields are captured at two different 20 milliseconds time slices. In the constructed range gated system, the illumination on the detector occurs in the order of nanoseconds so only one field can be used in the observed image. As a result, the effective resolution of the camera is 720x240 pixels.

3.2.1.2.2 Gated Intensifier

Gated intensifier has two main functionalities. First one is “gating” and the other is “intensifying”. Basically, “gating” is opening and closing to allow entering of light for a short period (a few nanoseconds) and “intensifying” is increasing the captured photons to allow obtaining an image using the CCD sensor. Figure 9 describes the working principle of a gated intensifier. In intensifier, three layers are located which are named Photocathode, Multi Channel Plate (MCP), and Phosphor Screen.

Light focused on the photocathode is converted into photoelectrons. Number of photoelectrons is proportional to the input light intensity. The voltage difference between the photocathode and MCP input surface accelerates these emitted electrons towards MCP. As shown in Figure 10, each channel on MCP serves as an independent electron multiplier. Input electrons hit the wall of each channel and produce secondary electrons. Through the channel, this process occurs many times and at the output side of the MCP, thousands of electrons are released for each input electron. These electrons again accelerate towards the phosphor screen and reconverted into light. Finally, an image is obtained which is intensified thousands times of the input light.

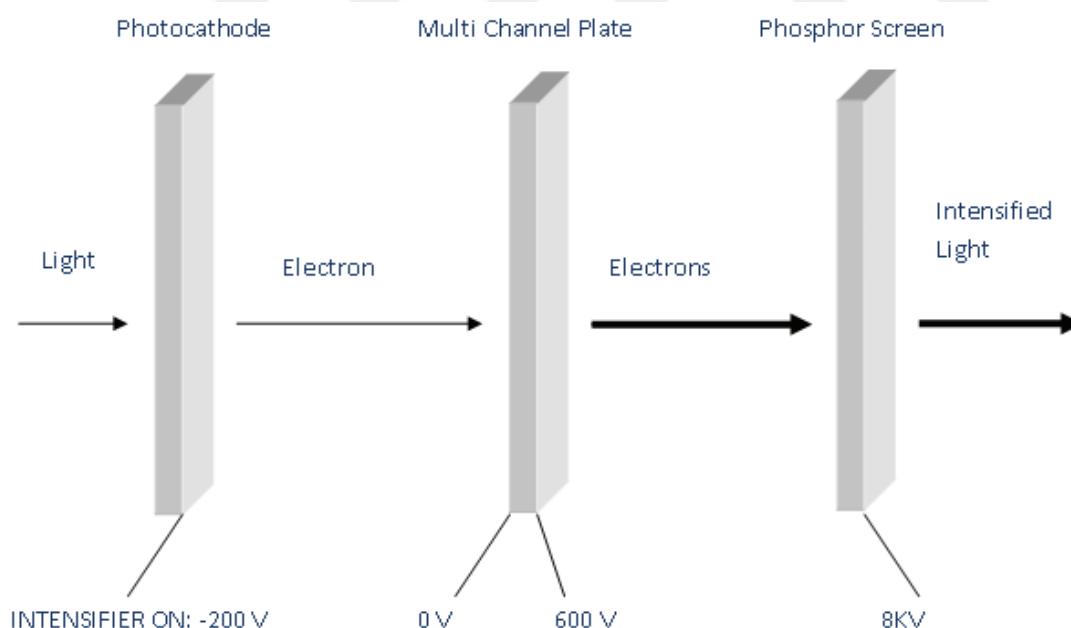


Figure 9 Working principle of a gated intensifier. Low amount of light enters to image intensifier. Photocathode converts photons into electrons, MCP multiplies the electrons, and Phosphor screen converts electrons to light. Thickness of the arrows indicates the amount of light or electrons

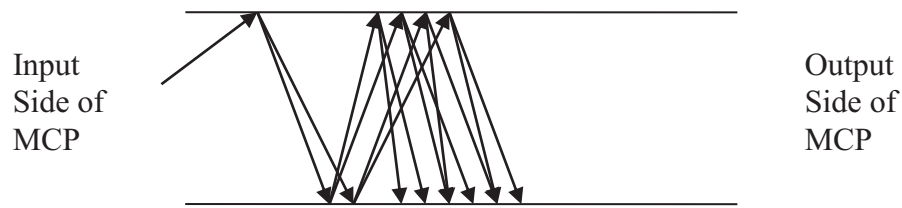


Figure 10 Electron multiplication in one channel of MCP

Second functionality of the gated intensifier is ‘gating’. Since the camera cannot have exposure times in the order of nanoseconds (it has a minimum of around 10 microseconds), gating is done by the intensifier. As shown in Figure 9, intensifier can be closed or opened by only changing the photocathode voltage value. When the voltage on the photocathode is negative, the voltage difference between the photocathode and MCP input surface accelerates the electrons towards MCP. However, if the photocathode voltage has a positive value, the electrons stick on the photocathode and do not travel towards MCP. In this case, intensifier acts as a shutter. If the voltage on the photocathode can be changed in nanoseconds, system has the gating capability in the order of nanoseconds.

3.2.1.3 Host Computer

Section 3.1.1 describes the timing relations between the laser unit and the camera unit. Triggering time of the laser pulse, exposure time of the camera and the gating times of the intensifier is set by the Host Computer.

Host computer also process the output images, proposed algorithm runs on the host computer and controls the system timings.

3.2.2 Construction of the System

For the construction of the system, following components are used;

- Laser Unit (developed at Aselsan, specifications are described in section 3.2.1.1)
- Intensified Camera Unit (details are described in this section)
- Host Computer (a standard notebook)

3.2.2.1 Intensified Camera Unit

Nanocam HF4 V 5N Camera System was procured to construct the system. The Nanocam HF4 V 5N is a fast shutter CCD camera system that includes a digital control unit to supply the gate signals to the camera. Appearance of the system is shown in Figure 11.



Figure 11 Camera system with the camera and the control unit taken from [42]

3.2.2.1.1 Intensifier

System includes a gateable intensifier with double multi channel plate. By this intensifier, system is an extremely sensitive camera ideal for very low light situations. The image intensifier increases the sensitivity of the camera up to 1.000.000 times than a standard CCD camera. Spectral sensitivity of the selected intensifier is shown in Figure 12. Since the used laser unit emits laser pulse at near

infrared wavelength, type N image intensifier is selected and used in HF4 V 5N Camera System.

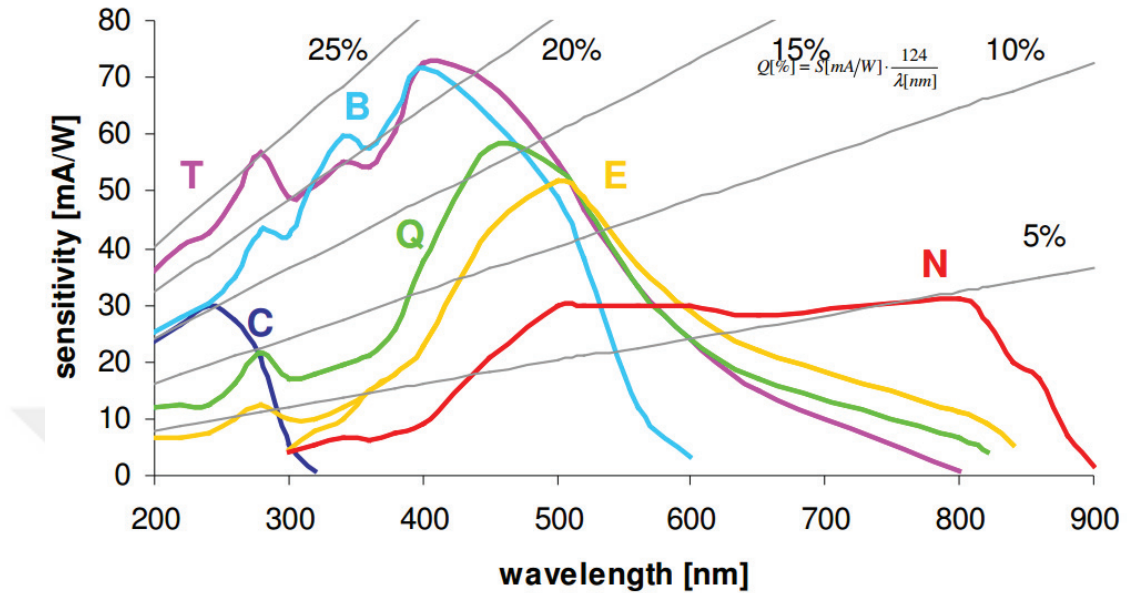


Figure 12 Spectral sensitivity of the intensifier taken from [42]

The image intensifier is gated with pulses generated by the pulse amplifier. Triggering of the pulse amplifier is possible by internal (V-synchronous) or external trigger pulses. The internal system pulse delay is approximately 150 ns with a jitter of max. 1 ns. An additional adjustable pulse delay time is available through a digital time delay unit with adjustable delay times between 0 and 65 ms in steps of 1 ns.

3.2.2.1.2 Sensor

The camera utilizes SONY type ICX 022/083 CCD detector, having 768x494 (H x V) pixels in the EIA version and 756x581 (H x V) pixels in the CCIR version.

3.2.2.1.3 Digital Control Unit

Control of the camera via the Digital Control Unit or a computer interface is available and gives access to all camera functions. A keyboard in the control unit allows operation of the whole camera system without using a PC. PC control via

RS232 interface is also available. Following camera settings are controlled in the scope of this project;

- MCP-gain (gain of the MCP image intensifier)
- Gate mode, Gate duration, Gate delay

3.2.3 Gated Imaging

Range gated imaging is based on the synchronization of the reflected laser pulse and the sensor gate time. Since only the reflected pulses from the objects that reach the camera at the gating time contributes to the image, the image includes both reflectance and the depth information. Reflected amount of the light from the surface contribute as the reflexivity and the round trip time of the light give information about the depth.

As shown in Figure 13, relative timings of the sensor gate and the laser pulse determine the imaged distance. The emitted laser pulse illuminates the related area on positive z direction. At the same time, sensor gate collects the laser pulses reflected from the objects at the positive z direction at only during the sensor gate time. The convolution of the laser pulse and the sensor gate contributes to the image.

In Figure 14, I_z describes the contribution of laser pulse and sensor gate to the image for each point in z direction.

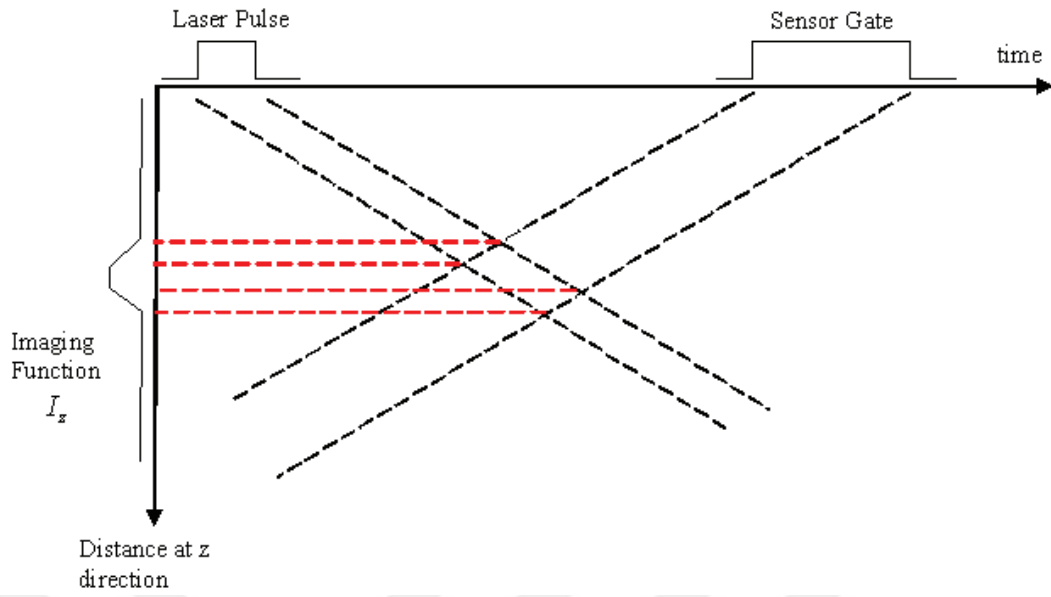


Figure 13 Distance-time model for gated viewing

As illustrated in Figure 14, depth-intensity profile can be calculated from the sensor gate and the laser pulse by taking the convolution of two functions. Effects of the gate delay time and round trip time of the laser must be included in the convolution operation.

$$I_z = \int_{-\infty}^{\infty} P(t-t_l) G(t-t_s) dt , \quad (\text{Equation 3.4})$$

where t_l is the round trip time of the laser pulse and P is the pulse function, t_s is the delay time for the sensor gate and G is the gating function.

By arranging the laser pulse, sensor gate and the delay time of the sensor, we can observe different I_z function - image pairs. Object distance information can be calculated by using these pairs.

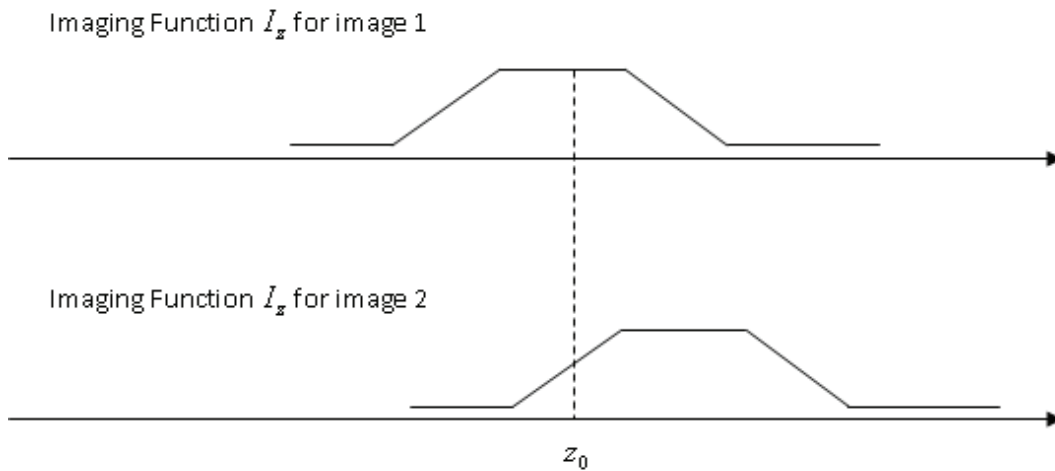


Figure 14 Imaging functions for two different parameter (sensor delay time, laser pulse width, sensor gate width) set

3.2.4 Tracking of Moving Objects

Tracking the position of a moving object is a critical problem in range gated camera systems. Since the moving object is illuminated using the distance information between the object and the system (displacement in z direction), it should also be estimated in real time in order to continue to get new range-gated images having the object of interest. By only getting images from one gate depth (Δt), the distance changes cannot be tracked. To track the object continuously, our system will get images from behind and ahead of the tracked objects.

Optical flow and particle filter are the methods used for object tracking in x and y directions. Details of these algorithms are described below.

3.2.4.1 Particle Filter Method

For object tracking, particle filter method is investigated and applied on sample videos. The procedure of the particle filter that is used in this research is as follows.

State Parameters:

Tracked object is considered to be located in a rectangle so the bounding box parameters of the tracked object are;

- Center X coordinate (θ_x)
- Center Y coordinate (θ_y)
- Width (θ_w)
- Height (θ_h)

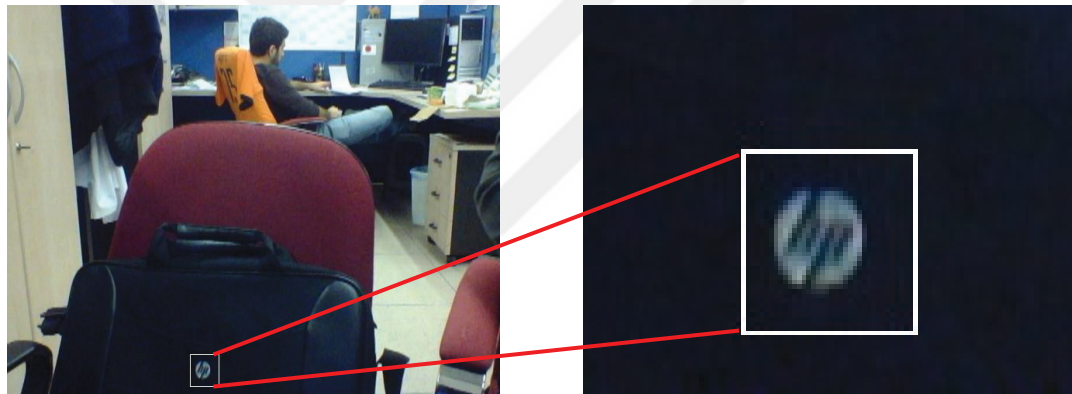


Figure 15 Selection of the object that is being tracked

Initialization:

Initial values for the first state can be output of automatic object detection or user defined value. In our case, initial object is selected by the user and based on the user selection; $\theta_x, \theta_y, \theta_w, \theta_h$ parameters of the system are initialized. In Figure 15, selection of the rectangle is shown.

Transition Model:

Transition Model specifies how objects move between frames. It is the model for predicting the next value of a state with given current state value. The basic method

for the transition model is a Gaussian window around current state values. However, there are more complex models that take into account velocity and acceleration information using the previous state values. In this thesis, output of the Optical flow algorithm is used to predict velocity information.

State transition model is defined as $\theta_t = F_t(\theta_{t-1}, O_t, G)$ where θ_{t-1} is the previous state, O_t is the output of the Optical flow algorithm and G is the Gaussian noise.

Observation Model:

Observation model specifies the likelihood of an object being in a specific state (i.e. at a specific location). Gray scale histogram based model can be used for range gated images. For the prior distribution, histogram of the image selected at Figure 15 is shown at Figure 16.

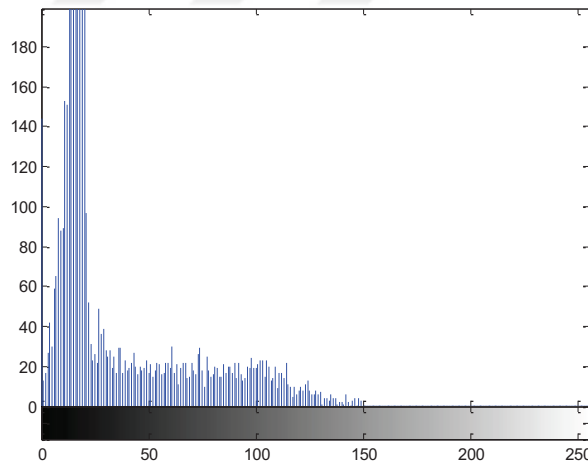


Figure 16 Histogram for the prior distribution

Likelihood is based on a distance metric between histogram of the prior distribution and the current observation. For each particle, histogram of the current observation is calculated and the likelihood of it with the prior distribution histogram is calculated. The less the distance between histograms, the higher the probability of particle be a

correct estimation. For each particle, likelihood function is determined and the particle with the highest likelihood represents the current state. If the likelihood value of the current state is smaller than a predefined value, X tracker is accepted as failed.

The next issue is to determine the particles for the next frame. The particles with low likelihoods are removed, the particles with high likelihoods are duplicated, and the weights of the particles are made equal. The resampling step prevents the existence of incorrect particles and the weight equalization prevents degeneration of the weights.

3.2.4.2 Optical Flow Method

Optical flow is the pattern of motion in an image caused by the relative motion between the camera and the scene. The optical flow method aims to calculate the motion of the pixels between two images which are taken at times t and $t + \Delta t$.

Intuitively, given a set of points in an image, optical flow methods aim to find those same points in another image. Formally, given a point (u_x, u_y) in image I_1 , optical flow finds the point $(u_x + \delta_x, u_y + \delta_y)$ in image I_2 that minimizes ε :

$$\varepsilon(\delta_x, \delta_y) = \sum_{x=u_x-w_x}^{u_x+w_x} \sum_{y=u_y-w_y}^{u_y+w_y} (I_1(x, y) - I_2(x + \delta_x, y + \delta_y)) \quad (\text{Equation 3.5})$$

However, “which points shall be tracked?” is another question. Intuitively, the corner points or the objects which have textures are the best points for using in optical flow. Shi and Tomasi [22] describe an effective method for the detection of tracking points.

In this thesis, Pyramidal Lucas Kanade Optical Flow algorithm is implemented by using the features found by Shi and Tomasi feature detector.

3.2.4.3 Combination of Optical Flow and Particle Filter Methods

Particle filter method gives satisfactory results if the related object moves slowly. However, sudden movements of the objects result in failure if Gaussian window around the current state is used as transition model. Using the velocity and acceleration information also has drawbacks in sudden movement changes. Results of the particle filter method are shown in Figure 17.

In order to get satisfactory results, we need to estimate the movement of the related object. Speed and acceleration information may give hints for the next state but this method has drawbacks at the sudden movement changes.

Kodama et al. [34] propose the technique for tracking the objects by combining the particle filter and optical flow. Optical flow is used in the state transition step of the Particle filter technique. In this study, same technique is used for tracking in X-Y directions.

Figure 18 shows the results of the particle filter with optical flow feedback. Movement observed from the optical flow algorithm is used as an input in the state transition model of the particle filter.

3.2.4.4 Tracking with Range Gated Images

The combination of optical flow and particle filter methods is used for object tracking with range gated images. For each consecutive image, optical flow is used in the state transition model of the particle filter.

Different to the RGB images, range gated images are grayscale images and have more noisy characteristics, which need additional investigation for the usage of optical flow and particle filter. As shown in Figure 19, due to the different nature of the range gated images, optical flow generated high number of incorrect matches. As only a limited distance is illuminated, there are many dark pixels around the object adversely affecting the optical flow algorithm. As it is seen in Figure 19, optical flow method generates many motion vectors in consecutive images. To alleviate the effect of these incorrect matches in range gated image, we used an outlier removal step after optical flow. Firstly, the magnitudes of motion vectors are analyzed to remove the outliers. Then, the outlier points in the second image are removed. The motion vector representing the crude motion of the object is obtained by averaging the motion vectors of the remaining features.

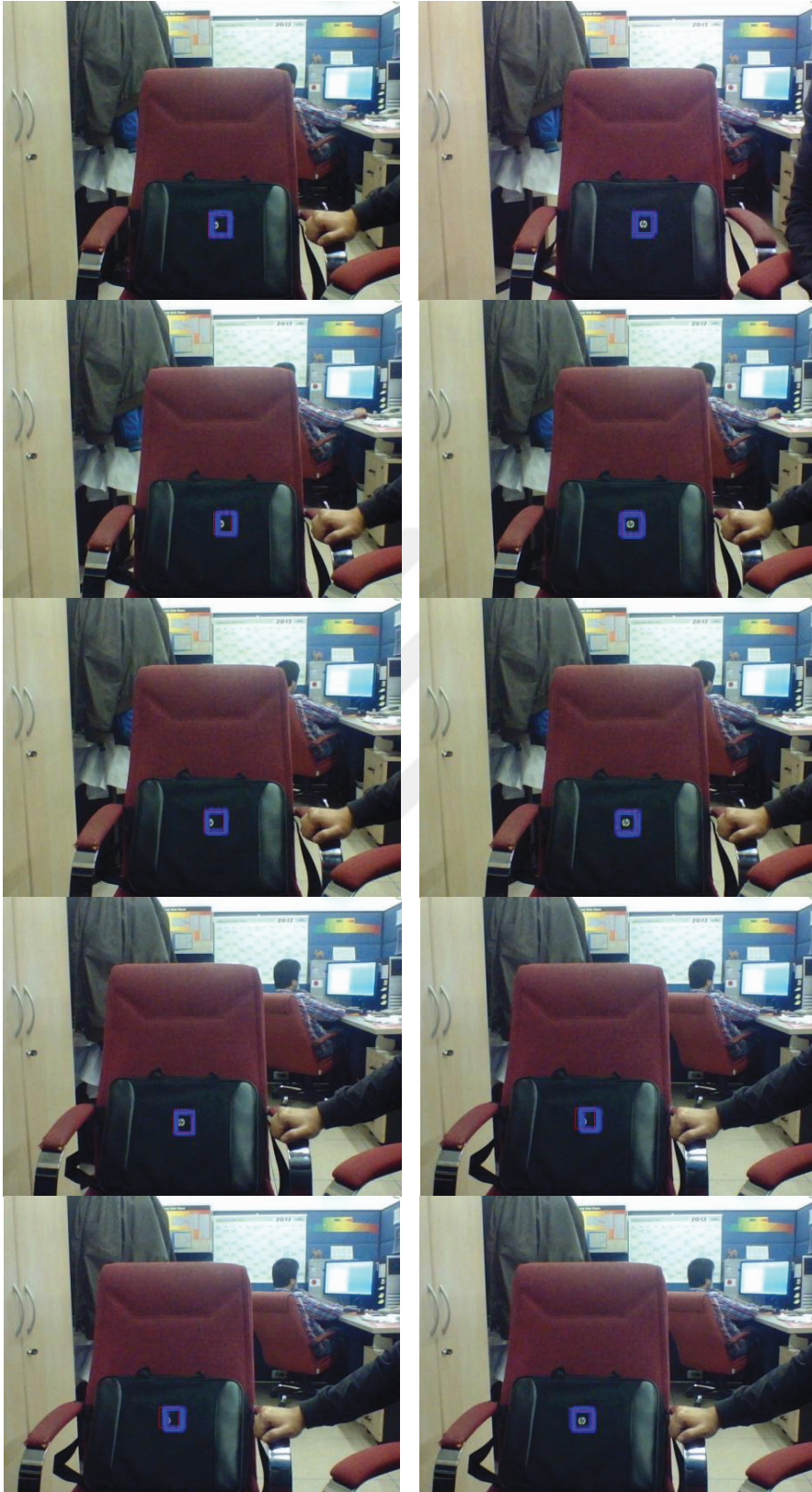


Figure 17 Result of the particle filter with Gaussian window around current state. Image order is from left to right, top to bottom



Figure 18 Optical flow (left) and Result of the particle filter with optical flow feedback (right)

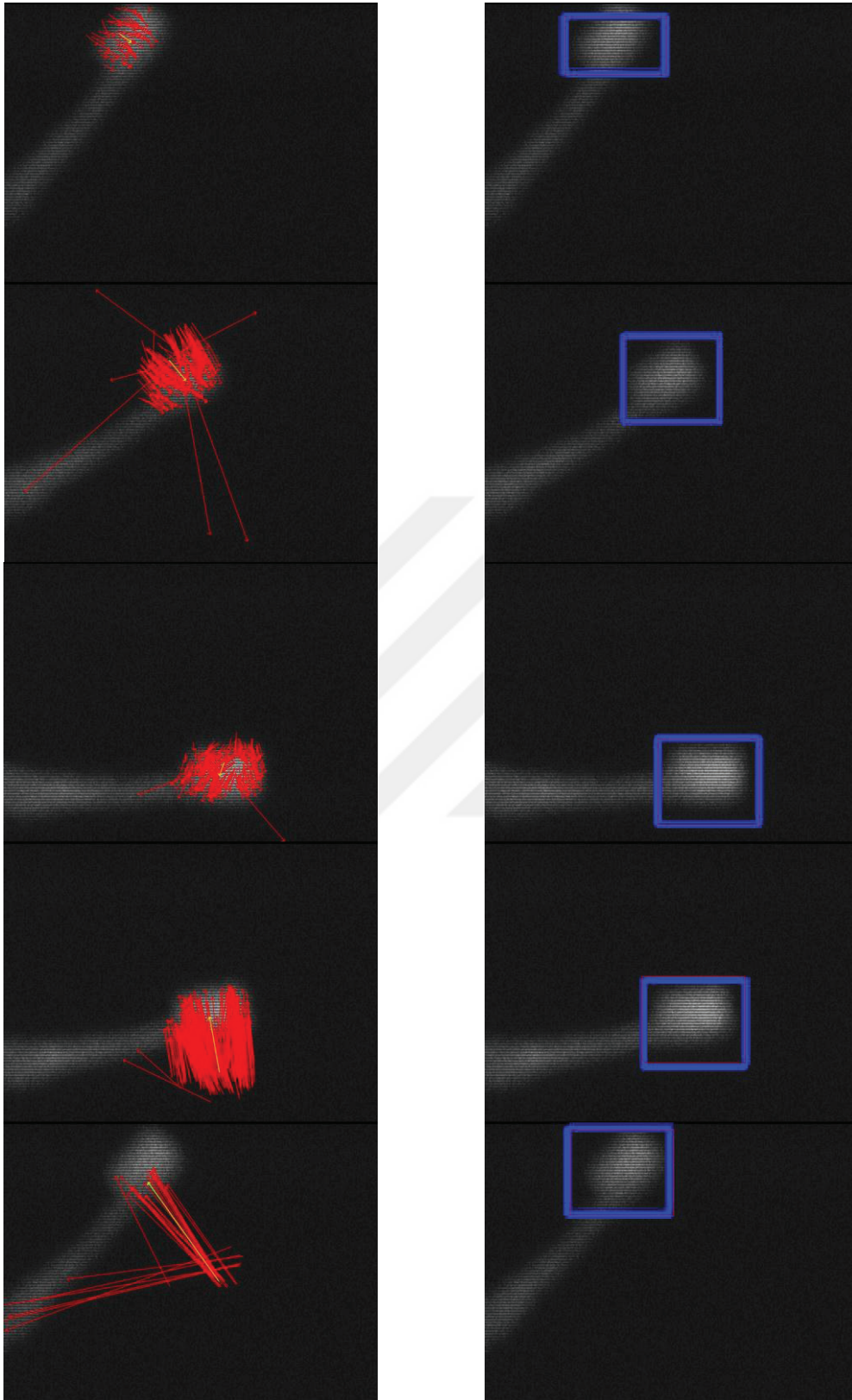


Figure 19 Optical flow output (left) and particle filter output with optical flow feedback (right)

CHAPTER 4

A ROBUST METHOD FOR REAL TIME OBJECT TRACKING IN RANGE GATED SYSTEM

In this chapter, details of the proposed algorithm for object tracking in a range gated system are provided. X-Y Tracking Module tracks the object at every frame. For each successful tracking, reference image is updated. If the tracking fails or a predefined time passes, Z Tracking Module intervenes to update the object distance. Figure 21 summarizes the whole algorithm and the following sections describe the details of the algorithm.

4.1 X-Y Tracking Module

For tracking the movement of the object in X and Y directions, an algorithm combination of Particle Filter and the Optical Flow described in Section 3.2.4.3.

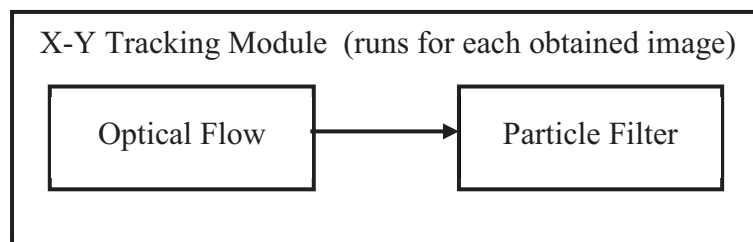


Figure 20 Summary of the X-Y Tracking Module. Optical flow determines the coarse movement and Particle filter makes the fine adjustment

Particle filter method without using the velocity information in transition model gives satisfactory results if the object is stationary or has small movement between consecutive frames. However, sudden movements of the objects result in failure if the Gaussian window around the current state is used as the transition model. Using

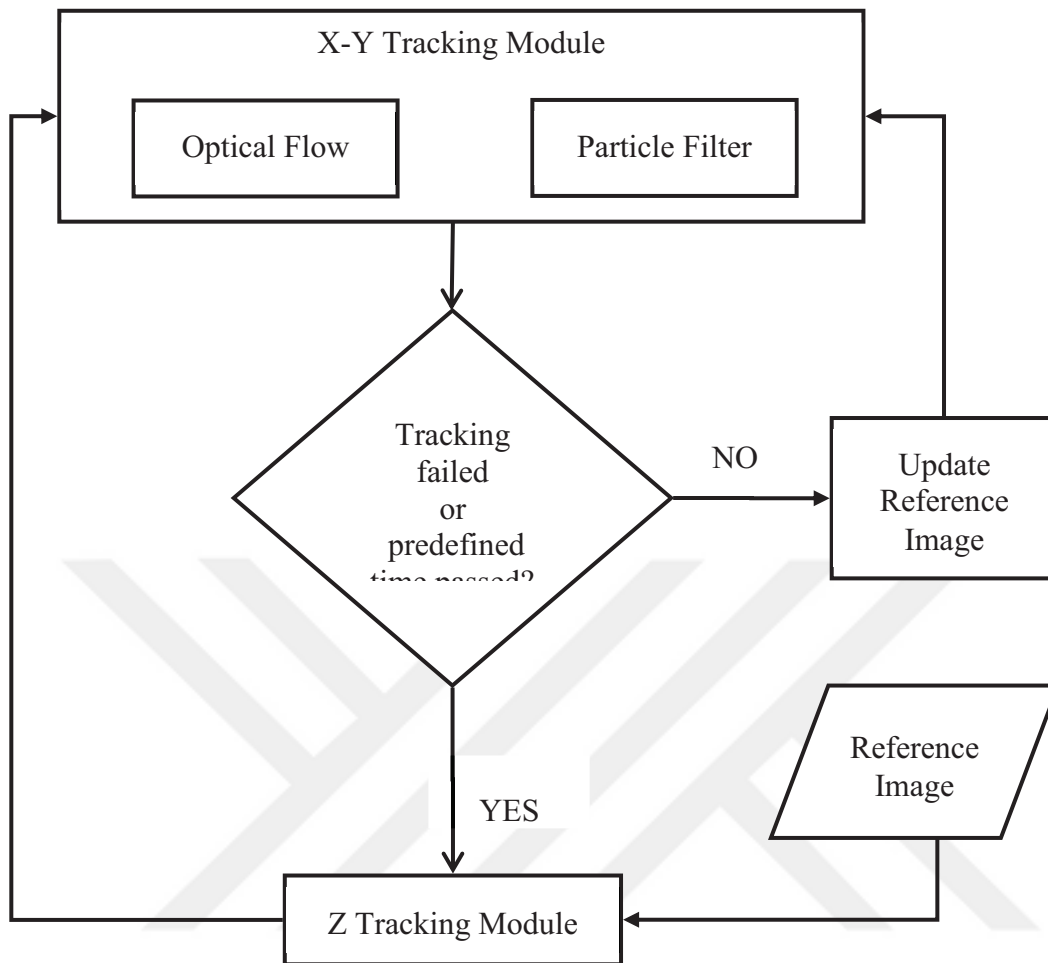


Figure 21 Summary of the Robust and Fast Image Tracking in Range Gated System

the velocity and acceleration information also has drawbacks in sudden movements of the object because velocity and acceleration feedback mislead the particle filter while their values are changing. In order to get satisfactory results, we need to estimate the movement of the related object. Kodama et al. [34] propose a technique for tracking the objects by combining particle filter and optical flow. Pyramidal Lucas Kanade Optical Flow algorithm is implemented by using the features found by Shi and Tomasi feature detector. Optical flow is used in the state transition step of the particle filter technique. In this study, we adopted this technique for tracking in X-Y directions.

4.2 Z-Tracking Module

While tracking the X-Y Movement of the object, tracking in the Z direction is a critical issue in range gated camera system. Since the moving object is illuminated by using the distance information between the object and the system, incorrect distance information results in the loss of the tracked object in the acquired image. Object distance d must be estimated in real time in order to continue to get new range-gated images at the object distance.

4.2.1 Ground-Truth for Object Location

For testing purposes, reference ground-truth distance of the object is calculated by analyzing the images taken from the system. For one position of the object, 50 images are taken for each delay time. Average image intensities for each delay time are shown in Figure 22. Since the reflected light reach maximum point at the correct object location, the delay time that has the maximum intensity value is used as the ground-truth location of the object. A polynomial is fitted to the average intensity values and the peak location is signed as the location of object.

4.2.2 Limitations

If the images of an object taken from the same delay value are investigated, fluctuations are observed on the brightness of the images. Images of a stable object with same delay time are shown in Figure 23. Main reasons of these fluctuations are due to inaccurate intensifier and laser timing precisions. While it can be argued that the changes in atmospheric attenuation and scattering might also affect, as the distance of the object is limited to 206 meters in this thesis, changes in atmospheric attenuation and scattering are not expected to have a significant effect.

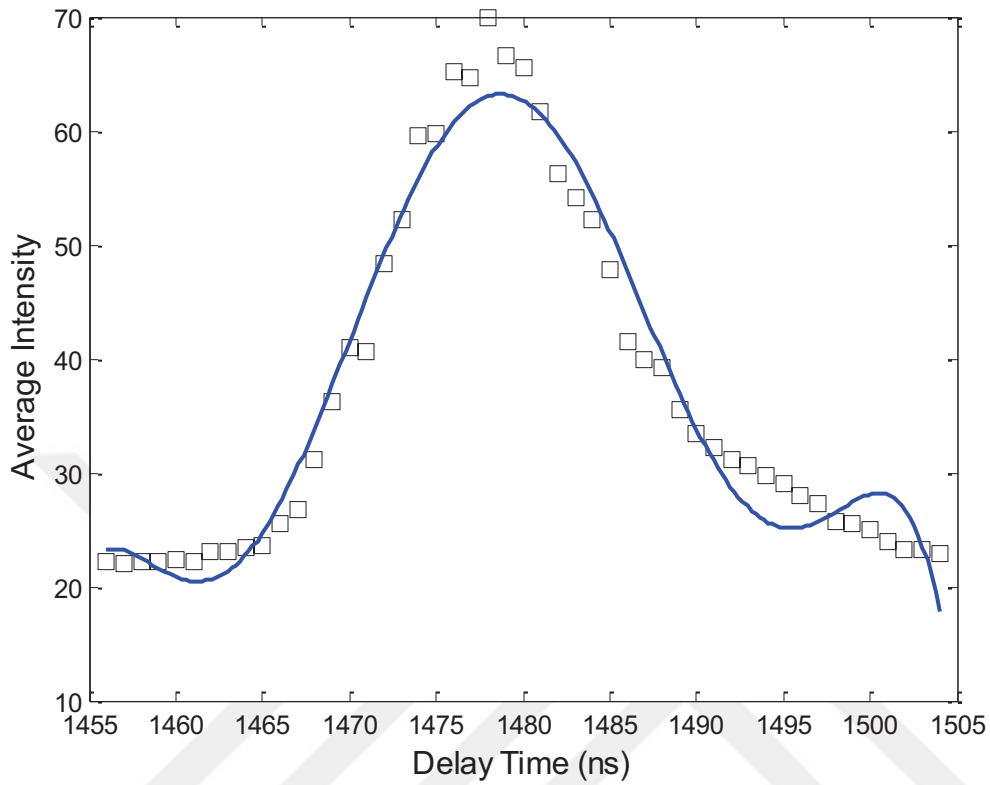


Figure 22 Average image intensities for each delay time are shown in black squares and the polynomial fitted to the black points is shown as line.

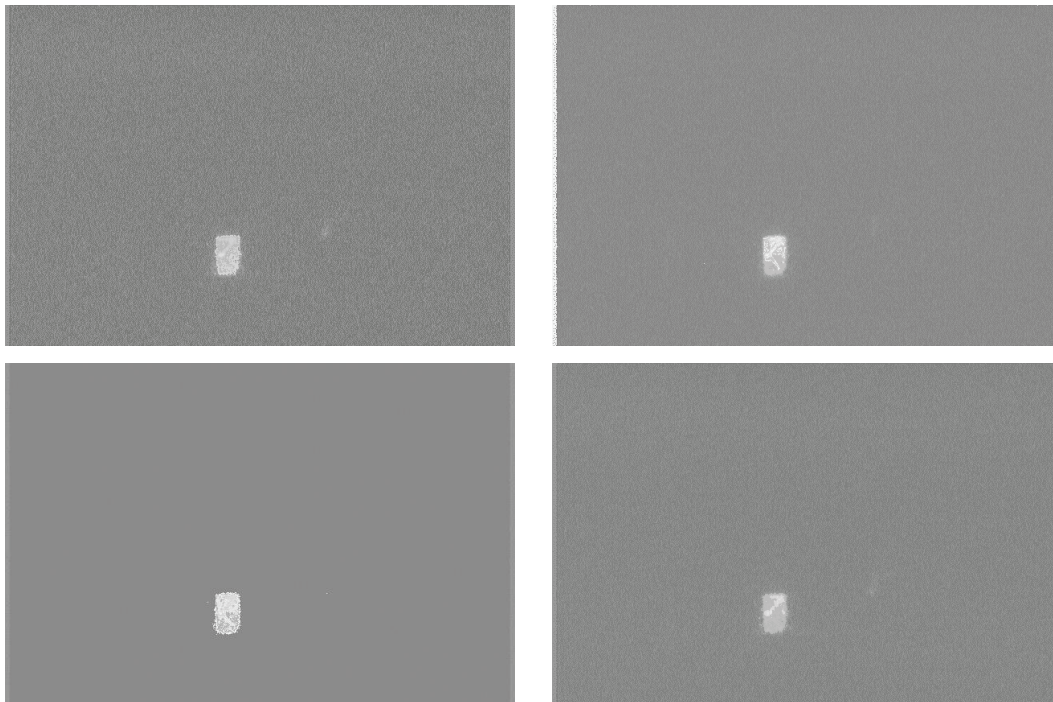


Figure 23 Four images taken with the same delay time, showing frame to frame brightness variation. For illustration purposes, dark pixels enhancement is applied by log transform.

After the measurements on laser unit, it is observed that the variance of the fluctuations in laser unit can be set to a Gaussian with variance=1ns.

While the camera producer states that the variance of the camera intensifier fluctuations is smaller than 1 ns, we take it as 1 ns to handle the worst case scenario.

Cumulative effect of two Gaussian noises with variance 1 ns results in a Gaussian noise with a variance 2 ns (which mean 1.41 ns standard deviation). Since light travels 30 cm/ns, in 1.41 ns, light travels 42.4 cm causing 21.2 cm standard deviation in the depth of view. In other words, 21.2 cm uncertainty in the z-axis on the system output arises from the imprecisions of the laser and camera units.

To analyze the system-wise effects of this uncertainty, noise characteristics of the system is analyzed on the acquired images. In Figure 24, the setup for system noise calculation is illustrated. Height of the observation tower is 12.5 meters; horizontal distance of the illuminated area to the tower is 201 meters.

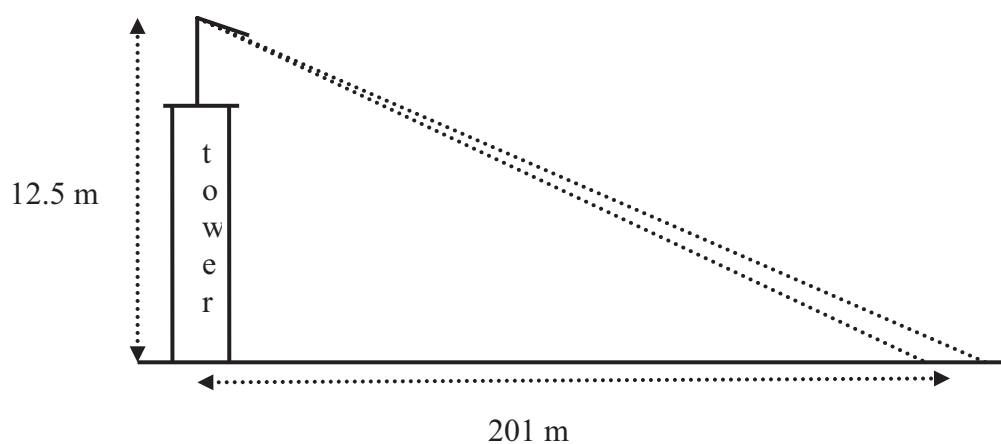


Figure 24 Setup for noise calculation

Since the objective FOV is 1.28 degrees and the vertical resolution of the camera is 480 pixels, field of view of one pixel is 0.0465 mrad. Images are taken from the

system and the centroids of the blobs are calculated for each image. In Figure 25, an image taken from the noise setup and the calculated centroid is illustrated.

For 13 images taken with the same system settings, illumination centroids were calculated. Illumination centroid is calculated as follows;

- Grayscale image shown in Figure 25 is converted to binary image with a predefined threshold. (For this study, threshold is selected as 38.25 by manual investigation on images).
- Centroid of the white blob is assigned as illumination centroid.

Figure 26 shows the distribution of the centroid locations. For each centroid location on image, exact location of the lightened area is calculated by the help of trigonometric calculations. Figure 27 shows the calculated lightened area distribution for these 13 centroids. Standard deviation of the illumination distance distribution was calculated as 302 millimetres. This noise value will be used in comparisons with system range accuracy standard deviation.

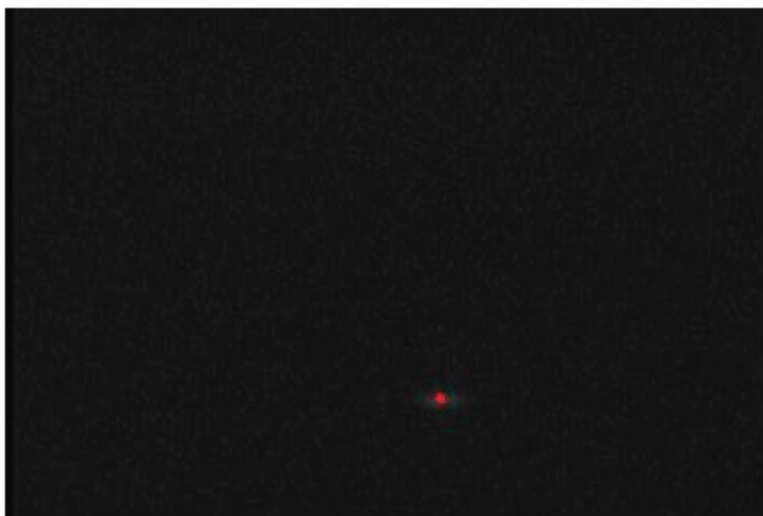


Figure 25 Image taken from the Noise Setup

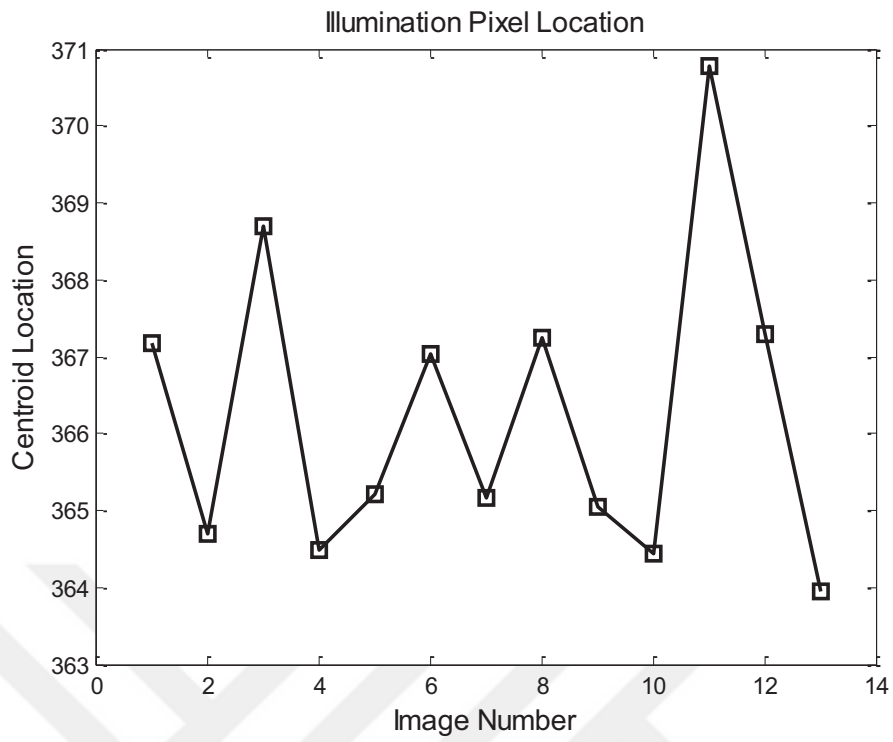


Figure 26 Illumination centroid distribution

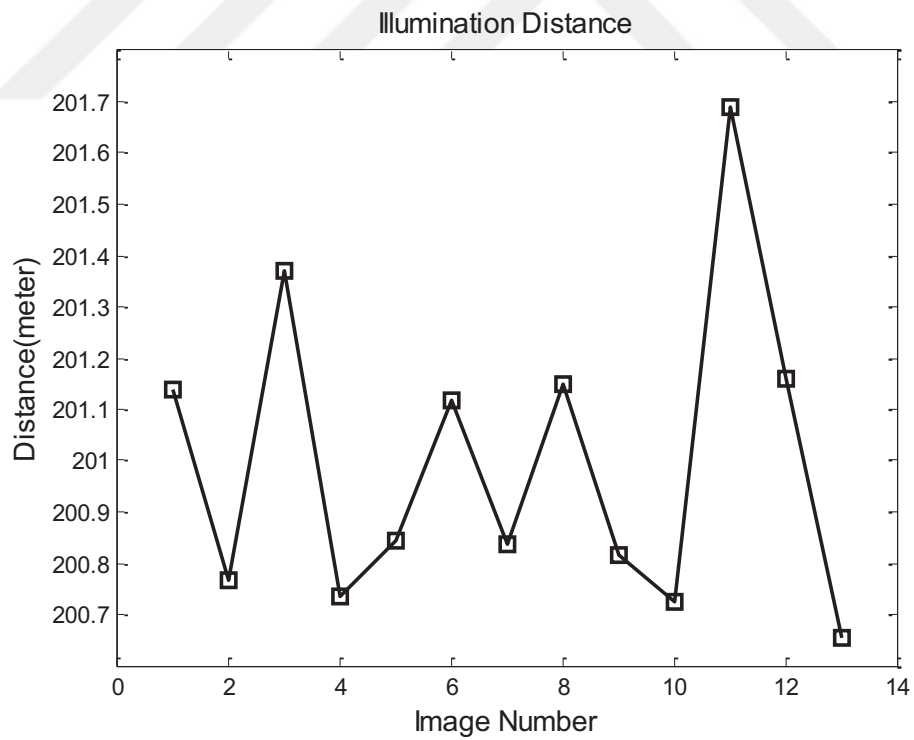


Figure 27 Illumination distance distribution

Another limitation regarding the experimental analysis of the system is the inaccuracy in the ground truth object distance calculation. The ground truth distance

of the object is calculated by taking samples with 1 ns precision at best that corresponds to 15 cm uncertainty.

4.2.3 Reference Image

After each successful tracking iteration, the reference image is updated with the new image in order to handle the shape changes of the tracked object. Since storing the whole set of reference images is difficult and the newer images have more weight on the reference image, formula specified in equation 4.1 is used for image storage.

$$REF = (1 - \alpha) \times REF + \alpha \times NEW, \quad (\text{Equation 4.1})$$

where REF stands for reference image and NEW stands for new image. α is a scalar between 0 and 1 and it is the weight of the new image while storing image. α is set to 0.3 for this project and it can be set to any value according to the object shape and movement characteristics. The optimum value of the α can be obtained by real time implementation and testing.

Changes in reference image with new successfully tracked images are shown in Figure 28. Reference image will be used at each step for the comparison with current image.

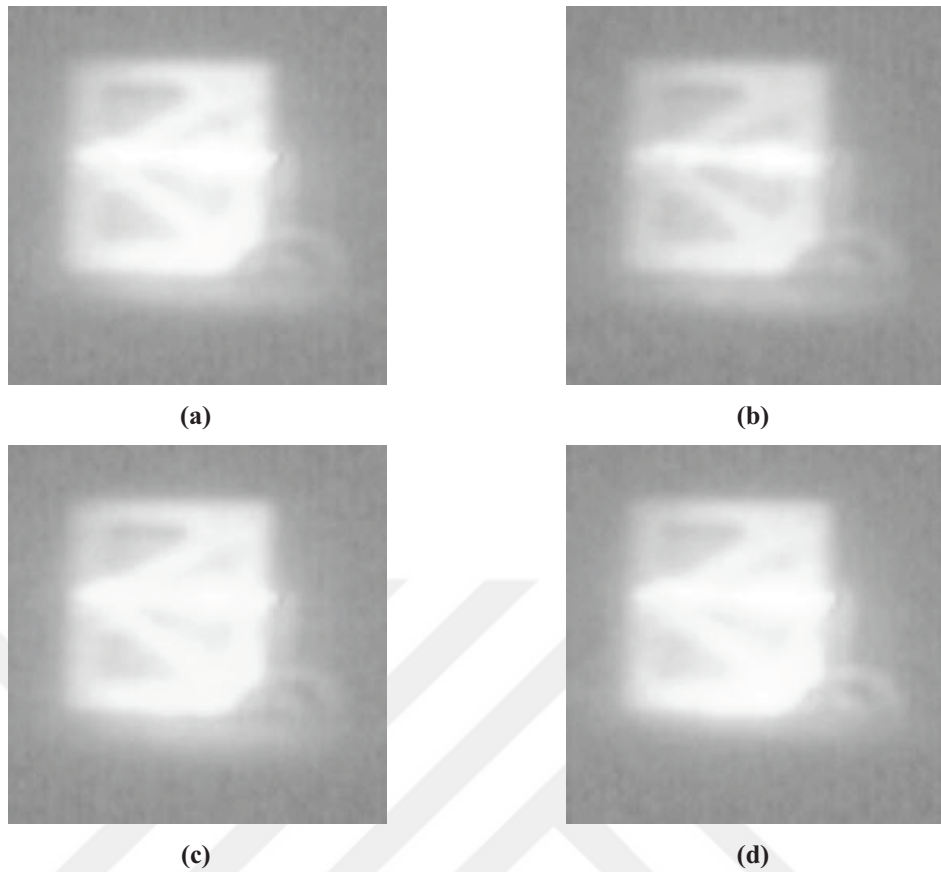


Figure 28 Changes in reference images with new tracked images. Initial reference image (a), and the updated reference images (b), (c) and (d). For illustration purposes, dark pixels enhancement is applied by log transform.

4.2.3.1 Partial Occlusion and System Robustness

If the object moves or another object occludes the current object, the reference image shall be updated to continue tracking without failure. Although Equation 4.1 supplies a robust algorithm to satisfy robustness while the object is stationary, reference image is blurred in moving or occluded object scenarios.

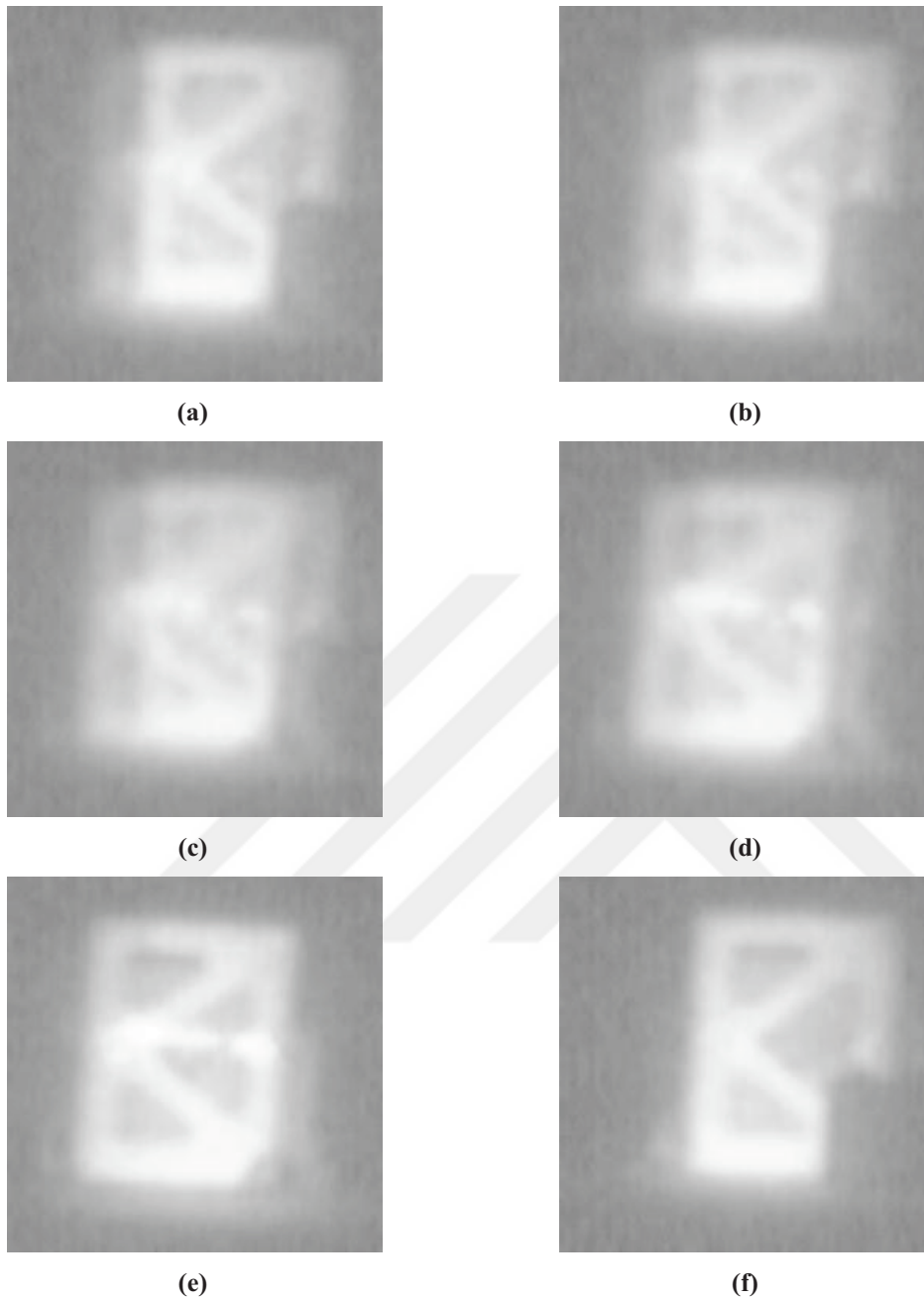


Figure 29 Changes in reference images. Object moves and partial occlusion occurs. (a) is the initial reference image, (b)-(e) are the reference images updated at each tracking iteration and (f) is the final reference image. For illustration purposes, dark pixels enhancement is applied by log transform.

Changes in the reference image with new 5 successful tracks are shown in Figure 29.

The result of the tracking in partial occlusion case is shown in section 5.1.4.

4.2.4 Algorithm

By only getting images from one gate depth (Δt), the distance changes cannot be tracked. To track the object continuously, our system gets images nearer from and further away of the tracked objects as shown in Figure 30.

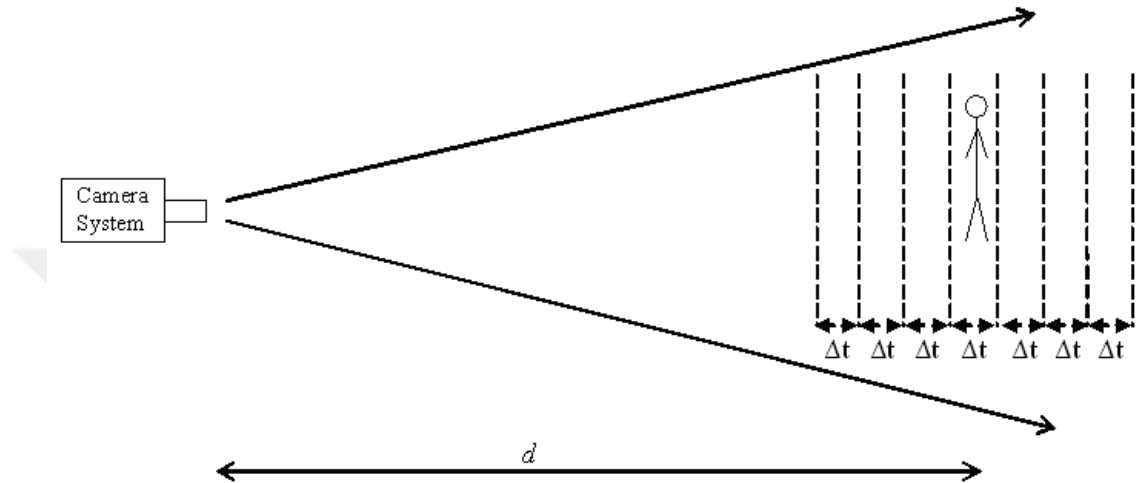


Figure 30 Image slices around the tracked object

The images taken from nearer and further away of the object are used to estimate the new location of the object. This distance is then used at the next iteration to determine the absolute location of the time slices. By updating the distance d at each cycle, system will have a capability of keeping track of the moving object even it moves towards or away from the camera. Figure 31 provides an illustration of the proposed object tracking algorithm for two iterations.

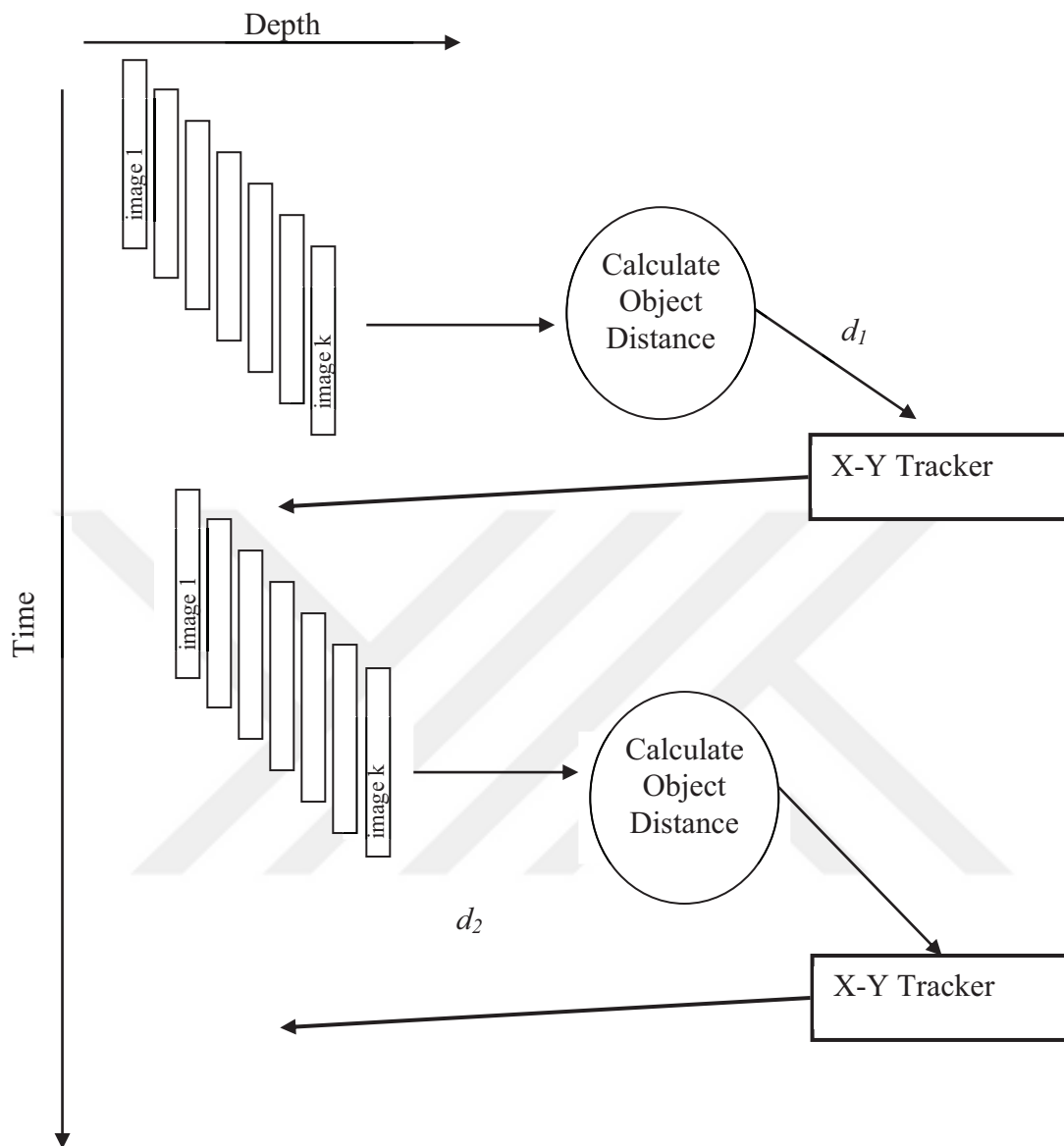


Figure 31 Summary of Object Tracking Algorithm

Object distance is calculated by the help of the methods described below. Histogram Comparison, Pixel-wise Comparison, Image Brightness, Mutual Information and Peak Signal to Noise Ratio methods are the mathematically defined metrics for image comparison. Beside these purely mathematical methods, a Human Visual System based method Structural Similarity is also tried for image comparison.

4.2.4.1 Histogram Comparison (HC)

In Figure 32, a number of sample images obtained using the system and their histograms are shown. For each candidate image, the histogram of the candidate image is compared with the histogram of the reference image to calculate a similarity value. Histogram Comparison (HC) of two images is calculated by summing up the intersection of two histograms. Each image is scored proportional to its likelihood with the reference image. Equation 4.2 gives the histogram comparison calculation of two images I_1 and I_2 . The reference image and all candidate images are scaled to the same width and height for all comparison methods.

$$HC = \sum_{k=0}^K \min(H_{I_1(k)}, H_{I_2(k)}), \quad (\text{Equation 4.2})$$

where H_{ref} is the histogram of the reference image, H_{cand} is the histogram of the candidate image, and K is the size of the histogram. K is equal to 255 in this study because the image data is 8-bit length for each pixel.

4.2.4.2 Pixel-wise Comparison (PwC)

Similarity between the current image and the reference image gives information about the object distance. If the delay time is closer to the actual object location, similarity will be higher and vice versa. Before the pixel-wise comparison, image is convolved with a Gaussian kernel to increase robustness to translation and image noise. Equation 4.3 gives the equation for calculating the likelihood between two images. Pixel-wise Comparison (PwC) is used as an indicator of the similarity between two images.

$$PwC(I_1, I_2) = \frac{1}{DIFF(I_1, I_2)},$$

(Equation 4.3)

$$DIFF(I_1, I_2) = \sum_{x=0}^M \sum_{y=0}^N (I_1(x, y) - I_2(x, y)),$$

where I_1 and I_2 are the two images, M and N are the row and column number of the images.

4.2.4.3 Image Brightness (IB)

Image Brightness (IB) is a good indicator of the distance of the object. If image brightness is high, the object is on the screen and the used delay is correct. If not, the distance of the object is not determined correctly.

$$IB(I) = \sum_{x=0}^M \sum_{y=0}^N I(x, y),$$

(Equation 4.4)

where I is the image, M and N are the height and width of the image respectively.

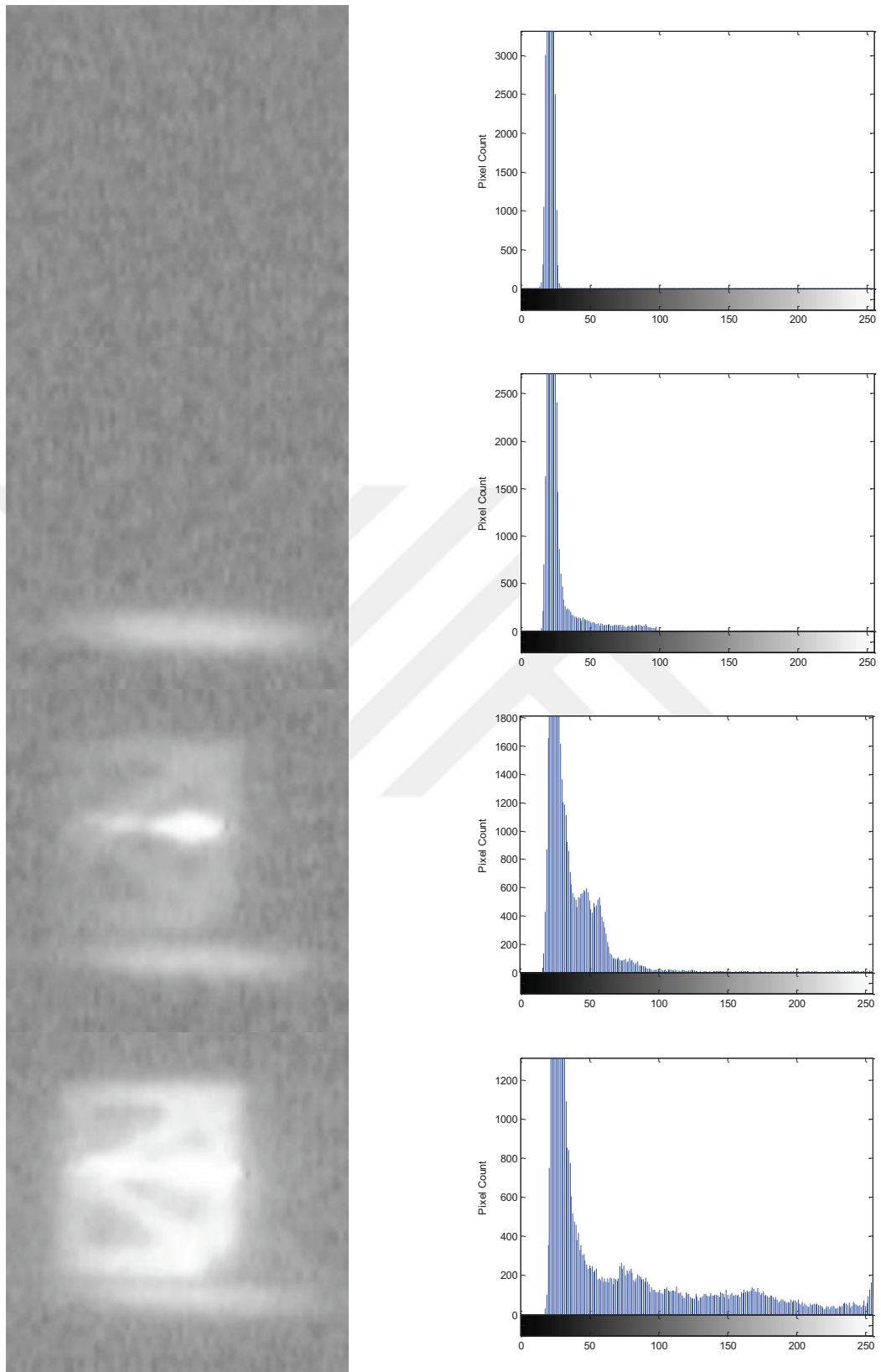


Figure 32 Sample images (left) and their corresponding histograms (right). For illustration purposes, dark pixels enhancement is applied by log transform.

4.2.4.4 Mutual Information (MI)

Mutual information has emerged as a similarity measure in the recent years, e.g. Russakoff et al. [41] who used Mutual Information for measuring the similarity between image regions. Mutual information of two images I_1 and I_2 is calculated by using both joint entropy $H(I_1, I_2)$ and the individual entropies $H(I_1)$ and $H(I_2)$.

$$H(I_1, I_2) = - \sum_{a,b} p_{I_1 I_2}(i_1, i_2) \log p_{I_1 I_2}(i_1, i_2), \quad (\text{Equation 4.5})$$

$$H(I_1) = - \sum_a p_{I_1}(i_1) \log p_{I_1}(i_1), \quad (\text{Equation 4.6})$$

where $p_{I_1 I_2}$ is the joint probability distribution of pixels associated with images I_1 and I_2 . By using the joint entropy and individual entropies defined in equation 4.5 and 4.6, mutual information is defined as:

$$MI(I_1, I_2) = H(I_1) + H(I_2) - H(I_1, I_2). \quad (\text{Equation 4.7})$$

4.2.4.5 Peak Signal to Noise Ratio (PSNR)

Peak Signal to Noise Ratio (PSNR) is commonly used to measure the quality of an original image and reconstructed image. PSNR is calculated using the Mean Squared Error (MSE) by;

$$MSE = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (I_1(x, y) - I_2(x, y))^2, \quad (\text{Equation 4.8})$$

$$PSNR = 10 \log_{10} \left(\frac{255^2}{MSE} \right), \quad (\text{Equation 4.9})$$

where I_1 and I_2 are the two images, M and N are the row and column number of the images.

4.2.4.6 Structural Similarity (SSIM)

Wang et al. [47] proposed an alternative complementary framework for quality assessment based on the degradation of structural information. To explore the structural information in an image, the luminance and contrast effects are removed from the image. Finally, luminance comparison, contrast comparison and the structure comparison are combined to obtain similarity measure between images. Summary of the structural similarity approach is shown in Figure 33.

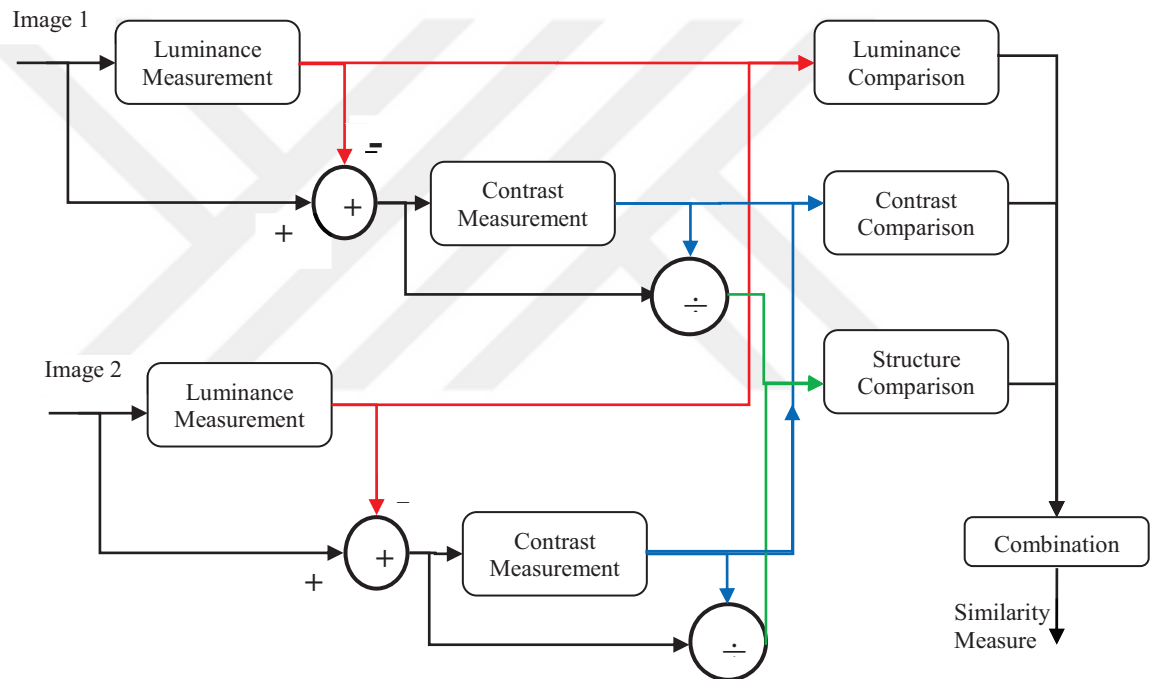


Figure 33 Diagram of structural similarity measurement system

Wang et al. [47] define the structural information in an image that represent the structure of objects in the scene, independent of the average luminance and contrast. They separate the task of similarity measurement into three comparisons: luminance, contrast and structure.

Luminance of each image is calculated as the mean intensity;

$$\mu_{I_1} = \frac{1}{N} \sum_{i=1}^N x_{iI_1}, \quad (\text{Equation 4.10})$$

where N is the image size. The luminance comparison function $l(I_1, I_2)$ is then a function of μ_{I_1} and μ_{I_2} as

$$l(I_1, I_2) = \frac{2\mu_{I_1}\mu_{I_2} + C_1}{\mu_{I_1}^2 + \mu_{I_2}^2 + C_1}, \quad (\text{Equation 4.11})$$

where the constant C_1 is included to avoid instability.

Secondly, Wang et al. [47] use the standard deviation as an estimate of the signal contrast. Standard deviation for a signal is calculated as;

$$\sigma_{I_1} = \left(\frac{1}{N} \sum_{i=1}^N (x_{iI_1} - \mu_{I_1})^2 \right)^{\frac{1}{2}}. \quad (\text{Equation 4.12})$$

The contrast comparison $c(I_1, I_2)$ is then a function of σ_{I_1} and σ_{I_2} as

$$c(I_1, I_2) = \frac{2\sigma_{I_1}\sigma_{I_2} + C_2}{\sigma_{I_1}^2 + \sigma_{I_2}^2 + C_2}, \quad (\text{Equation 4.13})$$

where the constant C_2 is also included to avoid instability.

Third, the structure comparison $s(I_1, I_2)$ is calculated by normalized signals $(I_1 - \mu_{I_1})/\sigma_{I_1}$ and $(I_2 - \mu_{I_2})/\sigma_{I_2}$. The correlation between these two is the measure to quantify the structural similarity. They stated that the correlation between $(I_1 - \mu_{I_1})/\sigma_{I_1}$ and $(I_2 - \mu_{I_2})/\sigma_{I_2}$ is equivalent to the correlation between I_1 and I_2 .

Thus, the structure comparison is defined as;

$$s(I_1, I_2) = \frac{\sigma_{I_1 I_2} + C_3}{\sigma_{I_1} \sigma_{I_2} + C_3}, \quad (\text{Equation 4.14})$$

where C_3 is added to avoid instability as in luminance and contrast measurements and $\sigma_{I_1 I_2}$ can be estimated as

$$\sigma_{I_1 I_2} = \frac{1}{N} \sum_{i=1}^N (I_{1i} - \mu_{I_1})(I_{2i} - \mu_{I_2}). \quad (\text{Equation 4.15})$$

Finally, the three comparisons metrics defined in equations 4.11, 4.13, and 4.14 are combined as;

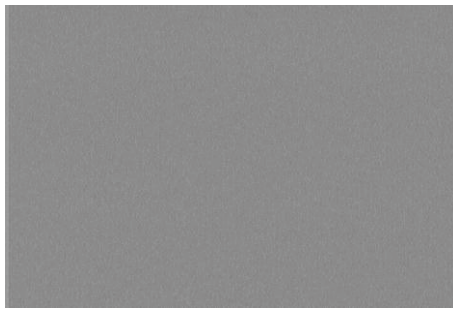
$$SSIM(I_1, I_2) = [l(I_1, I_2)]^\alpha [c(I_1, I_2)]^\beta [s(I_1, I_2)]^\gamma, \quad (\text{Equation 4.16})$$

where $\alpha > 0$, $\beta > 0$ and $\gamma > 0$ are parameters used to adjust relative importance of the methods. In this thesis α , β and γ values are all set to 1.

In this thesis, $c(I_1, I_2)$ is named as SSIM CNTRST, $s(I_1, I_2)$ is SSIM STR, $l(I_1, I_2)$ is SSIM LUM and $SSIM(I_1, I_2)$ is SSIM.

4.2.4.7 Algorithm Description

Performances of each of the methods stated above are investigated in this thesis. The object region calculated from X-Y Tracker is given as input to the Z-Tracker. In Figure 34, 11 images taken with 11 different delay times are shown. All 11 images are analyzed with Histogram Comparison, Pixel-wise Comparison, Image Brightness, Mutual Information, Peak Signal to Noise Ratio, and Structural Similarity point of views to determine the exact location of the object robustly. Performances of each stated methods are analyzed and the best one is proposed in this work.



(a)



(b)

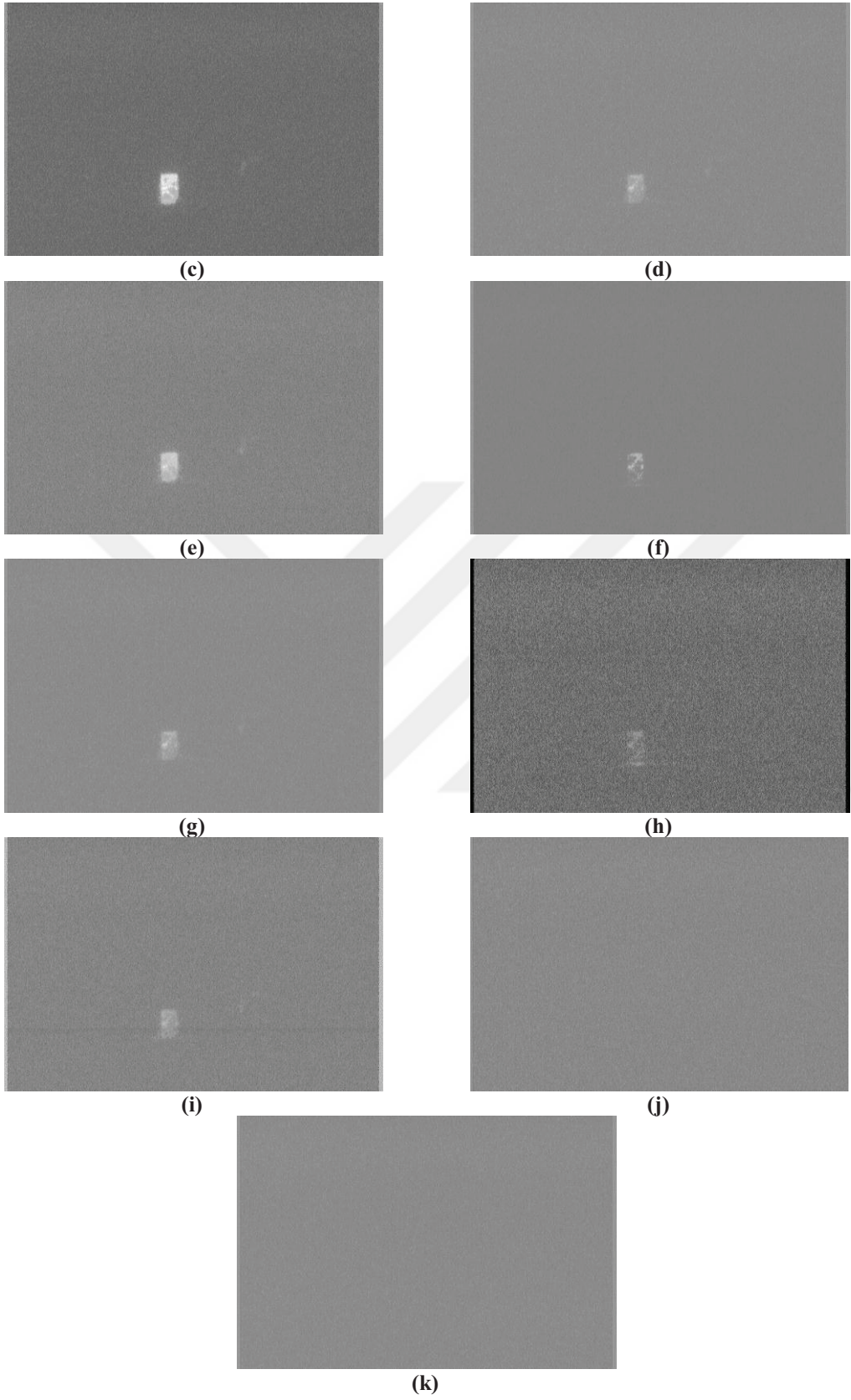


Figure 34 (a)-(k) 11 images taken with 11 different delay times. For illustration purposes, dark pixels enhancement is applied by log transform.

The object regions are selected from the 11 images. Object region is determined by the last successful X-Y tracker output. Some of the selected regions are shown in Figure 35.

These images are compared with the reference image by HC, PwC, IB, MI and SSIM metrics. Reference image is shown in Figure 36.

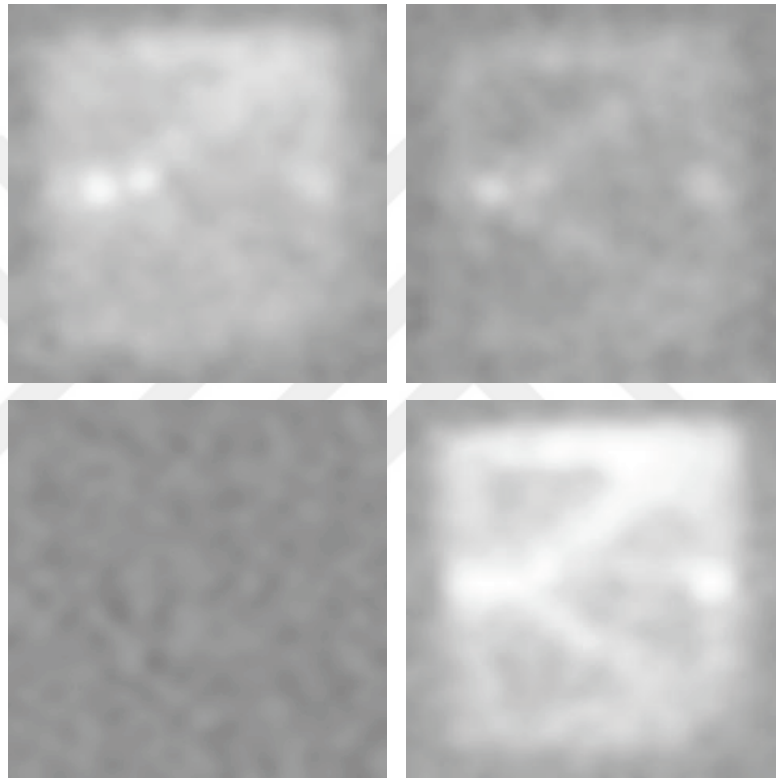


Figure 35 Some of the selected region images. For illustration purposes, dark pixels enhancement is applied by log transform.



Figure 36 Reference Image. For illustration purposes, dark pixels enhancement is applied by log transform.

In Figure 37, HC, PwC, IB, MI, PSNR and SSIM values for 11 images are shown. All values are normalized.

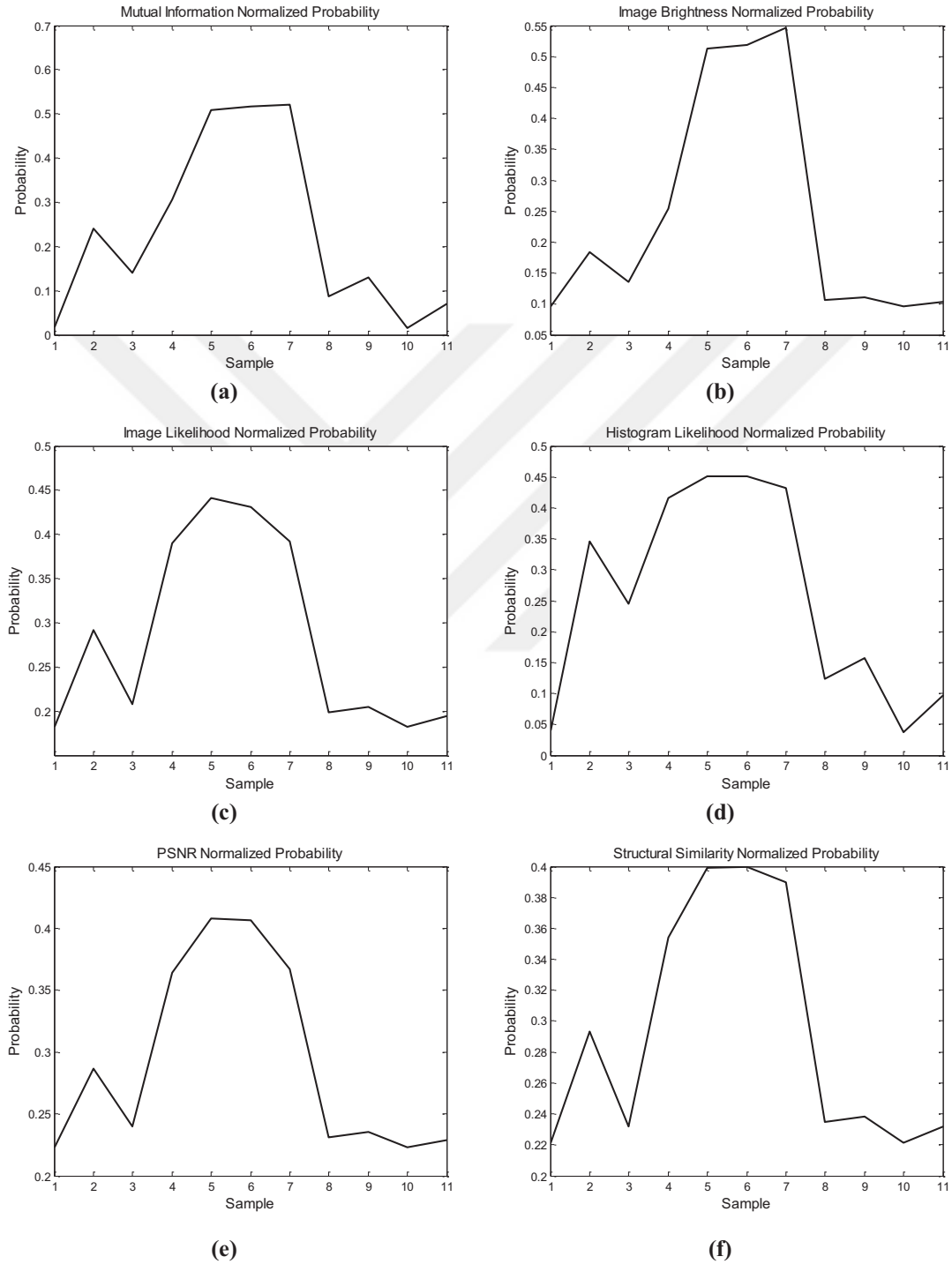


Figure 37 (a) Mutual Information, (b) Image Brightness, (c) Pixel-wise Comparison, (d) Histogram Comparison, (e) Peak Signal to Noise Ratio and (f) Structural Similarity metrics for 11 images

Next step is determining the location of the object by using these metrics. The location of the object could be determined by simply finding the maximum value of the metric and assigning the object location accordingly. However, it may cause errors and system failures if there is a high score (because of the noisy characteristic of the system) at incorrect locations. For example, in Figure 37 (b), according to IB method, the 7th image has the maximum likelihood and in Figure 37 (c), according to PwC method, the 5th image has the maximum likelihood. On the other hand, it is known that the object is located at the 6th location. To overcome this, mean value of the probability density function is used as the object location. Object locations are determined according to all approaches separately.

4.2.4.8 Interquartile Range

In the case object remains stationary and it is located at the same location at consecutive tracking iterations, getting the mean value of the probability density function of the 11 images accurately detects the location of the object. If the object moves then the probability density function is skewed as shown in Figure 38 where the correct object location is at around sample 7.33. Detecting the location by the mean value of the density tends to converge to a location near the previously known location (sample 6) because of the weight of the first four samples. To overcome this problem, it is desirable to eliminate the results which have low likelihood. In this example, eliminating the samples 1, 2, 3, 4, 10 and 11 would result in a better estimation of the location of the object.

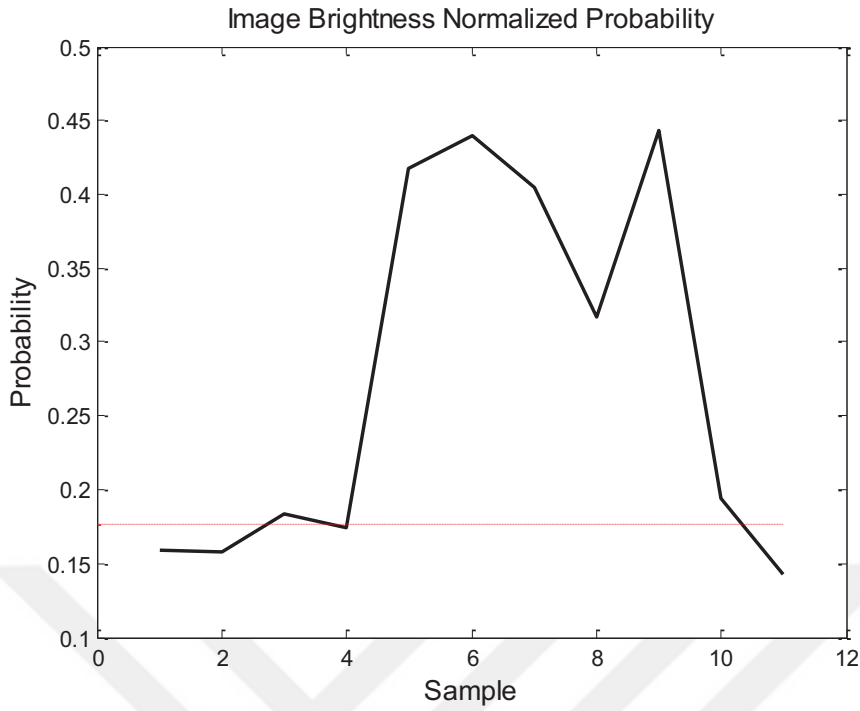


Figure 38 An example probability density function for 11 images in the case of object moving in one direction.

Interquartile Range is an easy but effective method for finding outliers in a data. Emeksiz and Temizel [40] used this technique to find the outliers in an object tracking setting. To apply Interquartile Range, the sample set is defined with the 11 probability values. Let X_s be the set of all probability values; $X_s = \{X_1, X_2, \dots, X_{11}\}$. After obtaining the sample set, the values are put in order in ascending order. So, the new set can be represented as $X_s' = \{X'_1, X'_2, \dots, X'_{11}\}$. Then, quartiles are found. These quartiles are Q_1 , Q_2 (median) and Q_3 . Quartiles are calculated with the equations 4.17, 4.18 and 4.19.

$$Q_1 = X'_3, \quad (\text{Equation 4.17})$$

$$Q_2 = X'_6, \quad (\text{Equation 4.18})$$

$$Q_3 = X'_9, \quad (\text{Equation 4.19})$$

The outliers o_1 and o_2 are calculated with equations 4.20 and 4.21 respectively.

$$o_1 = Q_1 - k(Q_3 - Q_1), \quad (\text{Equation 4.20})$$

$$o_2 = Q_3 + k(Q_3 - Q_1), \quad (\text{Equation 4.21})$$

where k is a constant.

In this thesis, a new threshold value o_3 is defined as in equation 4.22. The probability values lower than the outlier o_3 are ignored. The mean values of the remaining images are used to calculate mean value. Results are shown in Section 5.1.2.

$$o_3 = Q_2 - k(Q_3 - Q_1). \quad (\text{Equation 4.22})$$

4.2.5 Testing Environment

To develop Z-tracker algorithm, data acquired using the constructed setup. As shown in Figure 39, an object having a characteristic pattern was designed. On the floor, 22 locations were marked in 90 cm intervals. Ground truth object location is calculated for one location and the other locations are calculated by using the distance among marks. Images were taken from the tower shown in Figure 40.



Figure 39 Test object to be tracked

Once the object is located in one of these positions, 15 images were taken from each delay time. In other words,

15 images are taken with delay= $d-5\Delta t$

15 images are taken with delay= $d-3\Delta t$

15 images are taken with delay= $d-4\Delta t$

...

...

15 images are taken with delay= $d+3\Delta t$

15 images are taken with delay= $d+4\Delta t$

15 images are taken with delay= $d+5\Delta t$,

where d is the known object position. For test setup, Δt is selected as 3 nanoseconds which corresponds to a 45 cm distance window.



Figure 40 Image showing the test object and the tower



CHAPTER 5

EXPERIMENTS AND RESULTS

In this chapter, evaluation of the proposed method and the performances of the comparison algorithms are investigated. Test scenarios are created and the result of the tests are illustrated in this section.

5.1 Method Evaluation

In the scope of method evaluation, following two scenarios were implemented:

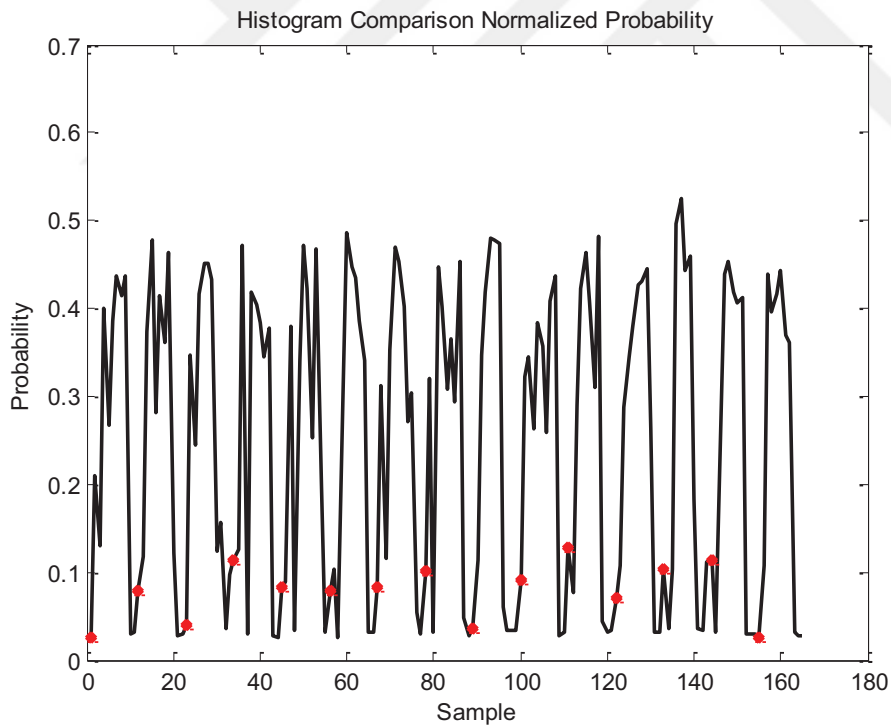
- Stationary object scenario: Object was located on the 1st marker, images were taken with 11 different gate delay time settings (15 images from each setting) covering a ± 225 cm distance centered on the object location.
- Moving object scenario: Object was located on the 5th marker, images were taken with 11 different gate delay time settings (15 images from each setting) covering a ± 225 cm distance centered on the 60 cm front of the object location. This test data simulates the scenario that object moved 60 cm from the last known position.

As a result, 15x11 images (15 set of images) were collected for the stationary object scenario and 15x11 images (15 set of images) for the moving object scenario.

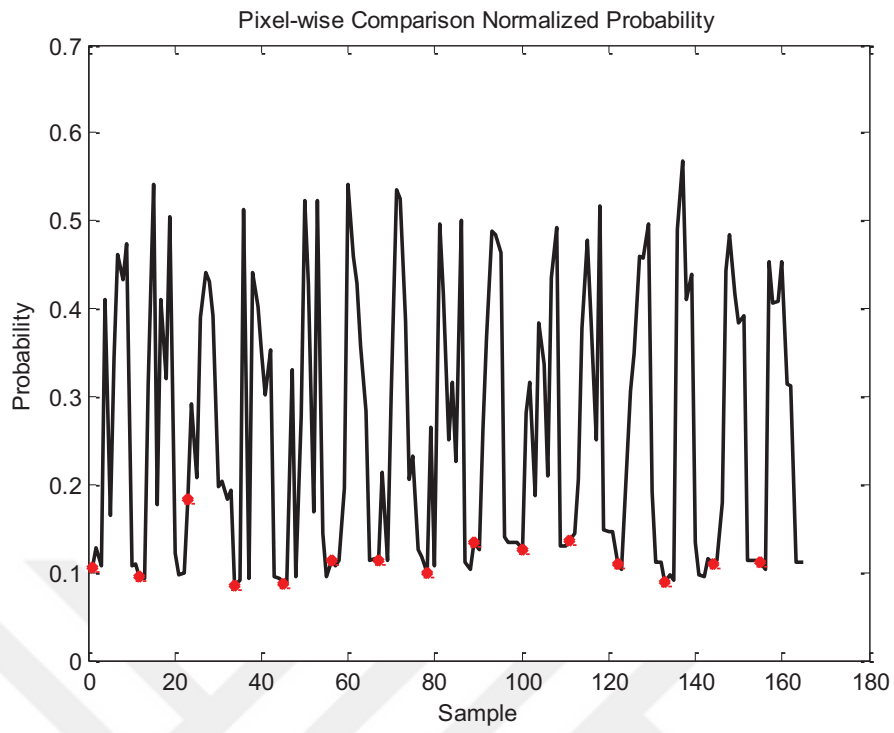
Beside the two scenarios, performances of the methods are investigated by Brightness Invariance, Translation Invariance and Artificial Object point of views.

5.1.1 Experiment 1: Stationary Object Scenario

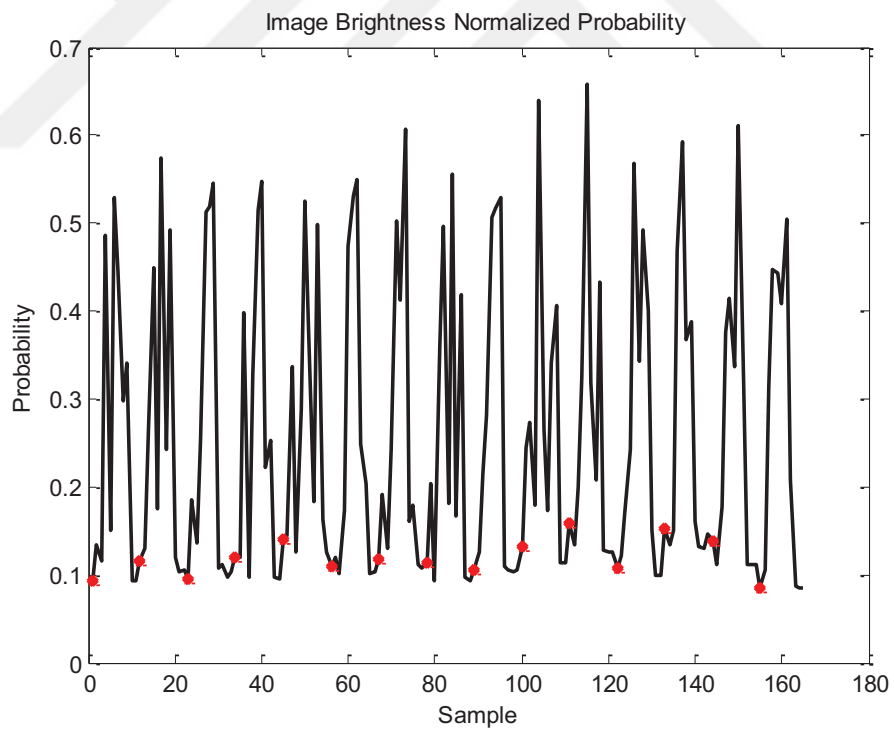
15 set of images (each set consists of 11 images taken with different gate delay times) are used for testing the performance of the algorithm for stationary object scenario. For each set of images, object location is calculated and the observed location is compared with the known position of the object. Figure 41 illustrates the probability density function of 15x11 images (15 set of images, 165 images in total). Images 1-11 are from the 1st set, 12-22 are from the 2nd set, 23-33 are from the 3rd set and so on. First images of each set are marked by a Red Dot sign. For each set of images, object location is estimated as described in Section 4.2.



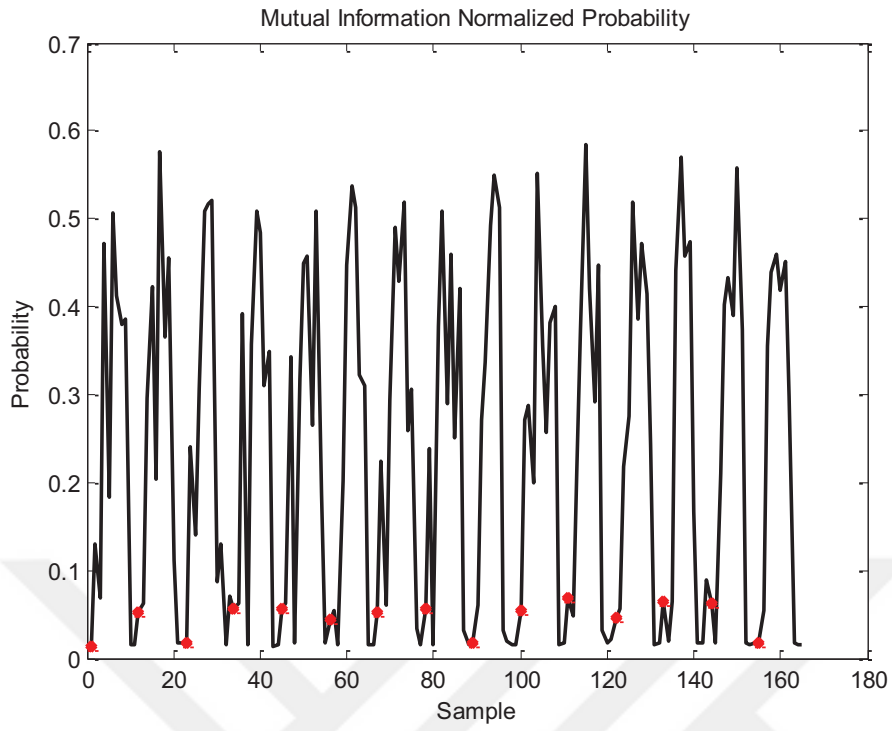
(a)



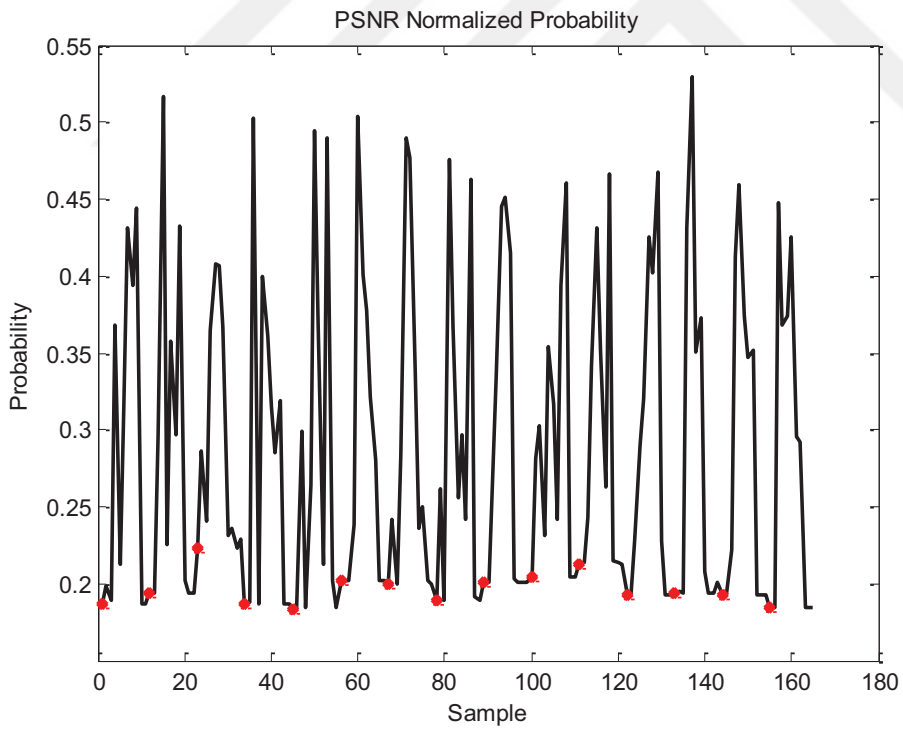
(b)



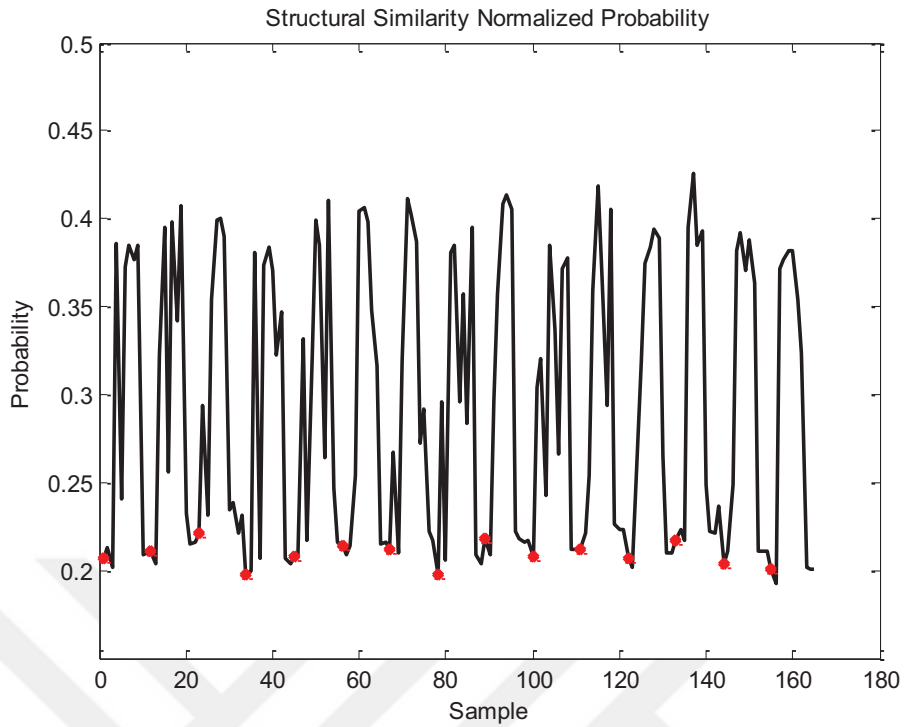
(c)



(d)



(e)



(f)

Figure 41 Histogram Comparison (a), Pixel-wise Comparison (b), Image Brightness (c), Mutual Information (d), Peak Signal to Noise Ratio (e) and the Structural Similarity (f) metrics for 15 set of images (each image set consists of 11 images captured with different delay times)

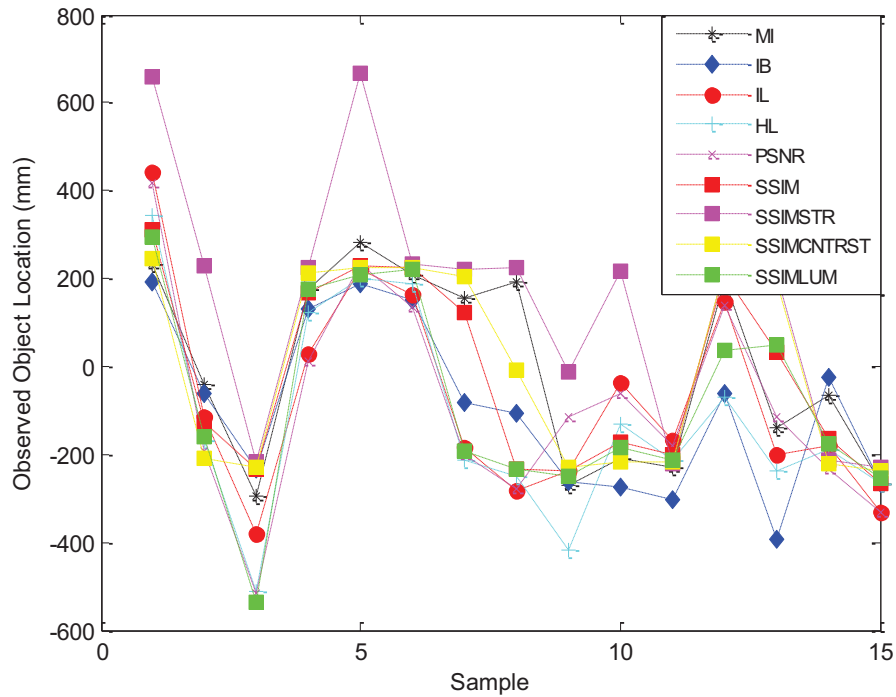


Figure 42 Error values (in mm) for Histogram Comparison, Pixel-wise Comparison, Image Brightness, Mutual Information and Structural Similarity methods for 15 set of images. (Interquartile range method with $k=0.02$ is applied) (Stationary object scenario)

Error values for each method and the system result is shown in Figure 42. As can be seen in Figure 42, these fluctuations affect all methods similarly. The main reasons for the oscillations in calculated errors are the fluctuations in image timings described in Section 4.2.2.

For the 15 set of images, specified methods are applied and the error values are calculated. Mean and standard deviation of the error values of 15 set of images are shown in Table 1.

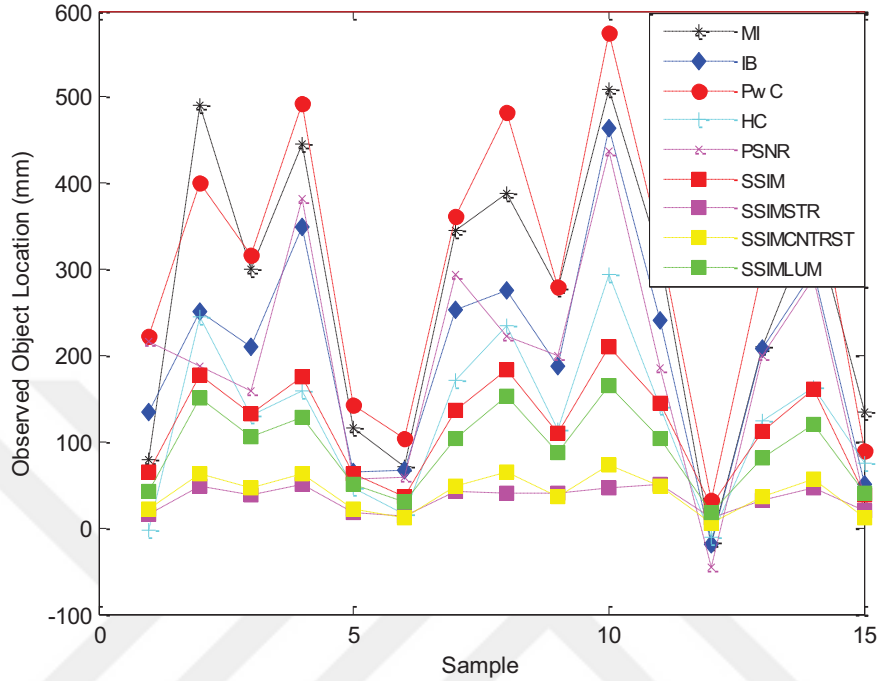
Table 1 Error mean and STD values for stationary object scenario. Bold is used to highlight the best result.

Method	Error Mean (mm)	Error STD (mm)	RMS Error (mm)
HC	-120.6	237.7	259.4
PwC	-74.5	231.9	236.1
IB	-94.2	193.1	209.0
MI	-4.7	217.2	209.9
PSNR	-86.4	239.8	247.2
SSIM	-23.0	212.3	206.3
SSIM STR	147.5	285.4	312.7
SSIM CNTRST	-3.0	221.0	213.5
SSIM LUM	-80.8	232.6	238.8

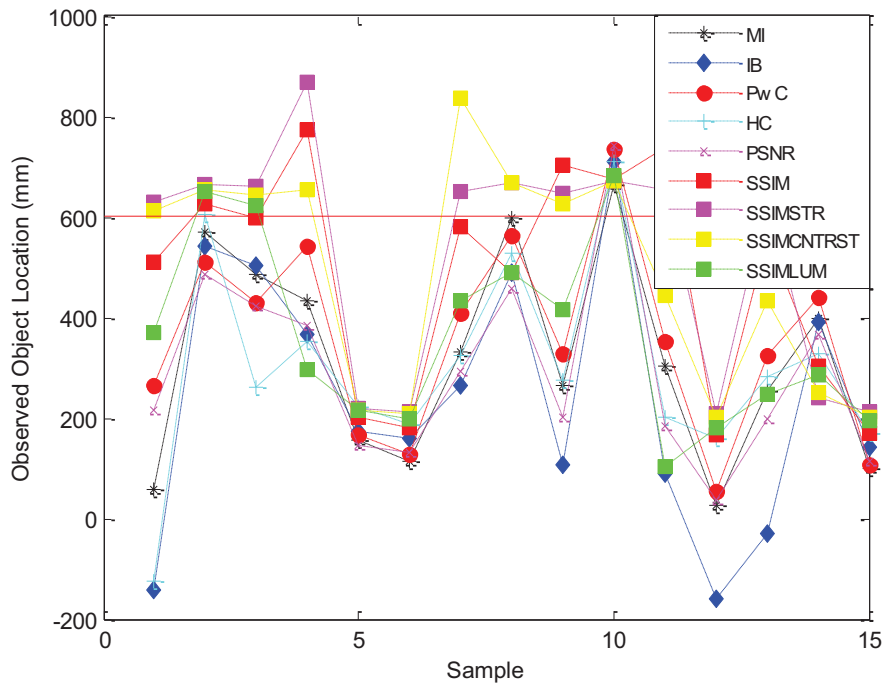
5.1.2 Experiment 2: Moving Object Scenario

To analyze the real time response of the system, object is located 600 mm away from the last known location of the object in z direction. However, the displacement of the object is calculated as 118 mm (+482 mm estimation error) for SSIM and 202.2 mm (+397.8 mm estimation error) for IB method. High estimation errors are observed also for all the other methods. The reason for calculating the object location nearer than the actual position is described in section 4.2.4.8. To overcome this problem, some of the samples that have probability smaller than a threshold value shall be excluded from the mean calculation. In this study, a modified version of interquartile range method is proposed to be used for elimination of these samples. Figure 43

illustrates the error in estimated object locations for 15 set of samples, with and without the effect of interquartile range method.



(a)



(b)

Figure 43 Observed object locations (in mm) for Histogram Comparison, Pixel-wise Comparison, Image Brightness, Mutual Information, Peak Signal to Noise Ratio and Structural Similarity methods for 15 set of images. Object is located at 600 mm away from the previous location. Interquartile range method is not applied (a) and applied with $k = 0.02$ (b).

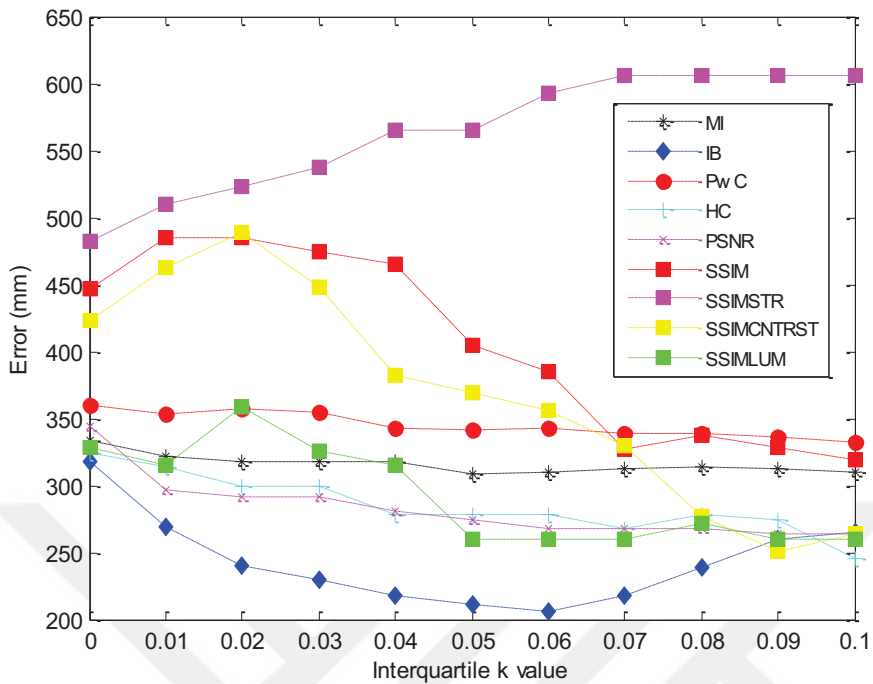


Figure 44 Observed object locations (in mm) for Histogram Comparison, Pixel-wise Comparison, Image Brightness, Mutual Information, Peak Signal to Noise Ratio and Structural Similarity methods for 15 set of images. Interquartile range method is applied with different k values. Object is located at 600 mm away from the previous location.

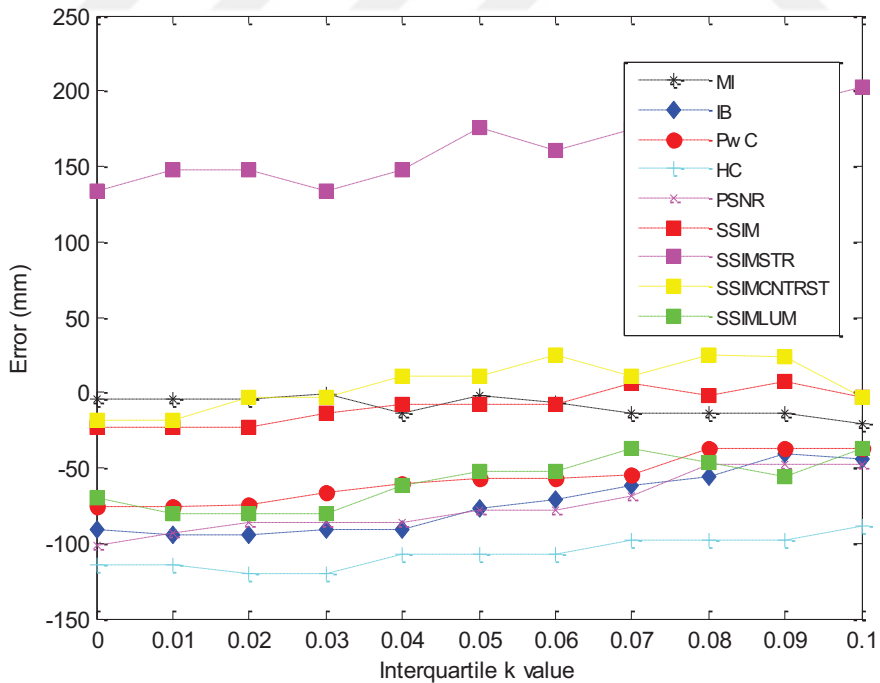


Figure 45 Error values (in mm) for Histogram Comparison, Pixel-wise Comparison, Image Brightness, Mutual Information, Peak Signal to Noise Ratio and Structural Similarity methods for 15 set of images. (Interquartile range method is applied with different k values) Object is located at the same previous location.

To determine the best interquartile k value, average object position for each k value is calculated and shown in Figure 44. For k=0.02, SSIM, SSIM LUM and SSIM CNTRST has the minimum error and for k=0.07, SSIM STR has the minimum error. For analysis, this two k values are used. Same analysis is applied to stationary object scenario as shown in Figure 45 and it is observed that changes in interquartile k value does not affect the error value significantly.

To investigate the effect of interquartile range method to stationary object scenario, error values are calculated with the same interquartile k values. Results are shown in Table 4. Obtained error values are used for calculating the success rates as described in Section 5.2.1.2. Table 3 and Table 5 show the RMS error for all methods for moving and stationary object scenarios.

Table 2 Error Mean and STD values for different methods. Object is located at 600 mm away from the previous location.

Method	No Interquartile		Interquartile with k = 0.02		Interquartile with k = 0.07	
	Error Mean(mm)	Error STD(mm)	Error Mean(mm)	Error STD(mm)	Error Mean(mm)	Error STD(mm)
HC	-473.2	90.8	-301.1	200.8	-332.0	226.8
PwC	-296.9	160.3	-243.0	191.1	-261.7	189.0
IB	-397.8	126.9	-359.2	256.7	-382.3	231.1
MI	-332.0	162.8	-282.8	204.4	-288.1	193.4
PSNR	-408.5	130.3	-308.5	184.8	-331.8	188.7
SSIM	-482.0	60.0	-115.0	221.3	-273.4	203.1
SSIM STR	-565.7	14.3	-76.4	230.1	6.5	247.5
SSIM CNTRST	-559.6	21.8	-111.3	220.7	-270.3	194.6
SSIM LUM	-508.1	47.7	-240.8	184.7	-339.7	238.2

Table 3 RMS Error values for different methods. Object is located at 600 mm away from the previous location. Bold is used to highlight the best result.

Method	No Interquartile RMS Error (mm)	Interquartile with k = 0.02 RMS Error (mm)	Interquartile with k = 0.07 RMS Error (mm)
HC	481.3	358.2	397.8
PwC	334.9	305.1	319.1
IB	416.3	436.5	442.7
MI	367.4	344.9	343.4
PSNR	427.4	356.4	378.5
SSIM	485.5	242.8	336.5
SSIM STR	565.8	235.1	239.2
SSIM CNTRST	560.0	240.6	329.3
SSIM LUM	510.1	299.7	410.3

Table 4 Error Mean and STD values for different methods. Object is located at the same previous location.

Method	No Interquartile		Interquartile with k = 0.02		Interquartile with k = 0.07	
	Error Mean(mm)	Error STD (mm)	Error Mean (mm)	Error STD (mm)	Error Mean (mm)	Error STD (mm)
HC	-113.7	181.4	-120.6	237.7	-97.6	239.9
PwC	-25.9	180.0	-74.5	231.9	-54.1	216.4
IB	-49.8	150.0	-94.2	193.1	-62.1	180.6
MI	-42.4	179.0	-4.7	217.2	-14.1	194.5
PSNR	-33.4	150.4	-86.4	239.8	-68.5	216.1
SSIM	-1.0	97.0	-23.0	212.3	5.8	183.0
SSIM STR	13.2	16.7	147.5	285.4	174.9	240.3
SSIM CNTRST	-1.3	18.1	-3.0	221.0	10.8	195.5
SSIM LUM	-17.6	84.5	-80.8	232.6	-37.6	200.5

Table 5 RMS Error values for different methods. Object is located at the same previous location. Bold is used to highlight the best result.

Method	No Interquartile RMS Error (mm)	Interquartile with k = 0.02 RMS Error (mm)	Interquartile with k = 0.07 RMS Error (mm)
HC	208.9	259.4	251.5
PwC	175.8	236.1	215.9
IB	153.3	209.0	185.2
MI	178.1	209.9	188.4
PSNR	149.1	247.2	219.7
SSIM	93.7	206.3	176.9
SSIM STR	20.8	312.7	290.7
SSIM CNTRST	17.6	213.5	189.2
SSIM LUM	83.5	238.8	197.4

5.1.3 SSIM STR Analysis

For both stationary and moving object scenarios, SSIM STR method has higher error in positive Z direction compared to the other methods. Besides that, SSIM STR method has drawbacks to translation of the object as described in section 5.1.6. Reasons and possible solutions are investigated in this section.

In Figure 46, a typical first image and a typical last image is shown. Ground at the front of the object is illuminated in first images. It should be stated that, this characteristic comes from the data collection setup. Since the camera and laser unit are located at tower and the object is located at ground, front of the object is illuminated for lower delay times but behind of the object is not for higher delay times. If the object and the system are located at same altitude level, this problem does not occur.

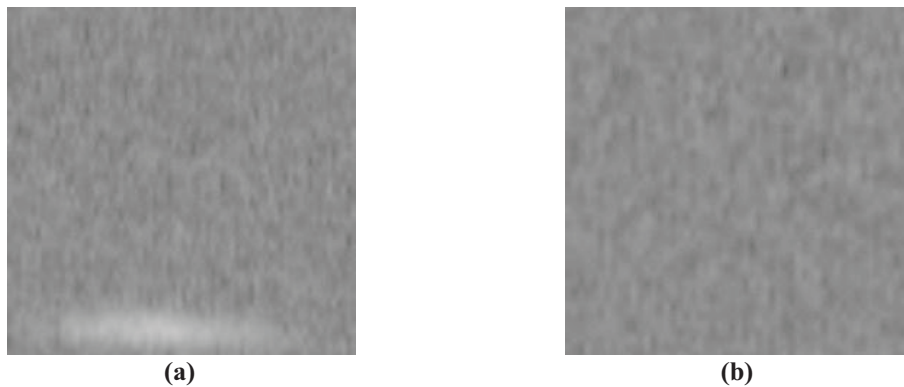


Figure 46 (a) is a typical first image and (b) is a typical last image in collected data set. For illustration purposes, dark pixels enhancement is applied by log transform.

Responses of the SSIM methods are illustrated in Figure 42. It is shown that last images have higher comparison value for SSIM STR compared to other methods. If a

comparison value is higher than the threshold, it contributes location calculations and thus, SSIM STR results in more error in positive Z direction.

Another analysis for the responses of methods to first and last images is shown in Figure 48. First and second images are the first images, third and fourth images are the last images. It is shown that, SSIM CNTRST and SSIM LUM methods has higher similarity values for first images but SSIM STR method has higher similarity values for last images.

In section 5.1.6, it is stated that SSIM STR has a negative performance value indicating that it cannot distinguish a translated object from a scene with no object present. The major drawback of the SSIM algorithm is that it is highly sensitive to translation, scaling, and rotation of objects in images. To overcome this, Wang et al. [49] and Sampat et al [50] proposed improved versions of the SSIM algorithm. Wang et al. [49] applies SSIM for different scales of the image and combines the result to obtain an extended index for image similarity. Sampat et al. [50] improved the SSIM against small rigid distortions such as rotation, translation and scaling. For a future work, the algorithms stated in [49] and [50] can be implemented to reach better results.

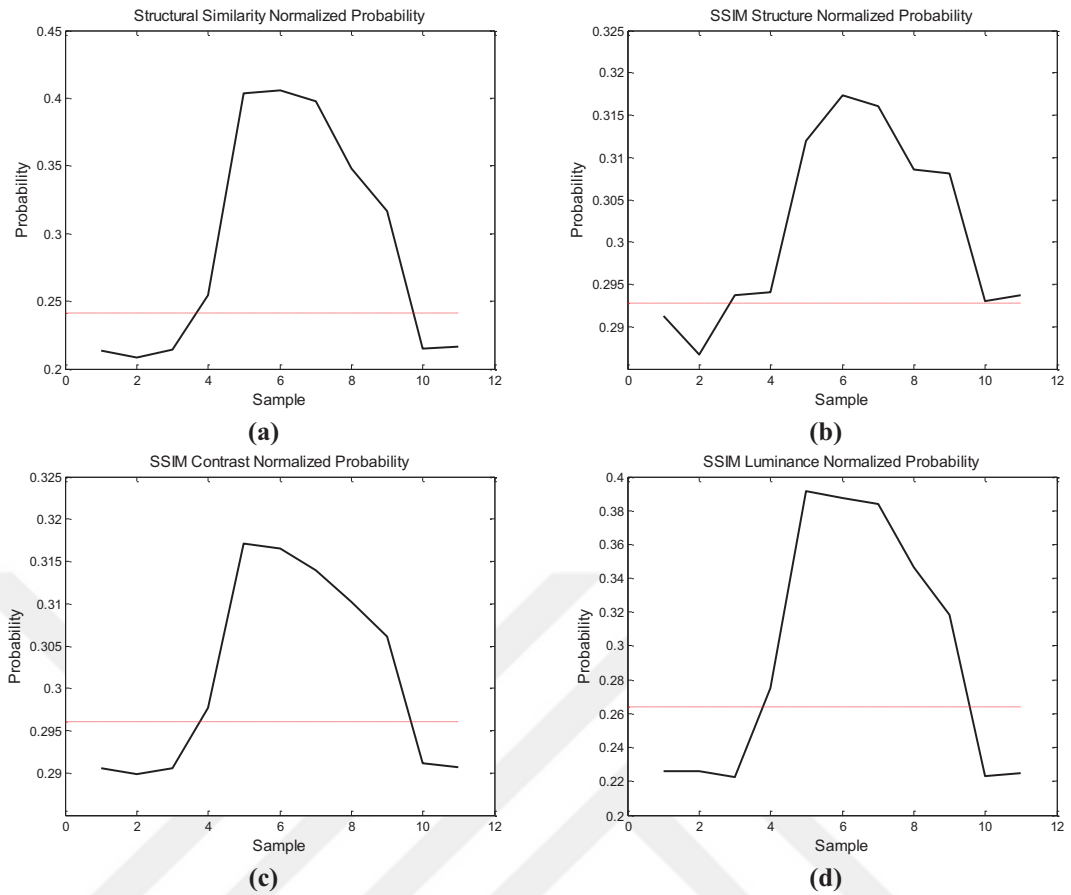


Figure 47 Comparison results for the SSIM methods. Red line is the interquartile threshold value for Interquartile with $k = 0.07$.

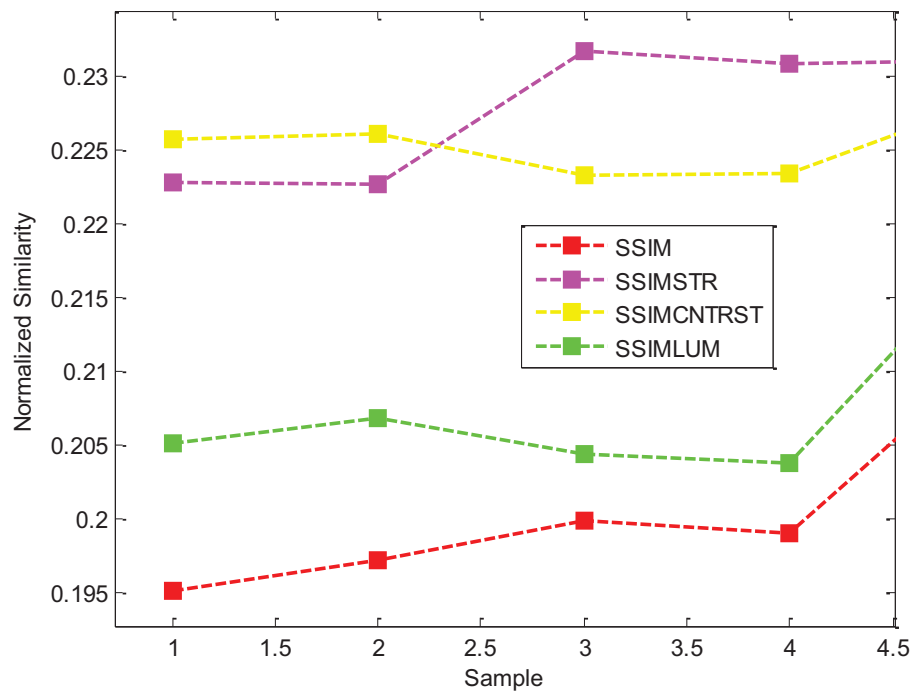


Figure 48 Comparison results taken from brightness analysis section. Images 1- 2 are the first images, 3-4 are the last images.

5.1.4 Reference Object Analysis

Performance of the reference object update method is investigated in this section. As described in section 4.2.3.1, reference image is updated at the end of each successful tracking step. An observed image is compared with the reference image for each updated reference image. Results of the comparisons are shown in Figure 49. Reference image 1 (sample 1 in Figure 49) is the first updated reference image after the object moves from one location to another. Up to reference image 5, reference image is a blurred image and the similarity between the selected image and the reference image is low. As a result, the reference object update method must be improved by considering the X-Y movement of the object between frames. Obtaining the exact location of the object for eliminating the effect of object movements can be a solution for further studies.

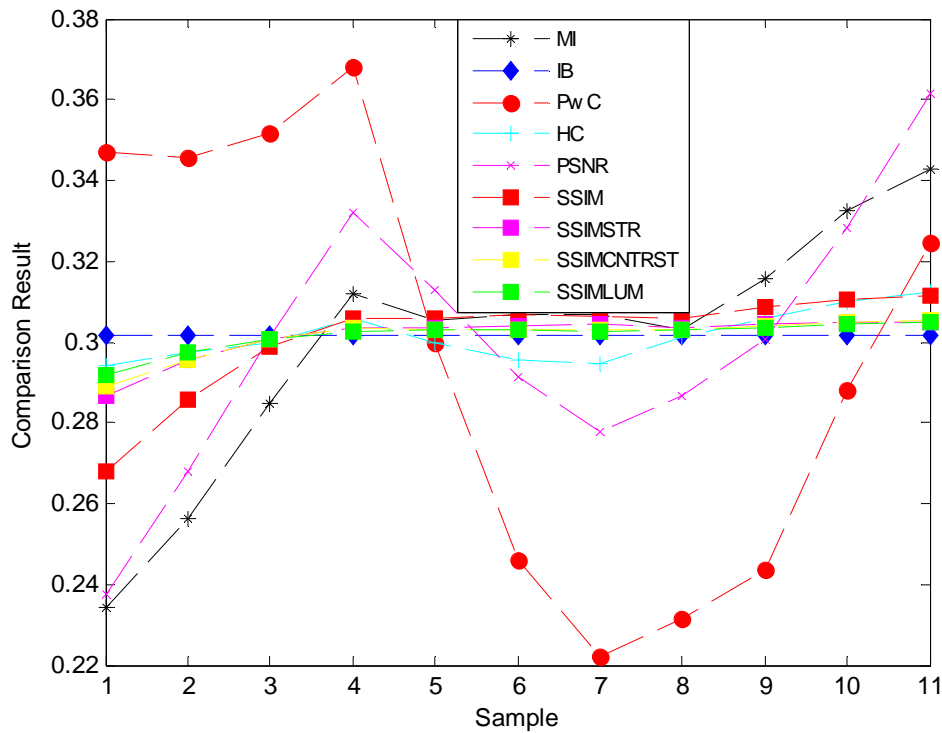


Figure 49 Comparison of the updated reference image and the selected object for each updated reference image

Another result that this analysis shows is about the performances of the methods. For the first a few sample, reference images are blurred and it is expected that the methods should get lower rates for the first few samples and get higher results for the other samples. However, PwC, HC, MI and PSNR methods do not behave in such a characteristic. Comparison results of all this methods have local maxima at 4th sample and decrease after this sample. Since IB method is not based on the comparison with reference image, this analysis is not applicable for IB method.

5.1.5 Brightness Invariance

Brightness invariance of the methods is investigated in this section. For this analysis, 4 images with no object and 14 images with object but different brightness levels are selected. Similarity analysis between the reference image and the 18 images is calculated for each method.

In Figure 51, samples 1-4 are the not object images and the other samples (5 to 18) are the images with different (increasing) brightness levels where 14th image having the most similar brightness value to the reference image.

Although IB method has the highest performance value, the reason is the high scores obtained from the higher brightness levels (i.e. images 13:16). Since high scores for bright images is not a desired result, IB method is not assigned as promoted method.

In Figure 51, none of the similarity result for images with different (increasing) brightness levels has lower value than the similarity result for not object images. This concludes that none of the method has catastrophic results.

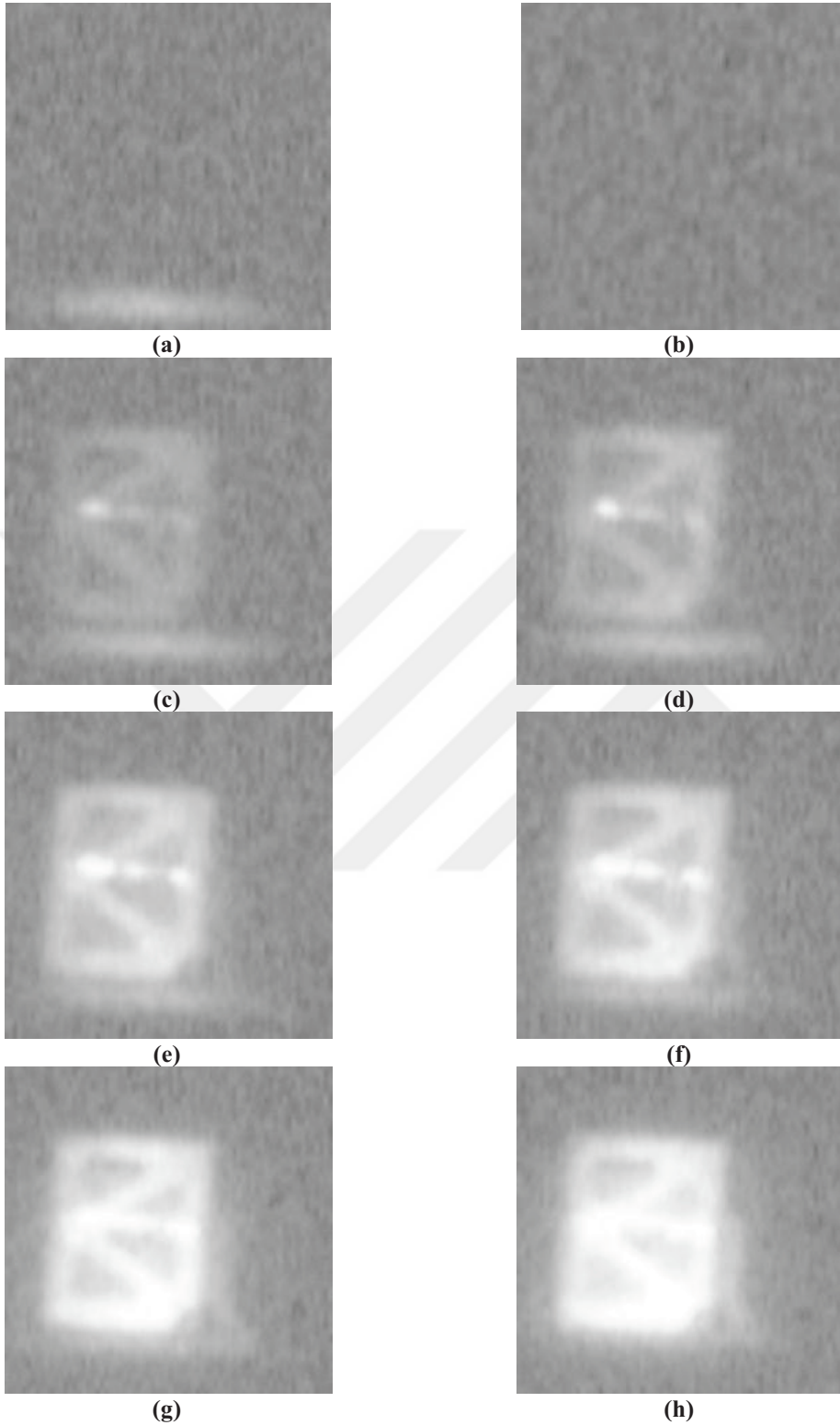
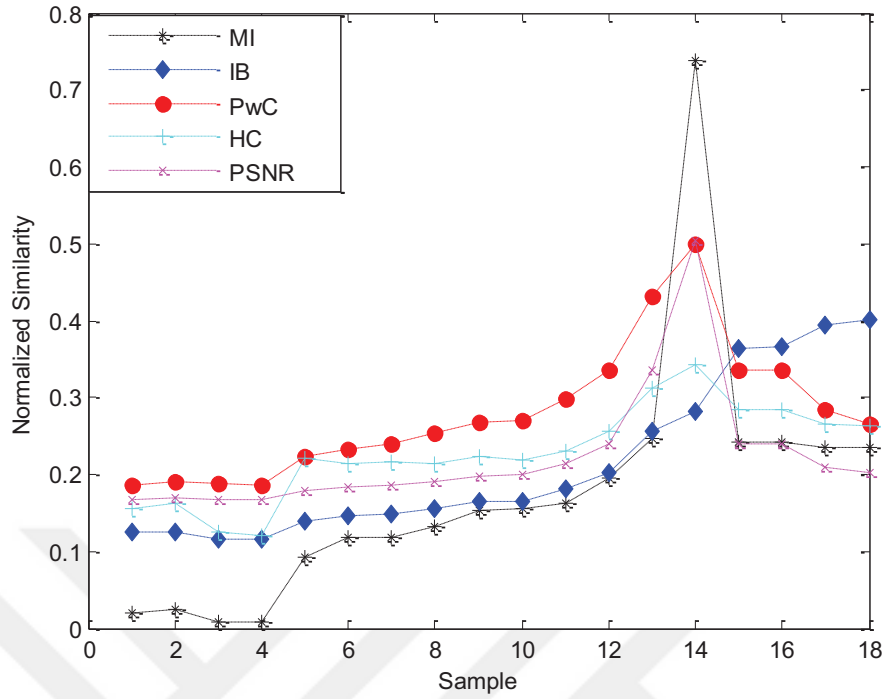
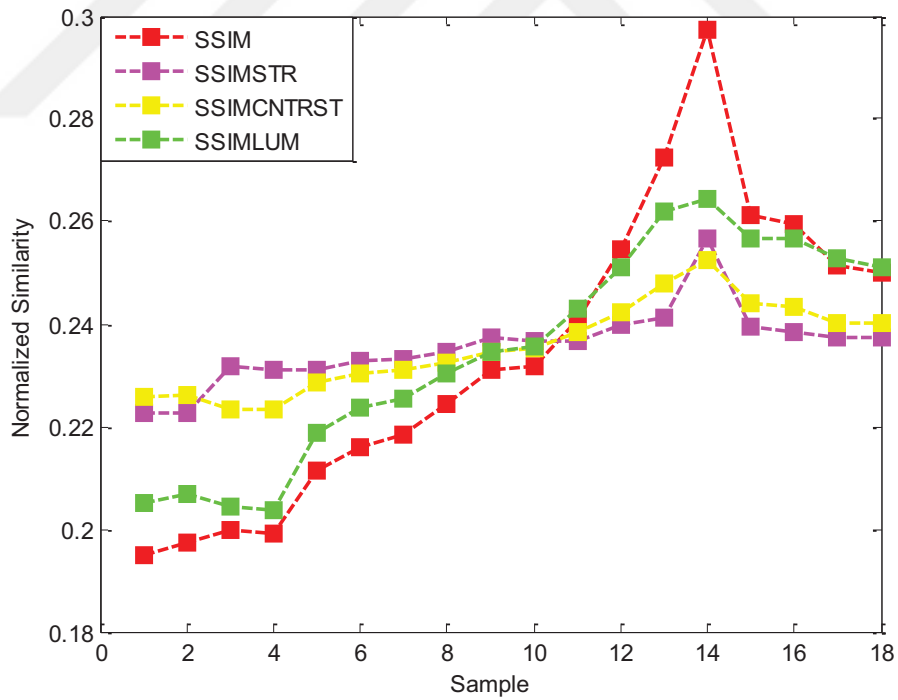


Figure 50 Images used for brightness invariance analysis. (a)-(b) are the images with no object, (c)-(h) are the images with object but different brightness levels. (f) is used as the reference image. For illustration purposes, dark pixels enhancement is applied by log transform



(a)



(b)

Figure 51 Similarity results between the reference image and the selected images for all methods. Results 1-4 are the not object images and the other results (5 to 18) are the images with different (increasing) brightness levels where 14th image having the most similar brightness value to the reference image. MI, IB, PwC, HC, and PSNR are shown in (a), other methods are shown in (b).

5.1.6 Translation Invariance

Translation invariance of the methods is investigated in this section. For this analysis, two images with no object and 10 images with object but different locations are selected. Similarity metric between the reference image and the 12 images is calculated for each method. Images without object and images with object but different locations are analyzed with all comparison methods and the results are shown in Figure 53.

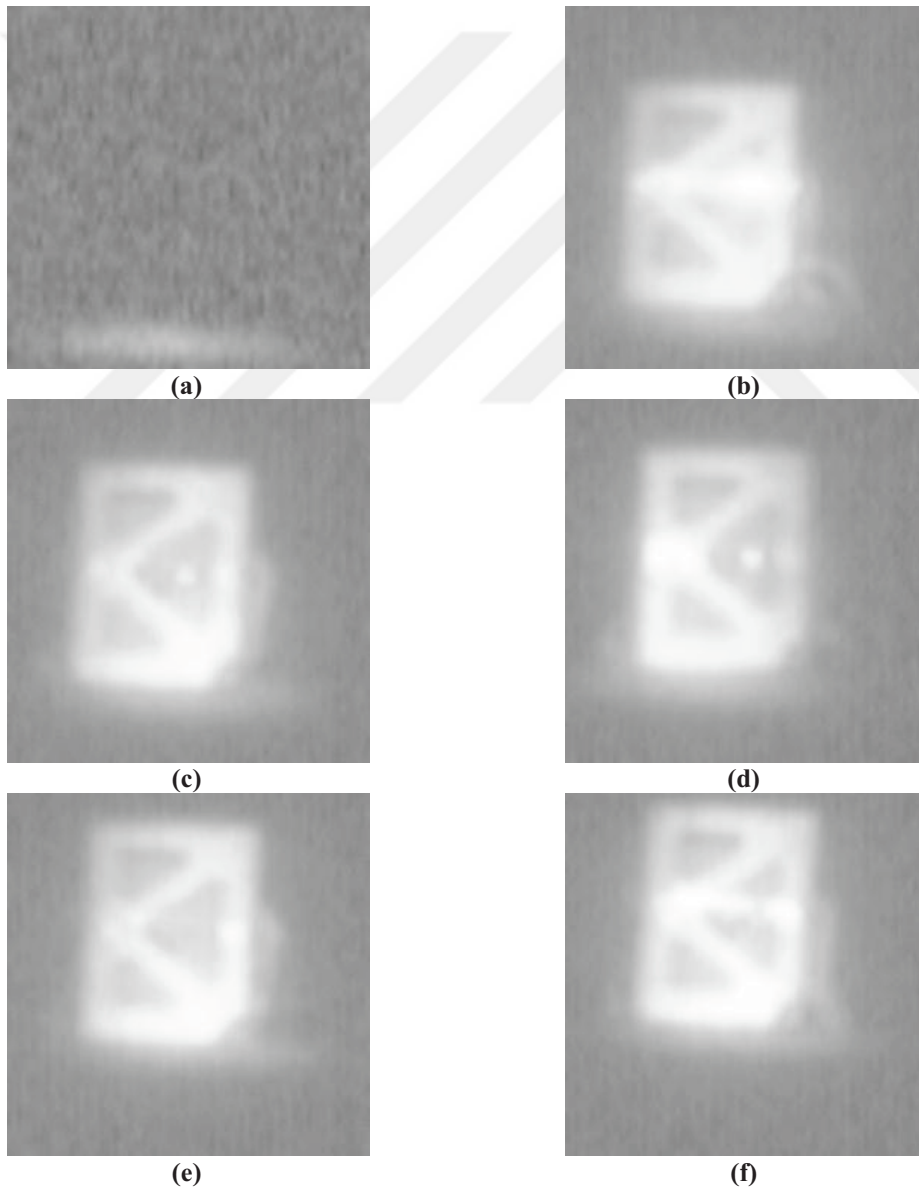
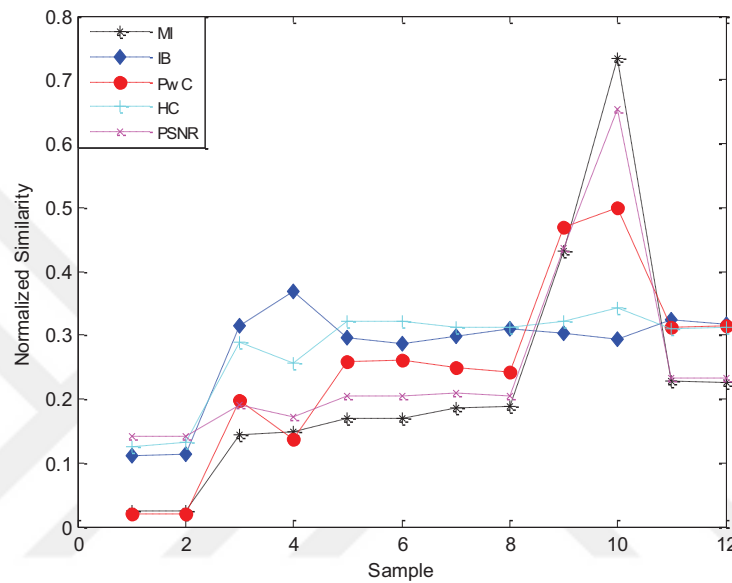
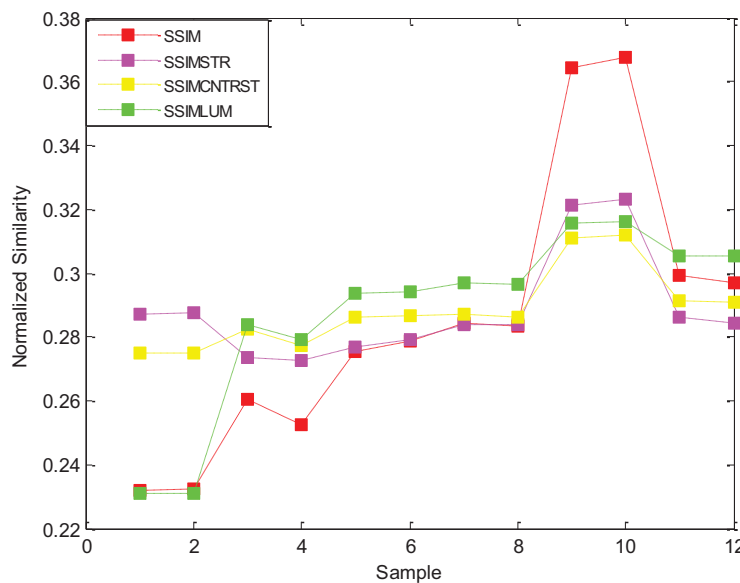


Figure 52 Images used for movement invariance analysis. (a) is an image with no object, (b)-(f) are the images with object but different locations. (e) is used as the reference image. For illustration purposes, dark pixels enhancement is applied by log transform.

According to these results, only SSIM STR has a negative performance value indicating that SSIM STR method cannot distinguish a moved object from a scene with no object present. SSIM LUM, IB and HC methods has better performances compared to the other methods. None of the method has catastrophic results.



(a)



(b)

Figure 53 Images not having the target object and images with the target object has gone through translation (i.e. present at different locations). Samples 1 and 2 are the results for images not having the target object; samples 3-11 having images of the object with different translation amounts and sample 12 is the result for the reference image. MI, IB, PwC, HC, and PSNR are shown in (a), other methods are shown in (b).

5.1.7 Experiments with Artificial Images

Response of each method to artificial images is investigated in this section. For testing issues, 8 artificial images are created. Two of them are illustrated Figure 54.

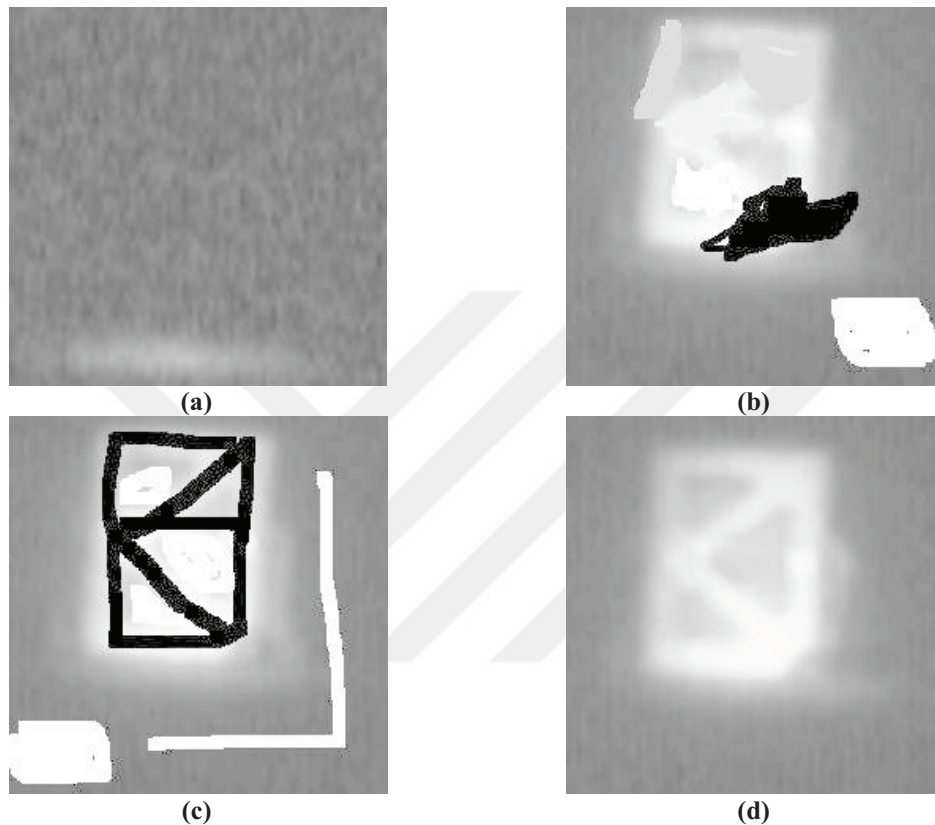


Figure 54 Images used for artificial image analysis. (a) is an image not having the target object, (b) and (c) are the two examples of the artificial images. (d) is used as the reference image. For illustration purposes, dark pixels enhancement is applied by log transform.

Assuming the criterion for the performance as; “similarity value for a virtual image should not be bigger than the similarity value for the reference object”, only the image brightness method perform weakly with artificial objects because the reference object has smaller similarity values compared to the artificial objects.

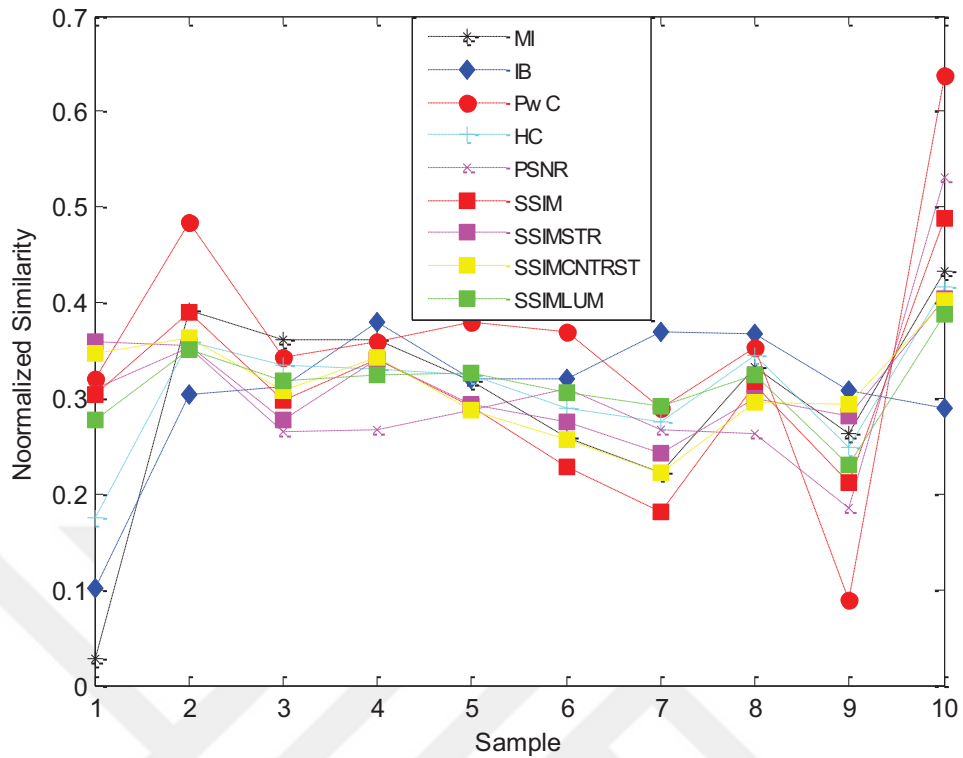


Figure 55 Comparison results for the images. Sample 1 corresponds to image not containing the object, Samples 2-9 for the artificial images and Sample 10 for the reference image.

5.1.8 Performance Analysis

For a system with 5 nanosecond gate time and 10 ns laser pulse as in our experimental setup, objects in 75 cm depth can be imaged with full illumination as described in Section 3.2.3. For SSIM method, the location of the object is determined with -2.3 cm average error and 21.2 cm standard deviation. Error values are assumed to be normally distributed with given mean and standard deviation. Figure 56 illustrates the distribution of error for SSIM method. Since the system can get images from 75 cm depth, error values between -37.5 cm and 37.5 cm can be tolerated. In Figure 56, the area inside the distances -37.5 cm and 37.5 cm (the area between the red dashed lines) is the correct tracking area. The area inside the red dashed lines is

92.09% of the total area. As a result, SSIM method tracks the object with 92.09% success. It is assumed that previous search found the object location correctly.

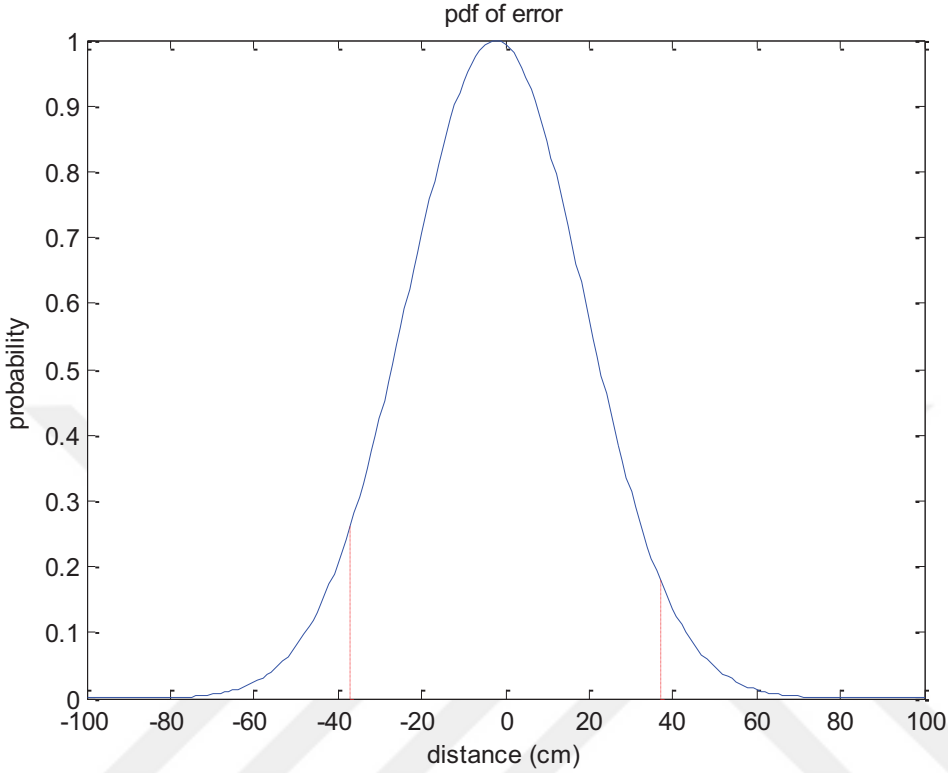


Figure 56 PDF of Error while the object is located at previous location

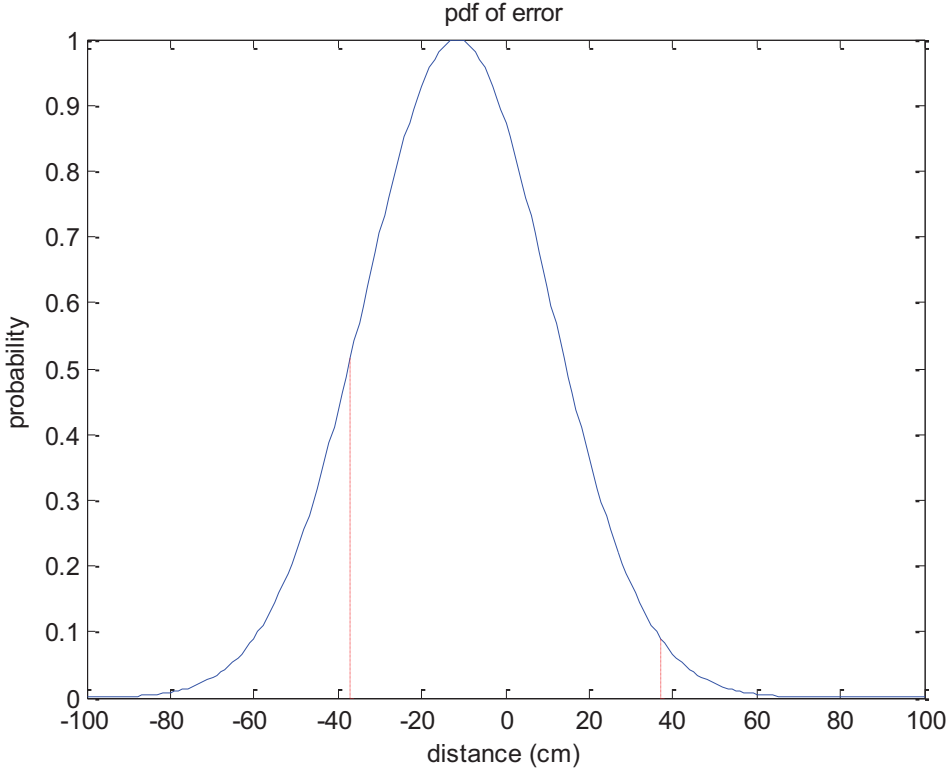


Figure 57 PDF of Error while the object is located 600 mm away from previous location

Figure 57 illustrates the distribution of error in moving object scenario where the error mean is -115 mm and standard deviation is 221 mm. In Figure 57, the area inside the distances -37.5 cm and 37.5 cm (the area between the red dashed lines) is the correct tracking area. The area inside these values is 86.66% of the total area. As a result, SSIM method tracks the object with 86.66% success if the object moves 60 cm between two consecutive searches. It is assumed that previous search found the object location correctly. In result, *Success Rate* is defined as the proportion of the area between the red dashed lines to the total area in Figure 56 and Figure 57.

Same success analysis is made for all the methods for stationary and moving object scenarios. Success rates of all the methods are shown in Table 6.

Table 6 Strong and weak sides of methods under different conditions

Method	Blurred Reference Image	Brightness Invariance	Translation Invariance	Artificial Object	Success Rate (%)	
					Stationary Object	Moving Object
HC	POOR	GOOD	GOOD	GOOD	83.92	64.32
PwC	POOR	GOOD	GOOD	GOOD	87.62	75.45
IB	N/A	GOOD	GOOD	POOR	91.95	52.24
MI	POOR	GOOD	GOOD	GOOD	91.57	67.34
PSNR	POOR	GOOD	GOOD	GOOD	85.84	64.04
SSIM	GOOD	GOOD	GOOD	GOOD	92.09	86.66
SSIM STR	GOOD	GOOD	POOR	GOOD	75.37	87.79
SSIM CNTRST	GOOD	GOOD	GOOD	GOOD	91.02	87.01
SSIM LUM	GOOD	GOOD	GOOD	GOOD	87.20	76.58

According to results of the all methods, SSIM, SSIM LUM and SSIM CNTRST methods can be used for object position estimation. Besides that, Human Visual System based method SSIM has better performance than the purely mathematically defined comparison metrics. Performances of these methods will also be investigated in Section 5.2 with high amount of test data.

5.2 Test Results and Conclusion

In this section, test scenarios are defined and performances of the methods with these scenarios are calculated. Best method is selected and proposed for system solution.

5.2.1 Test Scenarios

To calculate performances of comparison methods, three test images shown in Figure 58 are selected. Distances of the objects to observation tower are shown in Table 7.

Table 7 Object Distances

Object	Distance (meter)
Standing Man	222
Sitting Man	204
Antenna	315

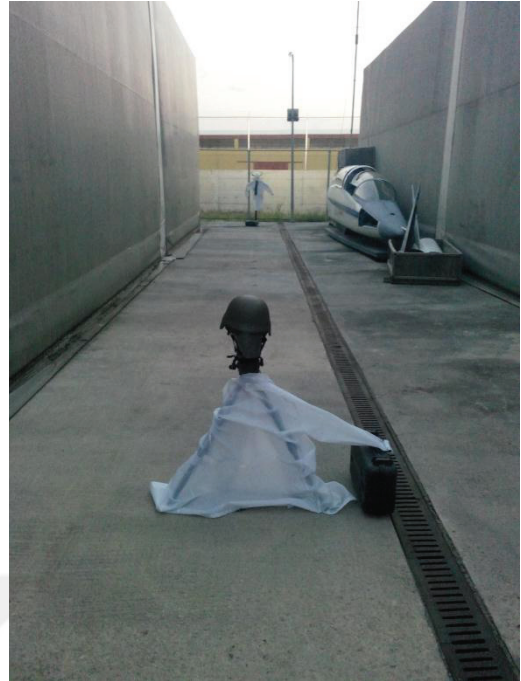
Five different data sets are collected with specified three objects. First data set is collected with Standing Man, second set is with Sitting Man and the others are collected with an Antenna.

Standing Man data set is collected with 1 ns step sizes (totaly; 40 images for 49 different delay time around the object) to allow step size analysis. Antenna data set is collected to test the performances of the algorithms of the object have a depth.

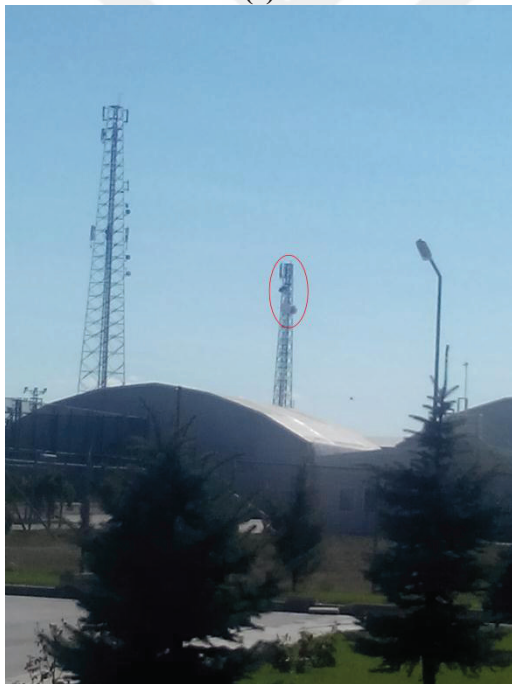
Reflectivities of the all objects are unknown and focus of the laser illuminator is arranged to get desired images. Whole data set is collected at summer night clear weather.



(a)



(b)



(c)



(d)

Figure 58 Objects used for performance tests of the comparison methods. (a) is named as “Standing Man”, (b) is “Sitting Man” and (c) is “antenna”. (d) is the images of all three objects taken from the observation tower.

5.2.1.1 Test Scenario 1 (Standing Man)

The object shown in Figure 58 (a) is located around 222 meters distance from observation tower. 19x40 set of images are collected where the object is located 19 different locations. An example of collected image is shown in Figure 59.

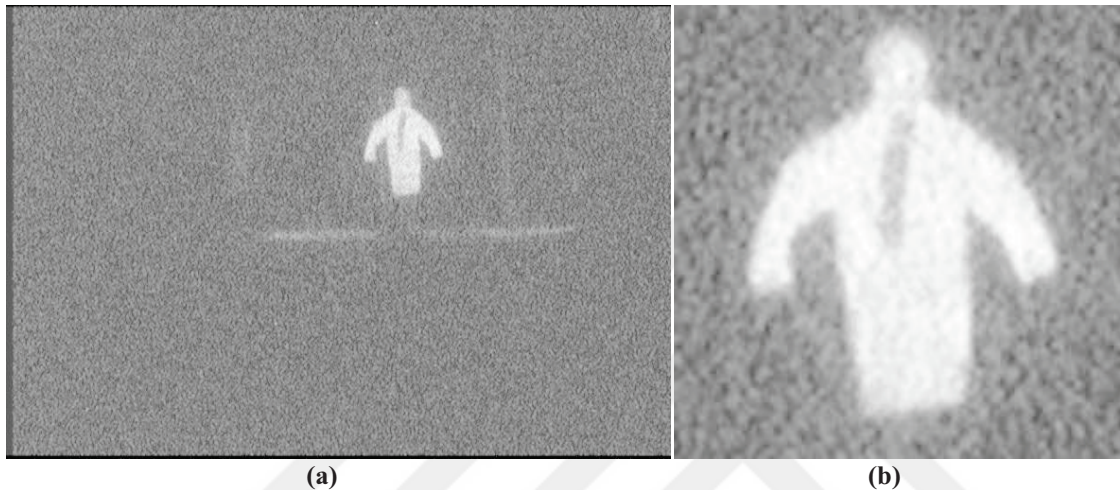


Figure 59 Image taken from standing man object (a) and the object region selected manually (b). For illustration purposes, dark pixels enhancement is applied by log transform.

For 19 different object locations, calculated locations and the error values are shown in Figure 60. As shown from the figures, if the object moves more than 1 meter from previous location, errors of the algorithms start to increase. For measuring the performances of the algorithms, locations between -1000 mm and 1000 mm are used.

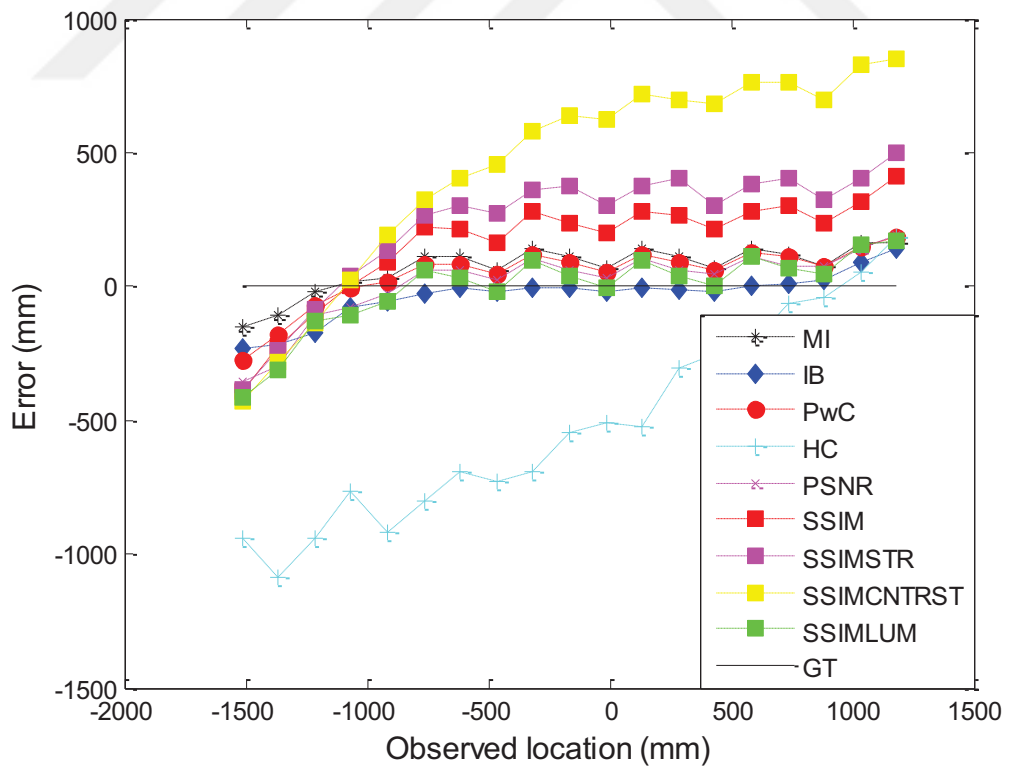
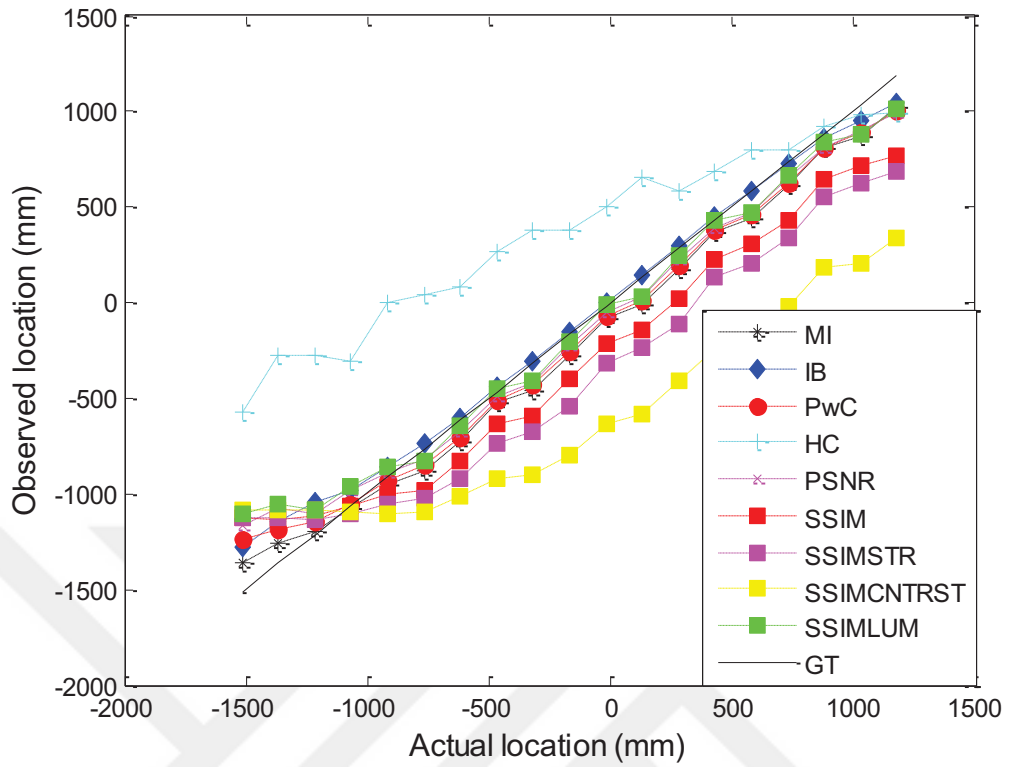


Figure 60 Observed locations (top) and the error values (bottom) for all methods for Standing Man Scenario. Black line is the ground-truth line.

5.2.1.2 Test Scenario 2 (Sitting Man)

The object shown in Figure 58 (b) is located around 204 meters distance from observation tower. 19x40 set of images are collected where the object is located at 19 different locations. An example of collected image is shown in Figure 61.



Figure 61 Image taken from sitting man object (a) and the object region selected manually (b). For illustration purposes, dark pixels enhancement is applied by log transform.

For 19 different object locations, calculated locations and the error values are shown in Figure 62. As shown from the figures, if the object moves more than 1 meter from previous location, errors of the algorithms start to increase. For measuring the performances of the algorithms, locations between -1000 mm and 1000 mm are used. Performances of the methods for both Standing Man and Sitting Man scenarios are shown in Table 8. Success analysis is shown in Section 5.2.2

Table 8 Error Mean and STD values for all methods for Standing Man and Sitting Man Scenarios. Object is located 15 locations between -1000 mm and 1000 mm. Bold is used to highlight the best result.

Method	Standing Man			Sitting Man		
	Error Mean (mm)	Error STD (mm)	RMS Error (mm)	Error Mean (mm)	Error STD (mm)	RMS Error (mm)
HC	483.6	512.2	704.1	44.5	285.4	288.5
PwC	-82.2	251.5	264.4	143.7	243.4	282.4
IB	11.1	256.1	256.1	28.9	226.2	227.8
MI	-99.3	225.4	246.1	101.2	218.6	240.7
PSNR	-59.4	258.8	265.3	149.8	261.5	301.1
SSIM	-228.9	276.5	358.7	486.4	911.3	1032.2
SSIM STR	-322.4	281.3	427.7	-104.2	1544.1	1546.2
SSIM CNTRST	-579.4	368.5	686.4	326.3	1038.7	1087.8
SSIM LUM	-39.5	269.5	272.2	103.7	271.7	290.6

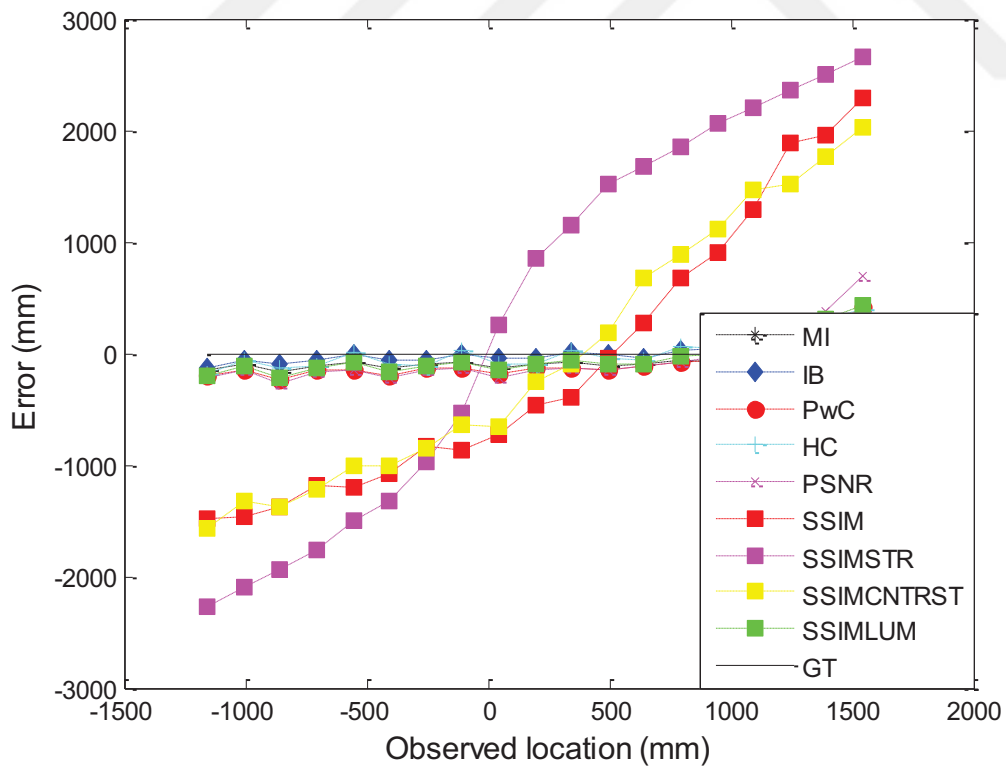
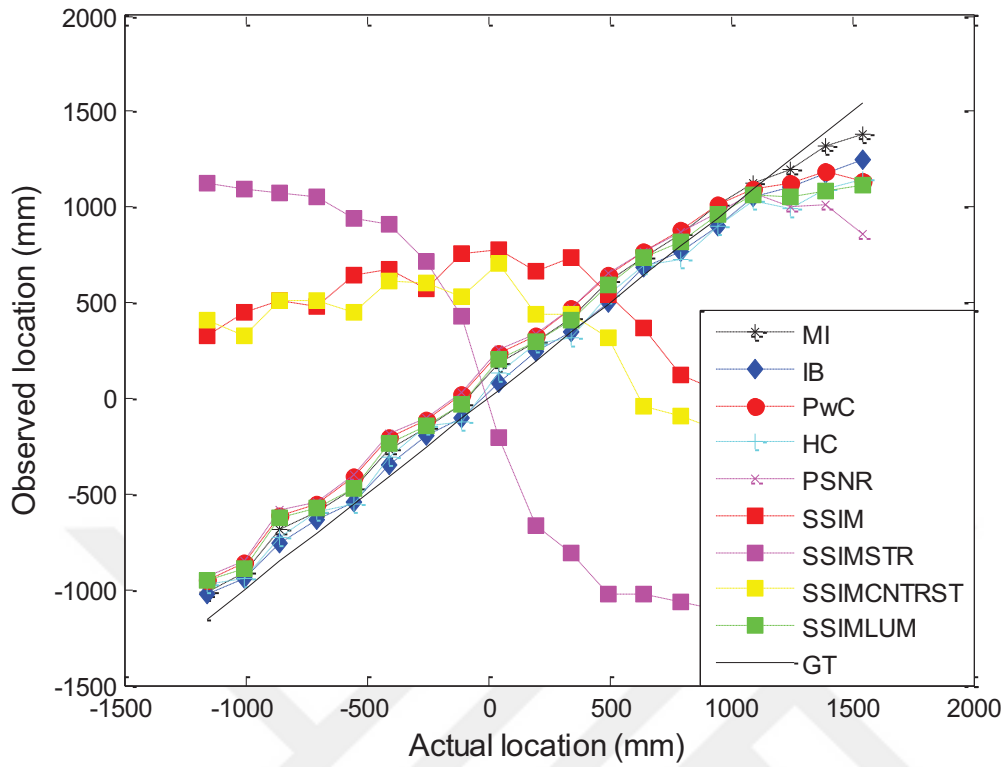


Figure 62 Observed locations (top) and the error values (bottom) for all methods. (Sitting Man Scenario)

5.2.1.3 Test Scenario 3 (Antenna-3ns StepSize)

The object shown in Figure 58 (c) locates around 315 meters distance from observation tower. 9x40 set of images are collected where the previous location of the object is set to 9 different locations. An example of collected image is shown in Figure 63.

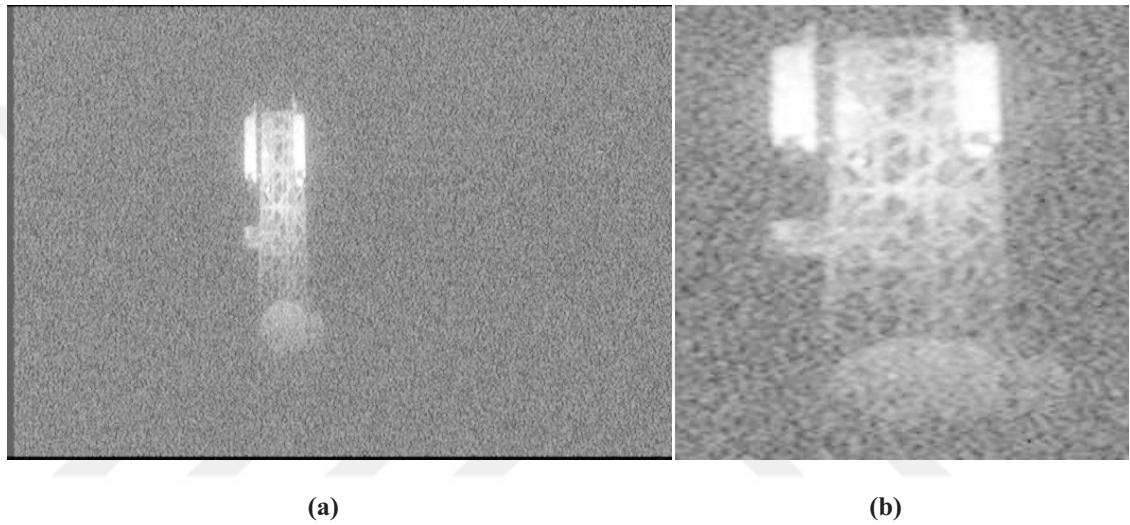


Figure 63 Image taken from Antenna (a) and the object region selected manually (b). For illustration purposes, dark pixels enhancement is applied by log transform.

For 9 different object locations, calculated locations and the error values are shown in Figure 64. As shown from the figures, if the object moves more than 1 meter from previous position, errors of the algorithms start to increase. For measuring the performances of the algorithms, locations between -1000 mm and 1000 mm are used.

5.2.1.4 Test Scenario 4 (Antenna-2ns StepSize)

The object used in Test Scenario 3 is used for this scenario. The only difference was the step size is set to 2ns. 9x40 set of images are collected where the object is located

at 9 different locations. For 9 different object locations, calculated locations and the error values are shown in Figure 65. As shown from the figures, if the object moves more than 1 meter from previous position, errors of the algorithms start to increase. For measuring the performances of the algorithms, locations between -1000 mm and 1000 mm are used.



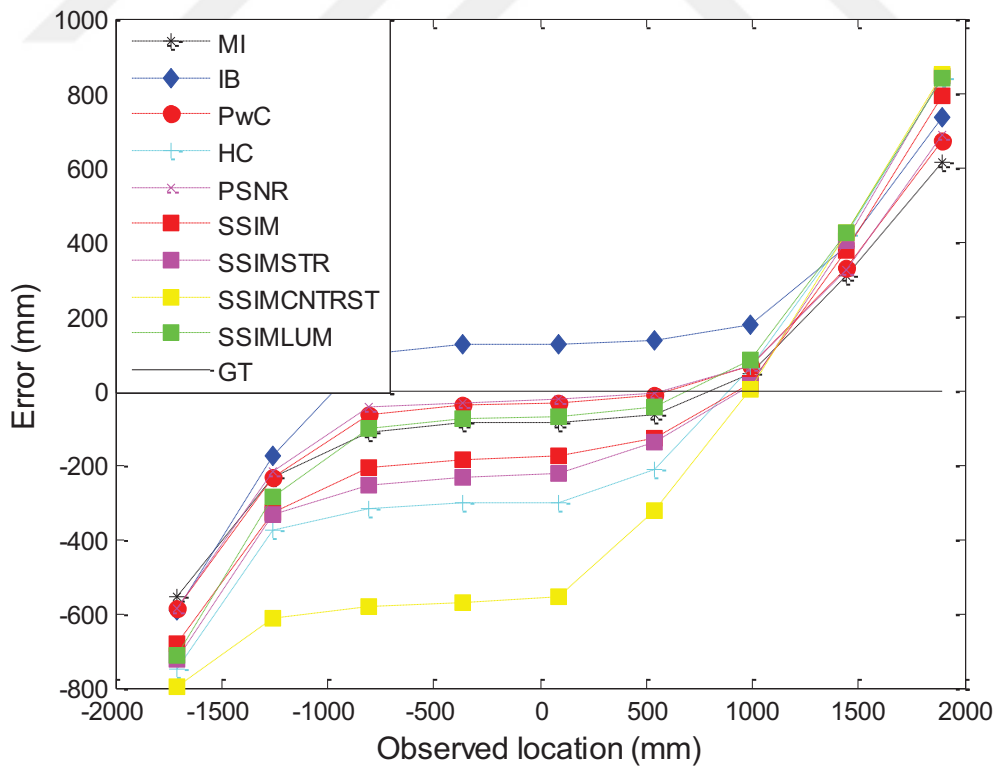
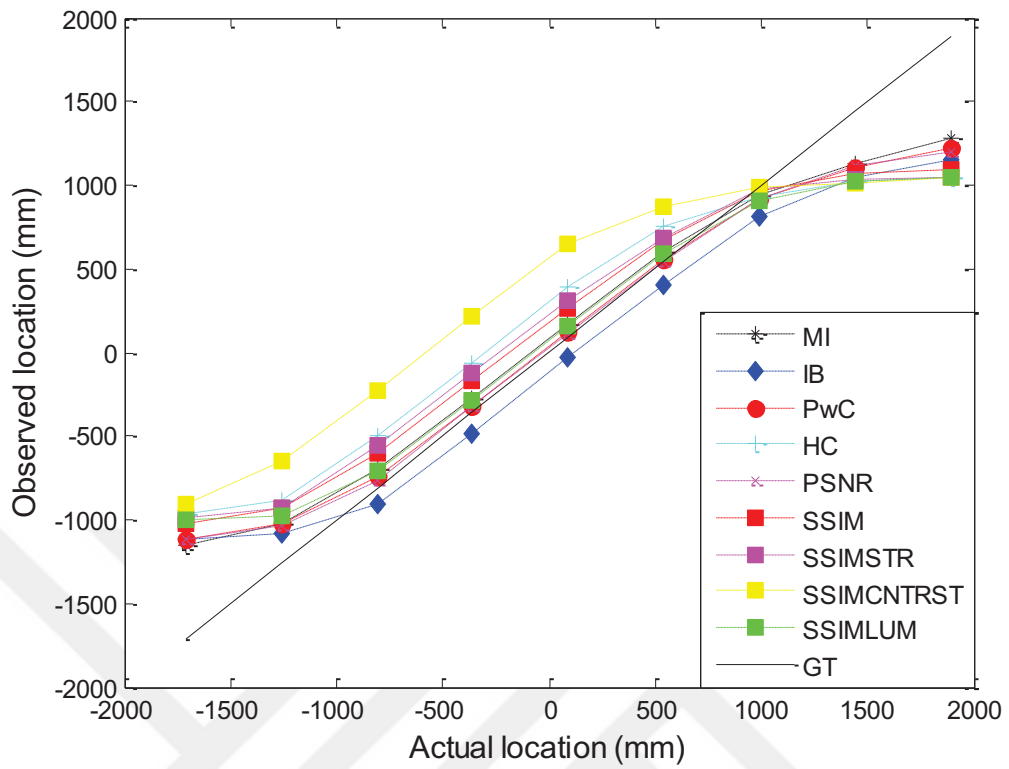


Figure 64 Observed locations (top) and the error values (bottom) for all methods. (Antenna-3ns StepSize Scenario)

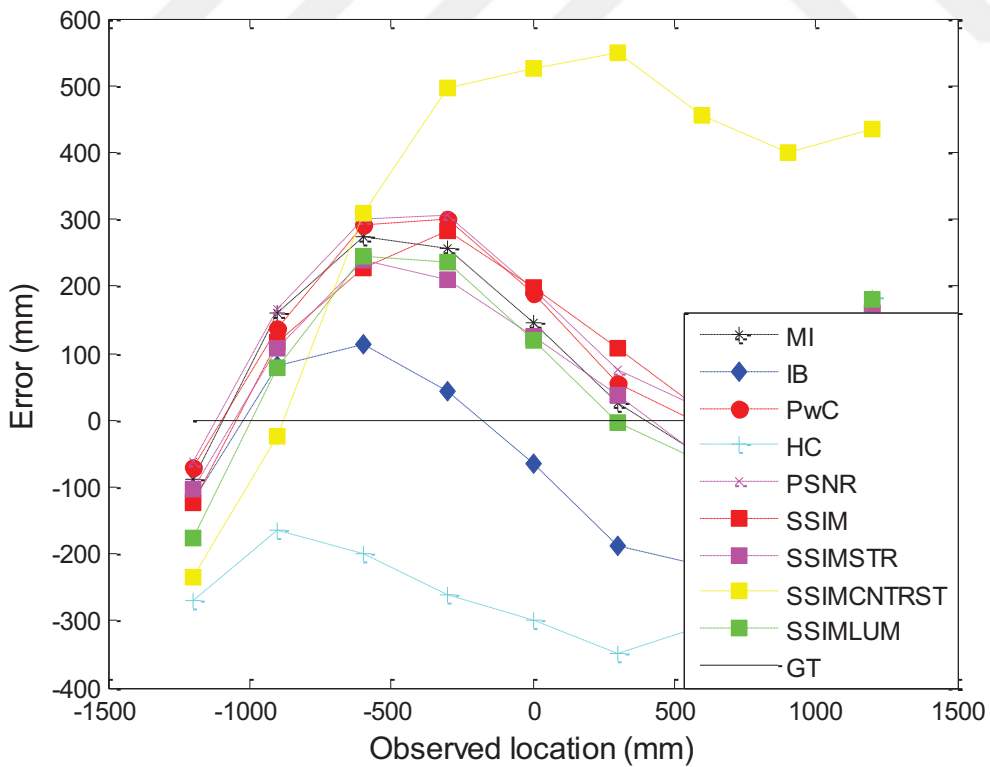
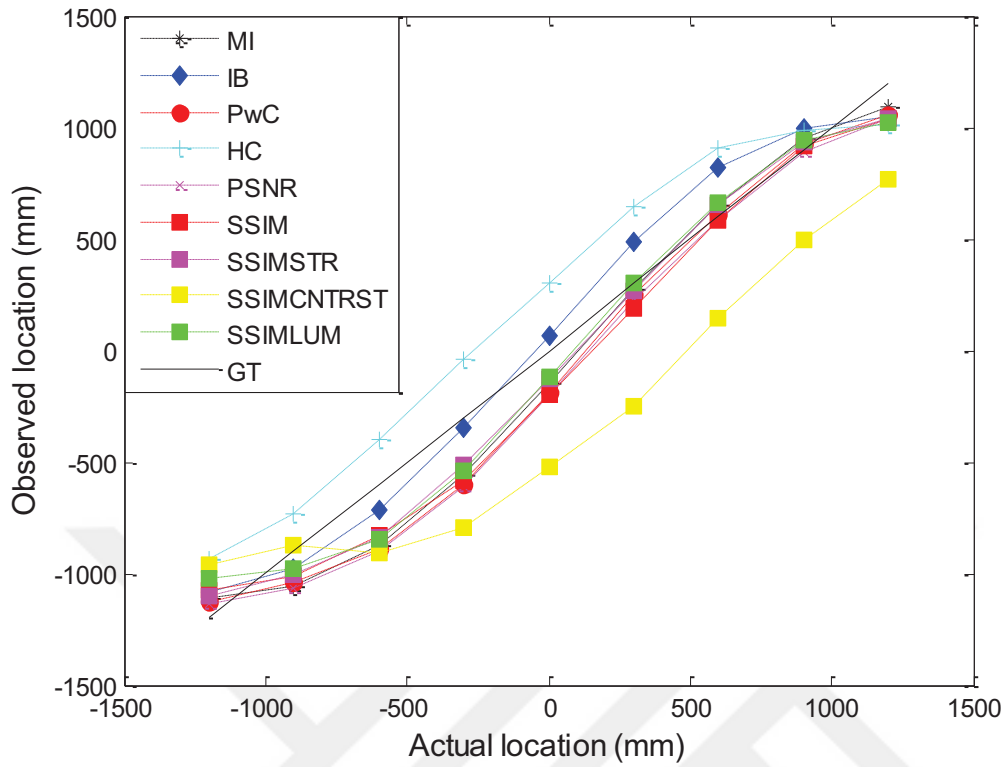


Figure 65 Observed locations (top) and the error values (bottom) for all methods. (Antenna-2ns StepSize Scenario)

5.2.1.5 Test Scenario 5 (Antenna-2ns StepSize Zoom Out)

The object used in Test Scenario 3 and 4 is used for this scenario. The only difference with Test scenario 4 was the lens of the camera is zoomed out. 9x40 set of images are collected where the object is located at 9 different locations. An example of collected image is shown in Figure 66.

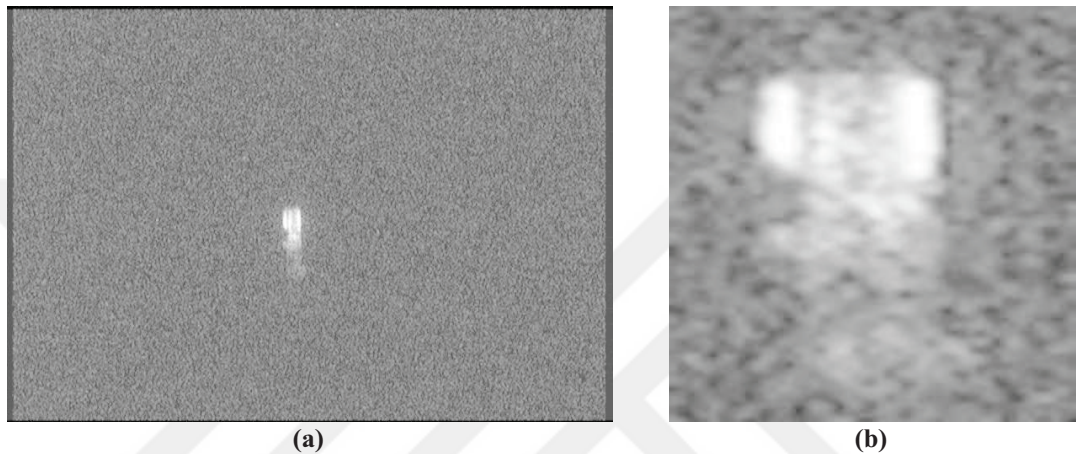
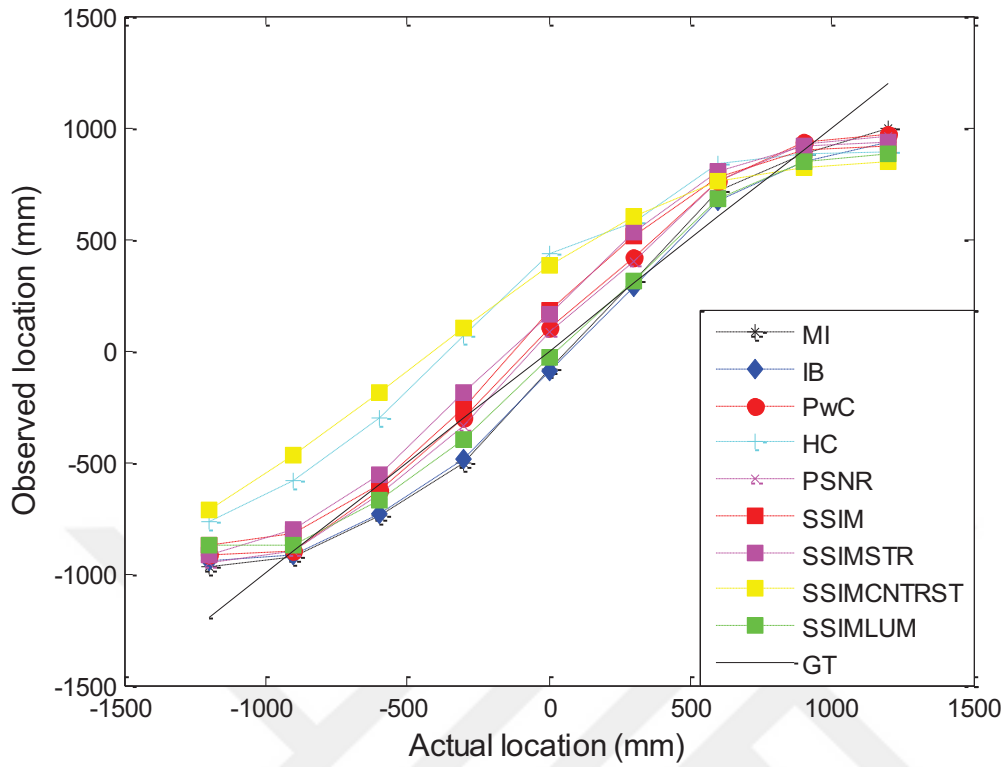


Figure 66 Image taken from Antenna (a) and the object region selected manually (b). For illustration purposes, dark pixels enhancement is applied by log transform.

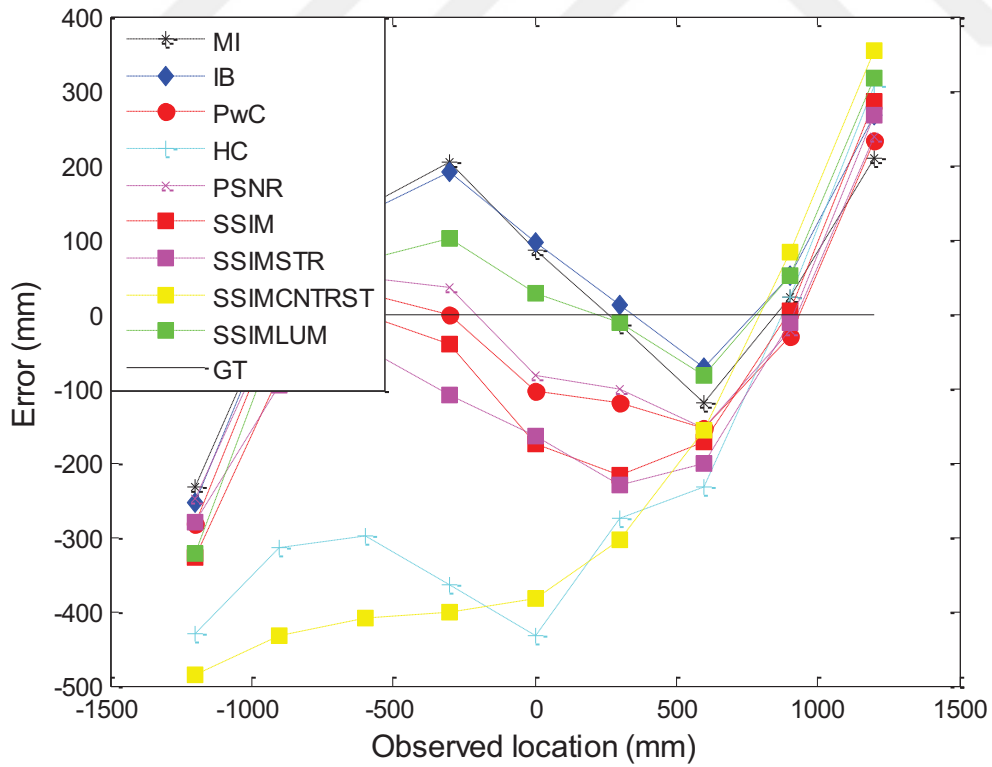
For 9 different object locations, calculated locations and the error values are shown in Figure 67. As shown from the figures, if the object moves more than 1 meter from previous location, errors of the algorithms start to increase. For measuring the performances of the algorithms, locations between -1000 mm and 1000 mm are analyzed.

Table 9 Error Mean and STD values for all methods for Three Antenna Scenarios. Object locations between -900 mm and 900 mm are used. Bold is used to highlight the best result.

Method	Antenna-3ns StepSize			Antenna-2ns StepSize			Antenna- 2ns StepSize Zoom Out		
	Error Mean (mm)	Error STD (mm)	RMS Error (mm)	Error Mean (mm)	Error STD (mm)	RMS Error (mm)	Error Mean (mm)	Error STD (mm)	RMS Error (mm)
HC	215.0	339.6	401.2	238.5	355.2	427.3	271.2	406.7	488.2
PwC	17.2	278.0	277.9	-134.8	306.2	334.0	55.1	329.3	333.3
IB	-131.2	211.3	248.3	46.8	310.1	313.1	-62.1	349.8	354.7
MI	60.9	271.3	277.4	-107.6	304.3	322.2	-49.0	323.3	326.4
PSNR	9.0	268.4	267.9	-152.0	313.1	347.6	40.2	323.4	325.4
SSIM	136.7	285.4	315.8	-132.6	313.9	340.3	98.1	349.0	362.0
SSIM STR	167.5	304.2	346.6	-89.8	280.4	294.0	123.0	339.7	360.7
SSIM CNTRST	405.9	375.7	552.4	-386.9	443.0	587.6	286.3	467.3	547.3
SSIM LUM	43.7	312.7	315.0	-82.0	311.6	321.7	-19.3	350.0	349.9



(a)



(b)

Figure 67 Observed locations (a) and the error values (b) for all methods. (Antenna-2ns StepSize Zoom Out Scenario)

5.2.2 Performance Analysis of Test Scenarios

For test setup (5 nanosecond camera gate time and laser pulse specified in Figure 68, illumination function (shown in Figure 69) is the convolution of the camera gate time and the laser pulse. Since the system can get images from FWHM/2 cm depth safely, error values between -FWHM/4 cm and FWHM/4 cm can be tolerated. FWHM value for the Imaging Function is 13.48 ns which corresponds 202.2cm. In result, if the error of the range estimation is between -50.55 cm and 50.55 cm, it is assumed as a successful estimation. Success rates of all methods for all scenarios are shown in Table 10.

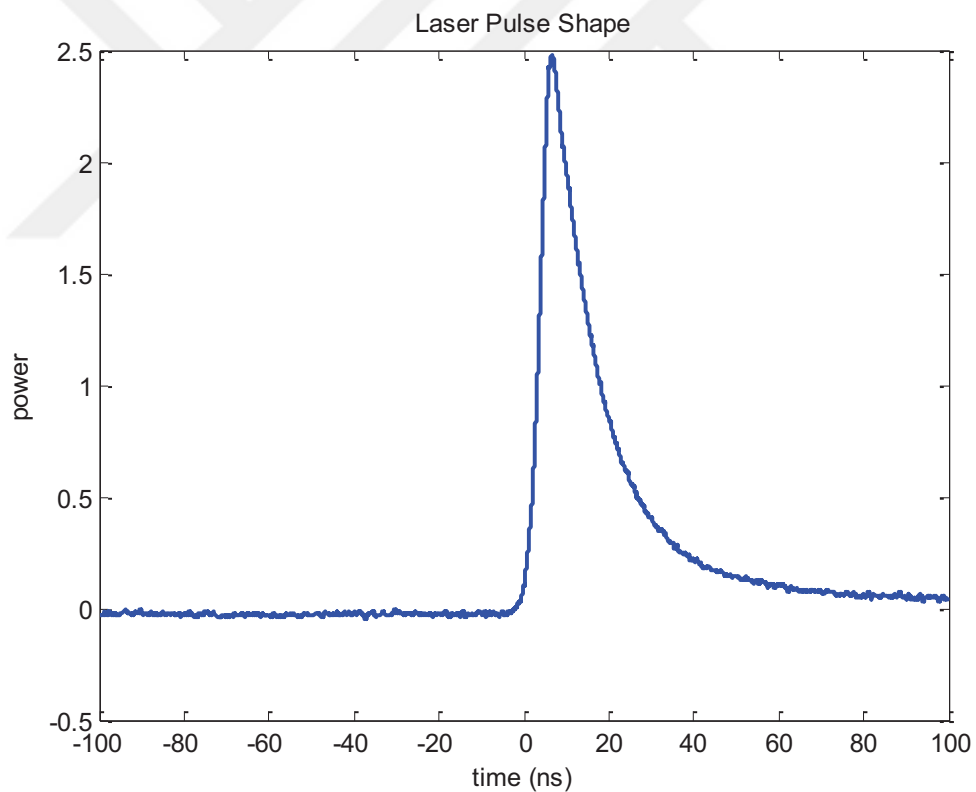


Figure 68 Laser pulse characteristic used in test setup. Power is not absolute value but a scale to show the signal characteristic.

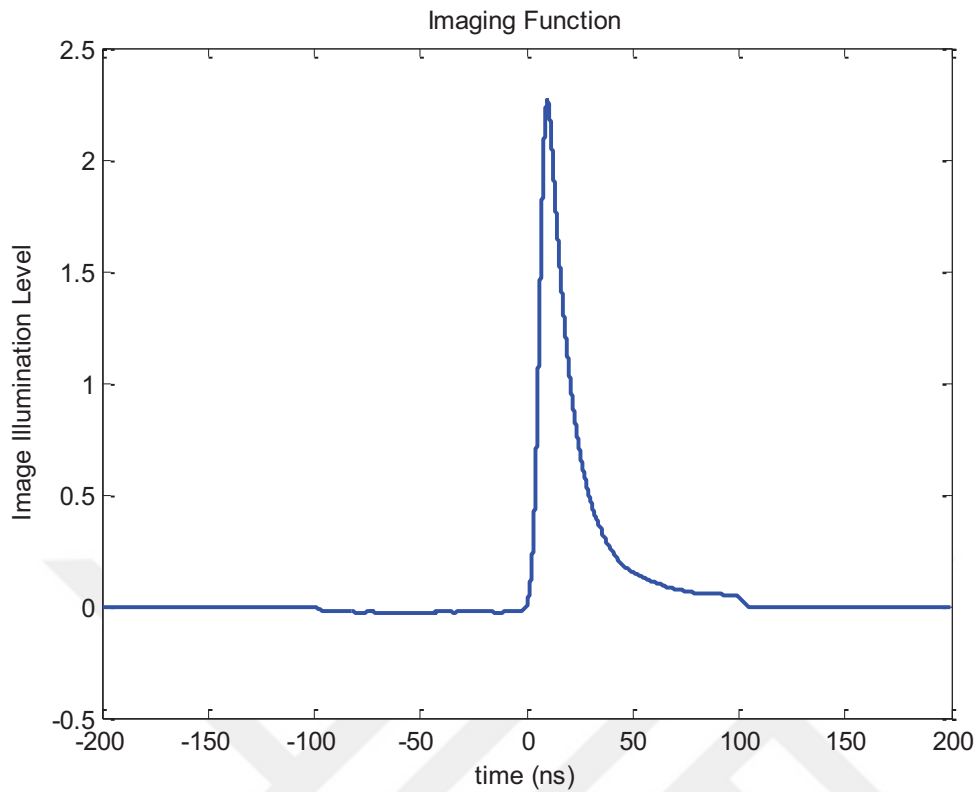


Figure 69 Imaging Function

Table 10 Success values for all methods and all test scenarios

Method	Success Rate (%)				
	Test Scenario 1 Standing Man	Test Scenario 2 Sitting Man	Test Scenario 3	Test Scenario 4	Test Scenario 5
HC	49.0	92.0	78.7	75.6	69.0
PwC	94.4	92.8	93.0	86.9	87.0
IB	95.1	97.3	96.0	89.3	84.5
MI	96.1	96.5	93.1	88.3	87.8
PSNR	94.3	90.7	94.0	85.3	87.9
SSIM	83.8	37.0	89.0	86.2	83.7
SSIM STR	74.1	25.6	85.3	91.4	83.8
SSIM CNTRST	41.9	35.7	59.7	58.4	63.5
SSIM LUM	93.6	91.8	89.1	88.3	85.1

5.2.3 Discussions on Test Results

According to performance results of the methods in Table 10, PwC, IB, MI and SSIM LUM methods can be used for object position estimation. However, only SSIM LUM method has no drawback according to Table 6 in Section 5.1.8. As a

result, SSIM LUM method is proposed for image comparison in range gated object tracking.

Main difference between Standing Man and Sitting Man scenarios is the noise level of images. It is observed that SSIM CTRST and SSIM STR methods have drawbacks in noisy images. The reason for the SSIM CTRST is observed as the difference noise characteristics of reference image and observed image. Since the reference image is somehow the combination of the previous observed images because of the method described in Section 4.2.3, it has less noise which affects the contrast characteristic directly. An observed image (which has same noise characteristic with other observed images) is used as reference image and results are shown in Figure 70 and Table 11. Although the SSIM CTRST method has successful results with new (noisy) reference image, SSIM STR continues to fail. Since the noise is in the nature of range gated images, SSIM STR and SSIM CTRST methods are not proposed for image comparison in range gated images.

Test Scenario 1 (Standing Man) is analyzed with different step sizes to understand the effect of step size on system performance. For 2 ns step size, 11 images around the object are collected with 2 ns delay between two images. As a result, totally 20 ns depth is searched to find the location of the object. 3 and 4 ns step sizes which means 30 and 40 ns search depth are also investigated. Comparisons are done for only proposed SSIM LUM method. Observed locations and error values are shown in Figure 71.

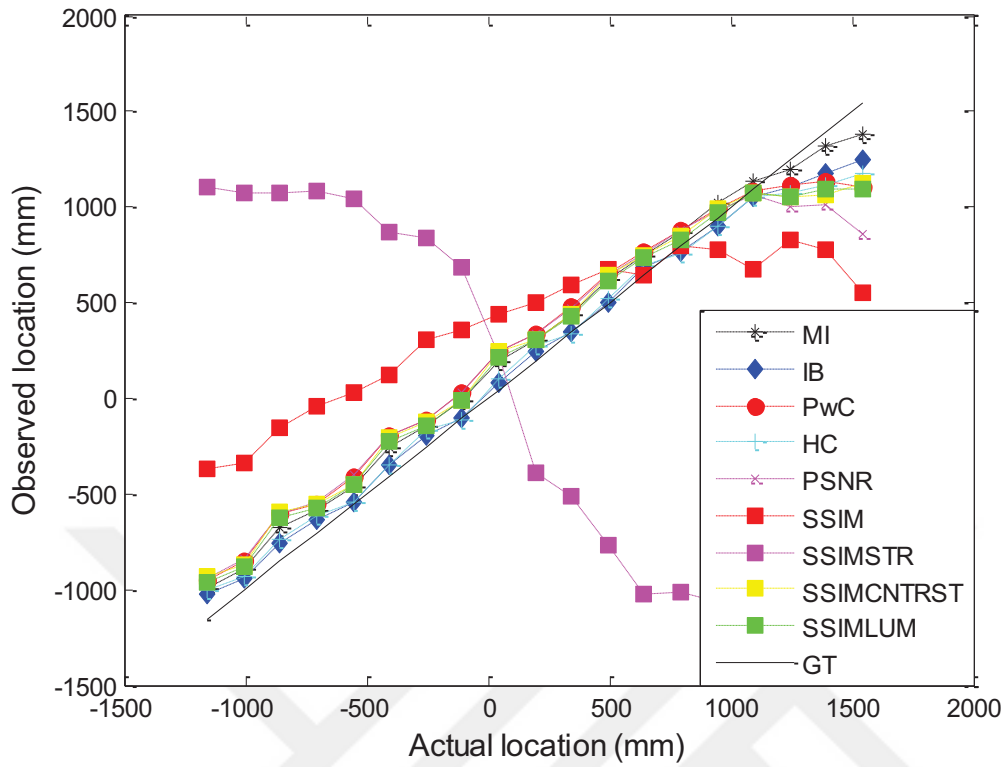


Figure 70 Observed locations for all methods. Sitting Man scenario is run but an observed image is used as reference object.

Table 11 Success rates for Sitting Man scenario with two types of reference image. Bold is used to highlight the changes in SSIM STR and SSIM CNTRST methods.

Method	Success Rate (%)	
	Reference image is an observed image	Reference image is calculated as described in Section 4.2.3
HC	92.0	95.1
PwC	92.8	92.7
IB	97.3	97.3
MI	96.5	95.9
PSNR	90.7	91.9
SSIM	37.0	60.2
SSIM STR	25.6	26.2
SSIM CNTRST	35.7	90.6
SSIM LUM	91.8	92.5

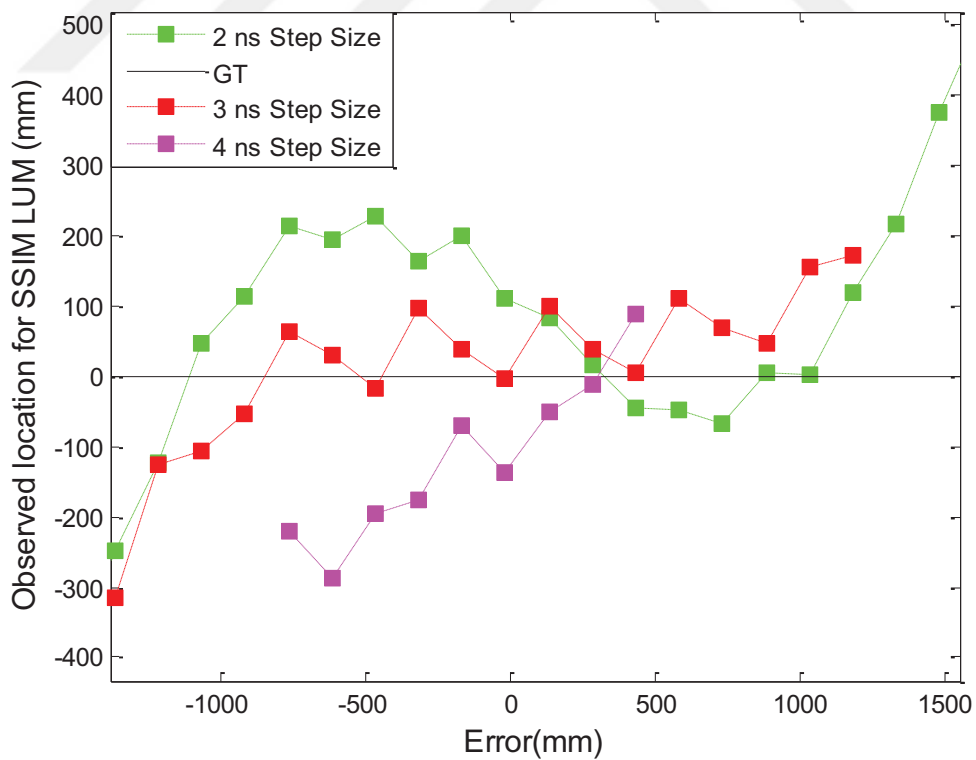
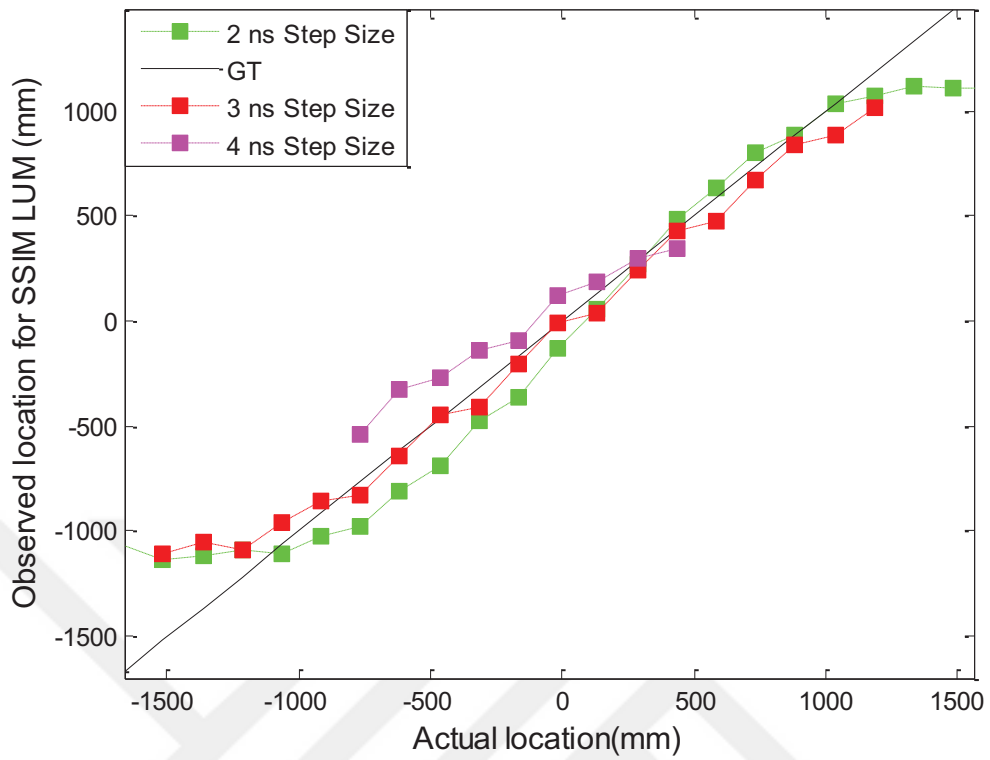


Figure 71 Observed locations (top) and the error values (bottom) for SSIM LUM method for 2ns, 3ns and 4ns step sizes.

Success rates for all step sizes are shown in Table 12. It is observed that step size = 3 ns results in best success ratio. It can be concluded that 30 ns search area and 3 ns resolution is the optimum point for our experimental setup.

Table 12 Success rates for Standing Man Scenario for 2, 3, 4 ns step sizes

Method	Success Rate (%)		
	Step size = 2 ns	Step size = 3 ns	Step size = 4ns
SSIM LUM	87.3	93.6	89.5

Difference between the Test Scenario 4 and Test Scenario 5 is the optical zoom levels which results on the quality difference between two images as shown in Figure 72. Although small amount of performance degradation occurs at zoomed out image, there is no significant failure for all comparison methods.

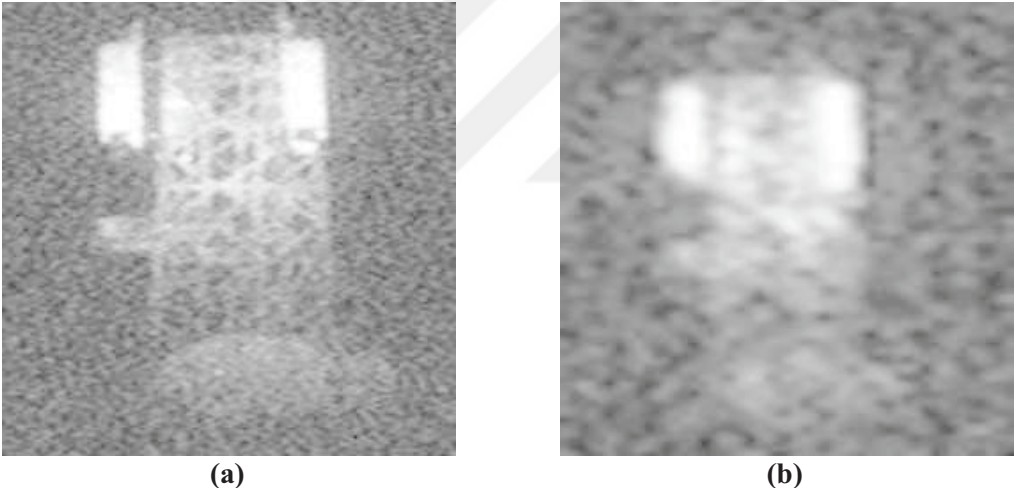


Figure 72 Image taken from Antenna (a) and the object region selected manually (b). For illustration purposes, dark pixels enhancement is applied by log transform.

CHAPTER 6

CONCLUSION AND FUTURE WORK

6.1 Conclusion

In this thesis, a range gated imaging system was constructed and a robust object tracking method is proposed for this system.

System noise characteristic was investigated and 302 mm fluctuation was observed on system output. The main factors of this fluctuation are the uncertainty of the laser timings and camera gate timings.

For calculating the ground truth object location, a method described in Section 4.2.1 is developed.

It is observed that SSIM LUM, IB, PwC, MI and PSNR are the prominent methods for object position estimation. SSIM CNTRST and SSIM STR have some drawbacks on noisy images and HC has weak points described in Table 6.

Range gated active imaging systems requires range information to illuminate the object. Although a laser range finder can be used for range estimation, it increases cost and complexity of the system.

Object tracking performance under partial occlusion and sudden X-Y movement cases were also investigated in the scope of this thesis. A method proposed to update the reference object but it must be improved by considering the X-Y movement of the object between frames.

6.2 Future Work

Used camera was a standard CCD camera with an analog video output. Output of the camera was interlaced video which means that the two fields are captured at two different time instances. In the constructed range gated system, the illumination on the detector occurs in the order of nanoseconds so only one field can be used in the observed image which halves the effective resolution. In the construction of a new system, use of progressive camera which captures the whole frame at a single time instance could be considered.

For future work, the system can be implemented in real time. By real time implementation, the effect of the accumulated error can be analyzed.

The reference object update method must be improved by considering the movement of the object between frames. Obtaining the exact location (in X-Y direction) of the object for eliminating the effect of object movements can be a solution for further studies. Besides that, noise characteristics of the reference image changes with proposed method which affects the performance of the comparison methods adversely. A better reference object update method can be developed by using the inputs of real time implementation.

REFERENCES

- [1] Middle East Technical University, Graduate School of Informatics Web site: <http://ii.metu.edu.tr/node/487>
- [2] Neumann, D. B. (1965). Range Gated Imaging System. *U.S. Patent 3.380.358*.
- [3] Kerpchar, M. (1972). Range Gated Image System Using Pulsed Illuminators. *U.S. Patent 3.689.156*.
- [4] Ulich, B.L., Keeler, R. N. & Phlipsen, K. (1990). Lidar System Incorporating Multiple Cameras for Obtaining a Plurality of Subimages. *U.S. Patent 4.967.270*.
- [5] Ulich, B.L., Keeler, R. N. & Phlipsen, K. (1991). Imaging Camera with Adaptive Range Gating. *U.S. Patent 5.029.009*.
- [6] Bonnier, D. & Larochelle, V. (1996). A range-gated active imaging system for search and rescue, and surveillance operations. *Proc. SPIE 2744 Infrared Technology and Applications XXII*, (pp. 134-145).
- [7] Repasi, E., Lutzmann, P., Steinvall, O., Elmqvist, M., Göhler, B. & Anstett, G. (2009). Advanced short-wavelength infrared range-gated imaging for ground applications in monostatic and bistatic configurations. *Applied Optics*, 48, 5956-5969.
- [8] Frecht, A. & Rothe, H. (1998). Comparison of imaging laser radars based on range gating under different weather conditions. *Proc. SPIE 3380 Conf. on Laser Radar Technology and Applications III*, 292.
- [9] David, O., Kopeika, N.S. & Weizer, B. (2006). Range gated active night vision system for automobiles. *Applied Optics*, 45, 7248-7254
- [10] Jin, W., Cao, F., Wang, X., Liu, G., Huang, Y., Qi, H. et al. (2008). Range-gated underwater laser imaging system based on intensified gate imaging technology. *Proc. SPIE 6621 Photoelectronic Imaging and Detection*.
- [11] Tan, C., Seet, G., Sluzek, A. & He D. (2005). A novel application of range-gated underwater laser imaging system (ULIS) in near-target turbid medium. *Optics and Lasers in Engineering*, 43, 995-1009
- [12] Driggers, R. G., Vollmerhausen, R. H., Devitt, N., Halford, C. & Barnard, K. J. (2003). Impact of speckle on laser range-gated shortwave infrared imaging system target identification performance. *Optical Engineering*, 42, 738-746
- [13] Zhang, S., Zhu, H.O, Wang, L. & Tang, S. (2010). Analysis on Detection Range of Laser Active Imaging System. *Academic Symposium on Optoelectronics Technology*, 236-239.

- [14] Gilles, J. (2007). Restoration algorithms and system performance evaluation for active imagers. *Proc. SPIE 6739 Electro-Optical Remote Sensing, Detection, and Photonic Technologies and Their Applications*.
- [15] Laurenzis, M., Bacher, E., Metzger, N., Schertzer, S. & Christnacher, F. (2012). Coding of range-gates with ambiguous sequences for extended three-dimensional imaging, *Proc. SPIE 8542, Electro-Optical Remote Sensing, Photonic Technologies, and Applications VI*.
- [16] Espinola, R.L., Jacobs, E.L., Halford, C.E., Vollmerhausen, R. & Tofsted, D.H. (2007). Modeling the target acquisition performance of active imaging systems. *Optics Express*, 15, 3816-3832.
- [17] Busck, J. & Heiselberg, H. (2004). Gated Viewing and High-Accuracy Three-dimensional Laser Radar. *Applied Optics*, 43, 4705-4710.
- [18] Laurenzis, M., Christnacher, F., Metzger, N., Bacher, E. & Zielenski, I. (2009). 3D range-gated imaging at infrared wavelengths with super-resolution depth mapping. *Proc. SPIE 7298, Infrared Technology and Applications XXXV*.
- [19] Laurenzis, M. & Bacher, E. (2011). Image coding for three-dimensional range-gated imaging. *Applied Optics*, 50, 3824-3828.
- [20] Andersson, P. (2006). Long-range three-dimensional imaging using range-gated laser radar images. *Optical Engineering*, 45, 034301.
- [21] Monnin, D. Schneider, A. L., Christnacher, F. & Lutz, Y. (2006). A 3D outdoor scene scanner based on a night vision range gated active imaging system. *Third International Symposium on 3D Data Processing, Visualization, and Transmission*, 938-945.
- [22] Steinvall, O., Andersson, P., Elmqvist, M., & Tulldahl, M. (2007). Overview of range gated imaging at FOI. *Proc. SPIE 6542, Infrared Technology and Applications XXXIII*, 654216.
- [23] Yilmaz, A., Javed, O. & Shah, M. (2006). *Object tracking: A survey*. *ACM Computing Survey*, 38, 4, 13.
- [24] Comaniciu, D., Ramesh, V. & Andmeer, P. (2003). Kernel-based object tracking. *Pattern Analysis and Machine Intelligence*, 25, 564-575.
- [25] Yilmaz, A., Li, X. & Shah, M. (2004). Contour based object tracking with occlusion handling in video acquired using mobile cameras. *IEEE Trans. Patt. Analy. Mach. Intell.* 26, 11, 1531-1536.
- [26] Ali, A. & Aggarwal, J. (2001). Segmentation and recognition of continuous human activity. *IEEE Workshop on Detection and Recognition of Events in Video*, 28-35.

- [27] Zhu, S. & Yuille, A. (1996). Region competition: unifying snakes, region growing, and bayes/mdl for multiband image segmentation. *IEEE Trans. Patt. Analy. Mach. Intell.* 18, 9, 884–900.
- [28] Paragios, N. & Deriche, R. (2002). Geodesic active regions and level set methods for supervised texture segmentation. *International Journal of Computer Vision*, 46, 3, 223–247.
- [29] Shi, J. & Tomasi, C. (1994). Good Features to Track. *9th IEEE Conference on Computer Vision and Pattern Recognition*, 593 – 600.
- [30] Guo-liang, W., Wang, L., Chun, Y. & Zhao-hua, H. (2007). Algorithm for Tracking of Fast Motion Objects with Adaptive Mean Shift. *Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing*, 359 - 363
- [31] Beauchemin, S.S. & Baron, J.L. (1995). The computation of optical flow. *ACM Computing Surveys*, 27, 433-466
- [32] Fleet, D. J. & Weiss, Y. (2005). Optical flow estimation. *Mathematical models for Computer Vision: The Handbook*, 239-257
- [33] Kristan, M., Pers, J., Kovacic, S. & Leonardis, A. (2009). A local motion-based probabilistic model for visual tracking. *Pattern Recognition*, 1, 4.
- [34] Kodama, T., Yamaguchi, T. & Harada, H. (2010). A method of object tracking based on particle filter and optical flow to avoid degeneration problem. *SICE Annual Conference*, 1529-1533.
- [35] Tung, T. & Matsuyama, T. (2008). Human motion tracking using a color-based particle filter driven by optical flow. *ECCV workshop on Machine Learning for Vision-based Motion Analysis ECCV'08*.
- [36] Spies, H., Jähne, B. & Barron, J. (2002). Range Flow Estimation. *Computer Vision and Image Understanding*, 85, 209-231.
- [37] Barron, J.L. & Spies, H. (2001). The Fusion of Image and Range Flow. *Lecture Notes in Computer Science*, 2032, 171-189
- [38] Wren, C., Azarbayejani, A. & Pentland, A.(1997). Pfunder: Real-time tracking of the human body. *IEEE Trans. Patt. Analy. Mach. Intell.* 19, 7, 780–785.
- [39] Mazin, A. A. A. & Miami, M. A. (2013). Effect of Atmospheric Attenuation on Laser Communications for Visible and Infrared Wavelengths. *Journal of Al-Nahrain University*, 16(3), 133-140.
- [40] Emeksiz, D. & Temizel, A. (2012). A continuous object tracking system with stationary and moving camera modes. *SPIE Security+ Defence, International Society for Optics and Photonics*, 854115.

- [41] Russakoff, D. B., Tomasi, C., Rohlfling, T. & Maurer Jr, C. R.(2004). Image similarity using mutual information of regions. *Computer Vision-ECCV*, 596-607.
- [42] ProxiVision GmbH Web site: <http://www.proxivision.de>
- [43] Beyan, C. & Temizel, A. (2015). A Multimodal Approach for Individual Tracking of People and Their Belongings. *Imaging Science Journal*, vol. 63, no. 4, pp. 192-202, doi:<http://dx.doi.org/10.1179/1743131X14Y.0000000101>, May 2015.
- [44] Beyan, C. & Temizel, A. (2012). Adaptive Mean-Shift for Automated Multi Object Tracking. *IET Computer Vision*, vol. 6, no. 1, pp. 1-12, doi:10.1049/iet-cvi.2011.0054.
- [45] Guler, P., Emeksiz, D., Teke, M., Temizel, A. & Taskaya Temizel T. (2013). “Real-time Multi-Camera Video Analytics System on GPU”, *Journal of Real-Time Image Processing*, doi: 10.1007/s11554-013-0337-2, March 2013.
- [46] Gurcan, İ. (2014). *Hybrid CPU-GPU Implementation of Tracking-Learning-Detection Algorithm*. Unpublished dissertation. METU, Ankara.
- [47] Wang, Z. Bovik, A.C., Sheikh, H. R. & Simoncelli, E. P. (2004). Image Quality Assessment: From Error Visibility to Structural Similarity, *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600-612
- [48] Göhler, B., Lutzmann, P. (2012). Range accuracy of a gated-viewing system as a function of the number of averaged images, *Proc. SPIE 8542, Electro-Optical Remote Sensing, Photonic Technologies, and Applications VI*, 854205.
- [49] Wang, Z. Z. Bovik, A.C. & Simoncelli, E. P. (2003) Multiscale structural similarity for image quality assessment. *Signals, Systems and Computers, Conference Record of the ThirtySeventh Asilomar Conference*, volume 2, pages 1398–1402.
- [50] Sampat, M.P., Wang, Z., Gupta S., Bovik, A.C., & Markey M.K. (2009) Complex wavelet structural similarity: a new image similarity index. *Image Processing, IEEE Transactions*, 18(11):2385–2401

CURRICULUM VITAE

PERSONAL INFORMATION

Surname, Name: Yağcıoğlu, Mustafa
Nationality: Turkish (TC)
Date and Place of Birth: 17 August 1981 , Antakya
Marital Status: Married
Phone: +90 505 389 73 71
email: musyagcioglu@gmail.com

EDUCATION

Degree	Institution	Year of Graduation
MS	METU Electrical and Electronics Engineering	2008
BS	BİLKENT Electrical and Electronics Engineering	2004
High School	Naim Atakaş Anadolu High School, Ankara	1999

WORK EXPERIENCE

Year	Place	Enrollment System
2008- Present	ASELSAN A.Ş.	Engineer Software
2004-2008	AYESAS	Engineer
2003 July	IBM Türk	Intern Engineering Student
2002 July	AYESAS	Intern Engineering Student

FOREIGN LANGUAGES

Advanced English, Fluent

TEZ FOTOKOPİ İZİN FORMU

ENSTİTÜ

Fen Bilimleri Enstitüsü

Sosyal Bilimler Enstitüsü

Uygulamalı Matematik Enstitüsü

Enformatik Enstitüsü

Deniz Bilimleri Enstitüsü

YAZARIN

Soyadı : Yağcıoğlu

Adı : Mustafa

Bölümü : Bilişim Sistemleri

TEZİN ADI: Object Tracking With Range Gated Camera System

TEZİN TÜRÜ : Yüksek Lisans

Doktora

1. Tezimin tamamı dünya çapında erişime açılsın ve kaynak gösterilmek şartıyla tezimin bir kısmı veya tamamının fotokopisi alınsın.
2. Tezimin tamamı yalnızca Orta Doğu Teknik Üniversitesi kullanıcılarının erişimine açılsın. (Bu seçenekle tezinizin fotokopisi ya da elektronik kopyası Kütüphane aracılığı ile ODTÜ dışına dağıtılmayacaktır.)
3. Tezim bir (1) yıl süreyle erişime kapalı olsun. (Bu seçenekle tezinizin fotokopisi ya da elektronik kopyası Kütüphane aracılığı ile ODTÜ dışına dağıtılmayacaktır.)

Yazarın imzası

Tarih 07.09.2015