UNIVERSITY OF GAZİANTEP
GRADUATE SCHOOL OF
NATURAL & APPLIED SCIENCES

NAVIGATION OF A MOBILE ROBOT USING STEREO VISION

M. Sc. THESIS
IN
MECHANICAL ENGINEERING

BY
ALİ KILIÇ
JANUARY 2010

Navigation of a Mobile Robot Using Stereo Vision

M.Sc. Thesis
In
Mechanical Engineering
University of Gaziantep

Supervisor
Prof. Dr. Sadettin KAPUCU

By

Ali KILIÇ

January 2010

T.C.
UNIVERSITY OF GAZİANTEP
GRADUATE SCHOOL OF
NATURAL & APPLIED SCIENCES
(Mechanical Engineering)

Name of the thesis: Navigation of a Mobile Robot Using Stereo Vision
Name of the student: Ali KILIÇ
Exam date: 08.01.2010

Approval of the Graduate School of Natural and Applied Sciences

Prof. Dr. Ramazan KOÇ
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

Prof. Dr. Sedat BAYSEÇ
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

Prof. Dr. Sadettin KAPUCU
Supervisor

<u>Examining Committee Members</u>	Signature
Prof. Dr. Sedat BAYSEÇ	______________________
Prof. Dr. İbrahim H. GÜZELBEY	______________________
Prof. Dr. L. Canan DÜLGER	______________________
Prof. Dr. Sadettin KAPUCU	______________________
Assoc. Prof. Dr. Hakan YAVUZ	______________________

ABSTRACT

NAVIGATION OF A MOBILE ROBOT USING STEREO VISION

KILIÇ Ali
M.Sc. in Mechanical Eng.
Supervisor: Prof. Dr. Sadettin KAPUCU
January 2010, 78 pages

Computer vision system is one of the main aplications in the mobile robot perception systems. Mobile robots may be able to navigate in an unknown environment acquiring visual information of their surroundings with the aim of estimating the position and orientation of an obstacle which is in front of it. Moreover, the pose of the robot has to be estimated as accurate as possible during navigation. The main purpose of this thesis is to develope stereo vision depth perception and obstacle detection systems for collision free navigation of a real hydraulically actuated experimental crawling mobile robot in an unknown indoor environment. Within the scope of this thesis, a robust stereo depth perception and obstacle detection algorithms are developed. To express our approach easily, a Stereo Vision Graphical User Interface is also prepared. In addition in this study for safely navigation of the mobile robot within its environments and reliably move from start point to destination point, a navigation algorithm, obstacle avoidance strategy and motion planning algorithms are developed. In order to implement the prepared perception algorithm and navigation algorithms, an experimental crawling vehicle was constructed and modified into a mobile robot and architecture details such as robot hardware and software were described in this thesis. To accomplish position control of mobile robot in two dimensional plane, kinematics of the mobile robot was analyzed and to control the robot locomotion system, a motion control algorithm was developed by the use of the kinematic data of the robot. In order to control the crawling robot easily and to combine all developed algorithms, Manual and Autonomous Navigation Graphical User Interfaces were prepared and implemented on the robot. After optimization and combining all algorithms and codes, behavior of robot was observed. With the helps of developed algorithms, the robot was able to reach destination point without collision with any obstacles.

Key Words: Stereo vision, Mobile robot navigation, Stereo vision obstacle detection, Depth perception via stereo vision, hydraulically actuated crawling robot.

ÖZET

MOBİL BİR ROBOTUN STEREO GÖRÜNTÜ TEKNİĞİ İLE DOLAŞIMI

KILIÇ Ali
Yüksek Lisans Tezi, Mak. Müh. Bölümü
Tez Yöneticisi: Prof. Dr. Sadettin KAPUCU
Ocak 2010, 78 sayfa

Bilgisayar görüntü sistemi mobil robot algılama sistemleri içinde başlıca uygulamalardandır. Mobil robotlar tanımlanmamış bir ortamda kendi çevresinden, önünde duran bir objenin konumunu ve yönünü hesaplamak amacıyla topladığı bilgiler ile hareket edebilmelidirler. Bunun yanı sıra, bir robot dolaşım sırasında her konumda kendi pozisyonunu hesaplayabilmelidir. Bu tezin başlıca amacı hidrolik güç ile hareketlendirilen deneysel bir paletli robotun iç mekânlarda hiçbir engele çarpmadan dolaşımı için stereo görüntü mesafe algılama ve engel tanıma algoritmaları geliştirmektir. Bu tez kapsamında stereo görüntü mesafe algılama ve engel tanıma algoritmaları geliştirilmiş kullanılan yaklaşımı daha kolay ifade edebilmek için Stereo Görüntü Kullanıcı Ara Yüzü geliştirilmiştir. Ayrıca bu çalışma da robotun kendi ortamında güvenli bir şekilde dolaşımı ve güvenilir bir şekilde başlangıç noktasından hedef noktasına hareket edebilmesi için bir dolaşım algoritması, engel kaçış stratejisi ve hareket planlama algoritması geliştirilmiştir. Hazırlanan algılama algoritması ve dolaşım algoritmasını deneyebilmek için deneysel paletli araç bir robot olarak değiştirilmiş ve geliştirilmiş ve bu tezin içinde robotun donanımsal ve yazılımsal mimarisi anlatılmıştır. Robotun iki boyutlu bir düzlem üzerinde konum kontrolünün yapılabilmesi için robotun kinematiği incelenmiş ve robot kinematiğinden alınan bilgilerin yardımıyla robotun hareket mekanizmasının kontrolü için hareket kontrol algoritması geliştirilmiştir. Robotu kolay bir şekilde kontrol edebilmek için ve bu çalışma sırasında geliştirilen algoritmaların bir araya getirilebilmesi için El ile Dolaşım ve Otonom Dolaşım Kullanıcı Ara Yüzleri geliştirilmiş ve robot üzerinde denenmiştir. Bütün algoritmaların ve kodların bir araya getirilip en iyileştirme yapılmasının ardından robotun hareketleri gözlenmiştir, hazırlanan algoritmaların yardımıyla robot başlangıç noktasından hedef noktasına hiçbir engele çarpmadan ulaşabilmiştir.

Anahtar Kelimeler: Stereo görüntü, Mobil robot dolaşımı, Stereo görüntü engel tanıma, Stereo görüntü ile mesafe algılama, Hidrolik hareketlendiricili paletli robot.

ACKNOWLEDGEMENTS

First of all I would like to thank to my supervisor Prof. Dr. Sadettin KAPUCU for his valuable guidance and suggestions during the preparation of this study. I would like to thank Prof. Dr. L. Canan DÜLGER for her positive criticism and suggestions. I would also like to thank Dr. Taylan DAŞ for his contributions and motivated supports. And finally, I would love to thank to my family for their endless supports.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

B	=	Distance between Stereo Camera Pair
C	=	Optical Center of Camera
d_x	=	Disparity in x Direction
d_y	=	Disparity in y Direction
f	=	Focal Length of Camera
i	=	Enumerator in x Direction
I_L	=	Left Image
I_R	=	Right Image
j	=	Enumerator in y Direction
L	=	Distance between Right Track Center and Left Track Center
P	=	Any Point of Any Object in Real World
p_l	=	Projection of Any Point P in Left Image Plane

p_r = Projection of Point Any *P* in Right Image Plane

R = Radius of Rotation of the Robot

$R(\theta)$ = Direction Cosine Vector

v_l = Horizontal Displacement of The Pixel Points with Respect to the Left Center of the Camera

v_r = Horizontal Displacement of The Pixel Points with Respect to the Right Center of the Camera

V_L = Left Track Velocity

V_R = Right Track Velocity

w_r = Rotational Velocity of Right Track

w_l = Rotational Velocity of Left Track

x = Image Coordinate in x Direction

y = Image Coordinate in y Direction

Z = Depth of Any Point in Robot Scene

ξ_I = Robot Position with Respect to Global Reference Frame

ξ_R = Robot Position with Respect to Robot Reference Frame

θ = Orientation of the Robot

ABBREVIATIONS

CCD	=	Charged Coupled Device
CMOS	=	Complementary Metal Oxide Semiconductor
DAQ	=	Data Acquisition Card
GPS	=	Global Positioning System
GUI	=	Graphical User Interface
IC Engine	=	Internal Combustion Engine
ICR	=	Instantaneous Center of Rotation
RGB	=	Red Green Blue (Color Image)
SAD	=	Sum of Absolute Difference
SSD	=	Sum of Square Difference
STP	=	Screw Terminal Panel
USB	=	Universal Serial Bus

CHAPTER 1

INTRODUCTION

1.1 Introduction

Mobile robots are a kind of robots which have the capability to navigate around their environment and are not fixed to one physical location in contrast with industrial robots. Navigation and motion planning in an unknown environment are important tasks for future generation mobile robots. The are several different fields in which it is possible to apply mobile robotic systems, such as applications in industrial environments, military and security applications, underwater and mining researches, the transport of dangerous substances, surveillance tasks and many others. In all these fields, it is necessary to equip the mobile robots with a sensor system which allows the vehicle to obtain information from the environment. The processing of these data, provided by the sensor system must be useful to facilitate actions and control planning of the mobile robot.

Navigation of mobile robots is an extensive topic and it covers a large spectrum of different technologies and applications. Its roots include many engineering and science disciplines, from mechanical, electrical and electronics engineering to computer science. Mobile robot navigation in an unknown environment is a very difficult problem to solve. The mobile robot must be able to safely navigate within its environment and reliably get from the starting point to the destination point. Additionally, mobile robot which can be placed in an unknown environment must discover unaided and collect sufficient information about its surroundings to be able to accomplish its task. For such a robot, the most important requirement is adequate sensing.

In order to achieve accurate navigation of the robot, it is very important to equip the robot with a sensory system, permitting the acquisition of information about the robot's surroundings. When considering an easy task of obstacle avoidance, the use of range systems such as ultrasonic and infrared sensors may be enough. However, with the aim of acquiring more accurate information and if we need more than only range information such as size and complete position of the obstacle the range systems will not be enough for sensing of robot’s environment. At that time it is related with the three dimensional perception systems as applied to mobile robots.

1.2 Three Dimensional Robot Perception

Spatial sensing is a technique to acquire three dimensional information of the environment. There are three main types of spatial sensors which are:

- Laser Range Finders
- Sonar Sensors
- Stereo Vision

1.2.1 Laser Range Finders

Laser range finders (i.e., 3D laser scanners) are sensors that measure the distance to objects in the closer surrounding by evaluating the time of flight of an emitted laser impulse. Laser rays are reflected by the objects in front of the laser source. An optical sensor is placed with a view angle slight different respect to the laser source as showed in Figure 1.1. From the difference between the expected point of incident and the measured one (where the laser ray hit the optical sensor) the depth information is evaluated. Laser scanners can send trillions of light photons toward an object and only receive a small percentage of those photons back via the optics that they use. The reflectivity of an object is based upon the object's color. [1] A white surface will reflect lots of light and a black surface will reflect only a small amount of light. Transparent objects such as glass will only refract the light and give erroneous dimensional information. In Figure 1.2 we can see typical laser range finder.

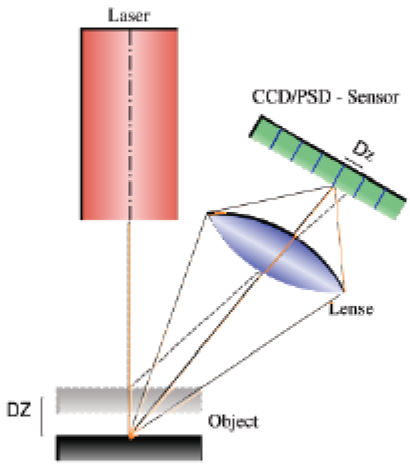


Figure 1.1 Schematics of A Laser Scanning Operation [1]

Figure 1.2 Sick Laser Scanner [2]

1.2.2 Sonar Sensors

Sonar sensors are employed to avoid obstacle. They have a small size and a low consummation, unfortunately their resolution is too low to reconstruct a dense depth map of the surrounding environment. To provide 3D information of the ambient with ultrasonic sensors, a scanning head composed of multiple sensors and motors are needed. A simple ultrasonic transmitter receiver pair is shown in Figure 1.3

Figure 1.3 A Simple Sonar Sensor [3]

1.2.3 Stereo Vision

Stereo vision is a widespread technique for inferring the three dimensional position of objects from two or more simultaneous views of a scene. Mobile robots can take advantages of a stereo vision system as a reliable and an effective way to extract range information from the surroundings. Accuracy of the results is generally adequate for applications such as depth perception and obstacle avoidance. Furthermore stereo vision is a passive sensing technique and there is no interference with other sensor devices when multiple robots are present in the same environment.

Figure 1.4 A Parallel Stereo Vision Camera [4]

Stereo vision systems do not need complex hardware two coupled video cameras are the minimal requirement. We can build our own binocular vision system, and fit its size to our design constraints. In stereo vision system the depth of information is not directly measured thought the hardware but it has to be extrapolated from the binocular images by helps of disparity.

Consequently, Sonar sensors are fast but inaccurate, whereas a laser scanning system is active, accurate, but rather slow. Stereo vision systems are passive and provide high resolution depth map. So, stereo vision can be used a preferable alternative to active sensing devices such as laser range finders and sonar sensors.

When mobile robots move in a real environment, recognition of surrounding objects is a big matter to deal with. Visual information is widely used for navigation and obstacle detection of mobile robots. In this research, we consider to use a stereo vision sensor to safely navigate our robots. Since stereo vision provides dense depth maps, this allows it to resolve small or narrow objects significantly better than sonar sensors. Laser range finders obtain high degree of accuracy but require a long time to scan an area.

In order to navigate a mobile robot in unknown environment, the obstacle, which is in front of the robot, must be detected and depth information of the obstacle must be calculated. However, in order to avoid obstacles, when there is enough information about obstacles and their distance from the robot, complete information about the object in front of the robot is not required. Obstacle avoiding is an essential task for mobile robots. This problem has been investigated to solve for many years by researchers and a lot of obstacle detections and avoiding systems have been proposed so far. Nevertheless designing an accurate and totally robust and reliable system remains a challenging task, in the real environments.

1.3 Research Objectives

The main objectives of this thesis are described in the following paragraphs.

The main concentration of this study is to have an autonomous robot that will be able to visually navigate an unknown environment and be able to keep away from the obstacles during navigation. Obstacles detection and avoidance have been performed by our robot using computer stereo vision system.

A navigation algorithm for a mobile robot has been developed using stereo vision by MATLAB programming language. In the navigation algorithm, the starting and destination coordinates are specified and robot travels from the starting point to the destination point without having a prior knowledge of the environment. Thus, robot has been able to avoid static obstacle in this manner.

Consequently, this thesis describes obstacle avoidance, in other words, collision free navigation using stereo vision perception system has been implemented on a hydraulically actuated crawling mobile robot.

1.4 Thesis Structure

This section outlines chapter by chapter the contents of the thesis. The body of this thesis is organized as follows:

Chapter 1 gives general information about mobile robot navigation using stereo vision and tries to explain why the stereo vision sensor is used instead of other spatial sensors. In addition, research objectives of this study and thesis structure are given in this part.

Chapter 2 presents a brief literature survey about stereo vision and mobile robot navigation using stereo vision system.

In Chapter 3, stereo vision perception system, stereo correlation algorithms, disparity map and depth map extracting have been discussed extensively. Additionally stereo camera calibration, pre-processing and post-processing operations such as image processing and image enhancement techniques are presented.

In Chapter 4, mobile robot navigation approach used in this thesis combination with stereo vision range data and also mobile robot navigation algorithms are presented.

In Chapter 5, mobile robot locomotion systems, robot architecture which is designed and developed in this study and hardware and software communication, and calibrations performed on the developed robot are presented.

In Chapter 6, the conclusion and discussion of the study is given highlighting the main issues and contributions and finally future works are discussed.

CHAPTER 2

LITERATURE SURVEY

2.1. Introduction

In the mobile robot navigation, there are two important requirements: perception and navigation. In this study, stereo vision has been used as a perception sensor and a brief literature review related with stereo vision is presented in section 2.2. Due to large number of available literature on the concept of mobile robot navigation, there is small number of studies closely related with mobile robot navigation using stereo vision and they are briefly presented in section 2.3.

2.2 Stereo Vision

In recent years, researchers have been giving special attention to computer vision systems which are capable of giving accurate three dimensional information on an observed environment, which leads to development of intelligent robots. Stereo vision systems are capable of extracting three dimensional features by passive sensing the environment [5, 6].

Generally, stereo vision system applications are composed of two camera configurations. In these systems, each camera delivers two dimensional representations of environment. Stereo vision systems are achieved by extracting three dimensional information by interpreting two or more two dimensional images of the environment. The interpreted three dimensional information creates a map that describes which point in the two dimensional image corresponds to the same point in the three dimensional scene. In addition detailed description of the stereo vision problem has been widely explained in [7, 8].

Recently, several stereo algorithms have been proposed to solve the problem of finding correspondence of the stereo images or calculating disparity maps. In mobile robot obstacle avoidance, simple and fast methods are “Sum of Absolute Differences” (SAD) and “Sum of Squared Differences” (SAD) which calculate absolute and square differences of the pixel intensities and to measure the similarity between the images. [9]

2.3 Mobile Robot Navigation

There are a few considerable researches in the field of mobile robot navigation using stereo vision. These studies are presented in the following paragraphs.

Real time or nearly real time stereo images are available from different stereo hardware systems. But raw stereo images are useless for robot obstacle avoidance and navigation tasks. These images require significant post-processing operation to extract three dimensional range informations. For example, researchers generally have applied some image processing and image enhancement techniques. Additionally researchers have applied various navigation and obstacle avoidance techniques which are developed for different sensors.

Don Murray et al. [10, 11] studied stereo vision based mapping and mobile robot navigation. They used occupancy grid mapping and potential field path planning techniques to form a robust cohesive robotic system for robotic mapping and navigation in both Spinoza [10] and Jose robots [11]. In these projects, trinocular, which is described as three camera stereo vision system, was used. Researchers used some techniques to improve the quality of stereo vision results on a working system and several example implementation results are given in related references.

Kumano and Ohya [12] proposed a new obstacle detection method. This method was based on stereo depth measurement but there was no corresponding point matching. Proposed method was fast enough for mobile robot navigation, but, not suitable robust for obstacle detection. For example, some ghost objects could not been detected during navigation.

Chin et al. [13] proposed an approach for robot navigation using distance transform methodology (DT). DT can be used in path planning for indoor robot navigation and also in performing obstacle avoidance simultaneously. But DT method is inefficient when performing both the tasks of path planning and obstacle avoidance. Usually, both have to be coupled and DT is normally only used for path planning. The DT methodology developed in this paper [9] can solve the navigation problem by optimizing the DT algorithm and reducing the processing area. Finally, the researchers performed simulation and actual tests on their autonomous mobile robot to verify the algorithm.

Sabe et al. [14] performed stereo vision system on a humanoid robot QRIO and also they developed a method for path planning and obstacle avoidance allowing the robot autonomously walk around in a home environment. Researchers' approach was based on floor extraction using Hough transform from image captured by stereo vision camera.

Stereo vision system also has been used for small mobile robot. It is a very challenging work and it requires small computing system instead of computer. Mingxiang and Yunde [15] studied on stereo vision system on programmable chip (SVSoC) which can be implemented by one Field Programmable Gates Array (FPGA). The researchers used three miniature CMOS cameras that are trinocular stereo systems with triangular configuration and all these systems were mounted on small hexapod robot for obstacle avoidance and navigation.

Recently stereo vision system has been tried to use on a passenger car as a driver assistant systems to perform with reliability to avoid any potential collision with the front vehicle. The feasibility of these systems in passenger car requires accurate and robust sensing performance. Huh et al. [16] developed an obstacle detection system using stereo vision sensors. Proposed method in this paper utilized stereo vision feature matching, epipolar constraint and feature collection in order to robustly detect initial corresponding pairs. After initial detection the system executed the tracking algorithm for the obstacles. After this operation system could detect a front obstacle, a leading vehicle then the position parameters of the obstacles and leading vehicles could be obtained. Researchers also implemented this system on a passenger car and their performances were verified experimentally on a Korean highway.

CHAPTER 3

STEREO VISION AND DEPTH PERCEPTION

3.1 Introduction

Stereo vision is the process in visual perception technique leading to the calculation of depth from the two different projections of the world onto image plane of the two cameras. In other words stereo vision is a technique that uses two cameras to measure distances from the images, similar to human depth perception. Stereo vision is a part of the field of computer vision. It is generally used in mobile robotics to detect obstacles and depth measurements. In computer stereo vision operation, two cameras take pictures of the same scene, but they are separated by a distance exactly like our eyes. A stereo vision corresponding algorithm compares the images while shifting the two images together over top of each other, to find the parts that match. In other words, it tries to find for every pixel in the left image the corresponding pixel in the right image. Correct corresponding pixel is defined as the pixel representing the same physical point in the scene. The distance between two corresponding pixels in image coordinates is called the disparity and is inversely proportional to distance. The disparity at which objects in the image, best match is used by the stereo vision triangulation algorithm to calculate their distance.

Computer stereo vision can be divided into three main consecutive stages which are;

- Pre Processing Operations
- Stereo Match (i.e. stereo corresponding)
- Post Processing Operations

Figure 3.1 shows overall stereo vision system and all operations which are developed and used in this study.

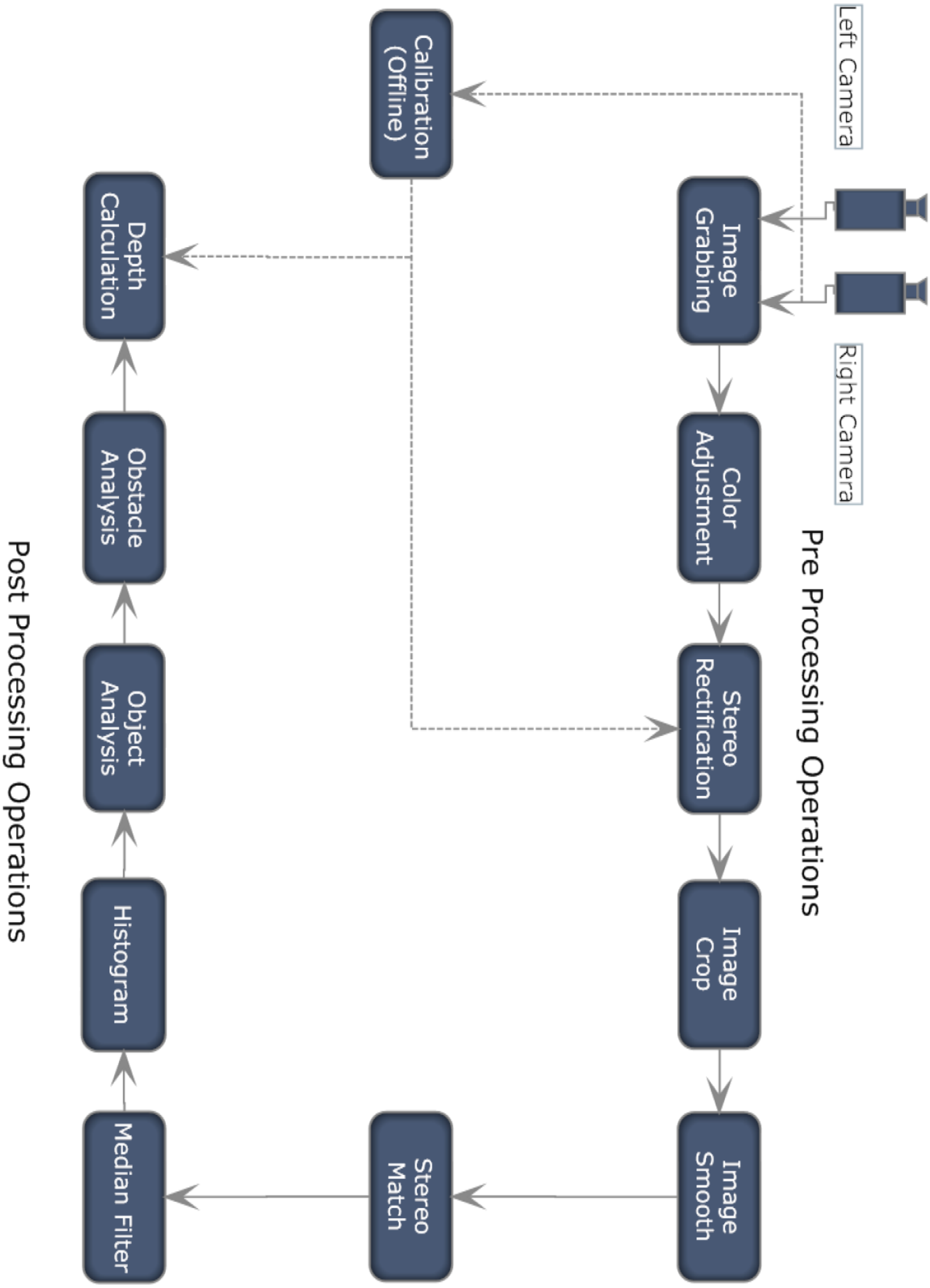


Figure 3.1 Overall Stereo Vision System and All Operations

In this study to understand stereo vision method easily for pre processing and post processing operation, a graphical user interface (GUI) has been prepared. Prepared GUI includes all pre processing, post processing operations and stereo match. In other words this GUI shows all steps of stereo vision depth estimation method which is proposed in this study. Figure 3.2 shows general appearance of Stereo Vision Depth Perception and Obstacle Detection User Interface.

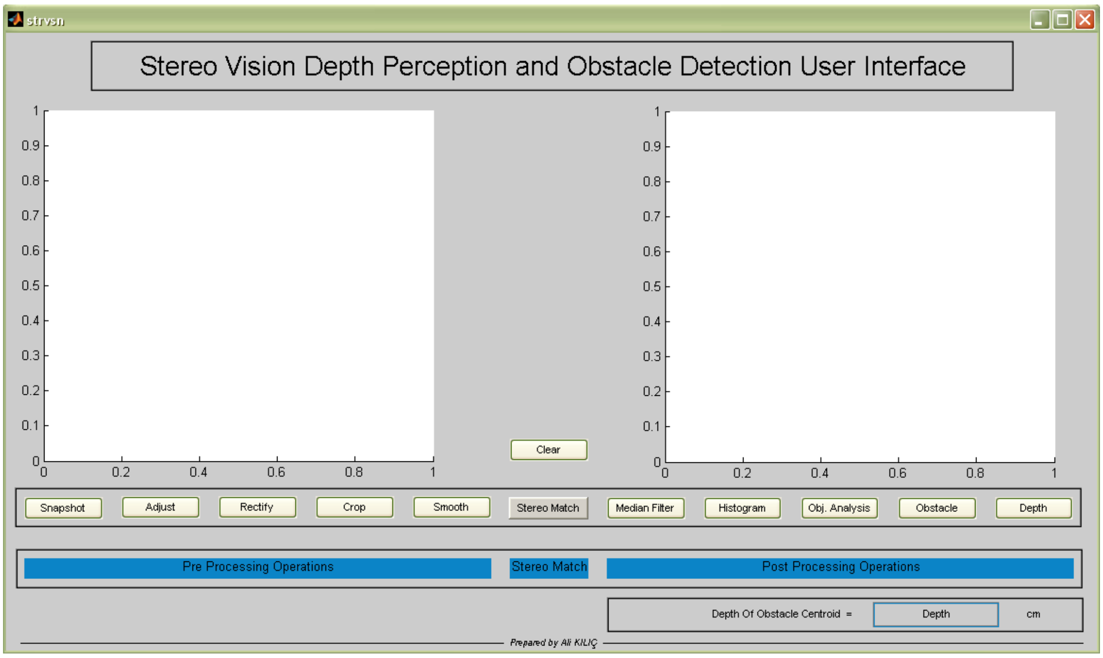


Figure 3.2 General Appearance of Stereo Vision Depth Perception and Obstacle Detection User Interface

3.2 Pre Processing Operations

Pre processing operations are all the operations that are done before stereo matching. These operations have crucial role in stereo vision depth perception, because raw images are not adequate enough for depth perception and images should be prepared for matching operation.

In this study five different pre processing operations are used which are;

- Image Grabbing
- Color Adjustment
- Stereo Rectification
- Image Crop
- Image Smooth

3.2.1 Image Grabbing

Image grabbing is the first operation of the stereo vision processes. This operation is used for transferring analog images into digital environment. In other words taken analog left and right pictures by help of stereo camera pair, are turned into digital images (i.e. arrays). Size and intensity of images are also set in this step. Detailed information about digital image can be found in Appendix A.

The binocular vision hardware (i.e. stereo vision cameras) has to be interfaced trough a driver in programming language in order to get synchronized images. The process uses two parallel cameras aligned at a known distance of separation. Each camera captures an image and these images are analyzed for same features.

In stereo vision image grabbing is performed via stereo rig. There are a lot of professional stereo rig in the market. But we don't need to use expensive hardware to be able to use in our robots. Today most webcams are of a reasonable quality and have a sufficiently high frame rate to be practical on slow moving robots. Because of this reason we decided to construct our own stereo rig.

3.2.1.1 Constructing Stereo Hardware

In constructing stereo vision hardware identical model USB webcams are used. When using the same model we ensure that the optics of both cameras as much as possible will be same and they will have the same field of view and focal lengths. Camera sensor is important criterion while selecting camera. In the market there are two types of optic sensors which are CCD (Charged Coupled Device) and CMOS (Complementary Metal Oxide Semiconductor). CMOS cameras are adequate in most situations, but under low illumination conditions such as artificial lighting CCD has superior performance characteristics. In this study another important criteria is field of view of the cameras. Wide field of view angle cameras have been used because a wide field of view means that the robot can see more than would otherwise at one time. When the mobile robot size is considered in this study the field of view would be better for obstacle avoidance. Considering this requirements, for this study two pieces Logitech QuickCam Pro are used. It has Carl Zeiss CCD sensor and 78 degree field of view.

In order to make a stereo rig a physical support is necessary to mount the cameras. The stability and the precision of the physical support are very important, because without a correct images alignment, the correspondence algorithm could not work properly. Regarding of this reason our camera support was cut by laser cutting machine and aluminum and Plexiglas strips were used as shown in Figure 3.3

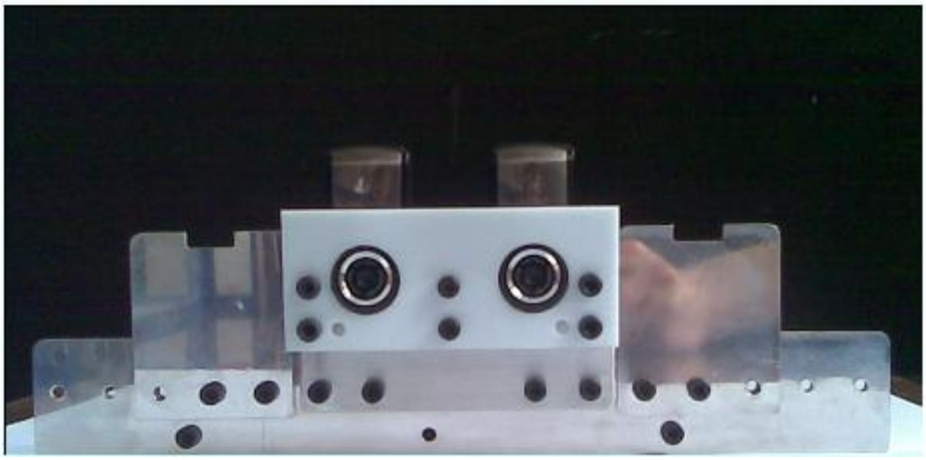


Figure 3.3 Constructed Stereo Vision Hardware

Example left and right images which are captured by prepared stereo vision hardware, can be seen Figure 3.4 in Stereo Vision GUI pane.

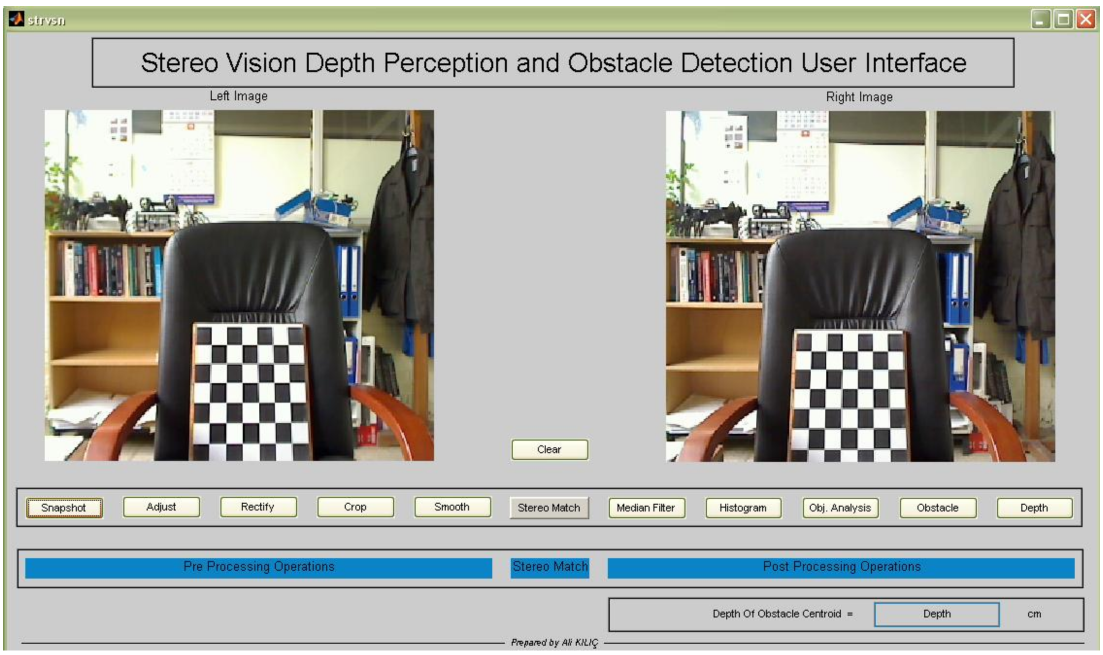


Figure 3.4 Captured Left and Right Images Using the Constructed Stereo Vision Hardware

3.2.2 Color Adjustment

Color adjustment operation is the second operation of the pre processing operations. This function is used for making contrast or brightness adjustment of stereo image pair. Contrast is the difference in brightness between two adjacent pixels. On the other hand, brightness refers to the overall lightness or darkness of an image [17]. This operation is necessary in order to improve the match quality, so stereo image pair should be set same contrast and same brightness level.

Figure 3.5 shows example of left and right images after color adjustment in Stereo Vision GUI pane.

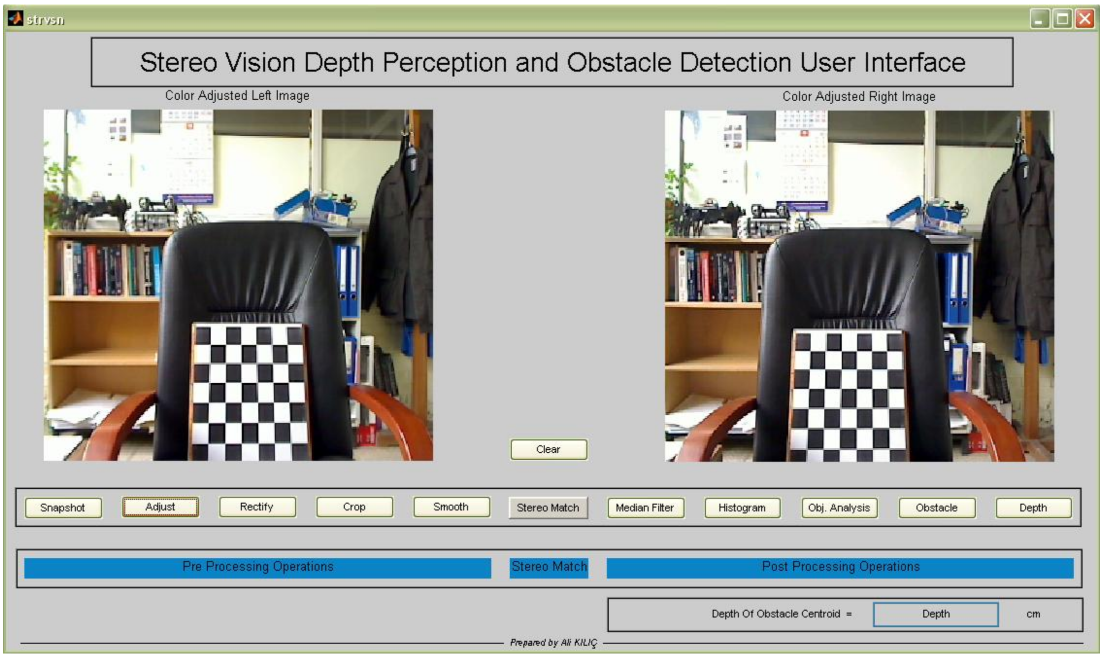


Figure 3.5 Left and Right Images after Color Adjustment

3.2.3 Stereo Rectification

In order to improve stereo correspondence in other words in order to get correct depth map, images require some pre-processing operations for example reduction of noise, adjustment of illuminations or a white balance of each cameras. But the most effective pre-processing operation is the calibration of the cameras and rectification of images to use of the epipolar constraint. Meaning of epipolar constraint is pixel in the left image the corresponding point in the right image lies on the same horizontal line (i.e. epipolar line). This constraint is used to reduce the searching area of the correspondence algorithms that calculates depth maps.

Camera calibration is an essential preliminary step of calculating depth from stereo vision. In order to use of the epipolar constraint the each camera of the stereo camera system need to be calibrated first, to get the intrinsic and extrinsic camera parameters, image rectifications can be done by helps of these parameters.

3.2.3.1 Camera Parameters

3.2.3.1.1 Pinhole Camera model

Every camera maps the three dimensional environment to a two dimensional image. The simplest camera model that models this mapping is the pinhole camera model.

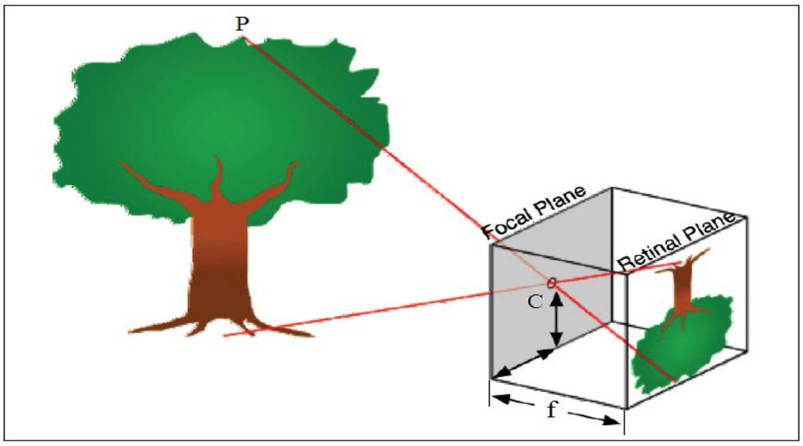


Figure 3.6 Pinhole Camera Model [18]

A pinhole camera consists of two planes, the retinal plane and the focal plane with the optical center C in the middle as shown in Figure 3.6. On the retinal plane the image is formed, the focal plane is parallel to the retinal plane on a distance f which called focal length. A three dimensional point P from the real world is mapped to the two dimensional image via a perspective projection. Pinhole cameras are characterized by two sets of parameters. Internal or intrinsic parameters, describe the internal geometry and the optical characteristics of the camera. Extrinsic or external parameters describe the camera position and orientation on the real world. To compute a comparison between two images captured from two different cameras, intrinsic and extrinsic parameters are essential.

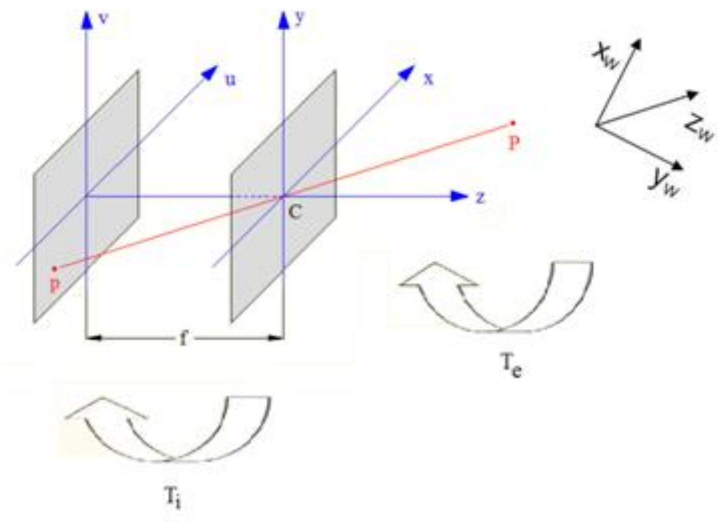


Figure 3.7 Intrinsic and Extrinsic Camera Parameters

As shown in Figure 3.7 the system for modeling two cameras consists of three different coordinate systems, the world reference frame (x_w, y_w, z_w), the camera frame (x, y, z) with the optical centre as origin and the image frame (u, v). A three dimensional point given in homogeneous world coordinates can be converted into the camera frame by a rotation r_{ij} and a translation t_j which is expressed by the extrinsic parameters T_e

$$T_e = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \tag{3.1}$$

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = T_e \times \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \tag{3.2}$$

Then this point is converted to the two dimensional image plane using the intrinsic parameters.

The intrinsic parameters are as follows:

f = focal length

(u_0, v_0) = center of the plane

(k_0, k_1) = the pixel size in mm

$$\propto = f/k_0 \tag{3.3}$$

$$\beta = f/k_1 \tag{3.4}$$

The transformation using the intrinsic parameters is as follows:

$$T_i = \begin{bmatrix} \propto & 0 & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3.5}$$

$$\begin{bmatrix} u \\ v \\ s \end{bmatrix} = T_i \times \begin{bmatrix} x \\ y \\ z \end{bmatrix} \tag{3.6}$$

Points on the focal plane, where z = 0 and s = 0 respectively, it cannot be transformed to image plane coordinates as division by zero is not defined and the straight line going through this point and the optical centre does not intersect with the image plane as it is parallel to the image plane. In summary, a point given in world coordinates is transformed onto a two dimensional image plane using the following equation

$$\begin{bmatrix} u \\ v \\ s \end{bmatrix} = T_i \times T_e \times \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \tag{3.7}$$

The knowledge of the intrinsic and extrinsic camera parameters allows for the rectification of images and after image rectification we ensure the epipolar constraint is ensured. The calculation of these parameters is needed for the camera calibration [19]

3.2.3.1.2 Camera Calibration

In order to assure the epipolar constraint with aligning horizontal image lines the two cameras need to be calibrated together. According to the chosen camera model the intrinsic and extrinsic camera parameters as described in previous part need to be acquired. Jean-Yves Bouguet’s developed a camera calibration toolbox using Matlab and it based on the Tsai / Lenz camera model and extends it and also this algorithm allows the stereo camera calibration. [20]

The Tsai / Lenz camera model improves the simple pinhole model that is described in previous part. It is more suitable for precise tasks as it takes circular distortion caused by the lens system into account. Points of the world reference frame are transformed to the image frame using the extrinsic and intrinsic parameters from equation 3.1 and 3.7

$$\begin{bmatrix} u \\ v \\ s \end{bmatrix} = \begin{bmatrix} \propto & 0 & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \times \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \qquad (3.8)$$

In order to get pixel coordinates u' and v', u and v are divided by s. Then the radial distortion first and second order (K_0, K_1) are taken into account by

$$u' = \frac{u}{s} \qquad (3.9)$$

$$v' = \frac{v}{s} \qquad (3.10)$$

$$d_1 = \left(\frac{u-u_0}{\alpha}\right) + \left(\frac{v-v_0}{\beta}\right) \qquad (3.11)$$

$$d_2 = {d_1}^2 \qquad (3.12)$$

$$u_{undistorted} = u' + (u' - u_0)(k_0 d_1 + k_1 d_2) \qquad (3.13)$$

$$v_{undistorted} = v' + (v' - v_0)(k_0 d_1 + k_1 d_2) \qquad (3.14)$$

Knowing the intrinsic and extrinsic parameters of each camera separately is sufficient to undistort the appropriate images but it is not sufficient to assure the epipolar constraint. The epipolar constraint ensures that the epipolar lines which are shown in Figure 3.8 coincide with the horizontal scan lines, therefore corresponding points in both images are only horizontally shifted, which reduces the searching area.

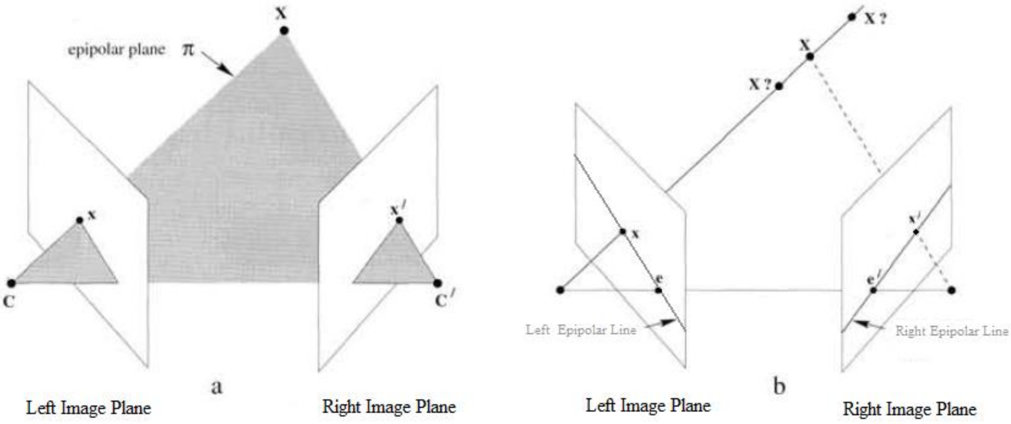


Figure 3.8 Epipolar Geometry [21]

Both intrinsic and extrinsic parameters of the camera system can be obtained from Bouguet’s camera calibration toolbox that uses a chess board as calibration pattern. In toolbox the process of camera calibration can be divided into three main steps:

- Image acquisition
- Extraction of the chess board corners in each image
- Computing the internal and external parameters

For image acquisition the two cameras in front of the robot are used to grab calibration pictures of a chess board from different distances and orientations simultaneously as shown in Figure 3.9

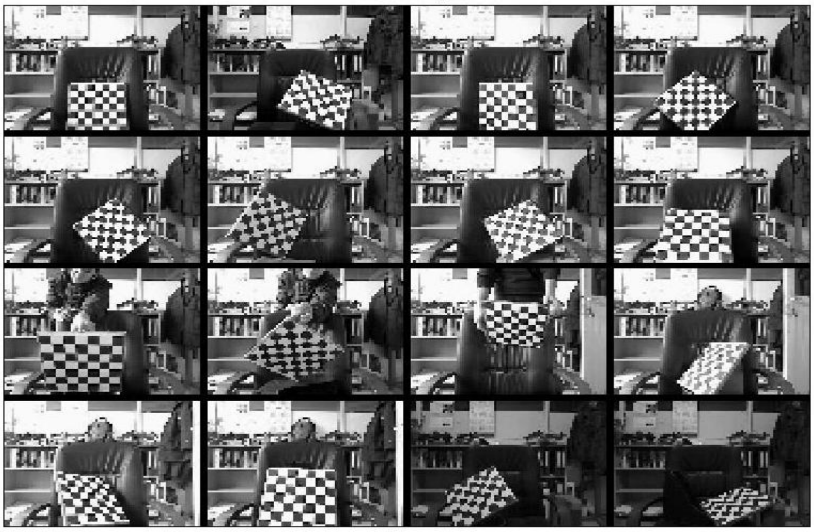


Figure 3.9 Left Camera Calibration Images from Different Distances and Orientations

In stereo camera calibration firstly each camera calibrated separately. The internal and external parameters are calculated by the algorithm. After that stereo system calibration is necessary. Fortunately Bouguet's camera calibration toolbox supports the stereo system calibration and stereo image rectification. During stereo calibration algorithm calculate the global extrinsic parameters. In other words the world coordinates of the chess board corners relative to each camera are known and the position of the right camera with respect to the left camera is calculated.

Figure 3.10 shows positions and orientations of the chess board and cameras all together.

Global Extrinsic Parameters (position of right camera with respect to left camera):

Rotation vector: Om = [0.01731 0.00485 -0.00284]

Translation vector: T = [59.93889 -1.84089 -4.87161]

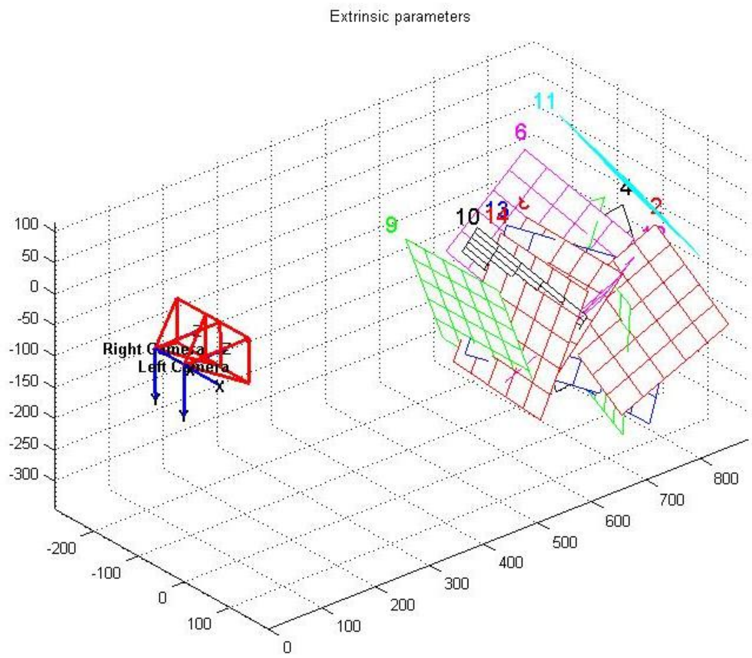


Figure 3.10 Global Extrinsic Parameters

3.2.3.1.3 Image Rectification

Image rectification is the correction of the distortion according to the calibration parameters calculated in the camera calibration step. After all intrinsic and extrinsic camera parameters are calculated both stereo images can be rectified according to these parameters. This way both images accomplish the epipolar constraint, it means that the corresponding pixels lie on the same horizontal line in both images as shown in Figure 3.11. Therefore Bouguet's algorithm pre-computes the necessary indices and blending coefficients to enable a quick rectification subsequently. [21]

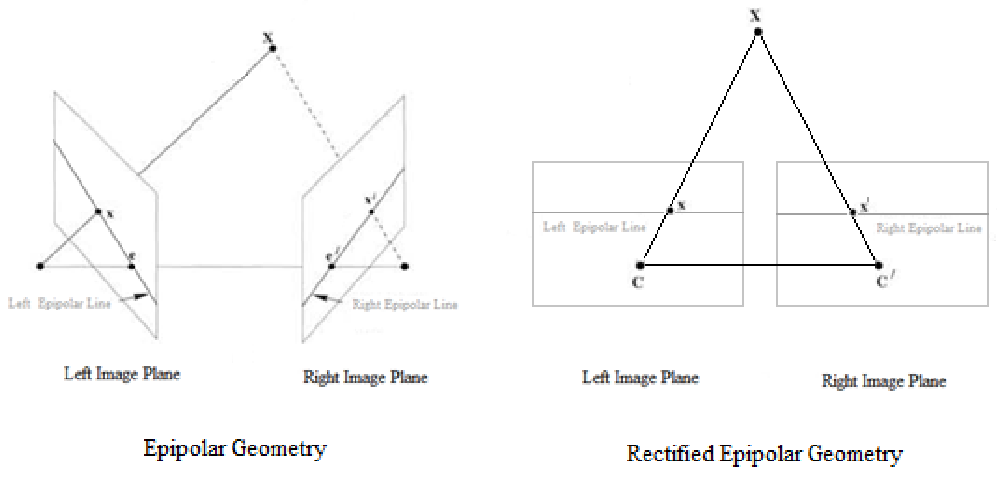


Figure 3.11 Epipolar Geometry Rectifications

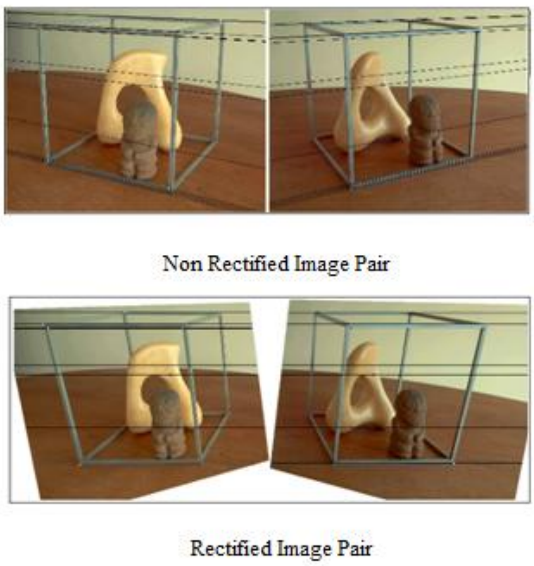
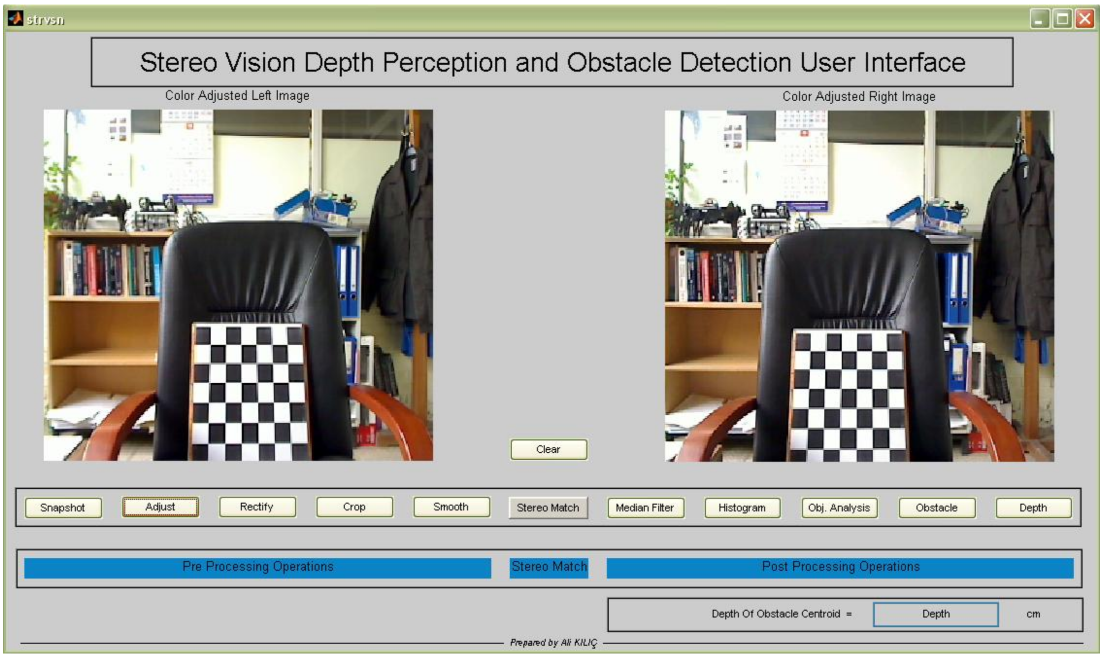


Figure 3.12 Sample Non Rectified and Rectified Image Pairs with Epipolar Lines [22]

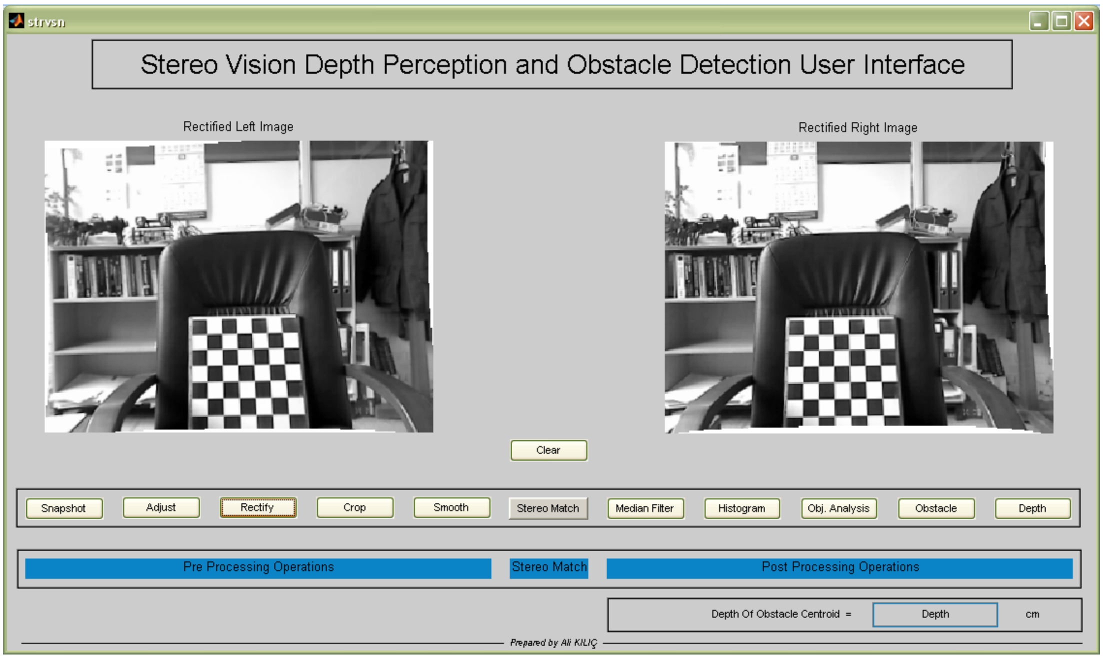
The rectified and unrectified images from the binocular vision hardware system is depicted in Figure 3.13 a and b. The same images are shown in Stereo Vision GUI pane in Figure 3.13 c and d.

(a)

(b)



(c)



(d)

Figure 3.13 Stereo Image Pair from Binocular Vision Hardware

a) Unrectified Left and Right Images

b) Rectified Left and Right Images

c) Unrectified Left and Right Images in GUI Pane

d) Rectified Left and Right Images in GUI Pane

3.2.4 Image Crop

After rectification operation some useless areas appear on both left and right images as seen in Figure 3.12 and Figure 3.13. For reducing computation time of stereo match, useless areas which are on the images are cut out by using image crop operation. Figure 3.14 shows left and right images in stereo vision GUI pane after crop operation.

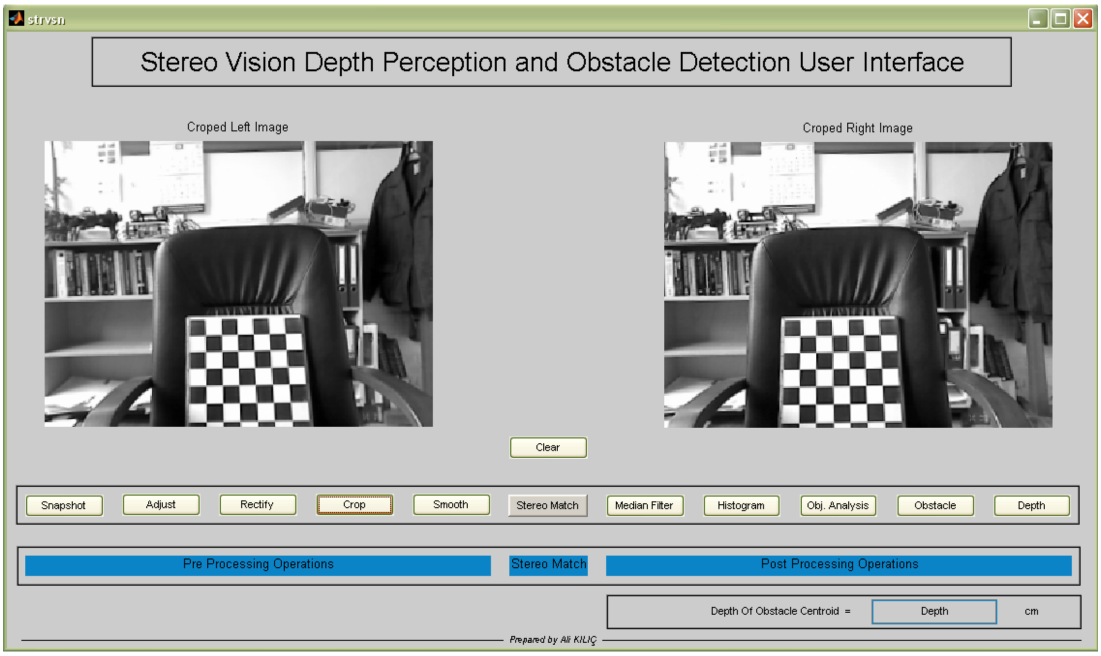


Figure 3.14 Stereo Image Pair after Image Crop Operation in GUI Pane. Useless areas on left and right images shown in Figure 3.13 are cut.

3.2.5 Image Smooth

Image smoothing operation (i.e. image blurring) is the last operation of pre processing operation and is used to make blur stereo image pair because image blurring reduces the sharp color changes. According to experimental results sharp color changes reduce stereo match quality. Because of that reason image smoothing operation were added into pre processing operations. Figure 3.15 shows stereo image pair after image blurring operation in stereo vision GUI pane.

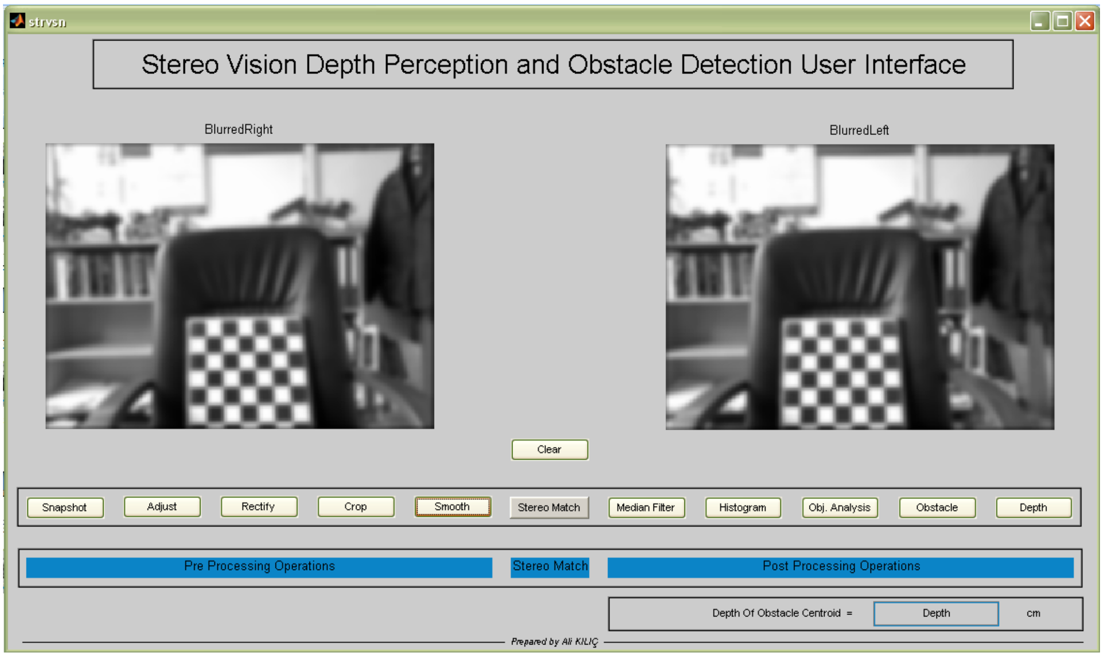


Figure 3.15 Stereo Image Pair after Blurring Operation

3.3 Stereo Match

Stereo vision correspondence analysis tries to solve the problem of finding which pixels or objects in one image correspond to a pixels or objects in the other image. This is stereo correspondence and also called as stereo matching. Stereo corresponding algorithms require a measure of similarity in order to find correspondences between the left and right image. The stereo correspondence algorithms can roughly be divided into feature based and area based, also known as region based or intensity based. [19]

Feature based algorithms extract features (e.g. edges, angles, curves, etc.) from images and try to match them in two or more views. They are very efficient but as drawback they produce poor depth maps.

Area based algorithms solve the correspondence problem for every single pixel in the image. Therefore they take color values and / or intensities into account as well as a certain pixel neighborhood. A block consisting of the middle pixel and its surrounding neighbors will then be matched to the best corresponding block in the second image.

In a robotic application generally an area based stereo algorithm have been used. Because feature based ones do not generate dense depth map and have difficulty to match smooth surfaces. Also the area based algorithm is faster than feature based. The main challenge of adopting an area based correspondence algorithm is its computational cost. An area based algorithm produces a dense depth map, which means that for each pixel of an image the algorithm tries to find its mutual pixel on the other view. This process is quite computationally expensive, but profiting on optimization techniques and by preprocessing operation such as image calibration and rectification. Finally we can reach a good compromise between the depth map density and its computation time.

In recent studies Sum of Absolute Differences (SAD) and Sum of Squared Differences (SSD) are the most widespread area or region based stereo correspondence algorithms in robotic and real time applications. Because they are faster than other algorithms and they give dense depth map (i.e. disparity map). According to these reason, SSD type stereo matching algorithm is used in this study. [19, 23]

$$SSD_x(x,y) = \sum_j \sum_i [I_L(x+i, y+j) - I_R(x+d_x+i, y+d_y+j)]^2 \qquad (3.15)$$

Where;

I_L = Left image (Reference image)

I_R = Right image

x = Image Coordinate in x direction

y = Image Coordinate in y direction

i = Enumerator in x direction

j = Enumerator in y direction

d_x = Disparity in x direction

d_y = disparity in y direction

In this study, during calculation of disparity between two stereo image pair, epipolar line constraint has been used. In other words stereo images are rectified with horizontal epipolar line. Because of that reason disparity in y direction is zero disparity only exist in x direction.

3.3.1 Disparity Map

A disparity map is a method for storing depth of each pixel in an image. Each pixel in the disparity map corresponds to the same pixel in a reference image. Determine the disparity of a physical point in multiple point of view projections. By repeating this process for all points of the 3D scene the correspondence phase computes a disparity map.

In Figure 3.16 are shown left and right pictures, which are taken from our stereo hardware, and in Figure 3.17 histogram equalized disparity map is shown. In disparity map objects which are on red pixels are close to cameras and objects which are on blue pixels farther away from cameras.

Figure 3.16 Stereo Image Pair

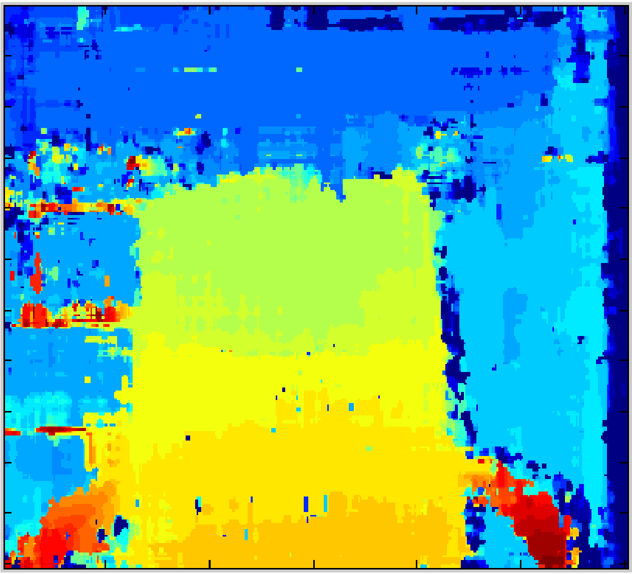


Figure 3.17 Disparity Map

3.4 Post Processing Operations

Post processing operations are all the operations that are done after stereo matching. These operations have crucial role in stereo vision depth perception because raw disparity map are not adequate enough and it is meaningless for depth perception and obstacle avoidance. Because of that reason disparity map should be passed through the following operations.

In this study five different operations are used which are;

- Median Filter
- Histogram
- Object Analysis
- Obstacle Properties
- Depth Calculation (Triangulation)

3.4.1 Median Filter

The median filter is a non-linear digital filtering technique, often used to remove noise from images or other signals. Median filtering is a common step in image processing. It is particularly useful to reduce speckle noise and salt and pepper noise [17]. After stereo matching produced, disparity map includes some noise due to mismatching. In order to remove this noise after stereo matching, median filter has been used.

Figure 3.18 shows left disparity map and median filtered disparity map in stereo vision GUI pane after median filter operation.

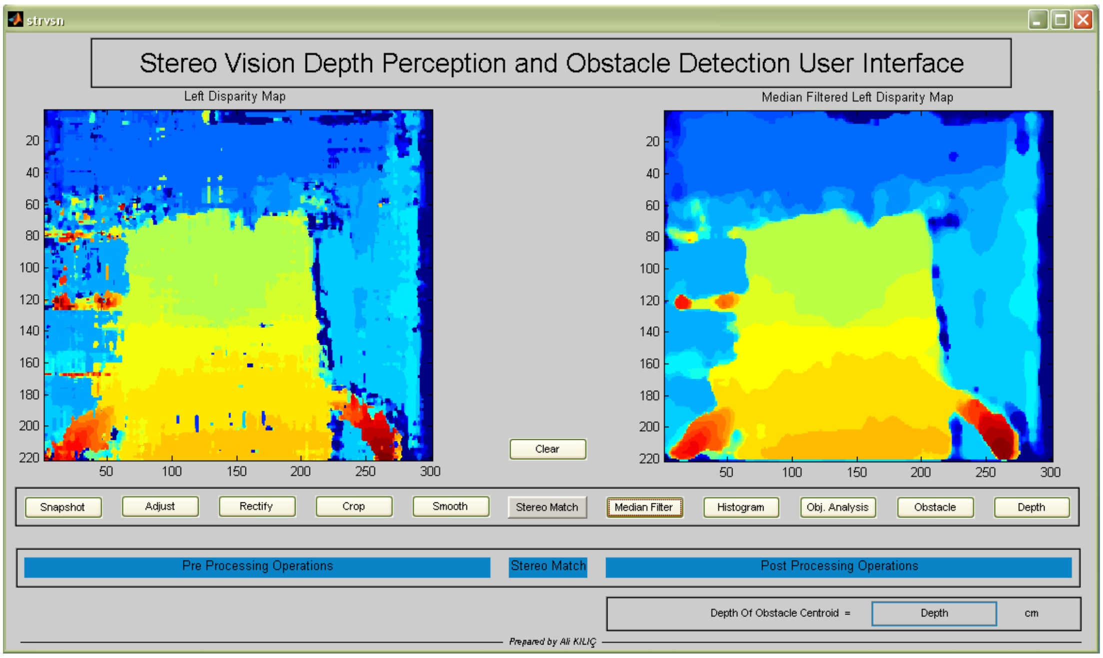


Figure 3.18 Left Disparity Map and Its Median Filtered Disparity Map

3.4.2 Histogram

An image histogram is type of histogram which acts as a graphical representation of the tonal distribution in a digital image. It plots the number of pixels for each tonal value. By looking at the histogram for a specific image a viewer will be able to judge the entire tonal distribution quickly. [24]

In disparity map each color represents a disparity level. Hence histogram of disparity map gives information about object or objects positions and by using this information an object is recognizable. In Figure 3.19 histogram of the median filtered disparity map shown in Figure 3.18 (left image on GUI pane) and Histogram of it after ghost object removed (right image on GUI pane). In the figure horizontal axis indicates disparity value 0 to 32 and vertical axis shows number of pixels.

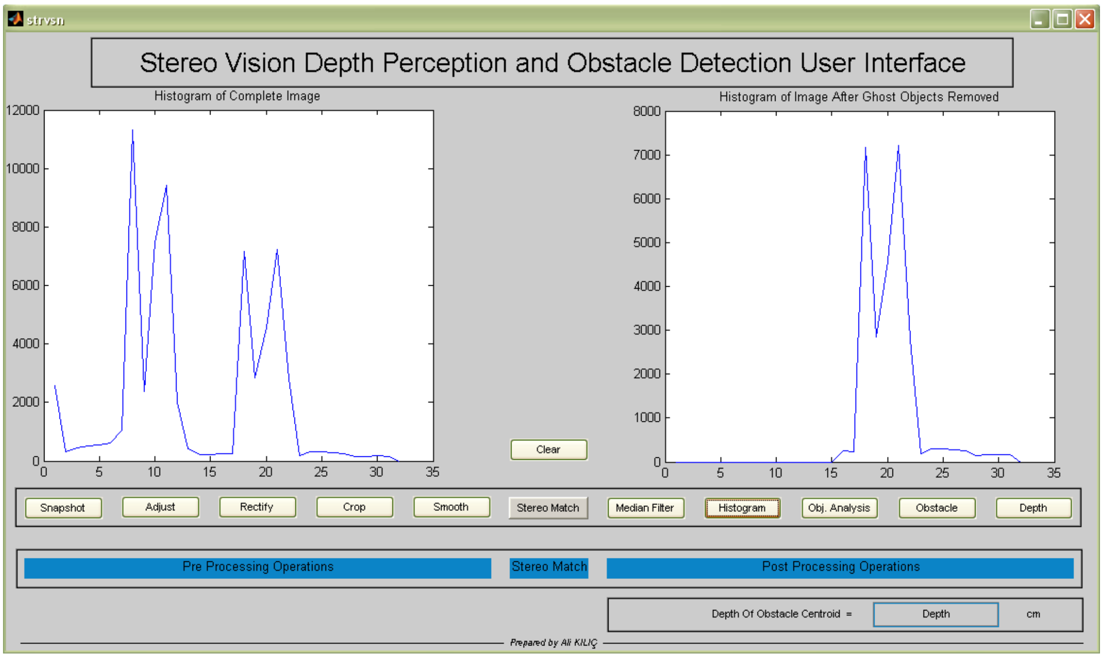


Figure 3.19

Histogram of Median Filtered Disparity Map Shown in Figure 3.18 (Left Image)

Histogram of It after Ghost Object Removed (Right Image)

In Figure 3.19 there are two different histogram displays. The left one is full histogram of disparity map after median filtering. In this study our stereo matching algorithm searches disparity at 5 bit level. In other words the closest object has a disparity value 32 and most far object has a 0 disparity value. If we look at the full histogram of the disparity map we can see three different groups. Of course the number of groups can be changeable. But the important group is last one every time because during robot navigation calculation is done for closest object. So the other objects can be canceled as seen histogram plot which is on right side on the stereo vision GUI pane.

The interpretation of histogram is main important operation for object detection and its disparity. After finding the closest object disparity level the contour of the object can be found easily.

3.4.3 Object Analysis

This function has been enhanced for finding the object position and object size in the disparity map. Because object position and size are important parameters for collision free robot navigation. This function firstly takes closest object depth information from histogram and then according to disparity level function combines neighbor disparity level and construct the near shape of the object.

In Figure 3.20 the left image shows calculated object position and size, the right image shows disparity levels of pixels.

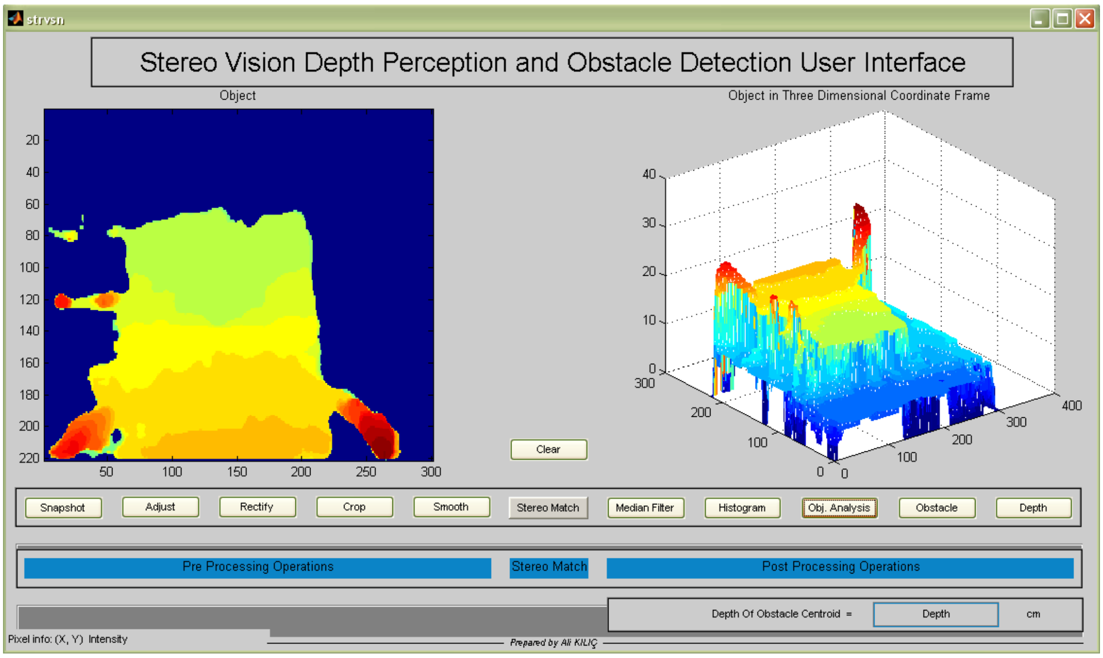


Figure 3.20 The Closest Object Analysis

3.4.4 Obstacle Analysis

Object analysis function is used to calculate area and centroid of the object which is found in object analysis step. Area of the object helps to get an idea about the obstacle size. Centroid of the object is used to calculate depth of obstacle instead of calculation of each pixel of the obstacle and also is used to determine position of the obstacle in the image.

Figure 3.21 shows calculated obstacle centroid and contour of the object. As shown in the left figure there is a blue box which shows the obstacle centroid.

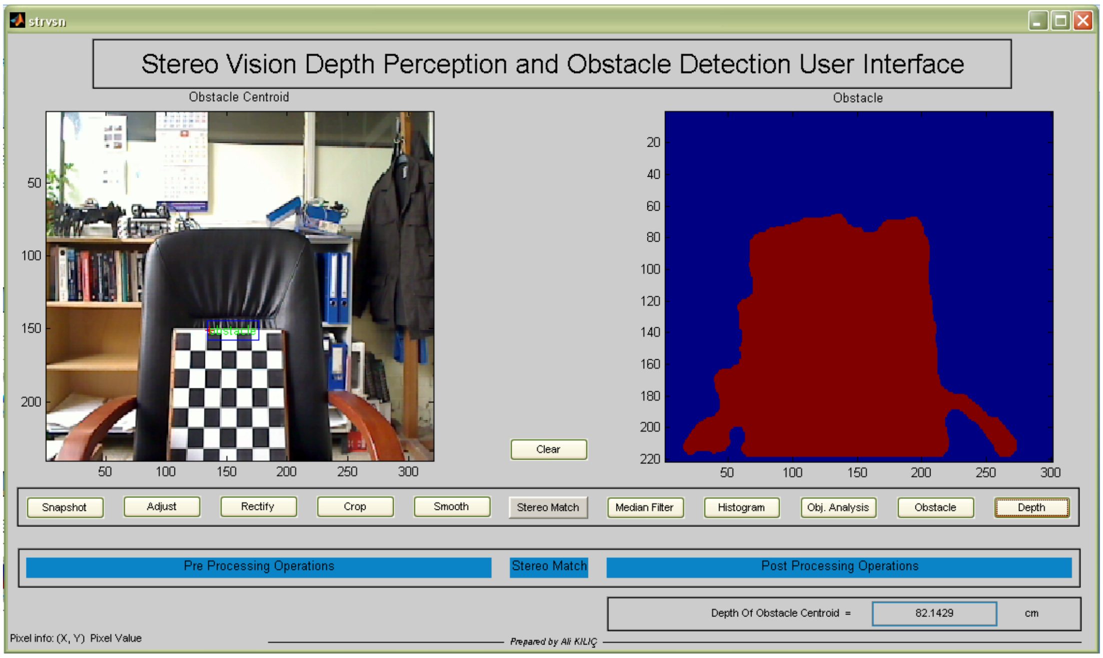


Figure 3.21 Detected Obstacle and Its Centroid on The Left. Detected Obstacle Contour of the Object on The Right

3.4.5 Depth Calculation

In depth calculation step, real depth of the obstacle centroid is calculated. Disparity value of obstacle centroid has been calculated in stereo match step. Using disparity value of centroid depth of the obstacle can be calculated via triangulation method and general information about triangulation and its mathematics are given in following section, it is shown in Figure 3.22

3.4.5.1 Triangulation

The technique for measuring depth information given two offset images is called triangulation. Figure 3.22 shows also geometry of triangulation and parallel optical stereo vision notations. Triangulation in stereo analysis is the task of computing the 3D position of any points in the images, given the disparity map and the known stereo geometry. In other words triangulation converts the disparity map into depth map and this map contains all depth information in the image. [20, 21]

For any point P of some object in the real world, p_l and p_r are pixel point representations of P in the left image plane and right image plane as taken by left and right cameras. f is the focal length of the camera (distance between retinal plane and focal plane). B is the offset distance between left and right cameras. v_l and v_r are the horizontal displacement of the pixel points with respect to the center of the camera. The disparity of the points p_l and p_r from image to image can be calculated by taking the difference of v_l and v_r. This is the equivalent of the horizontal shift of point p_l to p_r in the image planes. Using this disparity one can calculate the actual distance of the point in the real world from the images. The following formula can be derived from the geometric relation above and depth of the point P can be calculated by a triangulation formula which is:

$$Z = D = f\frac{B}{d}$$

Distance of point = (base offset) * (focal length of cameras) / (disparity)   (3.16)

Prepared stereo vision GUI also calculate depth of the obstacle and gives the depth value in centimeter in right down part of GUI pane. Calculated depth of the obstacle centroid in example stereo image pair is 82.14 cm as shown in Figure 3.21.

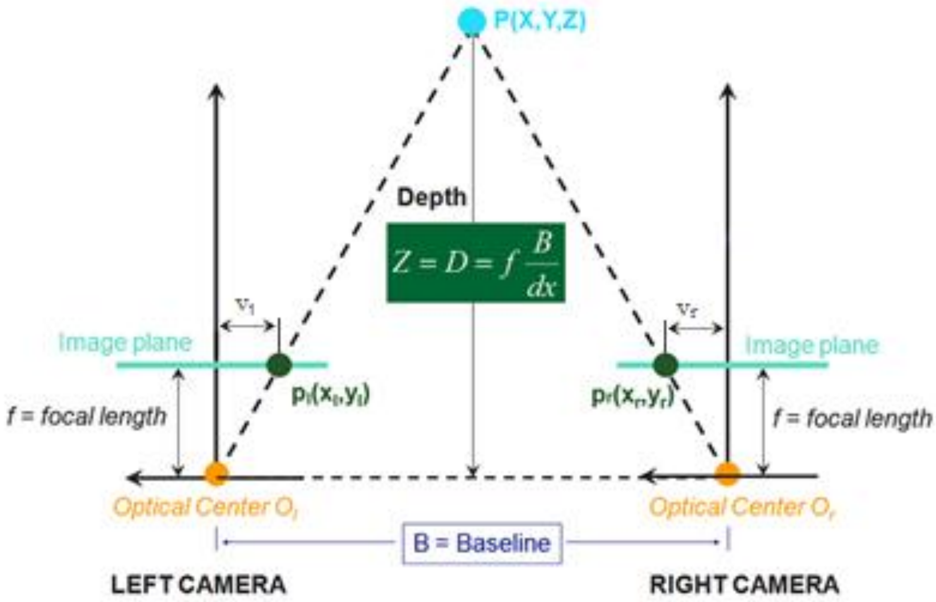


Figure 3.22 Stereo System with Parallel Optical Axis

CHAPTER 4

MOBILE ROBOT NAVIGATION

4.1. Introduction

For any mobile robotic system, the ability to navigate in its environment is one of the most important capabilities. Mobile Robots, which are equipped with computer vision, may be able to navigate around an unknown environment acquiring visual information of their surroundings with the aim of estimating the position of obstacles which stay in front of it. Actually mobile robot navigation consists of two main tasks. Firstly remaining operational that is avoiding dangerous situations such as, collisions, secondly going through the start point to destination point. In other words, robot navigation is an ability to determine its own position in its frame of reference and then to plan a path towards some goal location while avoiding collisions.

Navigation of mobile robots can be defined as the combination of the three fundamental abilities:

- Perception
- Localization
- Motion Control

Perception is the interpreting its sensor data (i.e. stereo image data) to meaningful depth data. Localization denotes the robot's ability to establish its own position and orientation within the start point or any reference. Also in that it requires the determination of the robot's current position and a position of a goal location, both within the same reference or coordinates frame. Motion control is the modulating of motor outputs to achieve the desired trajectory. Furthermore, interrelation between perception, localization and motion control can be seen in Figure 4.1 [25, 26]

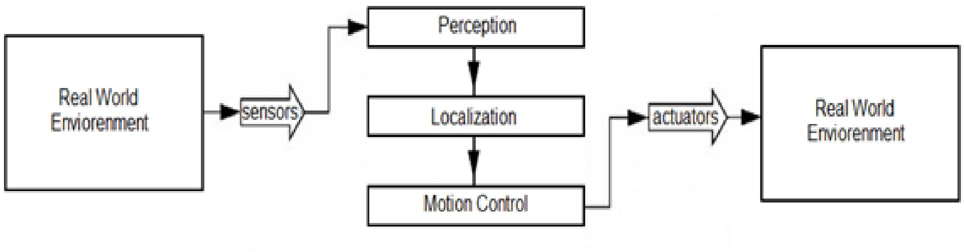


Figure 4.1 Mobile Robot Navigation

4.2 Perception

One of the most important tasks of an autonomous mobile robot of any kind is to acquire some useful information about its environment. This is done by taking measurements using various sensors and then extracting meaningful information from those measurements.

In previous chapter, mobile robot stereo vision perception system has been discussed. However, only depth perception is not adequate for mobile robot navigation. In order to safely navigation obstacle detection must be done by the robot perception system. Actually, not only obstacle detection but also position measurement of the robot has a crucial role during navigation. Because of that reason, Position measurement and motion planning sub system has been developed for safe navigation.
Perception system of mobile robots can be easily divided into two main groups;

- Depth Perception and Obstacle Detection
- Position Measurement of Robot

In this study obstacle detection of the robot includes a lot of calculations. The prepared stereo vision perception system gives high quality depth map, which is 240x320 pixels with 32 disparities, using two calibrated cameras with an algorithm named Sum of Squared Differences (SSD). But unfortunately using only depth map, accomplishing of mobile robot navigation is impossible. Hence an obstacle detection system has been developed and reported in previous chapter at the post processing part. So in this part, only position measurement of the robot will be discussed.

In order to perform robust and accurate robot localization, mobile robotics systems need some odometry sensors. There are a wide variety of odometry sensors used in mobile robots such as servo potentiometer or encoder for position measurement of wheels or tracks, tachometer for velocity measurement of wheels or tracks or GPS (Global Positioning System). GPS can give latitude, longitude and altitude of the robot but unfortunately for indoor applications GPS is useless. The mobile robot used in experiments is equipped with two servo potentiometers, two encoders and two tachometers.

During the experimental studies an experimental crawling robot, which is hydraulically actuated small scale tank, has been used. The name of the vehicle is Robotank. The research robot was developed and constructed at University of Gaziantep, Mechanical Engineering Department, Division of Theory of Machines and Mechanisms.

Our robot employs (i.e. Robotank), two servo potentiometers for position measurement of the tracks. They are enough for robust localization. Potentiometer has a capability to infinite turn. During linear movement each potentiometer gives a signal like a saw tooth. But unfortunately these signals are meaningless and due to the analog signal drawbacks they contain noise. In order to convert these analog voltage signals into the meaningful distance value, an odometry algorithm was developed. This algorithm is combination of some linear filter like mean filter, low pass filter, and also counting and summing algorithms. This algorithm calculates position of each track, turned angles and walked distances. Finally, measured and calculated parameters are sent to navigation algorithm and they are used for accurate localization.

4.3 Localization

Autonomous mobile robots need localization ability to move and to get to any goal location. During mobile robot navigation, both location and pose information of the mobile robot in its surroundings are very important. So robot current location, walked distance and distance between goal (destination) location and current location must be known by the robot navigation algorithm for accurate localization. Actually in order to sense robot current location, there are three methods. The first one is landmark sensing which can be usable with computer vision but it requires extra computation time and it needs landmarks. The second one is GPS (Global Positioning System). It is easy to use but it is useless for indoor applications. Third one is dead reckoning (i.e. odometry). It is very easy to use and of course it needs some calculations as it was discussed in previous section. But it is the most applicable method for our applications.

With the help of odemetry algorithm, mentioned in previous section robot is able to estimate its own position at any time and also it is able to calculate the distance between current location and goal location. Figure 4.2 shows data flow of obstacle avoidance algorithm operated during robot navigational motion.

The prepared robot navigation algorithm contains four different sub algorithms which are;

- Stereo Vision Depth Perception and Obstacle Detection Algorithm
- Odometry Algorithm
- Motion Planning Algorithm
- Motion Control Algorithm

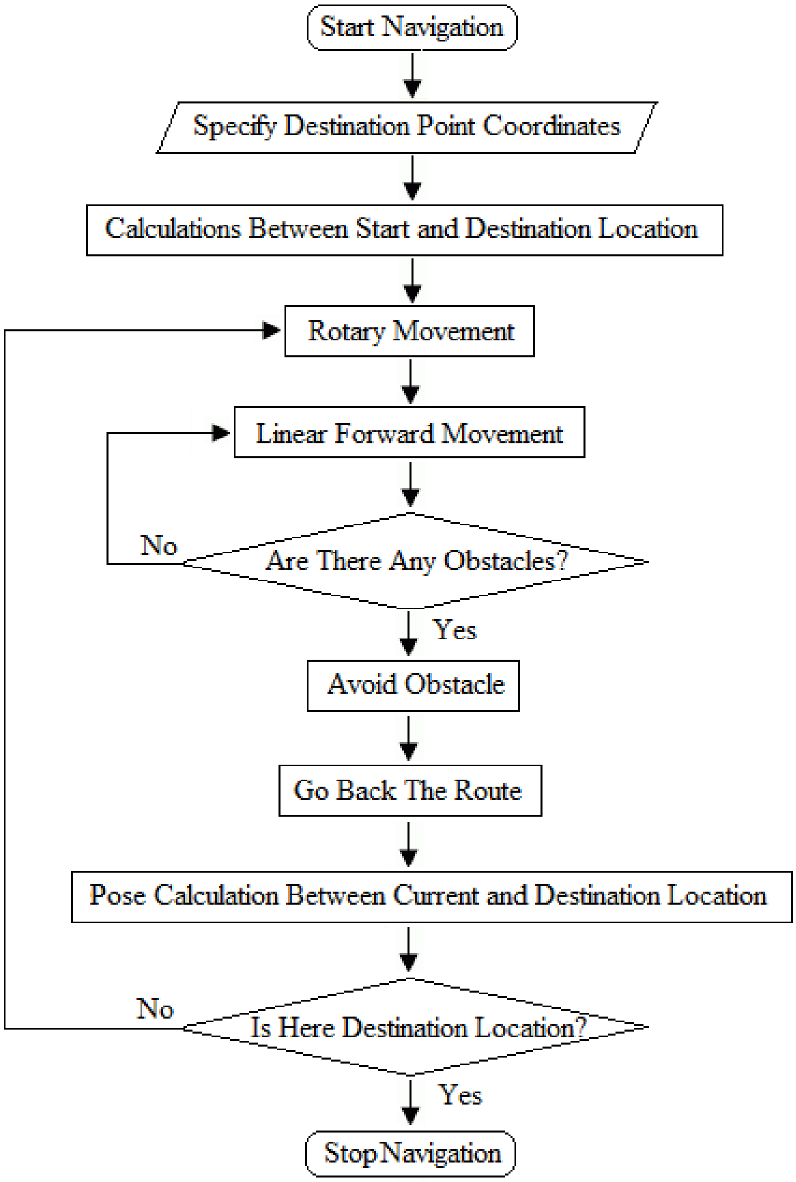


Figure 4.2 The Proposed Robotank Navigation Algorithm

Navigation algorithm of robot starts with specifying destination location coordinates. Navigation area of the robot is a two dimensional plane and it can be thought like a two dimensional Cartesian coordinate system as shown in Figure 4.3. Goal location coordinates has to be entered to the algorithm by the user both in x and y directions. In the next step navigation algorithm calculates the shortest distance between start point and destination point and turning angle. During this calculation the start point coordinates are taken into account as a zero in both x and y directions as shown in Figure 4.3. After the shortest distance and the turning angle calculated, by the help of robot motion control algorithm, the robot will turn up to calculated the turning angle and then robot starts to move in forward direction. During movement, stereo vision and depth perception algorithm checks the existence of obstacles. In the navigation algorithm "Are there any obstacles?" question is completely related with stereo vision depth perception and obstacle detection algorithm and detail of the algorithm can be seen in Figure 3.1. If the obstacle detection algorithm detects any obstacle, the robot tries to escape from collision and after escaping it tries to go back to the target route by the help of motion planning algorithm as shown in Figure 4.5. If there aren't any obstacles on the route then the robot keeps itself on the route and it keeps on linear movement. Finally at each position robot motion planning algorithm checks rest distance between current location and goal location by the helps of odometry algorithm. If the robot reaches the destination point then the navigation algorithm stops the navigation.

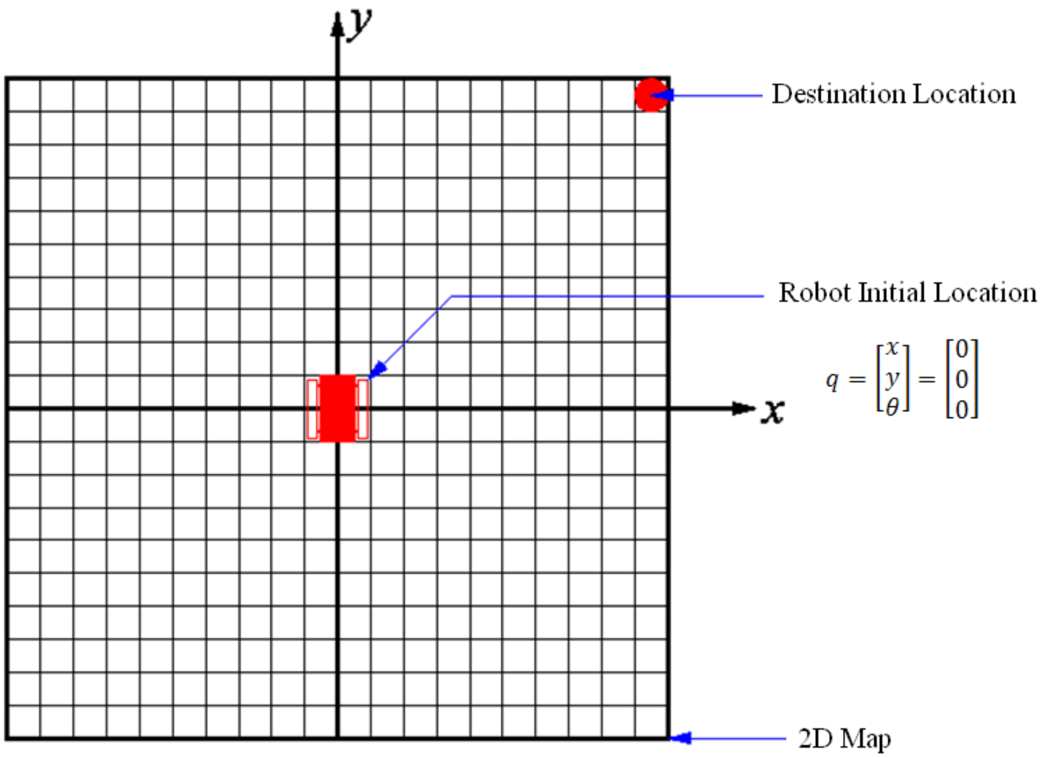


Figure 4.3 Robot Initial Position

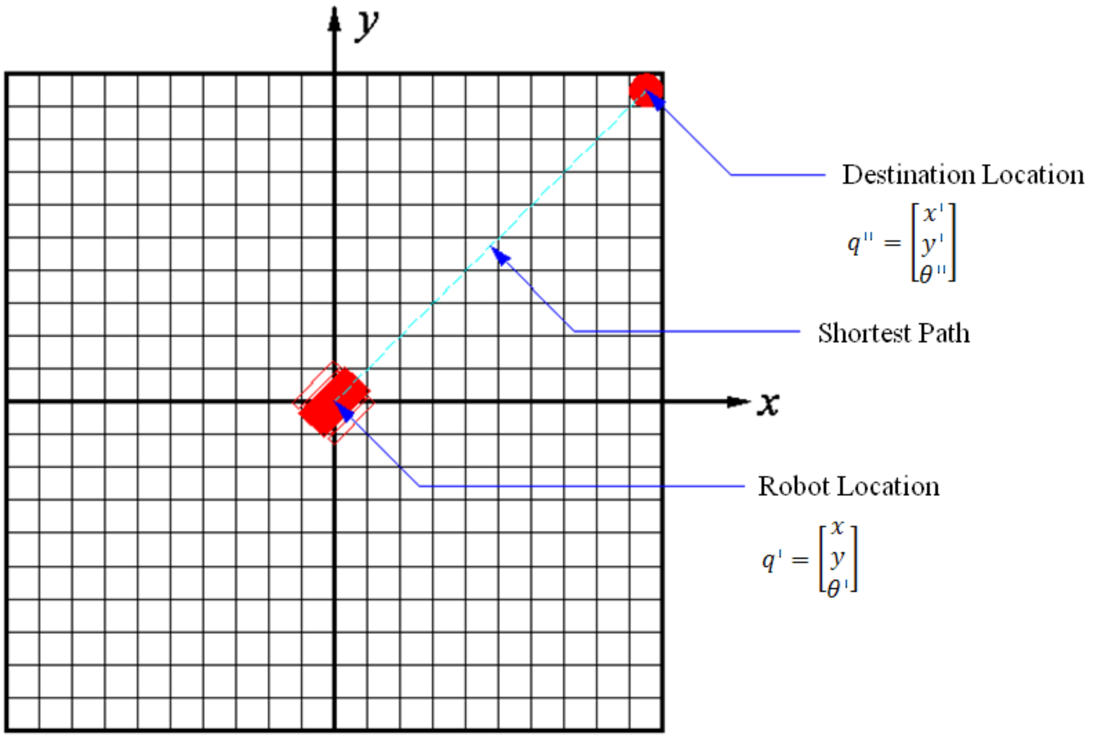


Figure 4.4 Robot After Directing Itself to the Destination

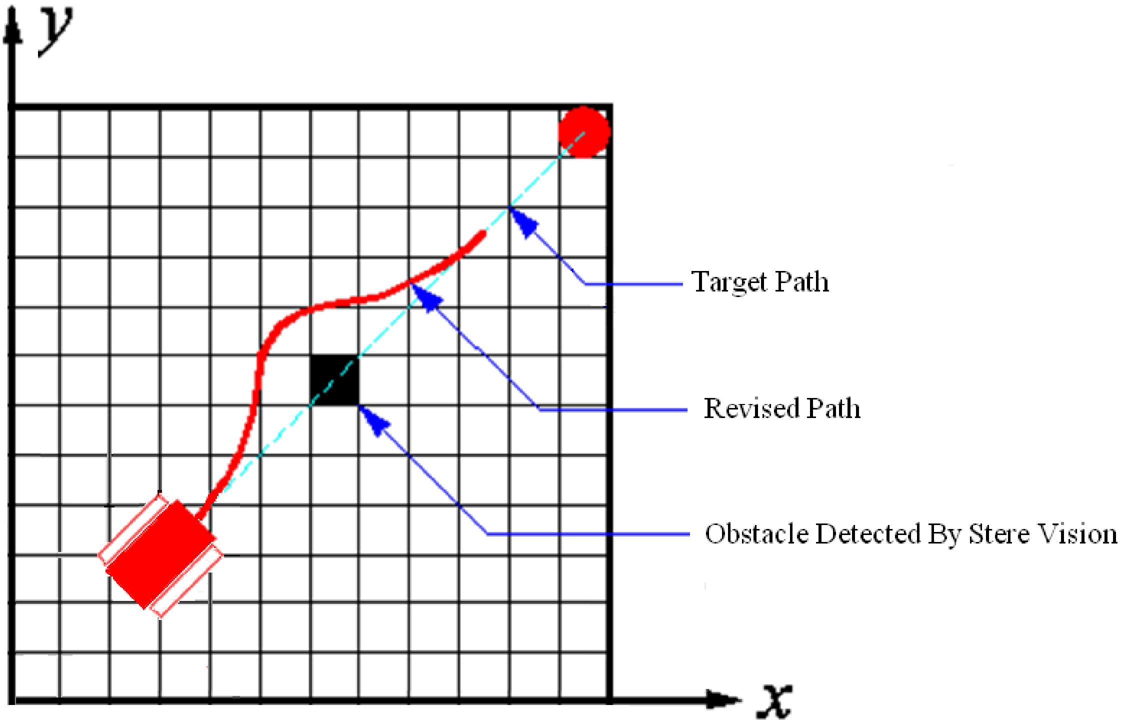


Figure 4.5 Robot During Navigation

4.4 Motion Control

Motion control of the robot is interconnection between perception algorithm and robot locomotion system. A basic motion planning problem is to produce a motion that connects a start position and a goal position, while avoiding collision with detected obstacles. Prepared stereo vision depth perception and obstacle avoidance algorithm (i.e. perception algorithm) gives depth map of the robot's current scene, and also if the algorithm detects any object, it is able to recognize as an obstacle, it is able to calculate the distance between robot and object and it is able to calculate centroid, occupied area and position of the obstacle. However, to escape the obstacle which is detected by the perception algorithm robot must have a motion control mechanism.

Motion control of Robotank contains two sub algorithms which are;

- Motion Planning Algorithm
- Motion Control Algorithm

At the beginning of the motion, motion planning algorithm calculates shortest distance and turning angle. During navigation if perception algorithm detects any obstacle, it sends the position information of obstacle to the motion planning algorithm. According to the obstacle information motion planning algorithm plans proper obstacle avoidance motion.

Motion Control Algorithm arranges track driving voltages according to the information which comes from the motion planning algorithm. In other words, the motion planning algorithm sends only distance and angle information to the motion control algorithm and then the motion control algorithm converts the distance and the angle information into voltage values by help of the odometry algorithm. Finally the motion control algorithm sends to voltage value to the robot locomotion system and also speed of the tracks are decided by this algorithm. Detailed information of robot locomotion system is discussed in chapter 5.

In order to use robot inextensive experimental stage easily, Robotank Manuel Navigation User Interface and Autonomous Navigation User Interface have been prepared. Robotank Manuel Navigation User Interface is used for positioning of robot by human guidance. In manuel navigation mode, the robot is same as hand controlled machine via computer. The manuel interface has five buttons for forward movement, backward movement, right turning, left turning and middle button to stop robot and also manuel interface is able to give total moved distance at the lower part of the pane. General appearance of manuel navigation interface can be seen in Figure 4.6.

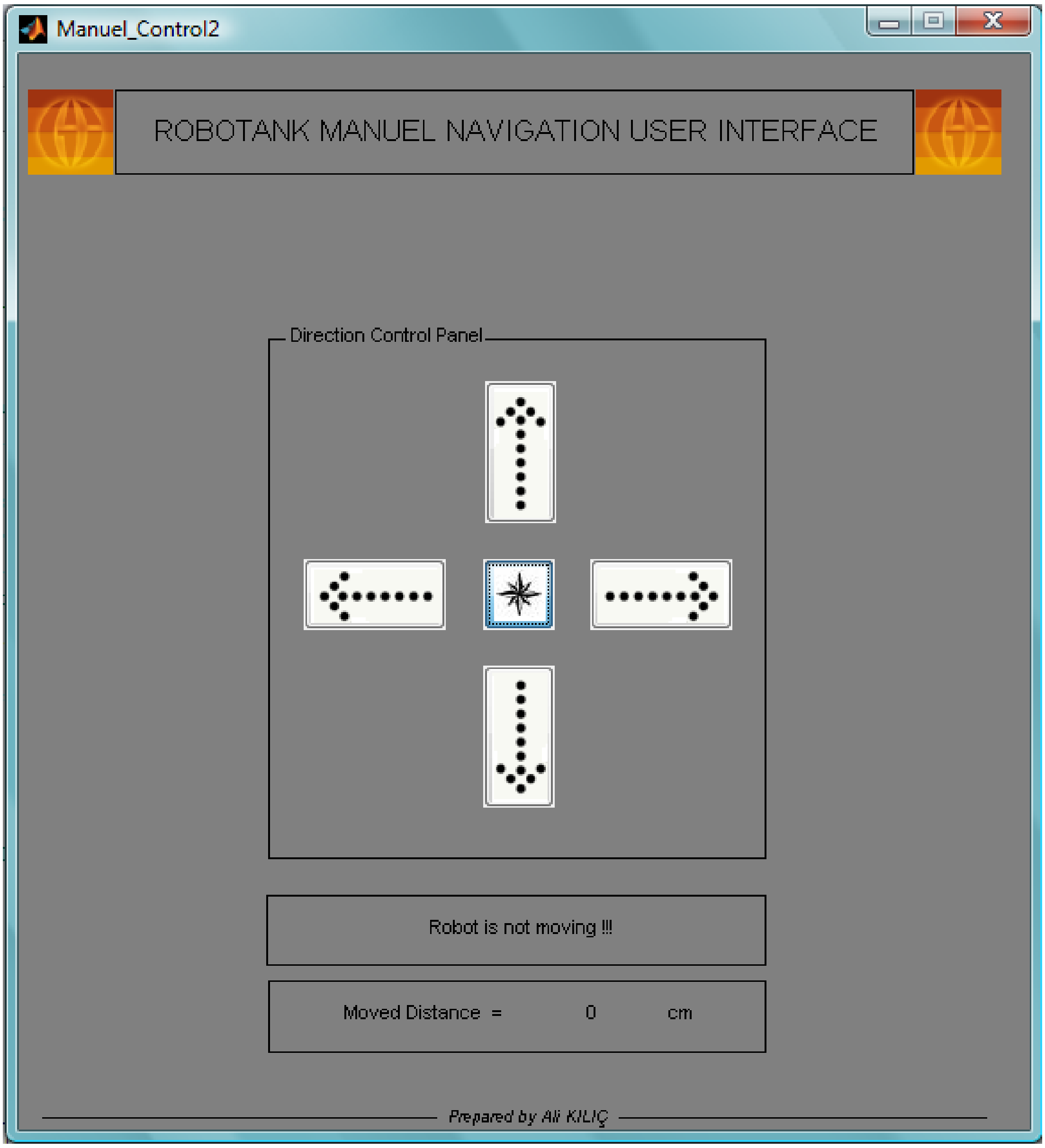


Figure 4.6 Robotank Manuel Navigation User Interface

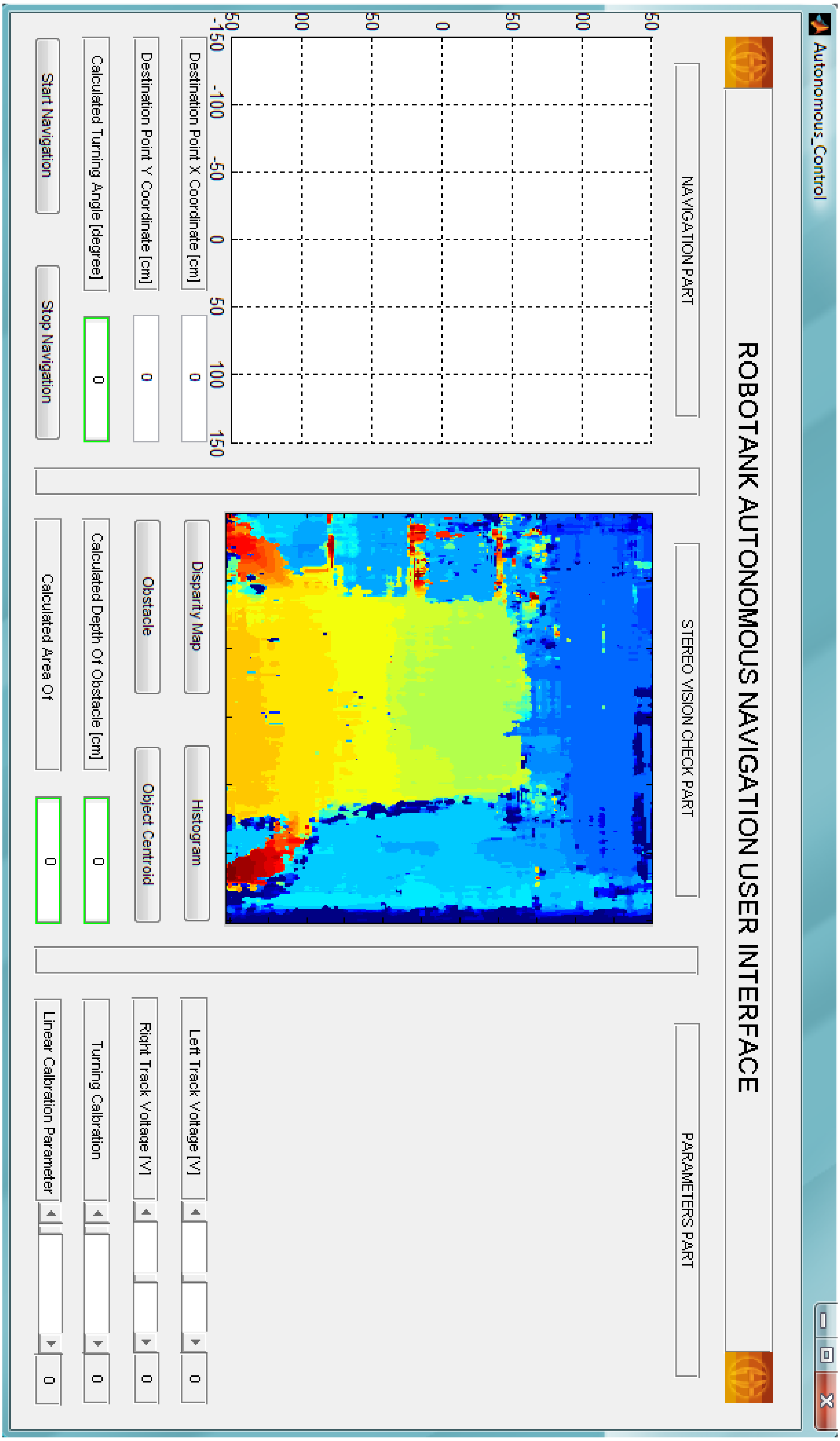


Figure 4.7 Robotank Autonomous Navigation User Interface

CHAPTER 5

MOBILE ROBOT LOCOMOTION

5.1 Introduction

A mobile robot needs locomotion mechanisms in order to move throughout its environment. In other words, robot locomotion is the study of how to design robot and control mechanisms to allow robot to move robustly. There are a large variety of possible ways to move, and so the selection of a robot's locomotion system is an important criterion of mobile robot design. Generally the robots that can walk, jump, run, slide, skate, swim, fly, and, of course, roll. Most of these locomotion mechanisms are duplication of biological systems and some of them can be found in Figure 5.1. There is only one exception that is wheel.

The wheeled locomotion is completely human invention and nature did not develop a fully rotating, actively powered joint, which is the mechanism necessary for wheeled locomotion. Wheeled robots are typically quite energy efficient and simple to control, but other forms of locomotion systems may be more appropriate for a number of reasons (e.g. traversing rough terrain, moving and interacting in human environments). [26] Due to mechanical complexity and hard controllability of the biological based locomotion systems are used rarely at indoor and roadway applications.

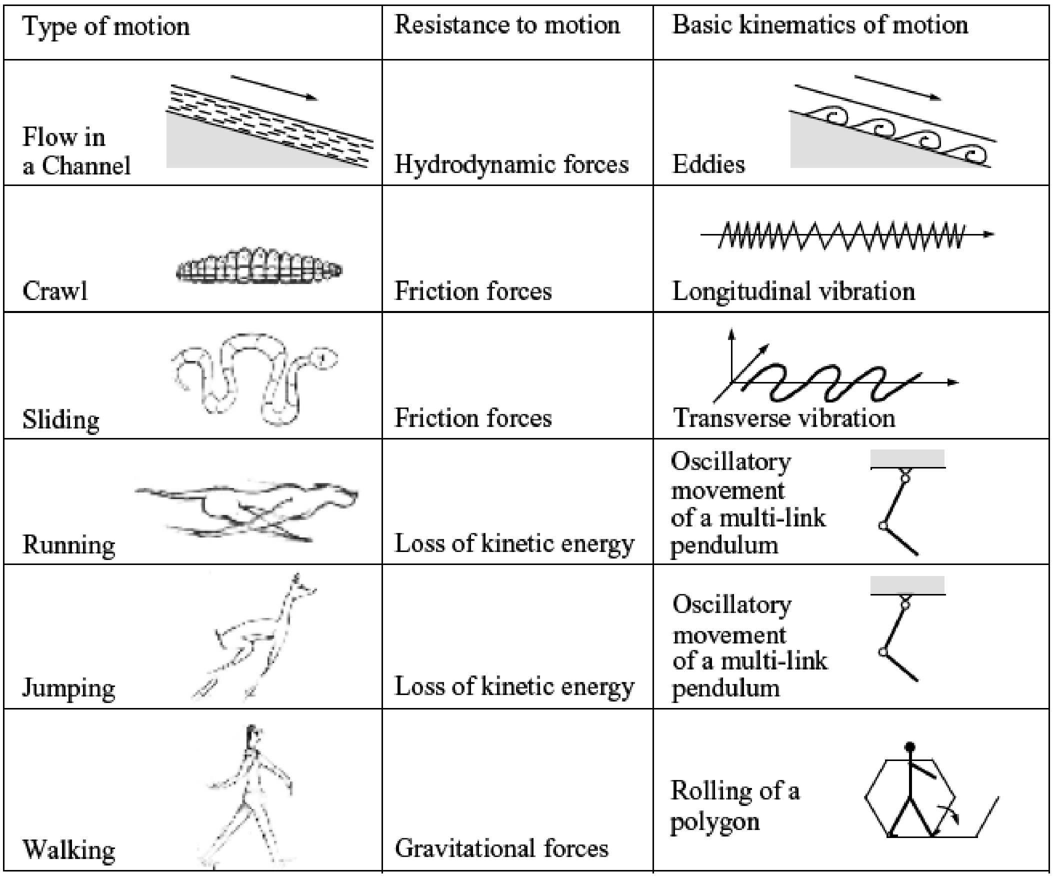


Figure 5.1 Locomotion mechanisms used in biological systems. [26]

The wheeled type locomotion systems need some steering system. There are a lot of wheeled mobile robot locomotion types. The most commons are Differential Drive, Tricycle, Synchronous Drive, Ackerman Steering and Omni Directional.

One improvement on wheels is caterpillar tracks. They are much more suited to uneven terrain and can give much better traction. This is the locomotion system commonly seen on tanks, cranes and excavators. Robotank is a type of hydraulically actuated crawling vehicle. Actuation of the robot is powered by the hydraulic motors which are coupled each track and velocities of the hydraulic motors are controlled by servo hydraulic valves. Our robot locomotion system consists of two tracks. But robot driving is same as differential driving. So the same kinematic formulations can be used.

5.2 Mobile Robot Kinematics

Kinematics is interested in behavior of mechanical systems. In mobile robotics, we need to understand the mechanical behavior of the robot both in order to design appropriate mobile robots for tasks and to understand how to create control software for mobile robot hardware. Robot kinematics is the study of the motion of robots. In a kinematic analysis the position, velocity and acceleration of the robot, are calculated without considering the forces that cause this motion.

Mobile Robot can be thought as a rigid body on wheels and operating on a horizontal plane. The total degree of freedom of a mobile robot chassis on the plane is three, two for position in the plane and one for orientation along the vertical axis, which is perpendicular to the plane.

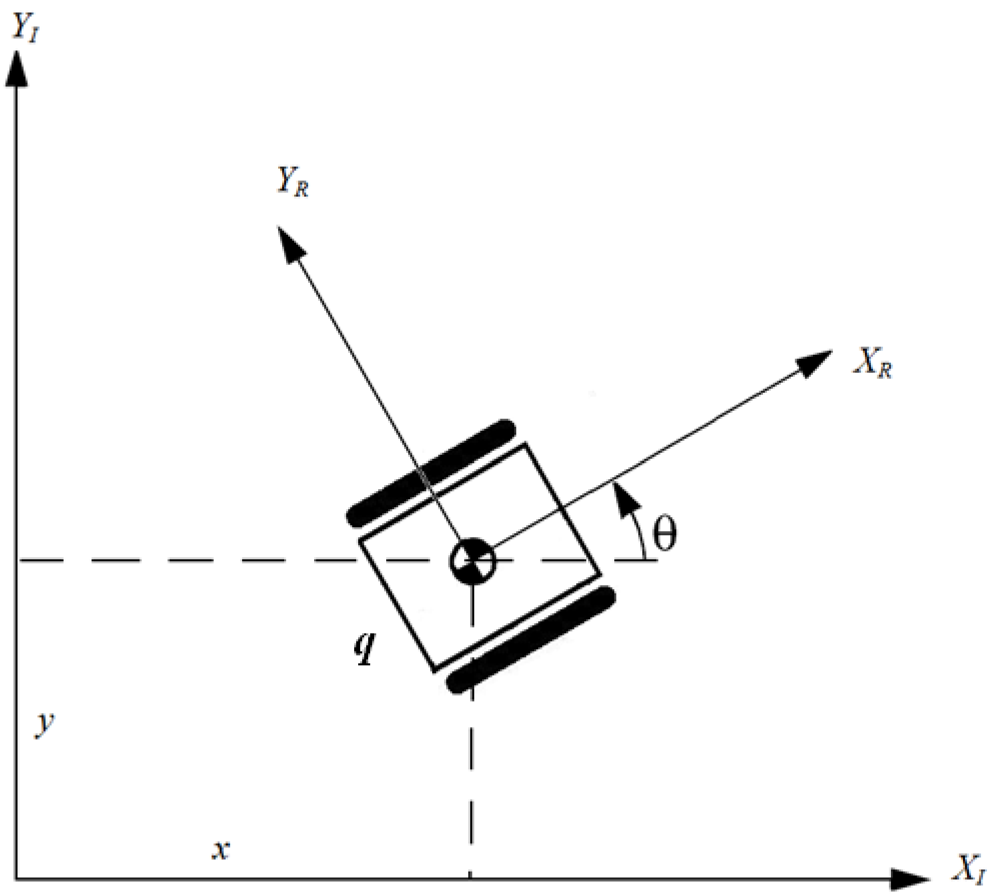


Figure 5.2 Global Reference Frame and Robot Reference Frame

In order to specify position of the robot on the plane we must establish a relationship between the global reference frame of the plane and the local reference frame of the robot, as seen in Figure 5.2. The axes X_I and Y_I define an arbitrary inertial frame on the plane as the global reference frame which has origin O [$X_{I,}$ Y_I]. To specify the position of the robot, point q is selected on the robot chassis as its position reference point. The axes X_R and Y_R define two axes relative to q on the robot chassis and are the robot's local reference frame axes. The position of q in the global reference frame is specified by coordinates x and y, and the angular difference between the global and local reference frames is given by Θ. The pose of the robot can be described easily as a vector with these three elements.

The robot position with respect to the global reference frame can be expressed as:

$$\xi_I = [x \quad y \quad \theta]^T \tag{5.1}$$

x, y: $position\ of\ the\ robot$

θ: $orientation\ of\ the\ robot$

Velocity of the robot can be expressed as:

$$\dot{\xi}_R = R(\theta)\dot{\xi}_I \tag{5.2}$$

$$\dot{\xi}_I = [\dot{x} \quad \dot{y} \quad \dot{\theta}]^T \tag{5.3}$$

$$R(\theta) = \begin{bmatrix} cos\theta & sin\theta & 0 \\ -sin\theta & cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{5.4}$$

$$\dot{\xi}_R = \begin{bmatrix} cos\theta & sin\theta & 0 \\ -sin\theta & cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \end{bmatrix} \tag{5.5}$$

There are several important assumptions to be used in the kinematic formulations which are;

- Movement on a horizontal plane
- Point contact of the tracks
- Tracks are not deformable
- Pure rolling
- No slipping, skidding or sliding
- No friction for rotation around contact point
- Steering axes orthogonal to the surface
- Wheels connected by rigid frame (chassis)

According to these assumptions, the robot kinematic equations can be found in following equations. In Figure 5.3, *ICR* refers to the instantaneous center of rotation, V_L refers to velocity of the left track, V_R refers to velocity of right track, L is the distance between right track center and left track center. R is the radius of rotation of the robot.

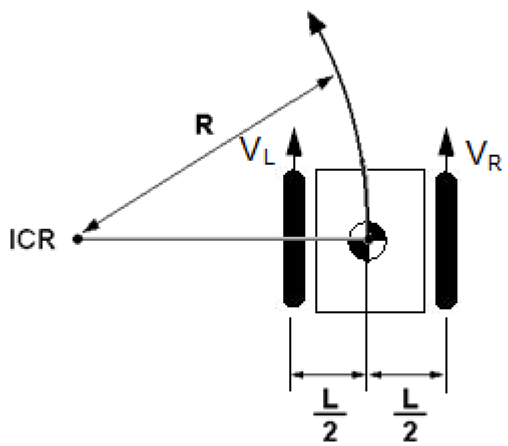


Figure 5.3 Velocity Directions of Tracks

Velocities of the tracks;

$$V_L = rw_L \tag{5.6}$$

$$V_R = rw_R \tag{5.7}$$

$$w = \frac{V_L - V_R}{2} \tag{5.8}$$

$$v = \frac{V_L + V_R}{2} \tag{5.9}$$

$$R = \frac{L}{2}\frac{V_L + V_R}{V_R - V_L} \tag{5.10}$$

For straight motion

$R = \infty \;\Rightarrow\; V_L = V_R$

For rotational motion

$R = 0 \;\Rightarrow\; V_L = -V_R$

In the previous equations r is the radius of main track gears, w_R and w_L are velocities of main track gears, w refers to rotational velocity of robot and v is the linear velocity of robot.

The rotational and translational velocity of our robots can be expressed as:

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} cos\theta & 0 \\ sin\theta & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix} \tag{5.11}$$

5.3 Mobile Robot Architecture

During the experimental studies an experimental crawling robot, which is hydraulically, has been used. Robotank is powered by four stroke air cooled single engine motor. The IC engine drives the pump shaft of hydraulic power pack unit with a time belt reduction. The specifications of IC engine and hydraulic pump are given in Table 5.1 and Table 5.2 respectively.

Table 5.1 IC Engine Specifications

Model	Robin EH25D
Displacement	251 cm^3
Power Output	6.4 HP at 3600 rpm

Table 5.2 Hydraulic Pump Specifications

Model	Hema 1P1-113 Gear Pump
Displacement	4.0 cm^3 / rpm
Max Pressure	250 bar
Max Speed	3000 rpm

Robotank has two independent hydraulic motors and each hydraulic motor is connected to tracks separately and its hydraulic circuit is given in Figure 5.4. Hence velocities of the tracks can be changed separately. Hydraulic motor can be defined as a motor which is able to convert to hydraulic power into mechanical power. Rotation speed of the hydraulic motor is proportionally dependent on flow rate of the fluid (i.e. oil). The torque which is produced by the hydraulic motor is proportionally dependent on the operating pressure. In our robot operating pressure of the hydraulic system is set manually by helps of pressure relief valve. But the flow rate, which is sent to hydraulic motor, is changed by direct operated servo solenoid valves.

Basically, servo valve is a type valve that produces an output that is proportional to the electronic control signal. For the any type of application there are many type of servo valve which are able to control direction, flow and pressure. In our robot flow rate and direction of the hydraulic motor are controlled by servo valves. The picture of valves which are used on the robot can be seen in Figure 5.5

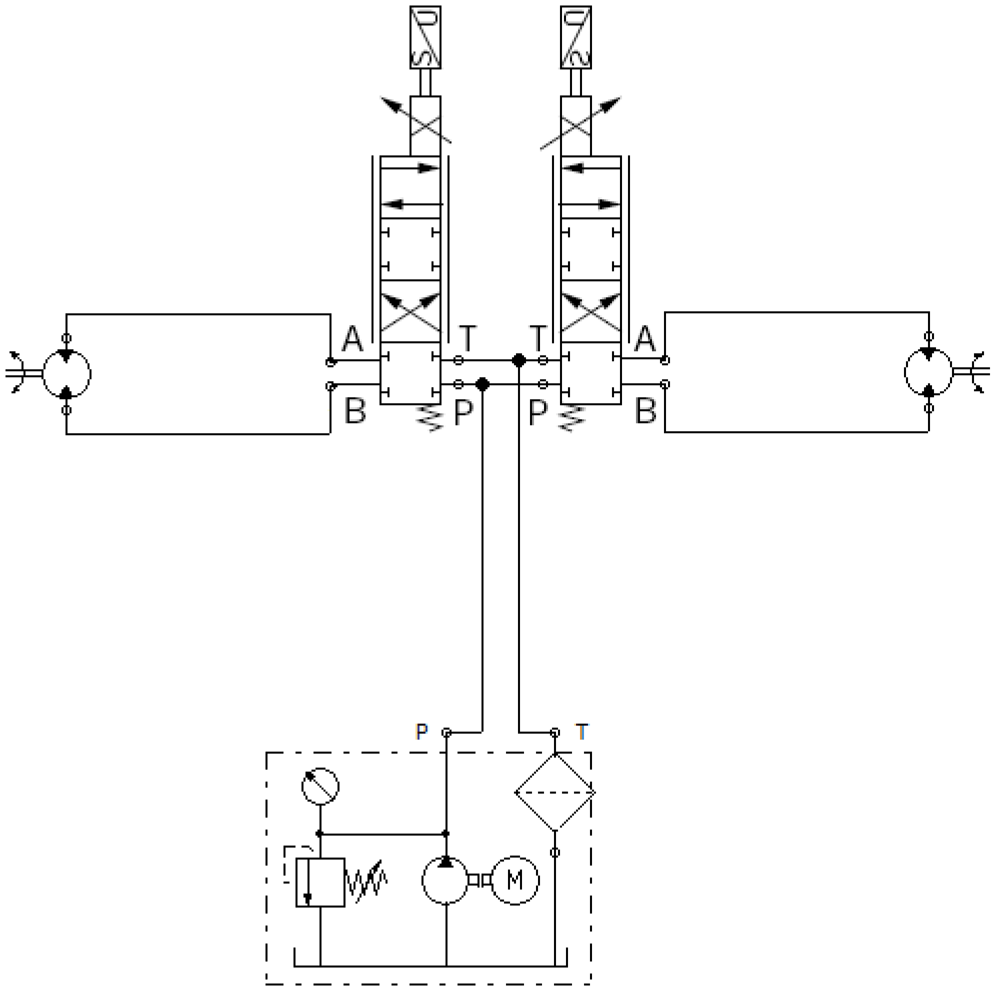


Figure 5.4 Hydraulic Circuit of Robot

Each hydraulic motor is controlled with one servo valve and flow rate passing through the hydraulic motor is proportional with the signal sent to the servo valve and it is also proportional to rotational speed of the motor. In order to operate the servo valve, the servo valve requires power amplifier. In our experimental robot, for each track 1 Bosch 0811 405 060 numbered power amplifier are used for controlling Bosch 0811 404 039 numbered Direct Operated Servo Solenoid Valve.

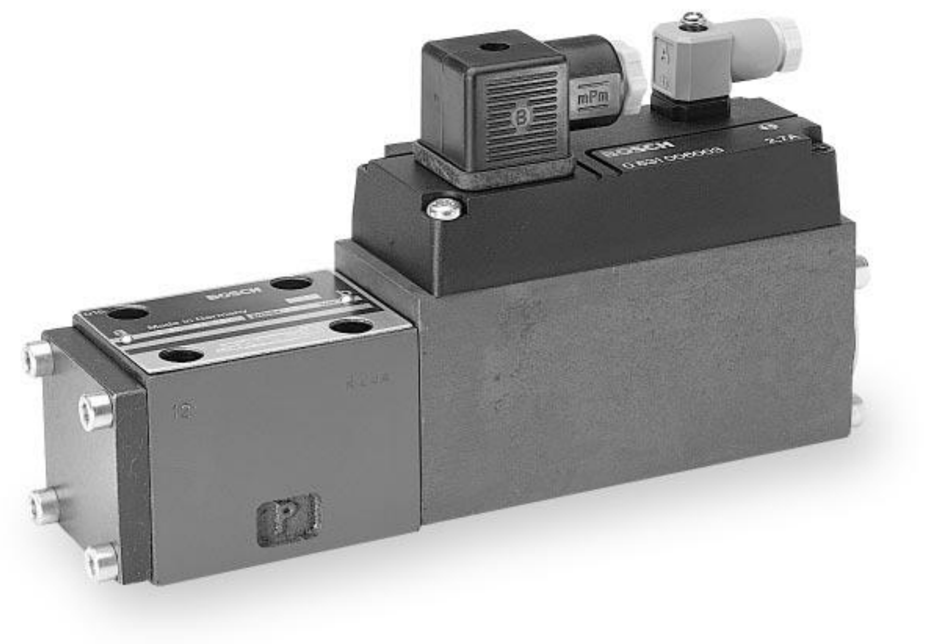


Figure 5.5 Direct Operated Servo Solenoid Valve [27]

The amplifier should be powered + 24 DCV 1.5 A power source. The amplifier needs a voltage between ±10 DCV and as a control input. For a certain pressure setting the flow rate of the hydraulic system is set by the valve spool position and eventually speeds of the hydraulic motors. So for each voltage input, spool position of amplifier changes and speed of the hydraulic motor changes. Which means that for a ± 10 DCV voltage value, valve spool is fully open and flow rate pass through the motors will be at maximum value. For 0 DCV voltage value, valve spool at neutral position and that is fully closed and no fluid pass through the hydraulic motors. By means of that property of hydraulic amplifier and valve, the flow rate of the fluid (i.e. speed of the tracks) can be controlled proportionally with input control voltage. For positive voltages, tracks will move in forward direction and for the negative voltages, tracks will move in backward directions.

The steering of the robot is possible by velocity differential of tracks, which is called as skid steering. Because of that reason for turning left or right, the control signals which are sent to amplifier must be different. For higher voltage value of the right track, the vehicle will turn left and for higher voltage value of the left track, the vehicle will turn right. Steering ratio of the robot can be changed easily by changing control signal. The value of the control signals are estimated in robot motion planning and motion control algorithms and detailed information about these algorithms can be found in Chapter 4.

To control the robot by computer, an interface is necessary between computer and hydraulic amplifiers of the robot. In our robot Data Translation DT 304 DAQ (Data Acquisition Card) is used for that purposes. DT 304 is a family of low cost multi function data acquisition board and is connected at PCI bus of computer. DT 304 has 2 analog output channels, 16 single ended 8 differential analog input channel and 23 digital input output channels. Analog channels of the board can acquire or send signal at 12 bit resolution and at 400 kS/s (i.e. kHz) sample rate. Analog output channels are able to generate voltage between -10 DCV to +10 DCV for controlling of tracks and analog input channels are able to measure between -10 DCV to +10 DCV for measuring position of tracks via servo potentiometers or for measuring velocity of tracks via tachometers. Figure 5.6 shows the general appearance of DT 304 Data acquisition card.

An interconnection between all electronic mechanic and hydraulic hardware, which are constructed on the robot, can be seen in Figure 5.8. Also general appearance of robot can be seen in Figure 5.9.

Figure 5.6 DT 304 Data Acquisition Card

The entire data acquisition cards need a screw terminal panel and connection cable to acquire or to send signals between card and equipments. In our application, Data Translation EP305 68-pin, 2 meter, shielded cable and Data Translation STP300 screw terminal panel are used and general appearance of the cable and screw terminal panel can be seen in Figure 5.7.

Figure 5.7 Data Translation STP 300 Screw Terminal Panel and EP305 Cable [28]

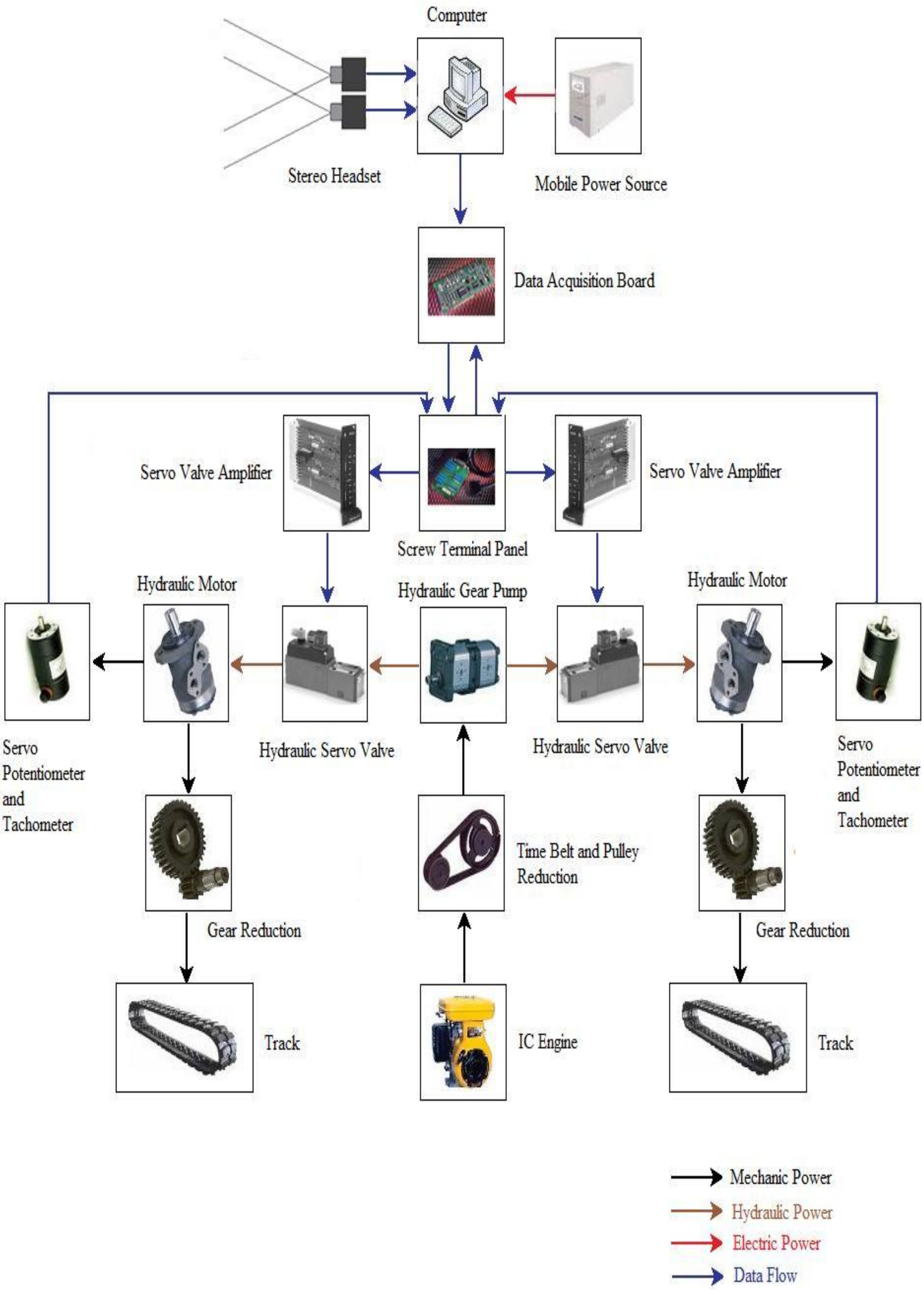


Figure 5.8 Hardware Architecture of Robot

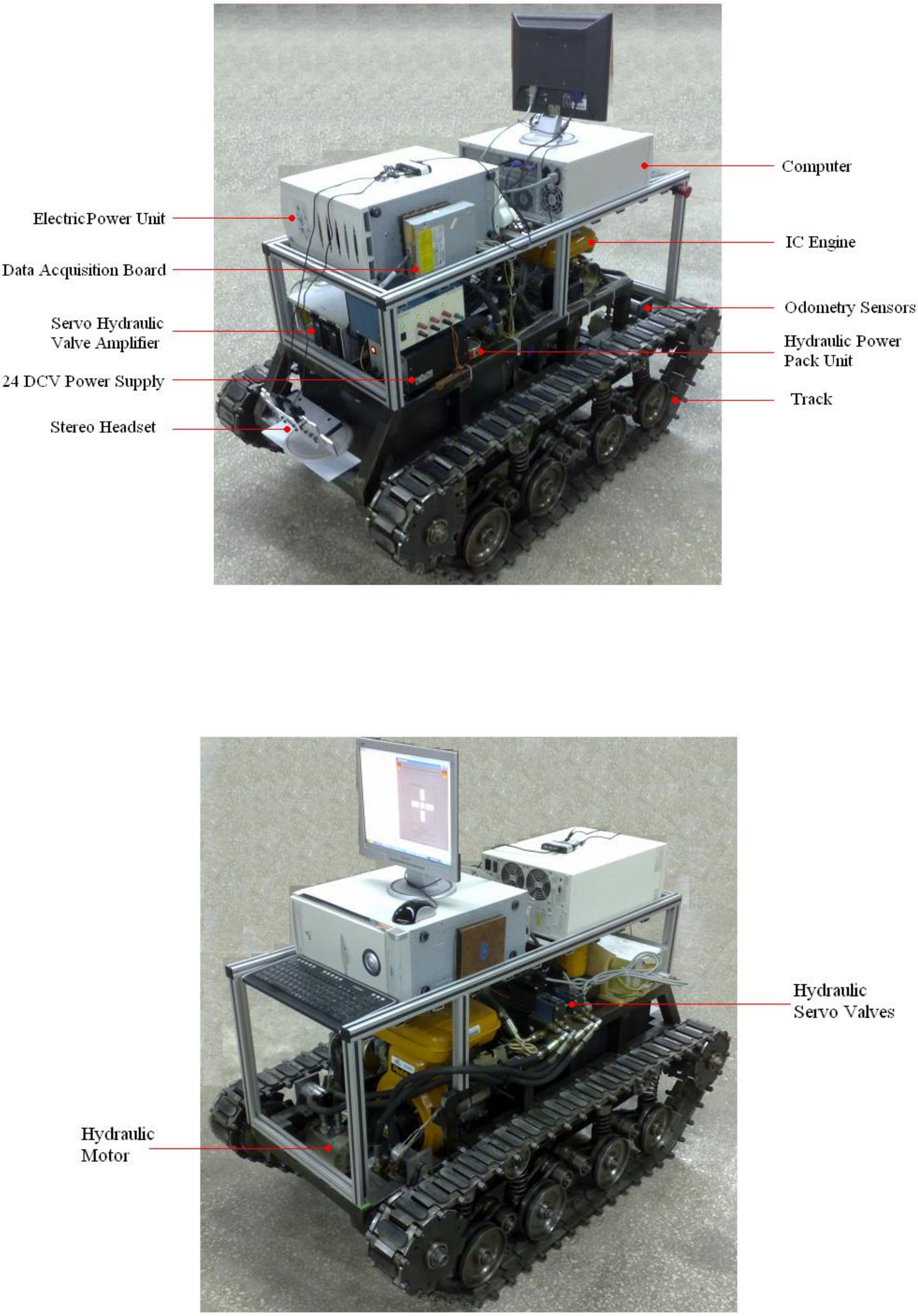


Figure 5.9 General Appearances of Robot

CHAPTER 6

CONCLUSION AND DISCUSSION

6.1 Conclusions and Discussions of Present Work

This thesis has focused on the study of developing stereo vision depth estimation and obstacle detection system from depth map for mobile robotic applications. Developed algorithm has been used for safely navigation (i.e. collision free) of real hydraulically actuated experimental crawling mobile robot in unknown indoor environment. Within the scope of this thesis four main works have been accomplished. These are;

- Developing stereo vision depth perception and obstacle detection algorithm.

- Constructing and calibrating stereo hardware

- Developing navigation, obstacle avoidance strategy, motion planning and motion control algorithms

- A crawling vehicle is turned into a robot by helps of our construction, modification and improvements.

In this study stereo vision depth perception method for mobile robot navigation is divided into three main groups which are pre processing operations, stereo matching and finally post processing operations. Interconnection between all these function are given in Figure 3.1

In proposed approach pre processing operation has five main functions which are image grabbing, color adjustment, stereo rectification, image crop and image smooth. All these pre processing functions are prepared to improve stereo match quality.

In order to accomplish safe robot navigation in unknown environment autonomous robot need to know not only depth map but also position, size and depth of the obstacle. The prepared stereo vision perception system gives high quality depth map, which is 240x320 pixels with 32 disparities, using two calibrated cameras with an algorithm Sum of Squared Differences (SSD). But unfortunately using only depth map, accomplishing of mobile robot navigation is impossible. At this point, post processing operations have been developed to extract depth, size and position information of the obstacle from disparity map. The proposed post processing operations has five main functions and it starts with median filter and it is going on sequentially with histogram of depth map, object analysis, obstacle analysis and it finishes with depth calculation.

In order to acquire sequentially image form robot surroundings, robot needs a stereo hardware system. To implement this operation a stereo hardware are designed and constructed with two CCD sensor USB webcams. Prepared stereo hardware is very cheap but it has some drawbacks in contrast to the professional stereo hardware. Firstly, during constructing stereo hardware it needs good design and good assembly work otherwise there will be some misalignment between two cameras and it will cause low stereo match quality. Secondly, our system does not need any frame grabber so it is very cheap solution. But, getting synchronize stereo image is very hard and passing time which is require for image grabbing is higher than the professional solutions. In contrast to professional stereo image grabbing hardware prepared stereo rig must be calibrated in offline by user.

In this study, a navigation task is defined which is collision free navigation from start point to any destination point which is defined by the user. To accomplish this task robot need some algorithms which are motion planning, obstacle avoidance and motion control algorithms and combination of all these algorithms is called as robot navigation algorithm. Motion planning algorithm is used for planning of robot motion, it calculates shortest distance between current position and destination position and required turning angle. By helps of obstacle avoidance algorithm, robot tries to escape obstacles according to the position and size of the obstacle. Motion control algorithm is interconnection software between robot locomotion system and navigation algorithm. In other words, motion control algorithm changes distance and angle value into track voltages and it does communication between control card and computer. In order to succeed position control of the mobile robot, kinematics of the robot is taken into account during writing codes of motion control algorithm.

To execute developed stereo perception algorithm and navigation algorithm an experimental crawling vehicle are modified and improved to build up a real autonomous robot. In order to control the robot easily and to combine all algorithms which are developed during this study, Manual and Autonomous Navigation Graphical User Interface are prepared and implemented on the robot. After optimization and combining all algorithms and codes, behavior of robot is observed, by use of developed algorithms, the robot is able to reach destination point without collision any obstacles.

6.2 Future Works

Mobile robots are able to navigate both indoor environments and outdoor environments. Indoor applications and outdoor applications need different navigation strategy. Due to the large amount of material and also rough terrain in outdoor environment, outdoor applications more difficult than indoor applications. In this study our stereo vision depth perception algorithm and navigation algorithm have been executed for indoor applications. As a further study, prepared algorithms can be tried in outdoor environments.

The use of a third camera in the stereo vision system is a very easy way to remove the false matching given by the correlation. By using the trinocular stereo (i.e. stereo vision system with three cameras) more robust depth perception can be done and three dimensional construction of scene is also possible.

Robot navigation in unknown environment is a very hard problem to solve and it completely depends on robot perception system. As a further study, different sensor systems can be added to robot to improve robot perception to accomplish extraordinary applications.

In this study internal engine powered crawling robot are used. Due to big size and some drawbacks of crawling robot, executing algorithm on the robot is very hard work in indoor environment. In my opinion small size mobile robot is more effective then big size robot especially in experimental usage and algorithm developments. Thus, a small size electrically actuated mobile robot can be constructed for future studies.

APPENDIX A

A.1 Digital Image

An image may be defined as a two dimensional function, *f(x, y)*, where *x* and *y* are spatial (plane) coordinates, and amplitude of *f* at any pair of coordinates *(x, y)* is called the intensity of the image at that point. [17] In other words, digital images are stored in digital environment as two-dimensional arrays (i.e., matrices), in which each element of the matrix corresponds to a single pixel in the displayed image. Pixel is derived from picture element and usually denotes a single dot on a computer display.

A.1.1 Pixel Coordinates

Generally, the most convenient method for expressing locations in an image is to use pixel coordinates. In this coordinate system, the image is treated as a grid of discrete elements, ordered from top to bottom and left to right, as illustrated by Figure A1.

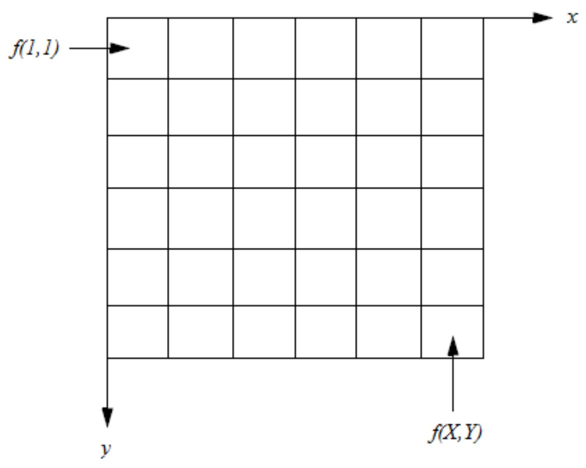


Figure A1 The Pixel Coordinate System

For pixel coordinates, the first component *x* (the row) increases downward, while the second component *y* (the column) increases to the right. Pixel coordinates are integer values and range between one and the length of the row or column.

Generally, in digital image processing four types image intensities are used. They are;

- True color images
- Indexed images
- Grayscale images
- Binary images

A.1.2 True Color Image

The precision with which a real-life image can be replicated has led to the commonly used term true color image. A true color image is an image in which each pixel is specified by three values one each for the red, green, and blue (RGB) components of the pixel's color. A RGB color space is illustrated in Figure A.2 and sample true color image can be seen in Figure A.3.

The eight corners of the cube (see Figure A.2) correspond to the three primary colors (Red, Green and Blue), the three secondary colors (Cyan, Magenta and Yellow) and black and white. All the different neutral grays are located on the diagonal of the cube that connects the black and the white vertices.

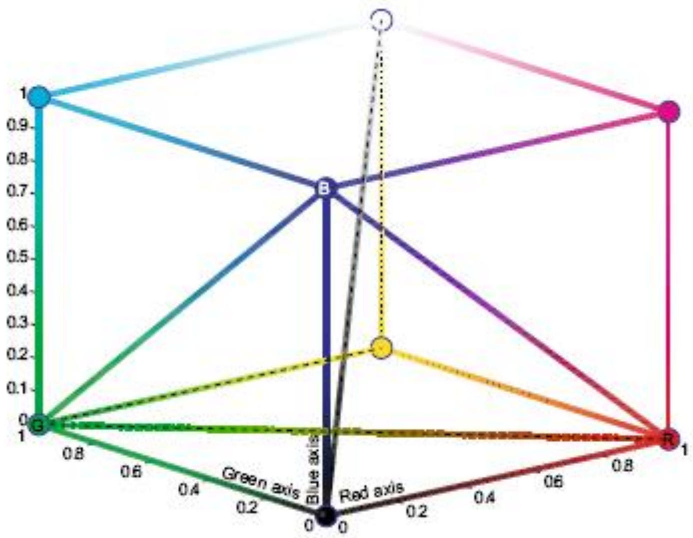


Figure A.2 RGB Color Space

Figure A.3 A Sample True Color Image

In the Figure A.3 the numbers on bottom and left side of the images shows pixel coordinates of the image both in *x* and *y* directions (see Figure A.1). The sample image has 384x512 pixels.

A.1.3 Indexed Image

Some color images are created using a limited palette of colors, typically 256 different colors. These images are referred to as indexed color images because the data for each pixel consists of a palette index indicating which of the colors in the palette applies to that pixel. [A2] A sample indexed image can be seen in Figure A.4

Figure A.4 A Sample Indexed Image

In the Figure A.4 the numbers on bottom and left side of the images shows pixel coordinates of the image both in *x* and *y* directions (see Figure A.1). The sample image has 384x512 pixels.

A.1.4 Grayscale Image

A grayscale image is made up of pixels each of which holds a single number corresponding to the gray level of the image at a particular location. These gray levels span the full range from black to white in a series of very fine steps, normally 256 different grays. A sample grayscale image can be seen in Figure A.5

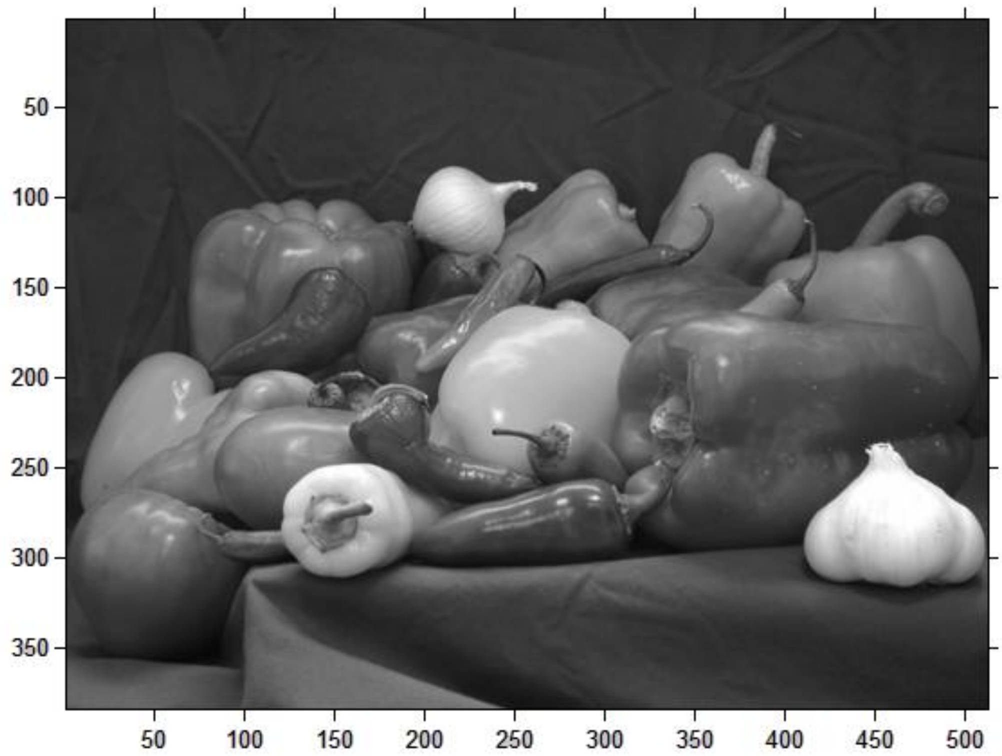


Figure A.5 A Sample Grayscale Image

In the Figure A.5 the numbers on bottom and left side of the images shows pixel coordinates of the image both in *x* and *y* directions (see Figure A.1). The sample image has 384x512 pixels.

A.1.5 Binary Image

In a the binary image, each pixel assumes one of only two discrete values 1 or 0 and these values are interpreted as white and black respectively. A sample binary image can be seen in Figure A.6

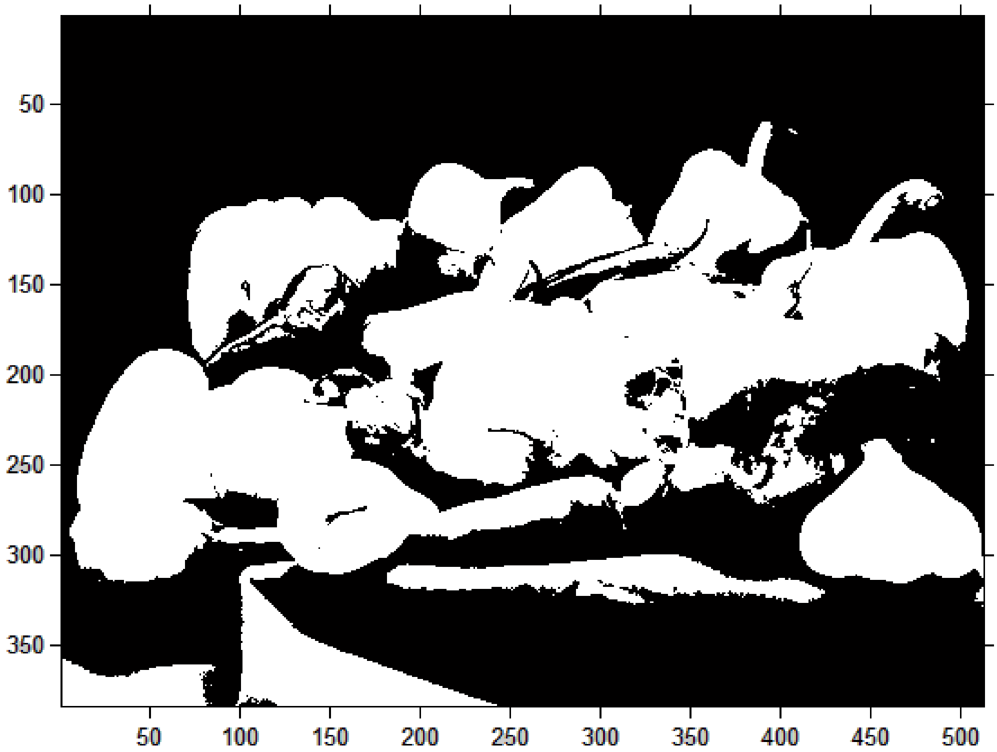


Figure A.6 A Sample Binary Image

In the Figure A.6 the numbers on bottom and left side of the images shows pixel coordinates of the image both in *x* and *y* directions (see Figure A.1). The sample image has 384x512 pixels.

REFERENCES

[1] http://en.wikipedia.org/wiki/3d_scanner

[2] http://www.pages.drexel.edu/~kws23/tutorials/sick/sick.html

[3] http://www.cs.uml.edu/~mbohan/robotics/sonar/

[4] http://www.ptgrey.com/products/stereo.asp?gclid=CMC_vpyu3pkCFYKB3god8ghoaQ

[5] Iocchi, L. and Konolige K. (1998)A Multiresolution Stereo Vision System for Mobile Robots *Proceeding of AI Workshop on New Trends in Robotics.*

[6] Calin, G. and Roda, V.O. (2007). Real Time Disparity Map Extraction in a Dual Head Stereo Vision System. *Latin American Applied Research,* **37** 21-24.

[7] Grusso, E., Sandini, G. and Tistarelli, M. (1989). 3D Object Reconstruction Using Stereo and Motion. *IEEE Transactions on Systems, Man and Cybernetics,* **19** 1465-1476.

[8] Grusso, E. and Tistarelli, M. (1995). Active/ Dynamic Stereo Vision. *IEEE Transactions. on Pattern Analysis and Machine Intelligence,* **17** 868-879.

[9] Scharstein, D. and Szeliski, R. (2004) A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms. *International Journal of Computer Vision,* **47** 7-42

[10] Murray, D. and Jennings, C.(1997) Stereo Vision Based Mapping and Navigation for Mobile Robots. *In Proceedings of the IEEE International Conference on Robotics and Automation* (ICRA '97), 1694-1699

[11] Don Murray and James J. Little.(2000) Using real-time stereo vision for mobile robot navigation. *Autonomous Robots,* **8** 161-171

[12] Kumano, M., Ohya, A. and Yuta, S.(2000) Obstacle Avoidance of Autonomous Mobile Robot using Stereo Vision Sensor, in *Proc.of 2nd International Symposium on Robotics and Automation, ISRA2000,*

[13] Chin, Y. T.,Wang, H., Tay, L.P., Wang, H. and Soh, W. Y. C.(2001) Vision Guided AGV Using Distance Transform. *Proc.of 32nd International Symposium on Robotics*

[14] Sabe, K., Fukuchi, M., Gutmann, J.-S., Ohashi, T., Kawamoto, K., and Yoshigahara, T. (2004) Obstacle Avoidance and Path Planning for Humanoid Robots Using Stereo Vision. *Int. Conf. on Robotics and Automation (ICRA)*

[15] Mingxiang, L. and Yunde, J. "Stereo Vision System on Programmable Chip (SVSoC) for Small Robot Navigation," *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '06)*, 1359–1365,

[16] Huh, K., Park, J., Hwang, J Daegun Hong, D. (2008) "A Stereo Vision-Based Obstacle Detection System in Vehicles", *Optics and Lasers in Engineering,* **46**, 168-178,

[17] Matlab Image Processing Toolbox Release Notes Version 6 (R2009a)

[18] http://en.wikivisual.com/index.php/Pinhole_camera

[19] Trucco E & Verri A; *Introductory Techniques for 3-D Computer Vision* (1998) Prentice Hall PTR

[20] Jean-Yves Bouguet. Camera Calibration Toolbox for MATLAB (http://www.vision.caltech.edu/bouguetj/calib doc/), 2007.

[21] Richard Hartley and Andrew Zisserman; *Multiple View Geometry in Computer Vision* (2003) Cambridge University Press

[22] C. Loop and Z. Zhang. *Computing rectifying homographies for stereo vision*. In Technical Report MSR-TR-99-21, 1999.

[23] Hannah M J; *Computer Matching of Areas on Stereo Images*, PhD Thesis, Stanford University (1974)

[24] http://en.wikipedia.org/wiki/Image_histogram

[25] http://en.wikipedia.org/wiki/Mobile_robot_navigation

[26] Siegwart R and Nourbakhsh I R; *Introduction to Autonomous Mobile Robots* (2004) MIT Press

[27] Bosch Core Product Catalog, ATUS AKY 018/2, January 1, 1997

[28] www.datatranslation.com/docs/specs/dt300_specs.pdf