2009

PERFORMANCE OF PAN-TILT TRACKER BASED ON THE PIN-HOLE LENS MODEL

Vikas Chandra Mehta
University of Kentucky, vikas.mehta@uky.edu

ABSTRACT OF THE THESIS

PERFORMANCE OF PAN-TILT TRACKER BASED ON THE PIN-HOLE LENS MODEL

In the modern day, recognition and tracking of face or the iris is potentially one of the most powerful ways of differentiating between an authentic person and an imposter. Our method uses stereo vision to track the 3-Dimensional coordinates of a target equivalent to a person's eyes and using a pan-tilt unit we target these areas for additional processing such as iris or facial imaging. One of the most important parts involved in tracking is the way the pan-tilt unit is calibrated. There have been techniques in the past where PTZ (Pan-tilt-zoom) digital camera has been used and calibrated using self calibration techniques involving a checker board calibration grid but the tracking error was found to be large in these techniques. We introduce a more accurate form of calibration of the pan-tilt unit using photogrammetric calibration technique and view the pan-tilt unit as an emulation of a Pinhole Lens Model to detect and track the target. The system is demonstrated on ideal targets.

KEYWORDS: Iris Tracking, Pan-tilt Control Unit, Pinhole Lens Model, Perspective Correction Coefficients, Calibration

VIKAS CHANDRA MEHTA

(04/10/09)

PERFORMANCE OF PAN-TILT TRACKER BASED ON THE PIN-HOLE LENS MODEL

By

Vikas Chandra Mehta

Dr. Lawrence G Hassebrook

Director of Thesis

Dr. YuMing Zhang

Director of Graduate Studies

(04/10/2009)

RULES FOR THE USE OF THESES

Unpublished theses submitted for the Master's degree and deposited in the University of Kentucky Library are as a rule open for inspection, but are to be used only with due regard to the rights of the authors. Bibliographical references may be noted, but quotations or summaries of parts may be published only with the usual scholarly acknowledgements.

Extensive copying or publication of the thesis in whole or in part also requires the consent of the Dean of the Graduate School of the University of Kentucky.

A library that borrows this thesis work for use by its patrons is expected to secure the signature of each user.

Name Date

_____

_____

_____

_____

_____

_____

THESIS

Vikas Chandra Mehta

The Graduate School

University of Kentucky

2009

PERFORMANCE OF PAN-TILT TRACKER BASED ON THE PIN-HOLE LENS MODEL

THESIS

A thesis submitted in the partial fulfillment of the requirements for the degree of Master of Science in Electrical Engineering in the College of Engineering at the University of Kentucky

By

Vikas Chandra Mehta

Lexington, Kentucky

Director: Dr. Lawrence G Hassebrook, Professor of Electrical Engineering

Lexington, Kentucky

2009

DEDICATION

To my Family

ACKNOWLEDGEMENTS

It is my pleasure to express my sincere gratitude to these wonderful persons who have bestowed upon me their guidance and support to the fullest and facilitated the successful completion of this thesis.

First of all, I thank Dr. Laurence Hassebrook, my advisor and thesis chair for giving me new fillip with his merited ideas and guidance throughout the completion of this work. I also thank the other members of my defense committee, Dr. Kevin Donohue and Dr. Robert Heath for their valuable comments and suggestions.

Further, I would like to express my thanks to all my friends and roommates, especially Vamshi Vankadara, Vidyadhar Rangojoo, Sandeep Marda, Vaibhav Badjatya, Chandan Gaddikopula and Krishna Suman for their support and encouragement. My special thanks to Vaibhav Badjatya who helped me in formatting my thesis document.

Last but not the least I would like to thank my parents and sister for their love, affection and moral support throughout my career till date.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF FILES

1. Thesis_Vikas.pdf…….1.67 MB

Chapter 1 Introduction

Face and iris recognition are important in many applications to provide identification of people and provide secure access to a building, room, or information. Most of the authentication systems today use passwords or identification number but these are at risk of being used by imposters. The use of biometrics such as face, iris, and finger prints or palm prints for the authentication purposes are very useful as they are unique and are attached to a person. Sometimes even the authentic person may differ because of the illumination level on the face or different facial expressions or the amount of pressure applied while pressing the finger or the palm during the scanning process. Thus perfect matching in presence of image variations of the same person is very important.

Pattern recognition and target tracking can be accomplished using correlation techniques. Advanced correlation filters have many advantages such as the shift invariance in the filtering, i.e. when the input images is translated by a certain amount, the filter output also shifts by the same amount and the shift can be estimated by the location of the peak response. The correlation filter can also be designed to achieve noise tolerance, discrimination, etc. The image matching is generally done by cross correlating an input image with a synthesized template and then analyzing the correlation output. Peaks are searched for in the correlation output and the heights of the peaks, the peak-to-side lobe ratio (PSR), etc are used to find if the input is from the authentic person or not. The location of the peaks indicates the positions of the objects.

This thesis describes the tracking of an object using two stereo cameras, a pan-tilt control unit, and a projector. A 2-plane calibration grid is used to calibrate both the cameras and the pan-tilt unit. Target detection is achieved using the correlation method. A ring filter is adopted to detect the target accurately. The least square approximation algorithm forms the basis of tracking the target. The pan-tilt unit is emulated as a pinhole lens model and based on that, the perspective coefficients of the pan-tilt unit are found. These coefficients along with the coefficients of the cameras and their coordinates helps in finding the pan-tilt coordinates thereby tracking the target.

1.1 Background

The matched spatial filters were first introduced by Vander Lugt in the 1960's for pattern recognition [1]. Due to the high signal-to-noise ratio (SNR) it can be used to detect known targets but the limitation was the degradation in the correlation response peak caused by image distortions such as rotations and scale variations. Moreover one matched filter was required for each combination of the face with different poses and expressions and this was practically not possible. Composite correlation filters were later proposed.

The Synthetic discriminant function filters (SDF) was introduced by Hester and Casasent [2]. SDF filter use a linear combination of reference images to create a composite image which is then correlated with the desired inputs. The weights for the linear combination are selected in such a way that the correlation output at the origin is the same for all images belonging to one class, i.e. if the correlation output values corresponding to the

training images of authentics are set to 1 and the output values due to the imposter training images are set to 0 then the correlation response will be 0 if the input image is of an authentic person. But the disadvantage of Equal Correlation Peak SDF filter is that it results in large side lobes and doesn't consider noise in the input image. Minimum variance SDF filters (MVSDF) were introduced to minimize output variance due to noise, and minimum average correlation energy filters (MACE) can be used to improve correlation peak sharpness [3]. MVSDF filters emphasize low frequencies, whereas MACE filters emphasize high frequencies. P. Re´fre´gier developed an optimal tradeoff between the MVSDF and MACE filters [4]. Initial SDF filters used hard constraints i.e., the correlation peak values were prespecified. Better performance may be obtained with non training images by using unconstrained correlation filters like maximum average correlation height filter [5].

In 1990 L. G. Hassebrook, B. V. K. Vijaya Kumar, and L. Hostetler introduced the Linear Phase Coefficient Composite Filter (LPCCF) [6]. It is based on the synthetic discriminant function filters. The LPCCF design uses a special form of training set matrix, i.e., a cyclic Toeplitz matrix; such that the training set selection and filter design are combined to yield a correlation matrix that is of cyclic Toeplitz form. It required certain selection criteria to yield training sets that produce the matrix that best approaches the cyclic Toeplitz form and also represents a range of possible distortions of the object. The disadvantage of the composite filter approach is that it is numerically intensive as it requires a 2D Fast Fourier Transform of the entire scene to perform the correlation.

Wei Su and L. G. Hassebrook introduced another approach called a Super Image Vector Inner Product (SIVIP) [7] [8]. It has distortion invariant properties of composite filters but doesn't require correlation. The SIVIP tracker is a combination of three different filtering techniques originating with LPCCFs. The LPCCF combines a complex phase response with a training set selection such that the phase varies with the distortion, such as rotation and if the distortion represented by the training set is translation then the LPCCF structure can be implemented without correlation. This approach led to the implementation of Vector Inner Product (VIP). The VIP uses translated combinations of the target image which when correlated with an input scene results in a pattern of multiple correlation peaks which could be arranged in an arbitrary pattern specific to the input target class, resulting in a morphological transform of the input. The disadvantage of the VIP is that it is sensitive to distortions and the problem with the phase response estimation is that it requires correlation. The SIVIP tracker solved all these problems. SIVIP uses a sequence of distortion specific super images implemented as Vector Inner Product (VIPs) organized into a decision tree spanning several distortions. A SIVIP is implemented by elementwise multiplying a super-image template by a partition of interest in the input scene and then summing the elementwise operations. The process begins with the generation of a super image set. These images are generated as a weighted sum of distorted target images. The weights are chosen to give a complex response whose phase indicates some characteristic about the target orientation and trajectory.

Most of the composite filters require a 2D FFT, and it requires $M^2 \log M$ operation for MxM images, whereas SIVIP requires only M^2 operations which is computationally more efficient than the conventional correlation methods.

1.2 Thesis Outline

This thesis consists of 6 chapters. The first chapter introduces different methods for detecting and tracking a target and describes the structure of the thesis. The second chapter describes the background on tracking, different calibration methods employed for tracking, the Pinhole Camera Model and the Least Square approximation method which deals with the mathematical part involved in tracking. Chapter three describes the calibration method used for calibrating the cameras and the pan-tilt control unit; process of correlation and detection of target. Chapter four introduces the tracking methodology employed and experiments involving the tracking of a target. Chapter five shows the experimental results and plots. Chapter six provides a concluding overview of the thesis and possible future work.

Chapter 2 Background on tracking

Object (face or iris) tracking has many vision applications such as person authentication face-based biometric, expression analysis, and human computer interface. Tracking a human face or iris is a challenging task as they are nonrigid and have a high degree of inconsistency. Tracking can be feature based or an appearance based. It can be divided into deterministic and stochastic approach. In visual tracking which are based on probabilistic analysis, a Bayesian perspective is used to estimate the current state of an object which changes over time using a sequence of observations made on the system [9]. Particle filtering methods also known as Sequential Monte Carlo (SMC) methods for stochastic approach were also introduced by several research groups [10]. These tracking methods have the advantage of not being subject to any limitation on linear systems nor require the noise to be Gaussian and proved to be more robust to distracting clutter. The Condensation algorithm also falls under the Sequential Monte Carlo methods has gained dominance in the visual tracking [11]. Stochastic tracking is more robust when compared to deterministic tracking as it has the potential for evading local minimum as the search directions are random. But the stochastic algorithms are more time-consuming than the deterministic ones.

Trackers which use classical 3-D vision techniques that provide tools for computing the pose, position and facial gestures suffer from the drifting problem since the facial features do not have enough stable local appearance due to many factors. Appearance-based technique which is a category of deterministic approaches has been widely used for

object tracking [12]. These techniques are easy to implement and are generally more robust than feature-based methods. Statistical facial textures have been adopted to overcome the problem of appearance changes. Gokturk, Bouguet and R.Grzeszczuk proposed a feature-based approach in which a two-stage method has been developed for 3-D tracking of pose and deformations in monocular image sequences [13]. The first stage of the system learns the space of all by applying Principle Component Analysis on real stereo tracking data. The second stage simultaneously tracks the pose and deformation of the face in the monocular image sequence using an optical flow formulation associated with the tracked features. Cascia, Sclaroff and Athitsos proposed an appearance-based approach in which the head is modeled as a texture mapped cylinder and the 3-D head pose is derived by registering the current texture map with a linear combination of texture warping templates and orthogonal illumination templates [14]. Fast and stable tracking was achieved via regularized, weighted least-squares minimization of the registration error. Jorgen Ahlberg proposed an appearance-based approach framework for tracking the face and facial features using active appearance models [15]. The disadvantage of statistical appearance-based tracking methods is that, they depend on the imaging conditions under which the learning is performed and thus, by changing these conditions, one has to repeat the whole learning process, which can be very tiresome. Online appearance models (OAM) for 2-D tracking have also been adopted by researchers. S. Zhou, R. Chellappa, and B. Mogghaddam developed a method adopting deterministic and stochastic principles to track the 2-D motion of faces using an affine transform [16]. The deterministic part was used for guiding random samples. The developed approach adopts adaptive observation models since the object appearance is

learned during the tracking phase. Thus, OAMs offers more flexibility compared to tracking approaches using statistical texture modeling.

2.1 Calibration Theory

Camera Calibration in 3D computer vision refers to determining the metric information (intrinsic and extrinsic parameters) from the 2D image. The intrinsic parameter includes the camera geometry and the optical characteristics such as the *x*-coordinate of the center of projection, the y-coordinate of the center of projection, the focal length, the aspect ratio and the angle between the optical axes. The extrinsic parameter includes the orientation of the camera with respect to some world coordinate system and also the 3D position of the camera. Basically, intrinsic parameters determine how light is projected through the lens onto the image plane of the sensor. There are different ways of calibration and they are based on perspective or projective camera models. The two basic techniques among them are the photogrammetric calibration and self calibration. In photogrammetric calibration technique, a target calibration object whose 3D space geometry is well known is observed. The target calibration object generally consists of 2 or 3 planes which are orthogonal to each other. Self calibration methods do not use any calibration frame, grid or object [17] [18]. This technique is used when no Euclidean information is present and the intrinsic parameters are varying. It uses only the information available in the images taken by the camera to find the intrinsic parameters. The problem with pure self calibration is the need for known units. Thus, some form of known distance must be introduced. A common method for doing this is a hybrid of a

fixed photogrammetric grid and no grid at all. This method uses a checkerboard pattern of known dimension which is randomly positioned in view of the camera and images are captured. From these images, the extrinsic parameters are estimated, and therefore the perspective coefficients are estimated.

Camera calibration can be divided into the following types [19] [20]:

1) Linear versus non-linear camera calibration:

   They are differentiated depending on the modeling of lens distortion. Linear techniques are fast but cannot handle lens distortion [21].

2) Intrinsic versus extrinsic camera calibration:

   Intrinsic calibration is used only to obtain the physical and optical parameters of the camera [22] [23], whereas extrinsic calibration deals with the measurement of the position and orientation of the camera in the scene [24] [25].

3) Implicit versus explicit calibration:

   Implicit calibration is the process of calibrating a camera without explicitly computing its physical parameters [26] [27]. The results are generally used for 3D measurement and the generation of image coordinates, but they cannot be used for camera modeling as the obtained parameters do not correspond to the physical parameters [28].

4) Methods using 3D rather than planar point arrays [29] [30].

5) Point based versus line based [31]:

   Point based methods are generally used in photogrammetry, whereas the line based method is used in plumbline calibration.

In a more specific classification, a combination of linear and non-linear techniques is also employed. A linear method is used to compute initial approximations for some of the parameters and the other parameters are computed iteratively [32] [33].

2.2 Pinhole Lens Model

A pinhole lens model is used for mapping the three dimensional world coordinate onto the two dimensional image plane. This type of mapping is called perspective projection. It has a small hole through which light enters before forming an inverted image on the image plane facing the hole. A non inverting model is commonly used where the image plane is placed in between the focal point of the camera and the object.

Figure 2.1 shows the perspective projection in pinhole lens model.

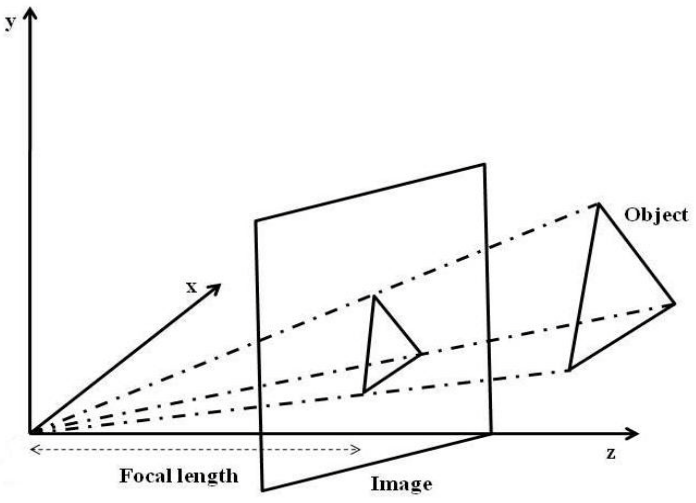

Figure 2.1: Perspective projection in pinhole camera model

In [33], Tsai describes the steps involved in the transformation of the 3D world coordinates into camera coordinates. If (x_w, y_w, z_w) are the object world coordinate system and (x, y, z) are the camera 3D coordinate system then the transformation from the object world coordinate system to camera 3D coordinate system is given by,

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = R \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} + T \,, \tag{2.1}$$

where R is the 3x3 rotation matrix

$$R = \begin{bmatrix} r_1 & r_2 & r_3 \\ r_4 & r_5 & r_6 \\ r_7 & r_8 & r_9 \end{bmatrix} \tag{2.2}$$

and T is the translation vector

$$T = \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix} \tag{2.3}$$

The 3D camera coordinates (x, y, z) are transformed into ideal image coordinates (x_u, y_u) using perspective projection with pinhole camera geometry.

$$X_u = f \frac{x}{z}, \tag{2.4}$$

$$Y_u = f \frac{y}{z}, \tag{2.5}$$

where f is the focal length

Distortion due to the radial lens is

$$X_d + D_x = X_u \,, \tag{2.6}$$

$$Y_d + D_y = Y_u \,, \tag{2.7}$$

where (X_d, Y_d) is the distorted or true image coordinate on the image plane and

$$D_x = X_d(k_1 r^2 + k_1 r^4 + ...) \,, \tag{2.8}$$

$$D_y = Y_d(k_1 r^2 + k_1 r^4 + ...) \,, \tag{2.9}$$

where k_i is the distortion coefficient and

$$r = \sqrt{X_d^2 + Y_d^2} \tag{2.10}$$

The transformation from the world coordinates to the camera coordinates can be expressed by homogenous equations such that

$$\begin{bmatrix} s.x^c \\ s.y^c \\ s \end{bmatrix} = \begin{bmatrix} s_x & 0 & x^{c0} \\ 0 & s_y & y^{c0} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r11 & r12 & r13 & t1 \\ r21 & r22 & r23 & t2 \\ r31 & r32 & r33 & t3 \end{bmatrix} \begin{bmatrix} X^w \\ Y^w \\ Z^w \\ 1 \end{bmatrix} \,, \tag{2.11}$$

where s is the scale factor.

Equation (2.11) can also be written as

$$\begin{bmatrix} s.x^c \\ s.y^c \\ s \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix} \begin{bmatrix} X^w \\ Y^w \\ Z^w \\ 1 \end{bmatrix} \,, \tag{2.12}$$

Therefore, the procedure to obtain the intrinsic and extrinsic parameters is equivalent to calculating the parameters $[m_{11}, m_{12},\ldots, m_{21}, m_{22},\ldots, m_{31},\ldots, m_{34}]^T$.

The affine transformation is included in the pinhole lens model.

In Chapter 3, the pan-tilt unit is viewed as an emulation of a pinhole camera model and based on that the perspective coefficients of the pan-tilt unit are calculated during calibration and are used for the tracking purpose.

2.3 Linear Relationship for Determining the Perspective Correction coefficients

The method of least square approximations is used for reconstructing the world coordinates from the triangulated data. The following equations give the mathematical relationships between the physical world coordinates of any point and the corresponding coordinate of that point as viewed by the camera [26].

Let $(x_{w,n}, y_{w,n}, z_{w,n})$ be the world coordinates of any one of the N grid points and $(x_{c,n}, y_{c,n})$ be the corresponding camera coordinate of that point as observed in the camera space. The Z_w dimension is assumed to be approximately parallel to the camera. The world to camera coordinate transforms for a pinhole lens are given by,

$$x_{c,n} = \frac{m_1 x_{w,n} + m_2 y_{w,n} + m_3 z_{w,n} + m_4}{m_9 x_{w,n} + m_{10} y_{w,n} + m_{11} z_{w,n} + m_{12}}, \tag{2.11}$$

and

$$y_{c,n} = \frac{m_5 x_{w,n} + m_6 y_{w,n} + m_7 z_{w,n} + m_8}{m_9 x_{w,n} + m_{10} y_{w,n} + m_{11} z_{w,n} + m_{12}} \quad , \tag{2.12}$$

where n=1, 2,… N known points.

The coefficient m_{12} is made equal to unity so that the transformation is linear at the world origin. A coefficient vector, not including the m_{12} coefficient is defined as 11x1 column vector such that

$$\underline{m_0} = [m_1 \; m_2 \; m_3 \; m_4 \; m_5 \; m_6 \; m_7 \; m_8 \; m_9 \; m_{10} \; m_{11}]^T, \tag{2.13}$$

Given a linear equation as

$$\underline{p}_0 = \underline{A}_0 \, \underline{m}_0 , \tag{2.14}$$

where $\underline{p}_0 = \begin{bmatrix} \underline{p}_{0\,zxc} \\ \underline{p}_{0\,zyc} \end{bmatrix}$, (2.15)

is a 2*N*x1 vector constructed from the augmentation of two *N*x1 vectors given by

$$\underline{p}_{0\,zxc} = \begin{bmatrix} m_{12} \; x_{c,n} & \ldots & m_{12} \; x_{c,N} \end{bmatrix}^T , \tag{2.16}$$

and

$$\underline{p}_{0\,zyc} = \begin{bmatrix} m_{12} \; y_{c,n} & \ldots & m_{12} \; y_{c,N} \end{bmatrix}^T , \tag{2.17}$$

The factor matrix $\underline{A}$ is 2*N*x11 and constructed by the augmentation of two matrices as

$$\underline{A}_0 = \begin{bmatrix} \underline{A}_{0\,xc} \\ \underline{A}_{0\,yc} \end{bmatrix}, \tag{2.18}$$

where,

$$\underline{A}_{0\,xc} = \begin{bmatrix} x_{w,1} & y_{w,1} & z_{w,1} & 1 & 0 & 0 & 0 & 0 & -x_{c,1}\,x_{w,1} & -x_{c,1}\,y_{w,1} & -x_{c,1}\,z_{w,1} \\ . & . & . & . & . & . & . & . & . & . & . \\ x_{w,N} & y_{w,N} & z_{w,N} & 1 & 0 & 0 & 0 & 0 & -x_{c,N}\,x_{w,N} & -x_{c,N}\,y_{w,N} & -x_{c,N}\,z_{w,N} \end{bmatrix}, \tag{2.19}$$

and

$$\underline{A}_{0\,yc} = \begin{bmatrix} 0 & 0 & 0 & 0 & x_{w,1} & y_{w,1} & z_{w,1} & 1 & -y_{c,1}\,x_{w,1} & -y_{c,1}\,y_{w,1} & -y_{c,1}\,z_{w,1} \\ . & . & . & . & . & . & . & . & . & . & . \\ 0 & 0 & 0 & 0 & x_{w,N} & y_{w,N} & z_{w,N} & 1 & -y_{c,N}\,x_{w,N} & -y_{c,N}\,y_{w,N} & -y_{c,N}\,z_{w,N} \end{bmatrix}, \tag{2.20}$$

The Least Squares Solution for the coefficient vector is

$$\underline{m}_0 = \left(A_0{}^T \; A_0 \right)^{-1} A_0{}^T \; \underline{p}_0 , \tag{2.21}$$

Chapter 3 Calibration Process

The following flowchart shows the system model and the general procedure to track an object.

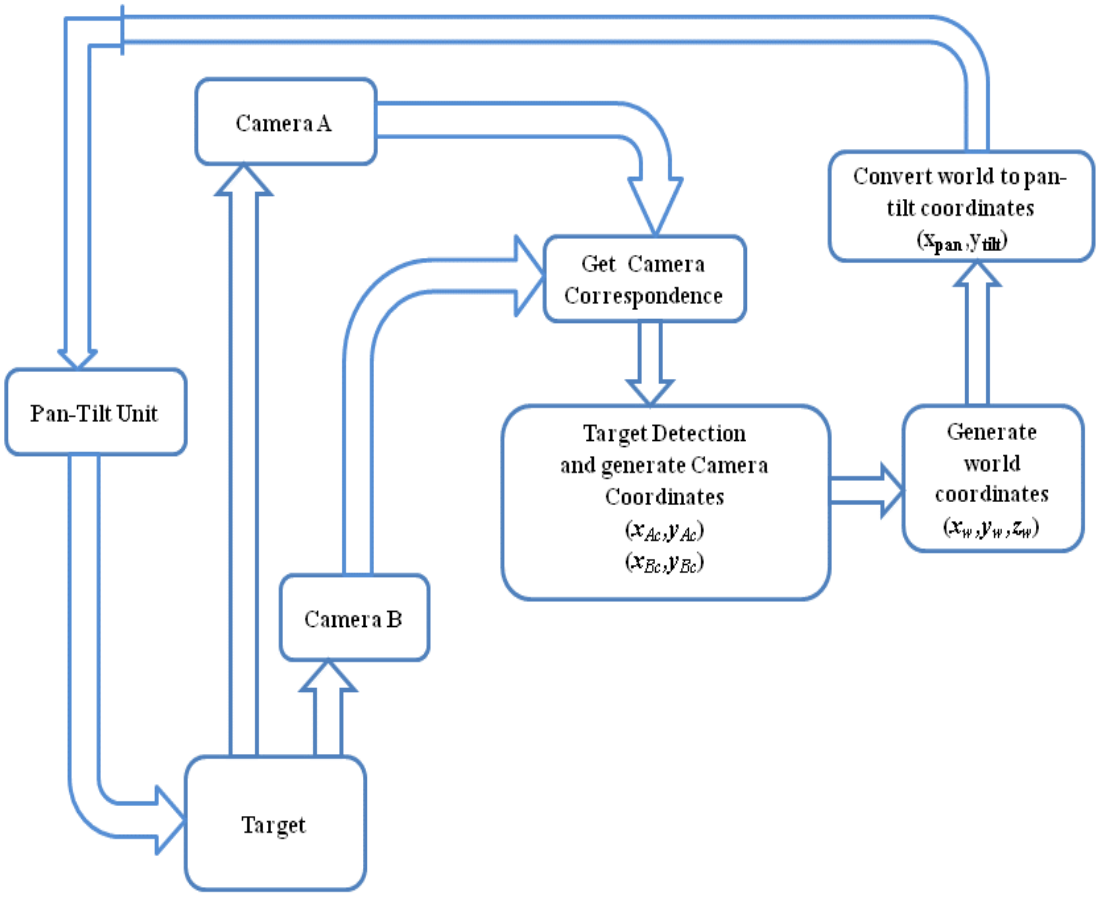

Figure 3.1: Flow Chart of the tracking system model.

After calibrating (explained in the next section) the cameras and the pan-tilt unit, the camera coordinates ($x_{Ac}, y_{Ac}, x_{Bc}, y_{Bc}$) of Camera A and Camera B are known. Using

these camera coordinates we find the world coordinates (x_w, y_w, z_w) of the object. These world coordinates are converted to pan-tilt angles and fed to the pan-tilt unit which calculates the pan and tilt required to track the target and pans and tilts in the direction of the object. The projector points a red spot on the object. This is the general procedure which we use for tracking an object (face and iris).

Figure 3.3 shows the apparatus setup which is equivalent to Figure 3.1

The following equations are used to find the world coordinates of the target.

Let $$D_n = \begin{bmatrix} x_{Ac,n} m_{A12} - m_{A4} \\ y_{Ac,n} m_{A12} - m_{A4} \\ x_{Bc,n} m_{B12} - m_{B4} \\ x_{Bc,n} m_{B12} - m_{B4} \end{bmatrix}, \tag{3.1}$$

$$C_n = \begin{bmatrix} m_{A1} - m_{A9} x_{Ac,n} & m_{A2} - m_{A10} x_{Ac,n} & m_{A3} - m_{A11} x_{Ac,n} \\ m_{A5} - m_{A9} y_{Ac,n} & m_{A6} - m_{A10} y_{Ac,n} & m_{A7} - m_{A11} y_{Ac,n} \\ m_{B1} - m_{B9} x_{Bc,n} & m_{B2} - m_{B10} x_{Bc,n} & m_{B3} - m_{B11} x_{Bc,n} \\ m_{B5} - m_{B9} x_{Bc,n} & m_{B6} - m_{B10} x_{Bc,n} & m_{B7} - m_{B11} x_{Bc,n} \end{bmatrix}, \tag{3.2}$$

where ($x_{Ac,n}, y_{Ac,n}, x_{Bc,n}, y_{Bc,n}$) are the camera coordinates of Camera A and B,

$(m_{A1}, m_{A2}, m_{A3}, m_{A4}, m_{A5}, m_{A6}, m_{A7})$ and $(m_{B1}, m_{B2}, m_{B3}, m_{B4}, m_{B5}, m_{B6}, m_{B7})$ are the coefficients of camera A and B which are explained later in this chapter.

We have,

$$D_n = C_n P_n \tag{3.3}$$

where P_n is the world coordinate

Multiplying both sides by C_n^T ,

$$C_n^T D_n = (C_n^T C_n) P_n^W , \tag{3.4}$$

$$P_n^W = R_{cc}^{-1} C_n^T D_n , \tag{3.5}$$

where,

$$R_{cc} = C_n^T C_n \tag{3.6}$$

Once the world coordinates of the target are found, they are used to find the pan-tilt coordinates. Equation 3.7 and 3.8 shows the formula for calculating the pan-tilt coordinates using the world coordinates and the '*m*' coefficients of the pan-tilt unit.

$$x_{p,n} = \frac{m_1 x_{w,n} + m_2 y_{w,n} + m_3 z_{w,n} + m_4}{m_9 x_{w,n} + m_{10} y_{w,n} + m_{11} z_{w,n} + m_{12}} \tag{3.7}$$

and

$$y_{t,n} = \frac{m_5 x_{w,n} + m_6 y_{w,n} + m_7 z_{w,n} + m_8}{m_9 x_{w,n} + m_{10} y_{w,n} + m_{11} z_{w,n} + m_{12}} \tag{3.8}$$

Equations (3.7) and (3.8) are nothing but the world to camera coordinate transforms for a pinhole lens. Here the pan-tilt unit is emulated as a pinhole camera model.

3.1 Apparatus

The technique used here is the photogrammetric calibration in which a 2-plane target calibration grid is used as shown in Figure 3.2. Both the planes are orthogonal to each other and consist of 9 rings each making to a total of 18 rings on the grid. A ring is equidistant to the immediate neighborhood ring. Two stereo cameras (scorpion), a projector, pan/tilt unit and a calibration grid are used for the calibration process. Figure 3.3 shows the apparatus used for the calibration process.

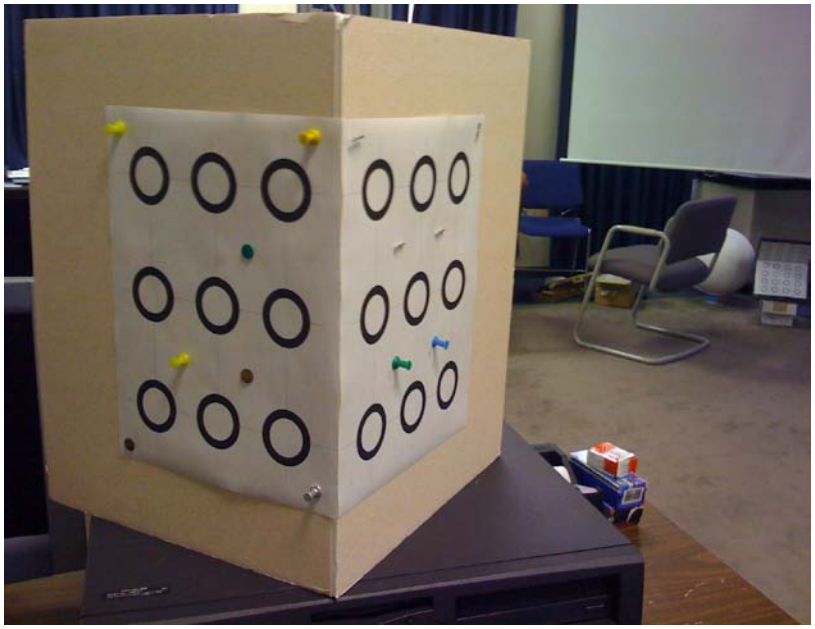

Figure 3.2: Calibration Grid for calibrating the cameras and the pan-tilt unit

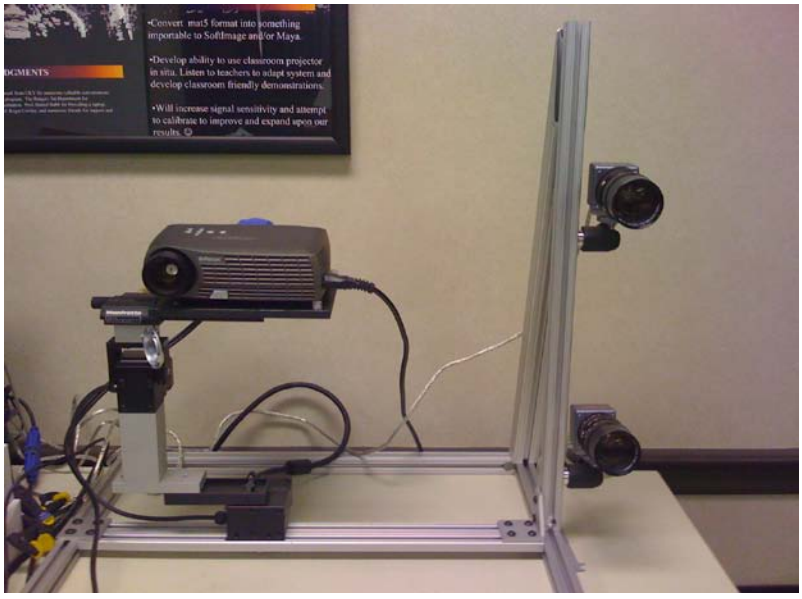

Figure 3.3: Apparatus Setup.

3.2 Calibration Method

The pan and tilt unit is programmed to pan left/right and tilt up/down. The projector is placed on the pan-tilt control unit in such a way that when the pan-tilt unit moves, the single projector point of projection also changes. This helps in projecting a spot at the required area. A Laser beam can also be used instead of a projector here but was less convenient. A white spot is projected with the help of the projector on the center of one of the rings of the calibration grid which is placed at a certain distance from the two stereo cameras. At this point the pan-tilt reading (the number of steps it moved from the zero position to the right or left and the number of steps it moved from top to bottom from the zero position) and the world coordinates of that particular ring are noted. Next the white spot is projected on to the center of the ring adjacent to the first one and the pan-tilt and world coordinates of that ring are noted. Similarly, the world coordinates and the pan-tilt reading of all the 18 rings in the calibration grid are noted. The pan and tilt readings of the calibration grid rings are represented as $x_{p,n}$ and $y_{t,n}$ respectively.

The pan-tilt spot projection can be viewed as a pinhole lens model as explained in section 2.2. We apply the pinhole lens model to the pan-tilt unit spot projection. The lens of the projector forms the center of projection. Therefore we are able to use the least square triangulation algorithm to calculate the *m* coefficients of the pan-tilt system.

Calculating *m* coefficients of the pan-tilt system:

m_p is a 11x1 column vector as shown below,

$$m_p = \begin{bmatrix} m_1 & m_2 & m_3 & m_4 & m_5 & m_6 & m_7 & m_8 & m_9 & m_{10} & m_{11} \end{bmatrix}^T, \tag{3.9}$$

The formula for calculating *m* is given by,

$$m_p = \left(A_p{}^T A_p\right)^{-1} A_p{}^T p_p \ , \tag{3.10}$$

where $A_p = \begin{bmatrix} A_{xp} \\ A_{yp} \end{bmatrix}$, (3.11)

and

$$A_{xp} = \begin{bmatrix} x_{w,1} & y_{w,1} & z_{w,1} & 1 & 0 & 0 & 0 & 0 & -x_{p,1}x_{w,1} & -x_{p,1}y_{w,1} & -x_{p,1}z_{w,1} \\ . & . & . & . & . & . & . & . & . & . & . \\ x_{w,N} & y_{w,N} & z_{w,N} & 1 & 0 & 0 & 0 & 0 & -x_{p,N}x_{w,N} & -x_{p,N}y_{w,N} & -x_{p,N}z_{w,N} \end{bmatrix}, \tag{3.12}$$

$$A_{yp} = \begin{bmatrix} 0 & 0 & 0 & 0 & x_{w,1} & y_{w,1} & z_{w,1} & 1 & -y_{t,1}x_{w,1} & -y_{t,1}y_{w,1} & -y_{t,1}z_{w,1} \\ . & . & . & . & . & . & . & . & . & . & . \\ 0 & 0 & 0 & 0 & x_{w,N} & y_{w,N} & z_{w,N} & 1 & -y_{t,N}x_{w,N} & -y_{t,N}y_{w,N} & -y_{t,N}z_{w,N} \end{bmatrix}, \tag{3.13}$$

and

$$p_p = A_p m_p, \tag{3.14}$$

where, $p_p = \begin{bmatrix} p_{zxp} \\ p_{zyt} \end{bmatrix}$, (3.15)

and

$$p_{zxp} = \begin{bmatrix} m_{12}x_{p,n} & \ldots & m_{12}x_{p,N} \end{bmatrix}^T, \tag{3.16}$$

$$p_{zyt} = \begin{bmatrix} m_{12}y_{t,n} & \ldots & m_{12}y_{t,N} \end{bmatrix}^T, \tag{3.17}$$

Table 3.1 shows the world coordinates $(x_{w,n}, y_{w,n}, z_{w,n})$ and the pan-tilt unit coordinates $(x_{p,n}, y_{t,n})$.

Table 3.1: World Coordinates and pan-tilt unit Coordinates.

Points	$x_{w,n}$	$y_{w,n}$	$z_{w,n}$	$x_{p,n}$	$y_{t,n}$
0	-31.82	0	31.82	1627	-1081
1	-76.368	0	76.368	1705	-1062
2	-120.92	0	120.92	1774	-1044
3	-31.82	90	31.82	1626	-831
4	-76.368	90	76.368	1701	-817
5	-120.92	90	120.92	1774	-812
6	-31.82	180	31.82	1624	-572
7	-76.368	180	76.368	1702	-574
8	-120.92	180	120.92	1770	-574
9	31.82	0	31.82	1445	-1089
10	76.368	0	76.368	1285	-1079
11	120.92	0	120.92	1127	-1068
12	31.82	90	31.82	1445	-833
13	76.368	90	76.368	1282	-828
14	120.92	90	120.92	1125	-823
15	31.82	180	31.82	1445	-573
16	76.368	180	76.368	1280	-573
17	120.92	180	120.92	1126	-573

Using the Table 3.1 and replacing $(x_{c,n}, y_{c,n})$ with $(x_{p,n}, y_{t,n})$ the *m* coefficients of pan-tilt unit are found to be,

m_0= [-3.255552 -0.043276 0.185681 1567.852407 0.139881 2.927261 -0.440951

-1097.618751 -0.000238 -0.000018 0.000742]

Similarly, using the equations 3.9 to 3.17, the *m* coefficients of the stereo cameras A and B are found.

The Table 3.2 and Table 3.3 show the world coordinates and the camera coordinates of camera A and camera B.

Table 3.2 World Coordinates and Camera Coordinates of Camera A.

Points	$x_{w,n}$	$y_{w,n}$	$z_{w,n}$	$x_{Ac,n}$	$y_{Ac,n}$
0	-31.82	0	31.82	546.27	651.18
1	-76.368	0	76.368	468.06	636.74
2	-120.92	0	120.92	396.91	623.4
3	-31.82	90	31.82	543.92	492.24
4	-76.368	90	76.368	466.77	483.15
5	-120.92	90	120.92	393.88	474.41
6	-31.82	180	31.82	541.71	330.6
7	-76.368	180	76.368	463.51	327.31
8	-120.92	180	120.92	392.55	323.92
9	31.82	0	31.82	664.23	649.05

Table 3.2, continued

Points	$x_{w,n}$	$y_{w,n}$	$z_{w,n}$	$x_{Ac,n}$	$y_{Ac,n}$
10	76.368	0	76.368	741.12	633.57
11	120.92	0	120.92	811.94	618.84
12	31.82	90	31.82	663.12	490.62
13	76.368	90	76.368	739.98	479.6
14	120.92	90	120.92	813.04	469.48
15	31.82	180	31.82	661.21	329.94
16	76.368	180	76.368	740.22	323.81
17	120.92	180	120.92	812.17	317.6

Table 3.3 World Coordinates and Camera Coordinates of Camera B.

Points	$x_{w,n}$	$y_{w,n}$	$z_{w,n}$	$x_{Bc,n}$	$y_{Bc,n}$
0	-31.82	0	31.82	490.3	769.17
1	-76.368	0	76.368	412.18	771.36
2	-120.92	0	120.92	341	773.88
3	-31.82	90	31.82	485.87	608.44
4	-76.368	90	76.368	410	616.26
5	-120.92	90	120.92	338.18	623.43

Table 3.3, continued

Points	$x_{w,n}$	$y_{w,n}$	$z_{w,n}$	$x_{Bc,n}$	$y_{Bc,n}$
6	-31.82	180	31.82	482.02	449.37
7	-76.368	180	76.368	406.07	462.55
8	-120.92	180	120.92	337.16	475.03
9	120.92	0	120.92	608.73	764.61
10	31.82	0	31.82	686.45	762.82
11	76.368	0	76.368	758.02	761.58
12	120.92	90	120.92	604.1	604.43
13	31.82	90	31.82	680.9	607.67
14	76.368	90	76.368	753.81	610.85
15	120.92	180	120.92	598.76	446.64
16	120.92	180	120.92	676.65	454.3
17	31.82	180	31.82	747.91	461.54

Using the Table 3.2 and Table 3.3, the *m* coefficients of the cameras are found to be.

m_A = [1.882042 -0.061885 0.492436 604.987292 -0.031149 -1.853129 0.166115 662.390418 -0.000016 -0.000075 0.000823].

m_B = [1.896437 -0.002000 0.462156 548.819154 -0.064282 -1.785649 0.635329 767.328187 -0.000009 0.000084 0.000828].

3.3 Ring Filter and Target detection

A ring filter is designed to detect the target. The basic methodology is to correlate the images of both the cameras A and B with the ring filter in real time and find the peak in the correlated image. Before the correlation in performed, the images of both the cameras are down sampled by 2 to speed up the process. The bmp image of the ring filter is shown in Figure 3.4.

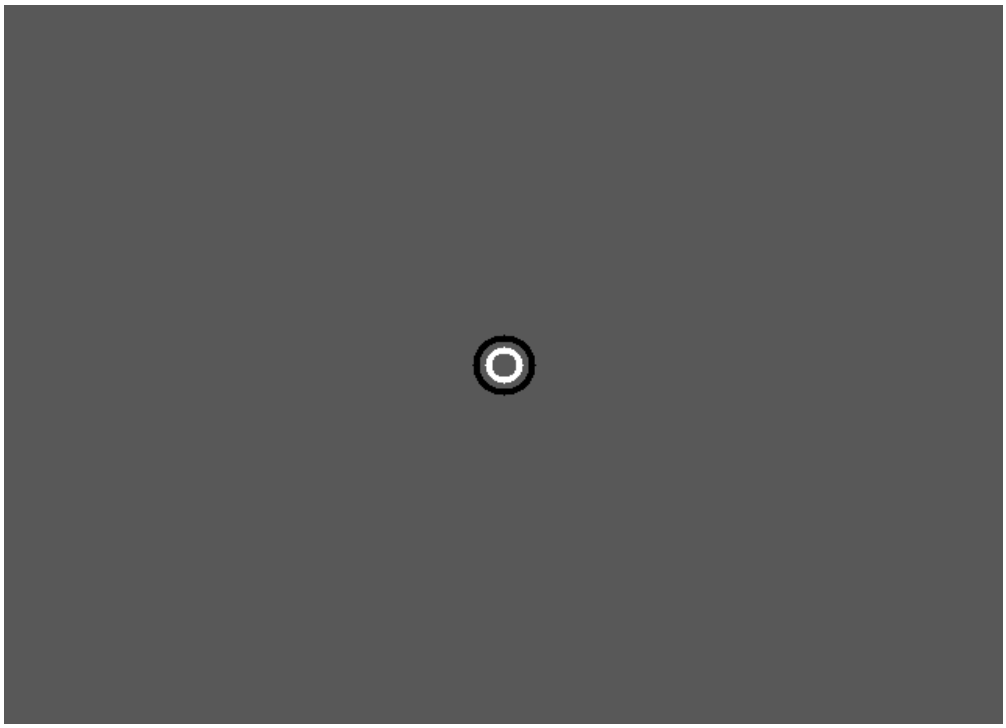

Figure 3.4: Bmp image of the Ring Filter.

The white portion on the image in Figure 3.4 is the positive target ring. When the ring filter is correlated with the camera image, the white portion forms the target to be

detected. The black portion is the negative suppression ring and forms the regions that are suppressed during correlation. The gray portion are don't care regions. The radius of the positive target ring and the negative suppression ring are chosen in such a way that the target is detected and the other regions are suppressed. The next step is to find the peak to side lobe ratio and the peak of the correlated image. The radius of the peak and the side lobe is set in such in a way that the radius of the side lobe is twice the radius of the peak in order to have a perfect detection of the target. After finding the peak in the correlated image, the coordinates of the peak are found. These coordinates are multiplied by 2 as the images were down sampled previously. Crosshairs are used to mark the peak which is nothing but the target.

An experiment is done in which a white circle in a dark background is used as the target and is to be detected.

Figure 3.5 shows the input images of the target taken from both the cameras A and B.

Figure 3.5: Input images of Camera A and Camera B.

The input images shown in Figure 3.5 are correlated with the ring filter shown in Figure 3.4. Figure 3.6 shows the correlated images of camera A and B respectively.

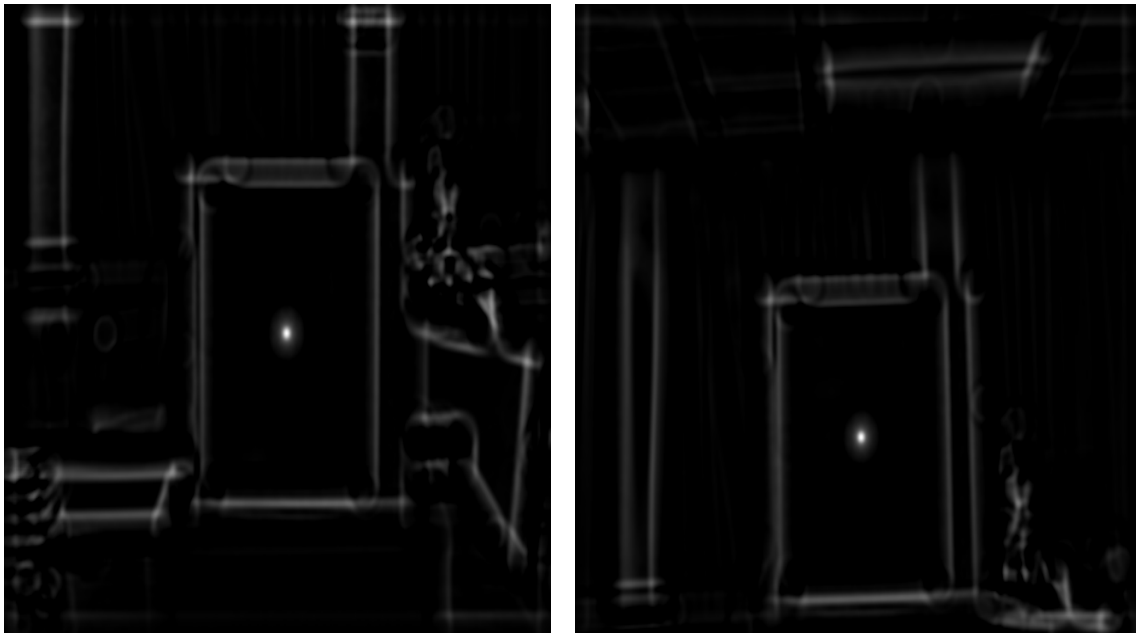

Figure 3.6: Result of the correlation of the input images in Figure 3.5 with the ring filter shown in Figure 3.4

After the correlation is done, the psr (peak to side lobe ratio) and peaks are found in the correlated image. Crosshairs are placed at the peak coordinates in order to indicate the target. Figure 3.7 shows the crosshair at the detected target.

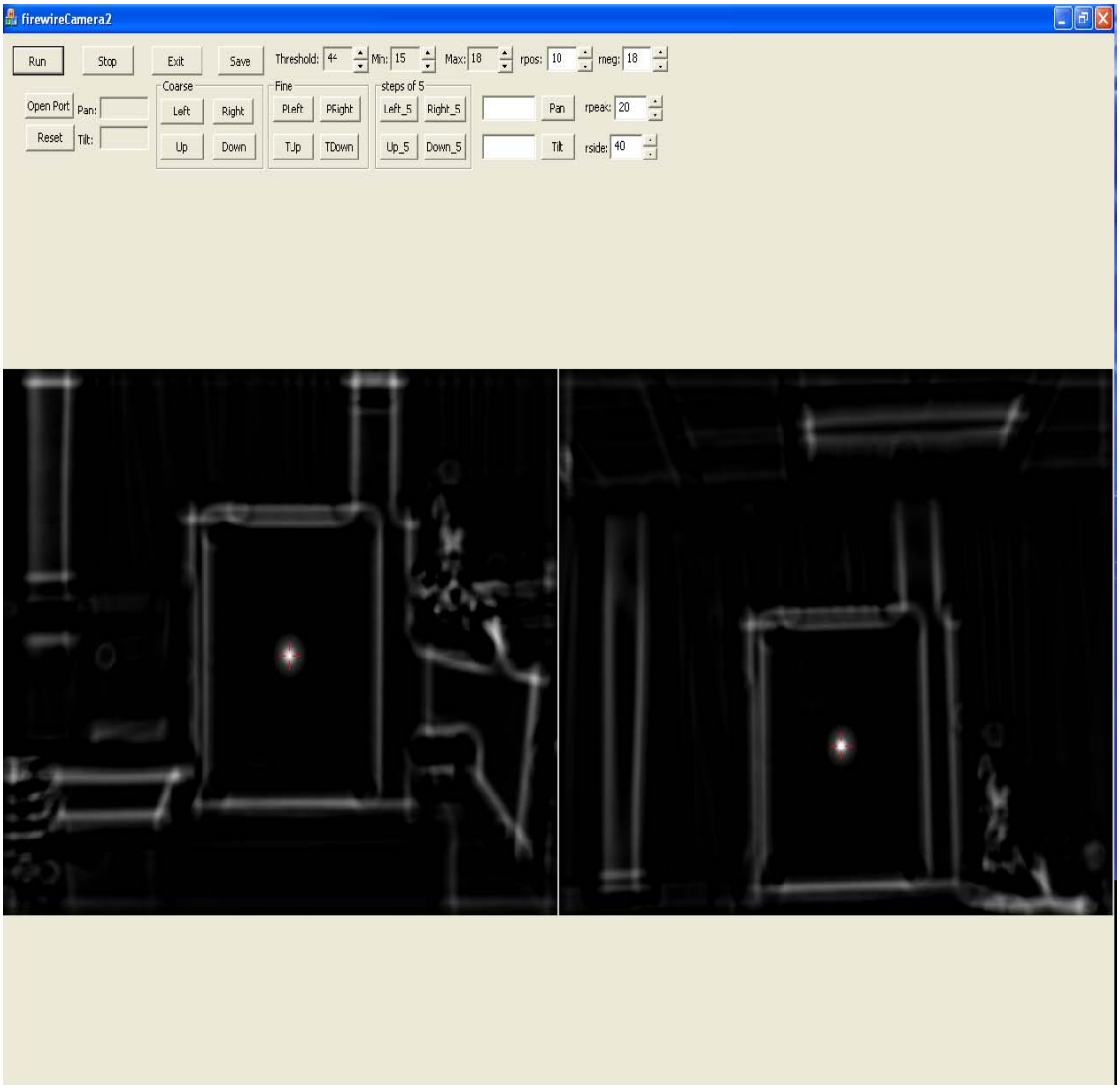

Figure 3.7: Target detection shown by the cross hair at the target.

Once the target is detected and the camera coordinates of the target have been found, the next step is to track the target using the pan-tilt unit and the projector. This is explained in Chapter 4.

Chapter 4 Tracking Methodology

4.1 Our Approach

After the target has been detected, the next step is to track the target within the range of the cameras. The coordinates of the center of the target i.e., the peak values of the target as found in Chapter 3, the '*m*' coefficients of both the cameras A and B and the '*m*' coefficients of the pan-tilt unit are used for tracking the target. The coordinates of the target are nothing but the camera coordinates of the target $(x_{Ac,n}, y_{Ac,n})$ and $(x_{Bc,n}, y_{Bc,n})$.

Three basic steps are involved in tracking the target:

1) The camera coordinates $(x_{Ac,n}, y_{Ac,n})$ and $(x_{Bc,n}, y_{Bc,n})$ and the '*m*' coefficients of camera A and B are used to find the world coordinates of the target. As described in Chapter 3, the 'm' coefficients of the cameras A and B were found to be

m_A = [1.882042 -0.061885 0.492436 604.987292 -0.031149 -1.853129 0.166115 662.390418 -0.000016 -0.000075 0.000823]

m_B = [1.896437 -0.002000 0.462156 548.819154 -0.064282 -1.785649 0.635329 767.328187 -0.000009 0.000084 0.000828]

Using the equations (3.1) to (3.7),

$$D_n = \begin{bmatrix} x_{Ac,n}\, m_{A12} - m_{A4} \\ y_{Ac,n}\, m_{A12} - m_{A4} \\ x_{Bc,n}\, m_{B12} - m_{B8} \\ y_{Bc,n}\, m_{B12} - m_{B8} \end{bmatrix}, \tag{4.1}$$

$$C_n = \begin{bmatrix} m_{A1} - m_{A9}\ x_{Ac,n} & m_{A2} - m_{A10}\ x_{Ac,n} & m_{A3} - m_{A11}\ x_{Ac,n} \\ m_{A5} - m_{A9}\ y_{Ac,n} & m_{A6} - m_{A10}\ y_{Ac,n} & m_{A7} - m_{A11}\ y_{Ac,n} \\ m_{B1} - m_{B9}\ x_{Bc,n} & m_{B2} - m_{B10}\ x_{Bc,n} & m_{B3} - m_{B11}\ x_{Bc,n} \\ m_{B5} - m_{B9}\ y_{Bc,n} & m_{B6} - m_{B10}\ y_{Bc,n} & m_{B7} - m_{B11}\ y_{Bc,n} \end{bmatrix}, \tag{4.2}$$

where $(x_{Ac,n}, y_{Ac,n}, x_{Bc,n}, y_{Bc,n})$ are the camera coordinates of Camera A and B,

$$D_n = C_n P_n \tag{4.3}$$

where P_n is the world coordinate

$$P_n^W = R_{cc}^{-1} C_n^T D_n, \tag{4.4}$$

where, $R_{cc} = C_n^T C_n$ (4.5)

Substituting the 'm' coefficients of the cameras A and B and the camera coordinates of both the cameras A and B in equations (4.1) to (4.4) gives the world coordinates of the target $P_n^W = \begin{bmatrix} x_{w,n} & y_{w,n} & z_{w,n} \end{bmatrix}^T$ (4.6)

The world coordinates are determined from the calibrated camera pair.

2) The world coordinates of the target $\begin{bmatrix} x_{w,n} & y_{w,n} & z_{w,n} \end{bmatrix}$ and the '*m*' coefficients of the pan-tilt unit are used to find the pan-tilt coordinates using the equations

$$x_{p,n} = \frac{m_1\ x_{w,n} + m_2\ y_{w,n} + m_3\ z_{w,n} + m_4}{m_9\ x_{w,n} + m_{10}\ y_{w,n} + m_{11}\ z_{w,n} + m_{12}} \tag{4.7}$$

and

$$y_{t,n} = \frac{m_5\ x_{w,n} + m_6\ y_{w,n} + m_7\ z_{w,n} + m_8}{m_9\ x_{w,n} + m_{10}\ y_{w,n} + m_{11}\ z_{w,n} + m_{12}} \tag{4.8}$$

3) The pan-tilt unit is moved to the spot position $(x_{p,n}, y_{t,n})$ and if the projector projects the spot on the target then the target tracking is accurate.

4.2 Experiment 1

In Chapter 3, Table 3.1, Table 3.2 and Table 3.3 shows the world coordinates, camera coordinates and the pan-tilt values. One way to test whether the system would eventually track the target is to use the camera coordinates of the calibration grid rings in Table 3.1 and Table 3.2 as the camera coordinates of the target. These coordinates along with the '*m*' coefficients of the cameras are substituted in the equation (4.1) and (4.2) to find D_n and C_n. The D_n and C_n are then used in equation (4.4) to get P_n which is nothing but the world coordinates. The world coordinate values in P_n should be same as the world coordinates in Table 3.1, Table 3.2 and Table 3.3. These world coordinates along with the '*m*' coefficients of the pan-tilt unit is then used in equation (4.7) and (4.8) to get the pan-tilt values. These pan-tilt values should be same as the pan-tilt values with respect to the world coordinates in Table 3.3.

The Table 4.1 shows the original world coordinates and the new world coordinates and the Table 4.2 shows the original pan-tilt values and the new pan-tilt values for the calibration grid rings.

Table 4.1: Original World Coordinates and New World Coordinates of the Calibration grid rings.

	Original World Coordinates			New World Coordinates		
Points	$x_{w,n}$	$y_{w,n}$	$z_{w,n}$	$x_{w,n}$	$y_{w,n}$	$z_{w,n}$
0	-31.82	0	31.82	-31.765	-0.035898	32.3
1	-76.368	0	76.368	-76.761	0.061406	75.922
2	-120.92	0	120.92	-120.81	-0.15025	120.64
3	-31.82	90	31.82	-32.092	89.857	32.214
4	-76.368	90	76.368	-76.278	89.888	76.583
5	-120.92	90	120.92	-121.08	90.036	121.34
6	-31.82	180	31.82	-32.215	180.08	31.968
7	-76.368	180	76.368	-76.742	180.06	75.659
8	-120.92	180	120.92	-120.26	180.14	121.06
9	31.82	0	31.82	32.222	0.50422	31.937
10	76.368	0	76.368	76.576	0.19487	75.564
11	120.92	0	120.92	120.47	-0.34533	121.22
12	31.82	90	31.82	32.189	90.007	31.683
13	76.368	90	76.368	76.301	89.984	76.437
14	120.92	90	120.92	121.13	89.852	121.19
15	31.82	180	31.82	31.745	179.53	31.994
16	76.368	180	76.368	76.665	179.88	75.746
17	120.92	180	120.92	120.71	180.45	121.15

Table 4.2: Original Pan-Tilt values and New Pan-Tilt Values.

	Original Pan-Tilt Values		New Pan-Tilt Values	
Points	$x_{p,n}$	$y_{t,n}$	$x_{p,n}$	$y_{t,n}$
0	1627	-1081	1626	-1082.3
1	1705	-1062	1704.7	-1062.4
2	1774	-1044	1773.9	-1044.6
3	1626	-831	1625.9	-828.53
4	1701	-817	1701.8	-818.91
5	1774	-812	1773	-809.65
6	1624	-572	1625.3	-573.01
7	1702	-574	1702.9	-573.84
8	1770	-574	1770.3	-574.39
9	1445	-1089	1445.7	-1088.3
10	1285	-1079	1283.9	-1078.8
11	1127	-1068	1128.9	-1069.6
12	1445	-833	1444.5	-831.8
13	1282	-828	1282.4	-826.71
14	1125	-823	1125.2	-822.17
15	1445	-573	1444.1	-574.36
16	1280	-573	1280.1	-573.83
17	1126	-573	1124.4	-572.77

The plot of the original world coordinates and the new world coordinates as shown in Table 4.1 are plotted. Figure to Figure 4.10 shows the plot between the original world coordinates and the new world coordinates:

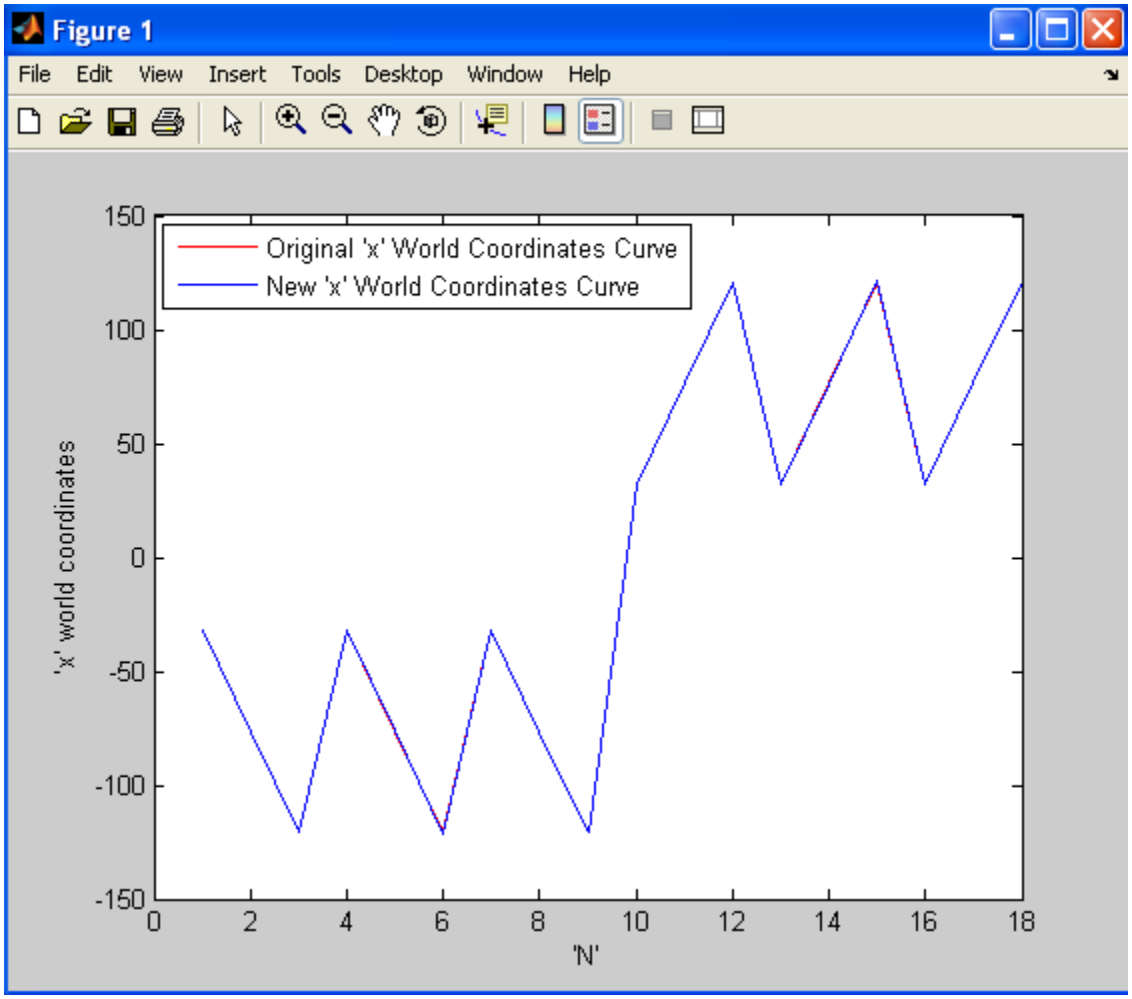

Figure 4.1: Plot of Original '*x*' World Coordinate Curve and New '*x*' World Coordinate Curve.

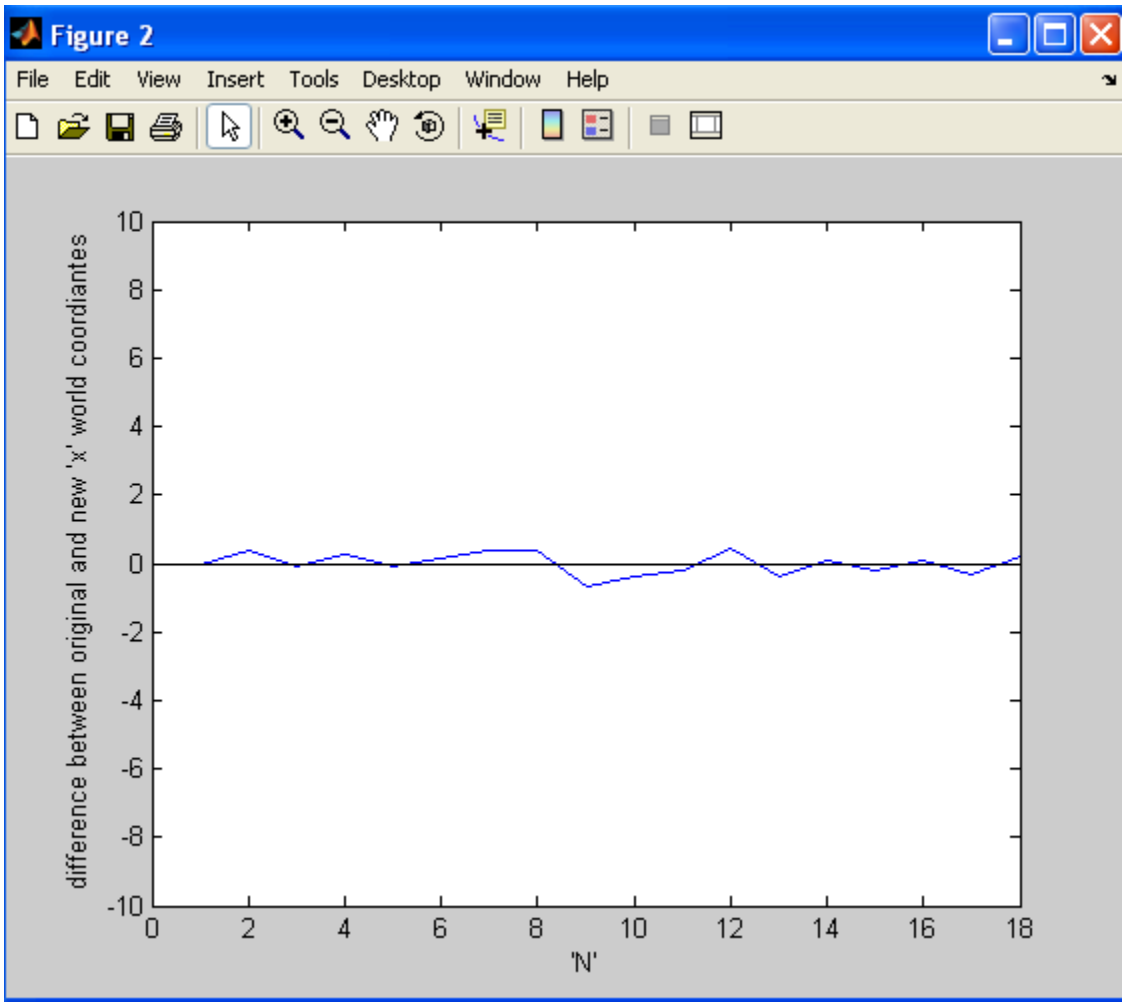

Figure 4.2: Plot of difference between Original and New '*x*' World Coordinates.

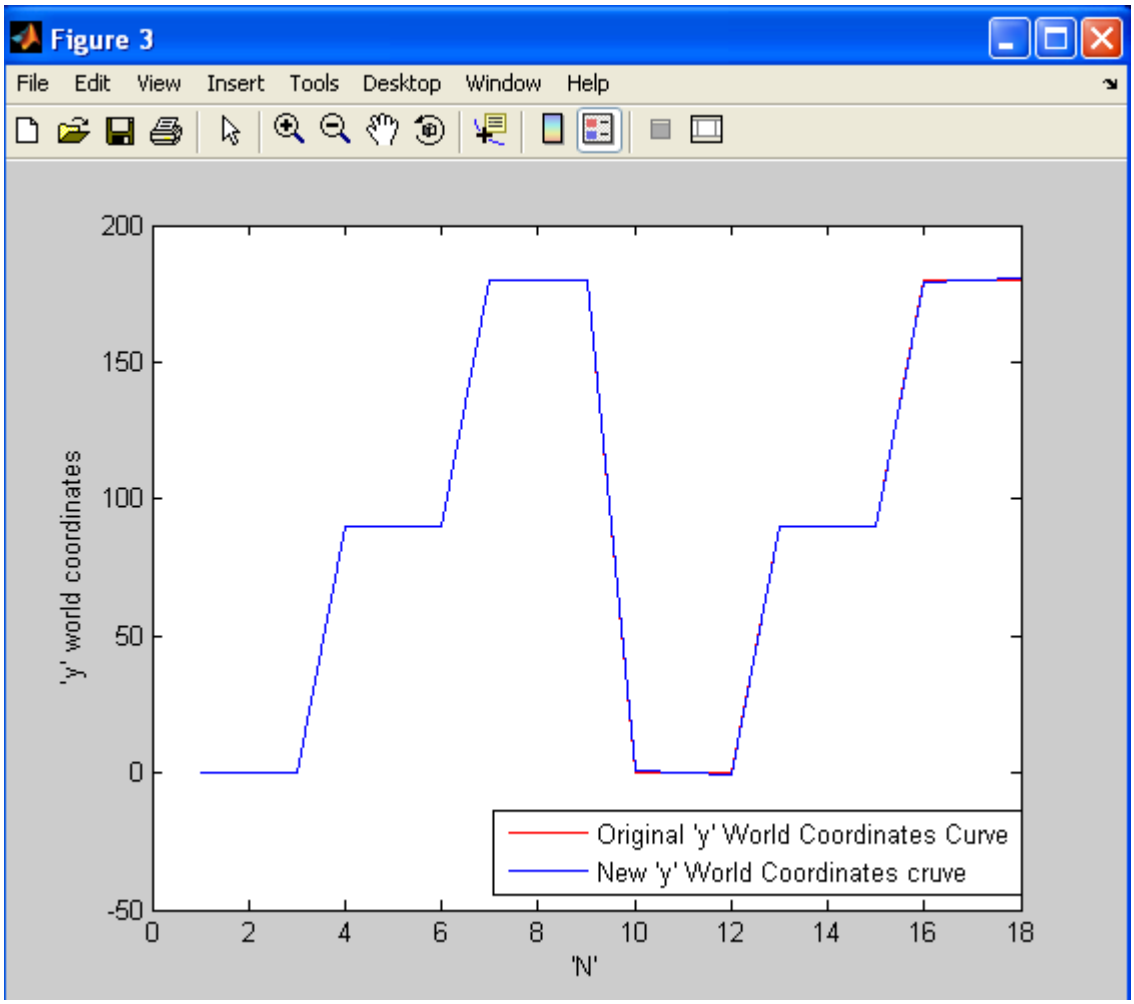

Figure 4.3: Plot of Original '*y*' World Coordinate Curve and New '*y*' World Coordinate Curve.

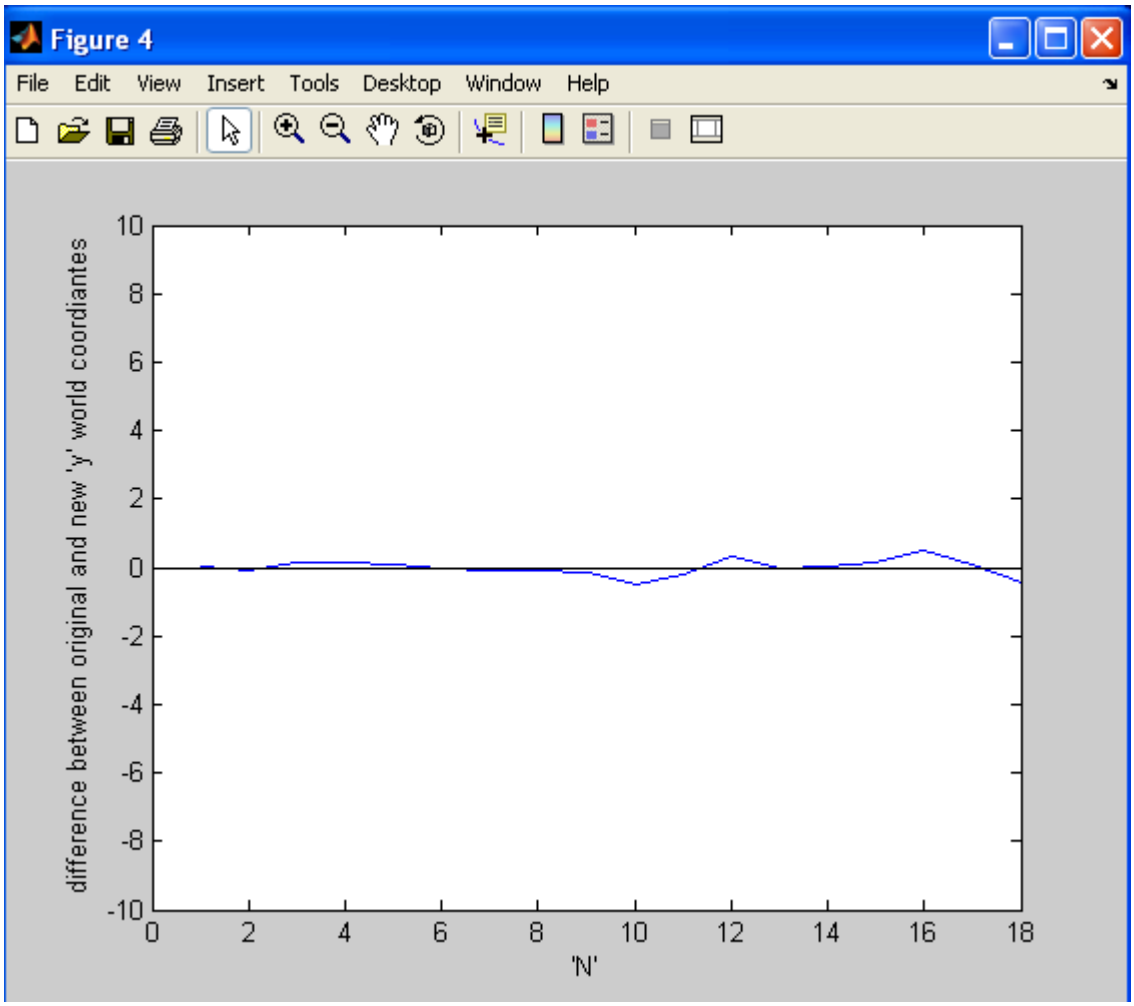

Figure 4.4: Plot of difference between Original and New '*y*' World Coordinates.

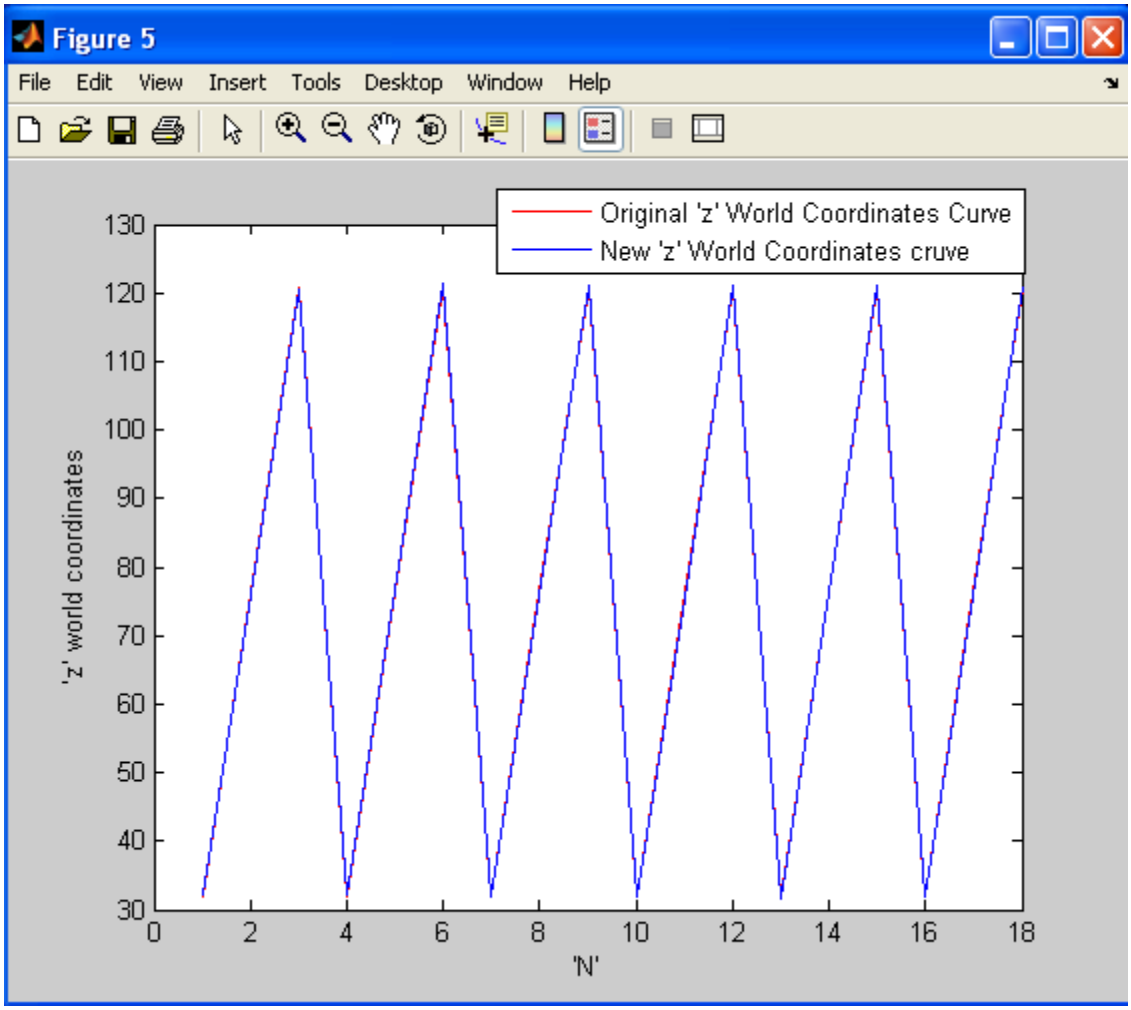

Figure 4.5: Plot of Original ‘z’ World Coordinate Curve and New ‘z’ World Coordinate Curve.

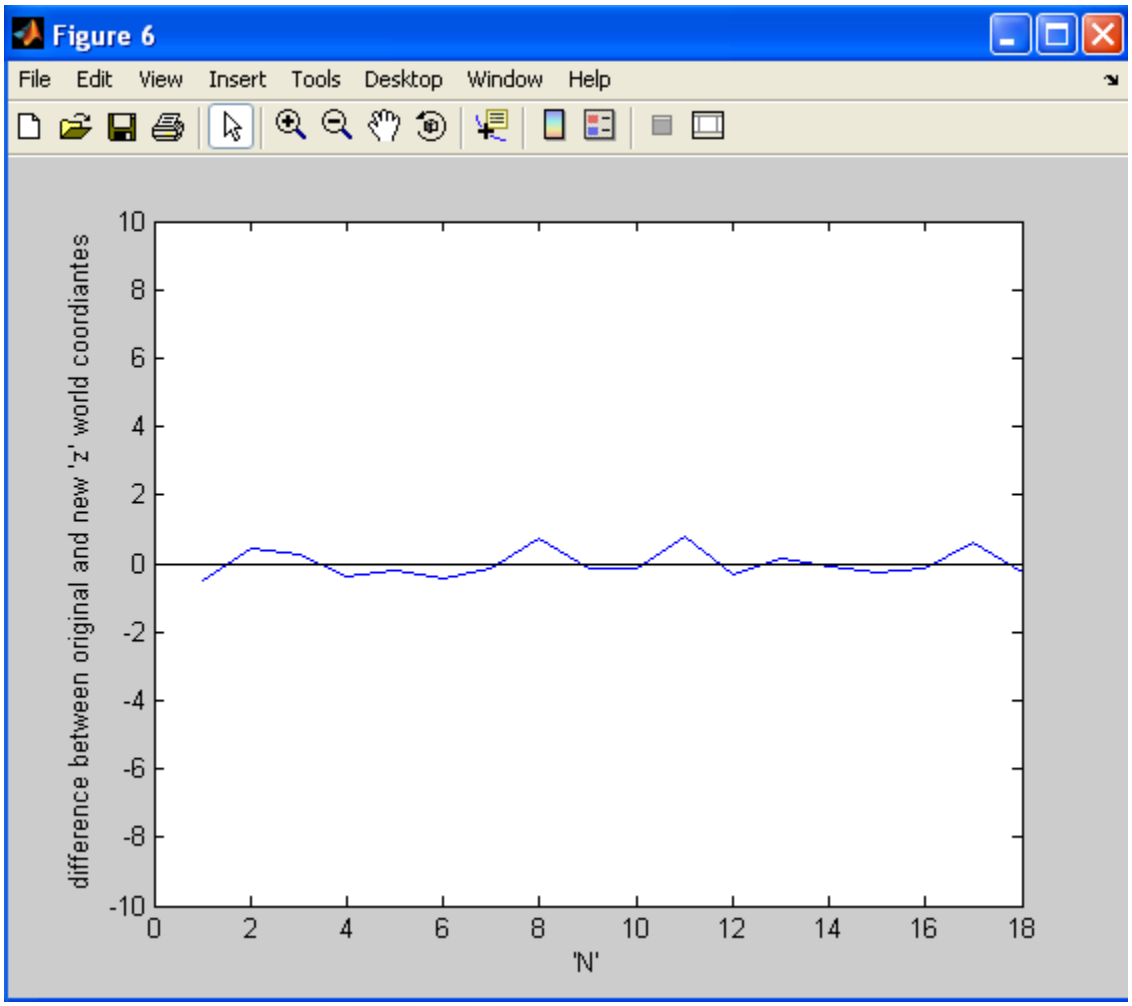

Figure 4.6: Plot of difference between Original and New ‘z’ World Coordinates.

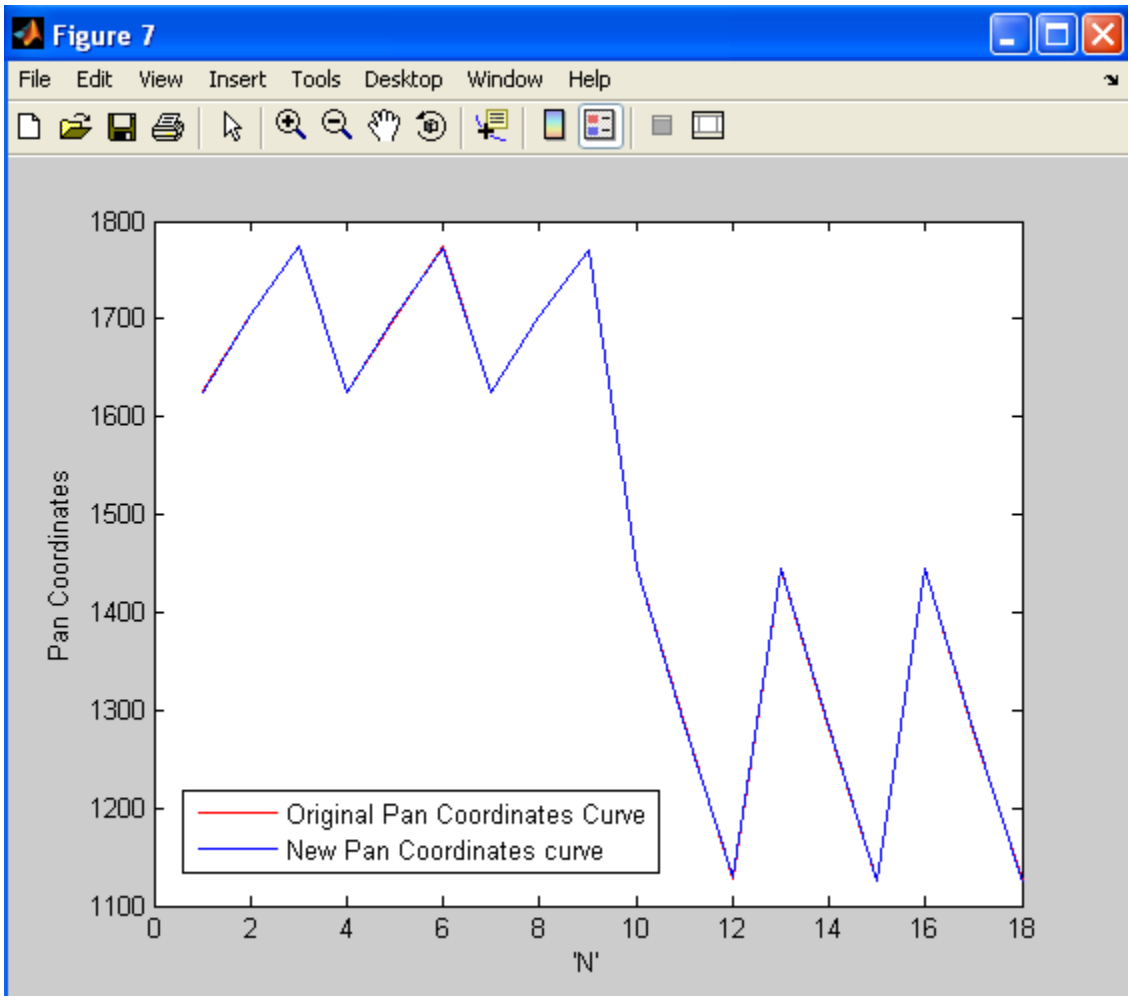

Figure 4.7: Plot of Original Pan Coordinate Curve and New Pan Coordinate Curve.

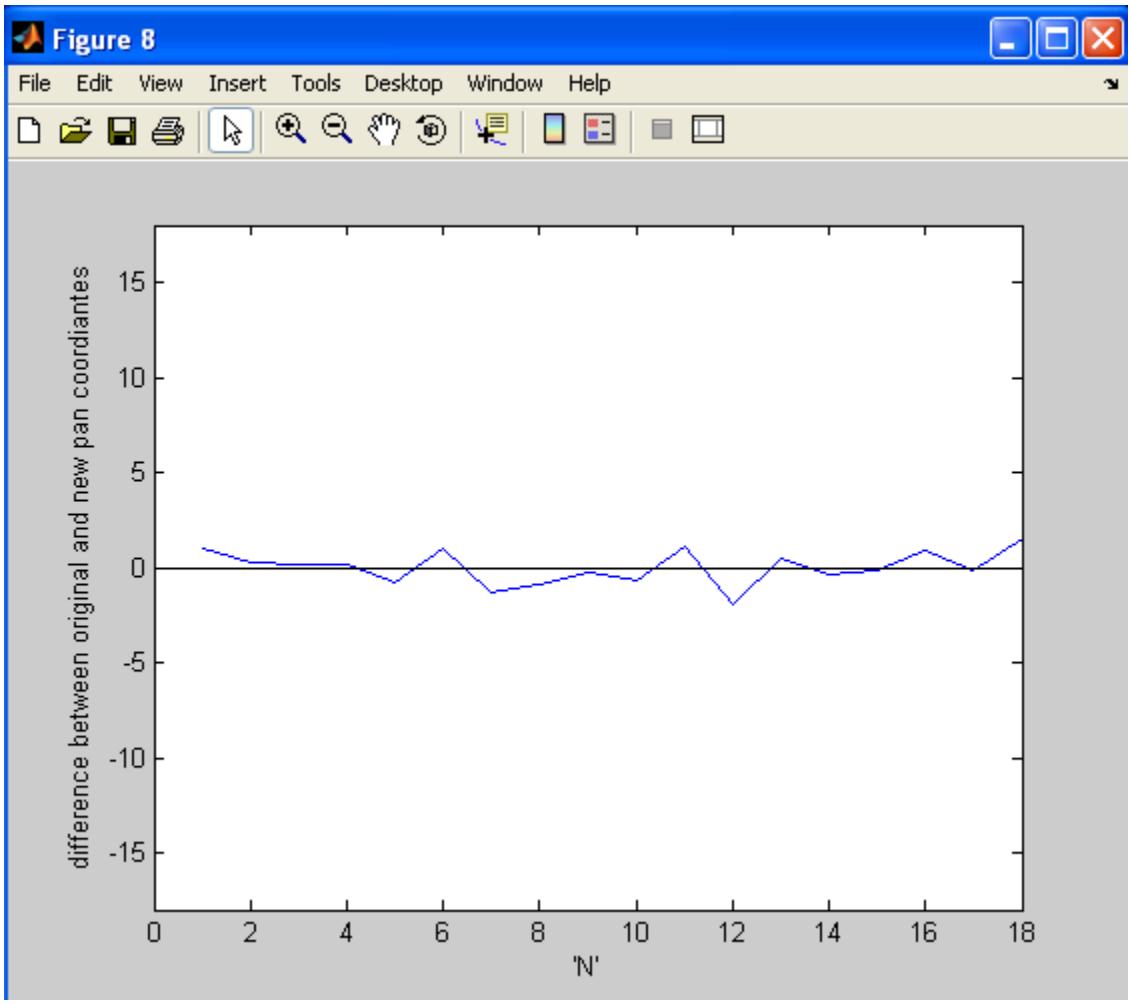

Figure 4.8: Plot of difference between Original and New Pan Coordinates.

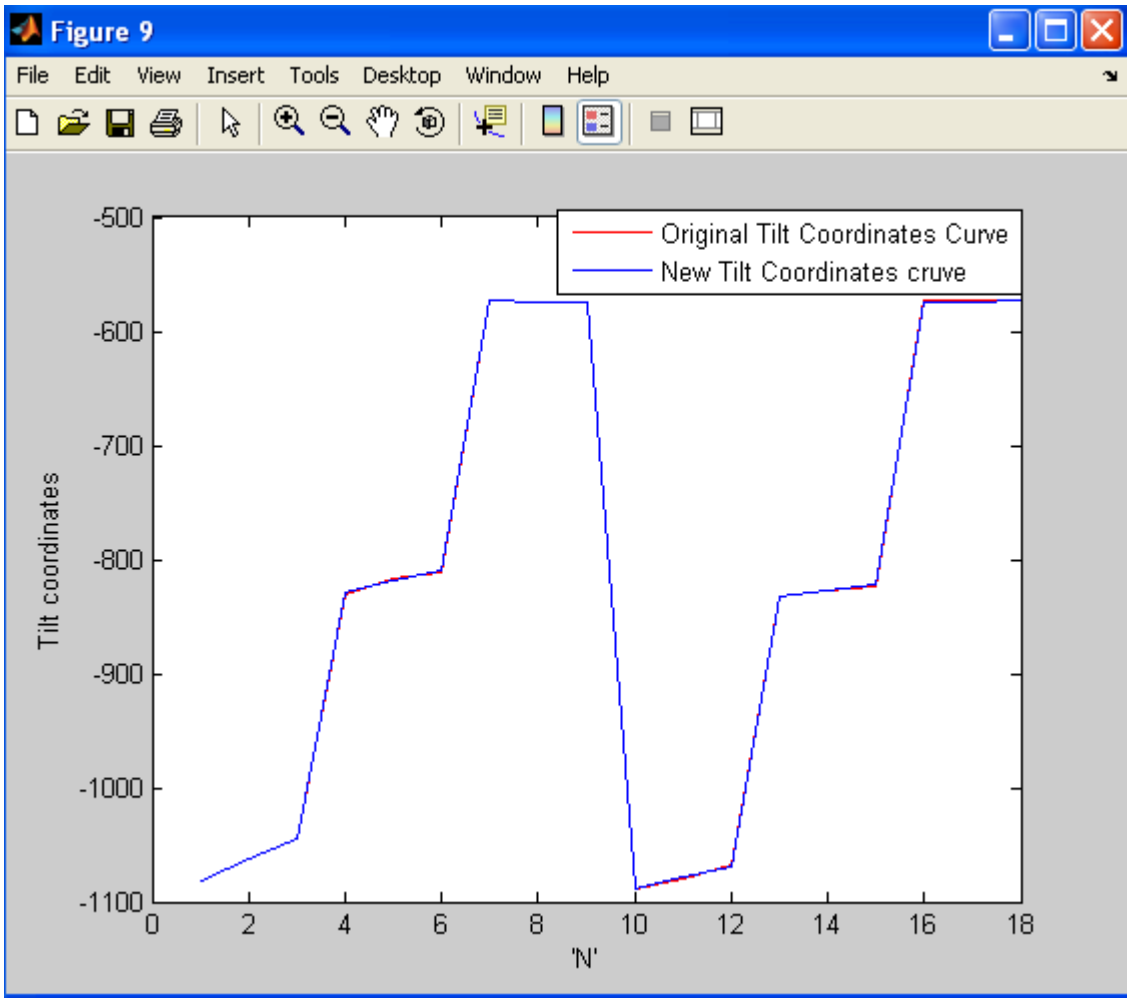

Figure 4.9: Plot of Original Tilt Coordinate Curve and New Tilt Coordinate Curve.

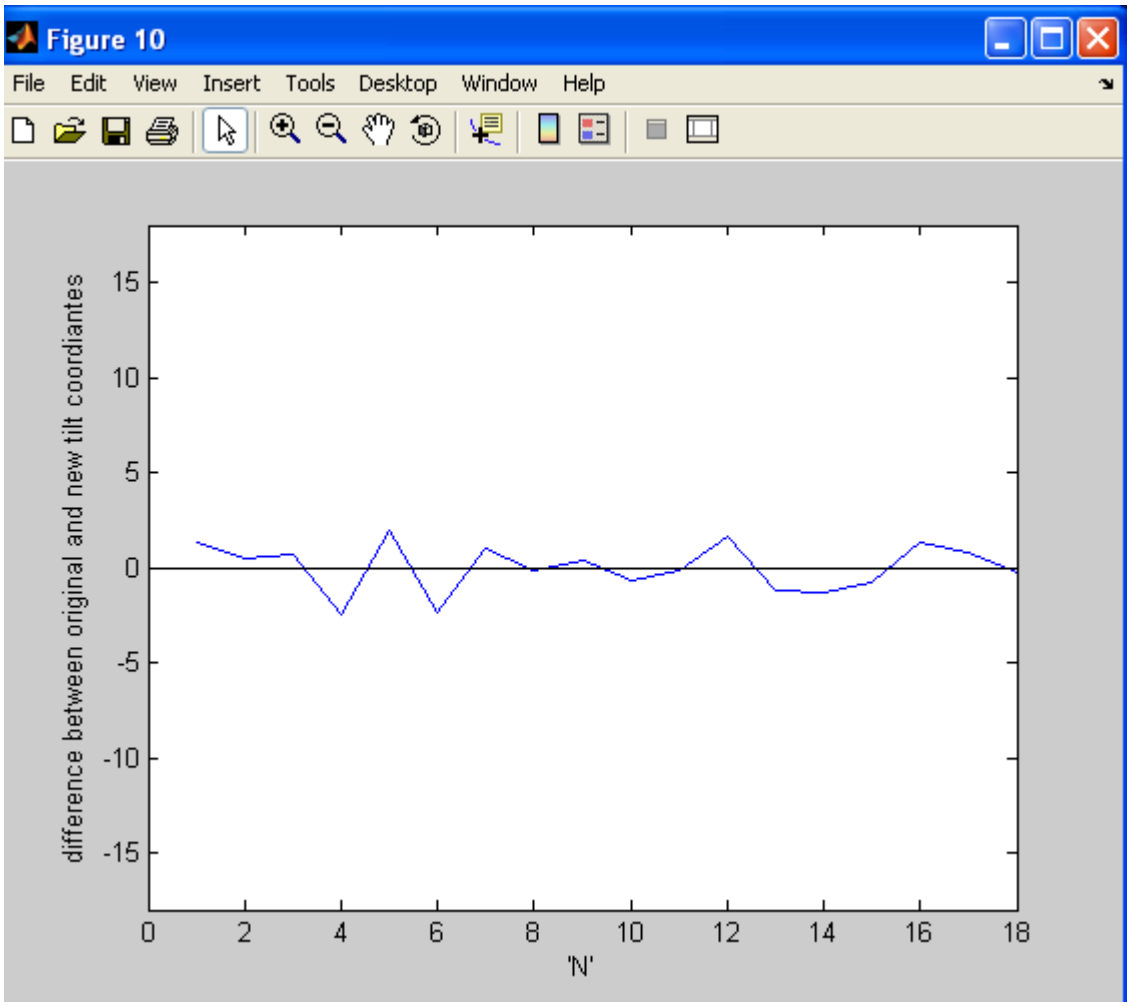

Figure 4.10: Plot of difference between Original and New Tilt Coordinates.

As seen in the Table 4.1 and Table 4.2 and Figure 4.1 to Figure 4.8, the difference between the original and new world coordinates, and original and new pan-tilt coordinates is very small. This shows that the system tracks the center of the calibration grid circles with minimum error.

4.3 Experiment 2

The same target as shown in Figure 3.5 is used for the tracking. The tracking functionality of the system is tested in two ways:

1) Moving the target from right to left in the '*x*' direction

2) Moving the target backwards in the '*z*' direction

4.3.1 Moving the target from right to left in the '*x*' direction

The target to be tracked is placed at a certain distance from the cameras (say initial position) and the projector and at a fixed height from the ground. Then the program to detect the target is run and we get the camera coordinates of the center of the target. These camera coordinates $(x_{Ac,n}, y_{Ac,n})$ and $(x_{Bc,n}, y_{Bc,n})$ are fed into the MATLAB program to get the world coordinates $(x_{w,n}, y_{w,n}, z_{w,n})$ and the pan-tilt coordinate $(x_{p,n}, y_{t,n})$ of the center of the target. Finally the pan-tilt coordinates are fed back to pan-tilt control unit which moves the projector in the direction of the target and the white spot from the projector falls on the target. If the white spot is not exactly at the center of the target, then the error in '*x*' and '*y*' is measured using a ruler. Also the distance from the center of the projector lens and the target is measured and noted. Next, the target is moved to the left by about 40cm so that the '*y*' and '*z*' world coordinates of the target are fixed and only the '*x*' world coordinates changes. Again the program is run to detect the target. The new camera coordinates are found and are used to get the new world

coordinates and the new pan-tilt coordinates. The pan-tilt unit is then moved to a new location and the projector projects the spot on the target. If the spot is within the circular target then system is able to track the target. Again, the error in '*x*' and '*y*' and the distance from the center of projector lens to the target is measured and noted. The target is again moved to the left by 40 cm and the whole process is continued for about 20 readings.

After the above process, the target is place at the initial position and moved backwards to about 20 cm and the whole process explained above is repeated. The tracking process is continued until the target is moved backwards to a distance of about 100cm from its initial position.

The result of this experiment is discussed in Chapter 5.

4.3.2 Moving the target backwards in the '*z*' direction

In this experiment, the target to be tracked is initially placed at a fixed position from the cameras and the projector and at a fixed height from the ground. The program to detect the target is run. We get the camera coordinates of the target. These camera coordinates along with the '*m*' coefficients of the cameras are fed into the MATLAB code to get the world coordinates and the pan-tilt coordinates of the target. The pan-tilt coordinates are fed to the pan-tilt control unit and the projector is moved toward the target. The error in '*x*' and '*y*' world coordinates and the distance from the projector lens to the target are measured. Next, the target is moved backwards in the '*z*' direction to a distance of about 10cm and the same process is repeated. The process is continued until the target is moved

to a distance of about 100 cm from its initial position. The result of this experiment is discussed in Chapter 5.

Chapter 5 Results

5.1 Introduction

This chapter provides an overview of the results of the experiments done in the research. The section 5.2 and 5.3 shows the results and plots of the experiment 2 explained in Chapter 4

5.2 Experiment 2.1 results

The Table 5.1 to Table 5.6 shows the error in 'x' and 'y' for the target at different positions. The distance 'd' of the target from the projector lens, 'x_w', 'y_w', 'z_w' are the world coordinates of the target, 'x_p','y_p' are the pan and tilt coordinates of the target respectively, 'x_{error}', 'y_{error}' are the error in 'x' and 'y' world coordinates of the target and 'x_{wn}' and 'y_{wn}' are the new 'x' and 'y' world coordinates .

Table 5.1: Error in 'x' and 'y' world coordinates when the target is at a distance of about 152 cm from the projector lens.

	d	x_w	y_w	z_w	x_p	y_t	x_{error}	y_{error}	x_{wn}	y_{wn}
0	152.4	-273.492	37.8041	12.1417	2291	-960	4	1	-269.492	36.80
40	149.1	-236.174	36.71	15.8665	2191	-965	4	1	-232.174	35.71
80	147.8	-197.221	37.6513	14.2262	2092	-966	3.5	0.5	-193.721	37.1513
120	146.8	-158.415	37.4876	12.5104	1992	-970	3	0.5	-155.415	36.99
160	145.3	-118.672	37.4601	16.0957	1881	-973	5	0.5	-113.672	36.9601
200	143.5	-81.5835	38.4301	14.5417	1782	-974	3	0.5	-78.5835	37.9301
240	142.2	-44.0998	38.2959	12.7322	1680	-978	2	0.5	-42.0998	37.7959
280	141	-4.6568	38.2697	16.504	1564	-981	2.5	0.5	-2.1568	37.77
320	139.7	35.9321	39.1462	15.07	1450	-982	2	0.5	37.9321	38.65
360	138.4	73.7055	38.97	13.558	1339	-987	2	0	75.7055	38.97
400	137.7	111.7889	39.0147	16.9501	1223	-989	3	0	114.7889	39.0147
440	136.70	149.9932	39.9478	15.3282	1108	-991	2.5	0	152.4932	39.9478
480	135.90	190.2161	40.8161	13.9163	984	-993	3	0	193.2161	40.8161
520	135.60	227.0429	40.6736	12.145	868	-997	2.5	0	229.5429	40.6736
560	135.10	263.1601	41.6777	10.1052	753	-998	2	0	265.1601	41.68
600	135.00	302.7113	41.6322	14.1318	621	-1001	4	0	306.7113	41.6322
640	134.50	339.047	41.754	17.273	499	-1004	5	-1	344.047	42.75

The standard deviation in 'x' and 'y' world coordinates are found to be 3.26 mm and 0.52 mm respectively.

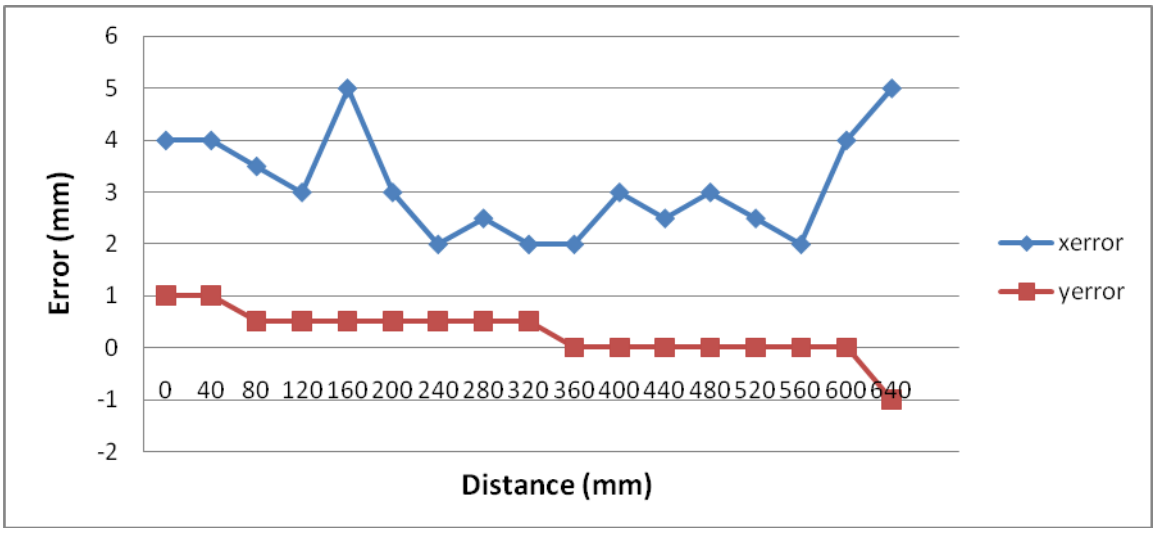

Figure 5.1: Plot of error in '*x*' and '*y*' world coordinates when the target is at a distance of about 152 cm from the projector.

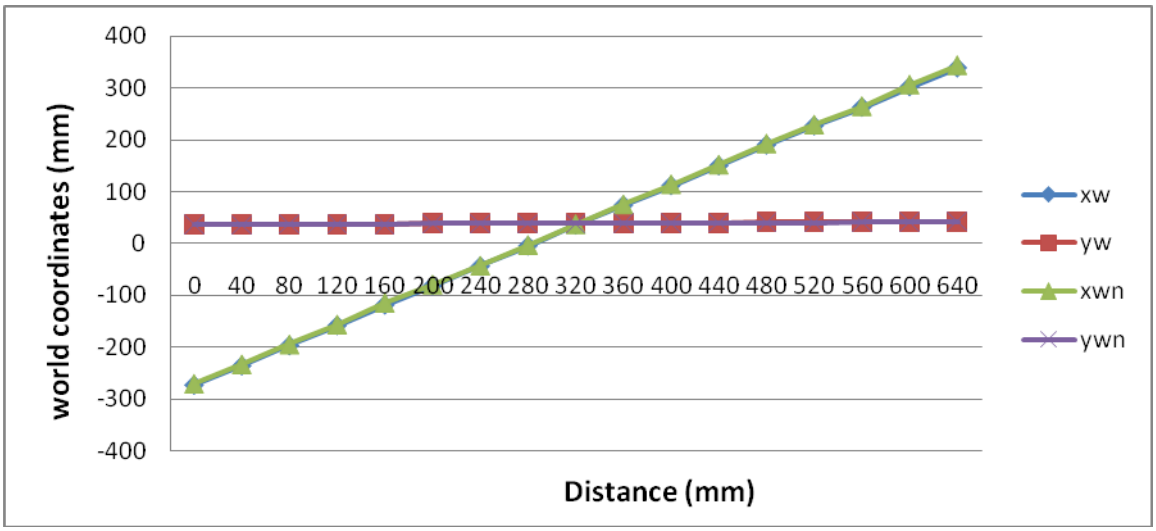

Figure 5.2: Plot of original and new '*x*' and '*y*' world coordinates.

Table 5.2: Experiment 1 results- Error in 'x' and 'y' world coordinates when the target is at a distance of about 170 cm from the projector lens.

	d	x_w	y_w	z_w	x_p	y_t	x_{error}	y_{error}	x_{wn}	y_{wn}
0	168.9	-276.978	32.5118	216.6926	2046	-928	5	1	-271.978	31.5118
40	167.6	-237.703	32.5104	213.5762	1960	-930	3	2	-234.703	30.5104
80	166.4	-199.653	32.677	217.7696	1867	-931	4	1.5	-195.653	31.177
120	165.6	-160.171	32.8087	222.1152	1771	-933	4	2	-156.171	30.8087
160	164.1	-120.499	32.7717	219.3215	1679	-936	3	1.5	-117.499	31.2717
200	162.8	-79.5567	32.8652	223.8232	1576	-938	5	1.5	-74.5567	31.3652
240	161.8	-40.0499	32.867	220.4528	1482	-941	3.5	1	-36.5499	31.867
280	160.5	-2.0119	33.0279	224.9347	1384	-943	3	1	0.9881	32.0279
320	159.5	37.9166	32.9875	221.9422	1285	-946	3	0	40.9166	32.9875
360	158.8	78.4417	33.2899	233.7511	1173	-946	6	1	84.4417	32.2899
400	158	116.5511	34.3992	223.2934	1080	-948	4	0.5	120.5511	33.8992
440	157	155.4219	34.4073	219.8971	978	-952	3	0	158.4219	34.4073
480	156.7	194.2166	35.6253	217.3298	847	-953	2.5	0	196.7166	35.6253
520	155.7	232.0359	35.8313	221.4199	768	-954	6	0	238.0359	35.8313
560	155.4	271.8728	36.0244	225.3591	656	-956	8	-0.5	279.8728	36.5244
600	155.4	311.0689	37.2575	222.4906	545	-956	8	-0.5	319.0689	37.7575
640	152.1	347.496	37.2746	219.285	440	-960	10	-2	357.496	39.2746
680	149.4	388.2403	37.411	223.6606	320	-962	12.5	-2	400.7403	39.411

The standard deviation in 'x' and 'y' world coordinates are found to be 5.85 mm and 1.23 mm respectively.

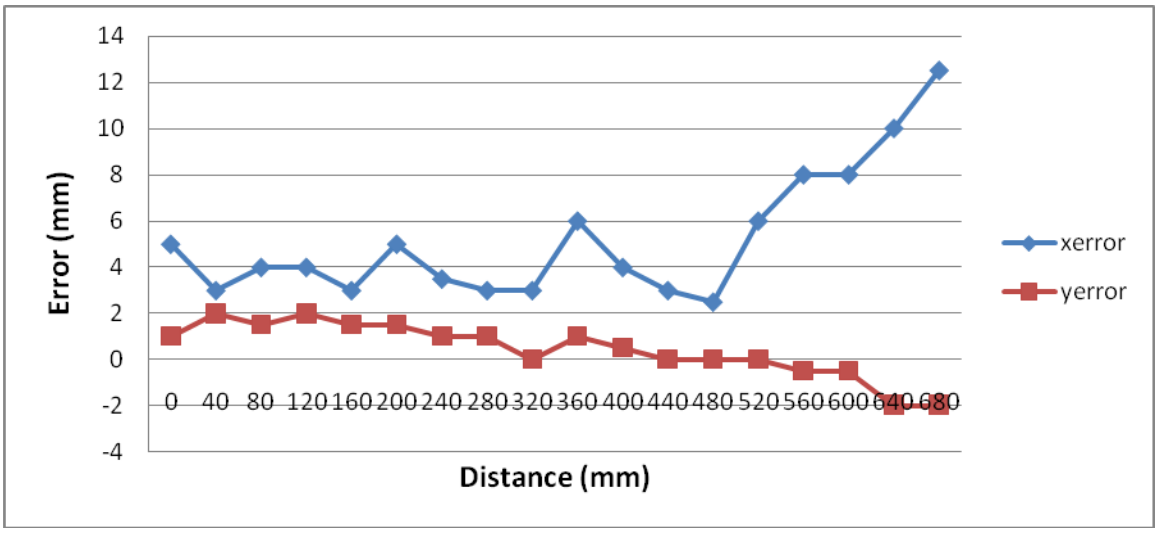

Figure 5.3: Plot of error in '*x*' and '*y*' world coordinates when the target is at a distance of about 170 cm from the projector.

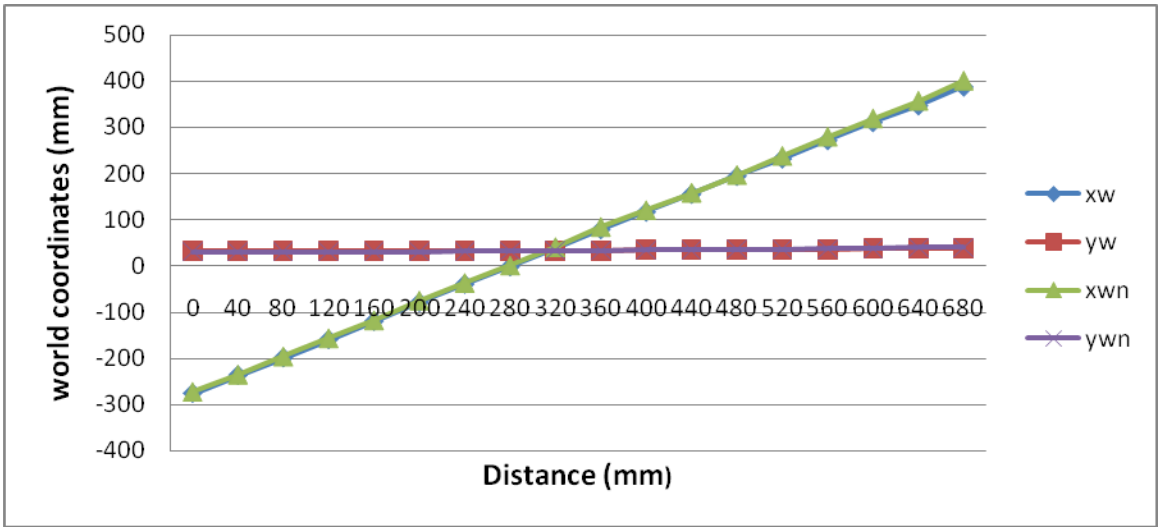

Figure 5.4: Plot of original and new '*x*' and '*y*' world coordinates.

Table 5.3: Error in 'x' and 'y' world coordinates when the target is at a distance of about 188 cm from the projector lens.

	d	x_w	y_w	z_w	x_p	y_t	x_{error}	y_{error}	x_{wn}	y_{wn}
0	186.7	-284.915	25.9979	430.25	1856	-902	9	1	-275.915	24.9979
40	185.4	-242.586	26.098	425.6839	1773	-905	7	0.5	-235.586	25.598
80	183.4	-203.865	26.3995	430.8357	1690	-905	6.5	0.5	-197.365	25.8995
120	182.4	-162.483	26.5367	425.8595	1606	-908	5	0	-157.483	26.5367
160	181.6	-121.411	26.7938	430.9363	1514	-909	6	0	-115.411	26.7938
200	181.1	-79.6408	26.8926	426.374	1427	-912	4	0	-75.6408	26.8926
240	180.3	-39.7876	27.1928	431.2795	1336	-913	5	0	-34.7876	27.1928
280	179.8	23.1443	28.1344	429.4081	1196	-915	4	0	27.1443	28.1344
320	179.1	44.1714	27.6579	442.0541	1142	-915	7	0	51.1714	27.6579
360	178.3	83.4505	27.8516	436.7384	1055	-918	7	0	90.4505	27.8516
400	177.8	125.3787	28.1131	441.8766	956	-919	8	0	133.3787	28.1131
440	177.8	162.7453	29.591	427.4133	873	-920	5	-0.5	167.7453	30.091
480	177.3	204.6232	29.8545	432.5021	771	-921	7	-0.5	211.6232	30.3545
520	176.3	244.6263	30.1359	437.8572	672	-922	10	-0.5	254.6263	30.6359
560	176.3	283.0211	30.3287	432.5606	578	-925	9.5	-2	292.5211	32.3287
600	176	322.0187	31.9194	427.9983	481	-925	9.5	-2	331.5187	33.9194
640	176.5	361.3715	32.2686	432.7578	381	-926	13	-2	374.3715	34.2686
680	176.8	402.3388	31.1038	437.6031	275	-930	16.5	-3	418.8388	34.1038
720	177.3	439.3446	32.4261	424.0344	177	-932	13	-3	452.3446	35.4261

The standard deviation in 'x' and 'y' world coordinates are found to be 8.6 mm and 1.3 mm respectively.

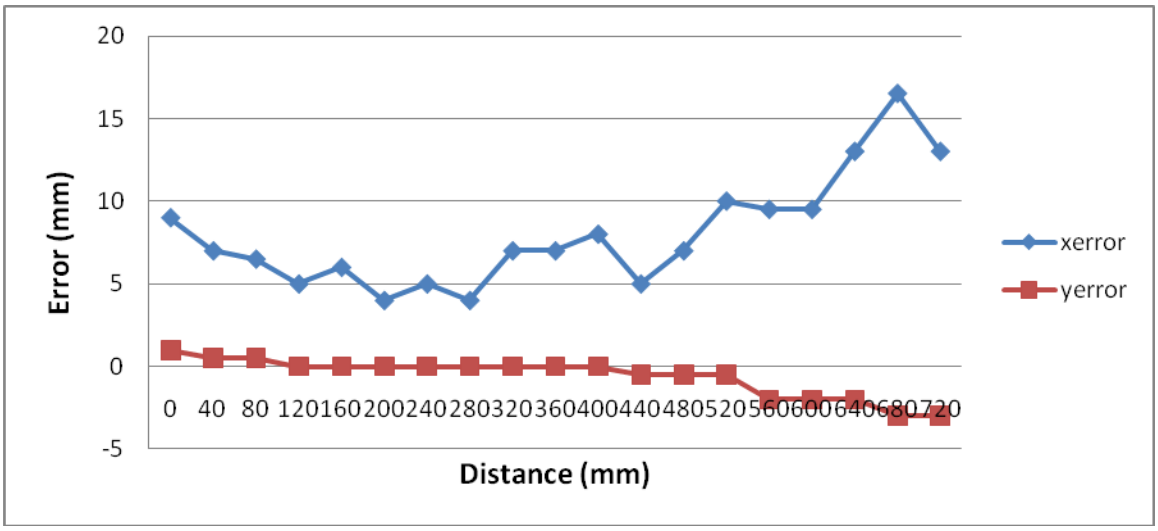

Figure 5.5: Plot of error in '*x*' and '*y*' world coordinates when the target is at a distance of about 188 cm from the projector lens.

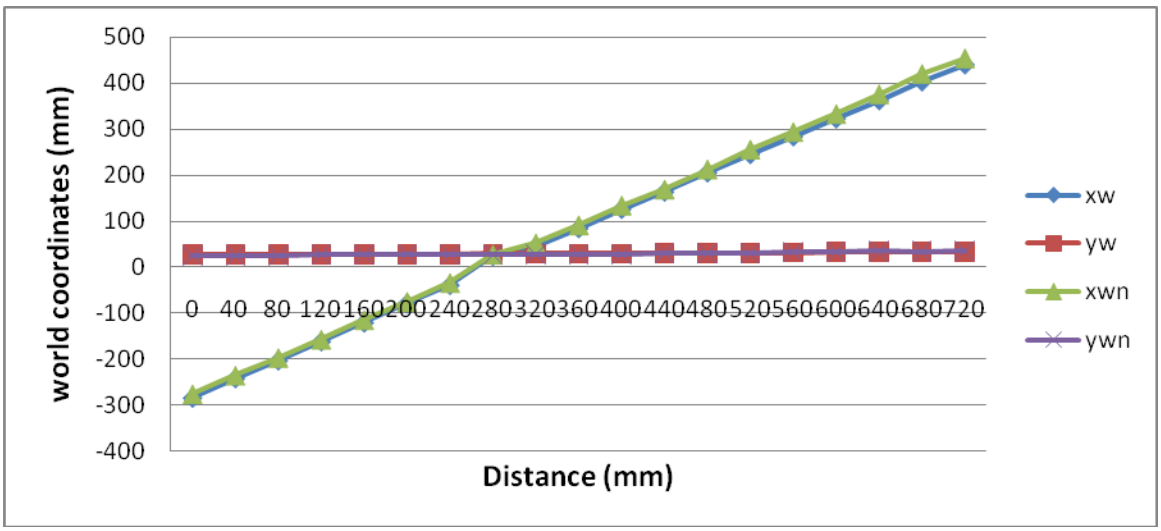

Figure 5.6: Plot of original and new '*x*' and '*y*' world coordinates.

Table 5.4: Error in 'x' and 'y' world coordinates when the target is at a distance of about 208 cm from the projector lens.

	d	x_w	y_w	z_w	x_p	y_t	x_{error}	y_{error}	x_{wn}	y_{wn}
0	207	-285.08	19.8055	629.3299	1702	-884	5	0	-280.08	19.8055
40	205.7	-247.153	20.31	634.607	1627	-884	5	1	-242.153	19.31
80	204.5	-204.263	19.0129	640.3793	1543	-888	6.5	0	-197.763	19.0129
120	203.2	-164.324	19.4708	645.9218	1463	-888	7	0	-157.324	19.4708
160	202.4	-122.567	19.7483	639.6743	1386	-890	6	0	-116.567	19.7483
200	201.2	-81.6315	20.2106	644.8759	1303	-890	7	0.5	-74.6315	19.7106
240	200.4	-41.0276	20.5353	638.1016	1226	-893	5	0	-36.0276	20.5353
280	199.6	-0.7545	20.9972	643.6201	1143	-893	6	0	5.2455	20.9972
320	198.6	41.0877	21.2706	637.0783	1060	-896	5	0	46.0877	21.2706
360	198.1	83.1514	21.6859	642.805	971	-896	8	0	91.1514	21.6859
400	197.6	121.3875	22.0511	635.8548	894	-898	7	0	128.3875	22.0511
440	196.9	163.6185	22.4675	641.5761	802	-899	9	0	172.6185	22.4675
480	196.3	204.7298	22.7333	635.0755	716	-901	9	-0.5	213.7298	23.2333
520	196.1	243.8786	23.2454	640.3668	629	-901	11	0	254.8786	23.2454
560	195.3	284.6309	23.5068	633.8903	540	-904	12	-2	296.6309	25.5068
600	195.1	321.8156	23.8598	627.0455	457	-907	11	-3	332.8156	26.8598
640	194.8	363.6343	24.2436	633.22	362	-907	15	-3	378.6343	27.2436
680	194.8	401.1522	24.6239	625.9077	275	-910	15	-4	416.1522	28.6239
720	194.8	442.1919	25.0869	631.3763	179	-910	18	-4	460.1919	29.0869

The standard deviation in 'x' and 'y' world coordinates are found to be 9.5 mm and 1.7 mm respectively.

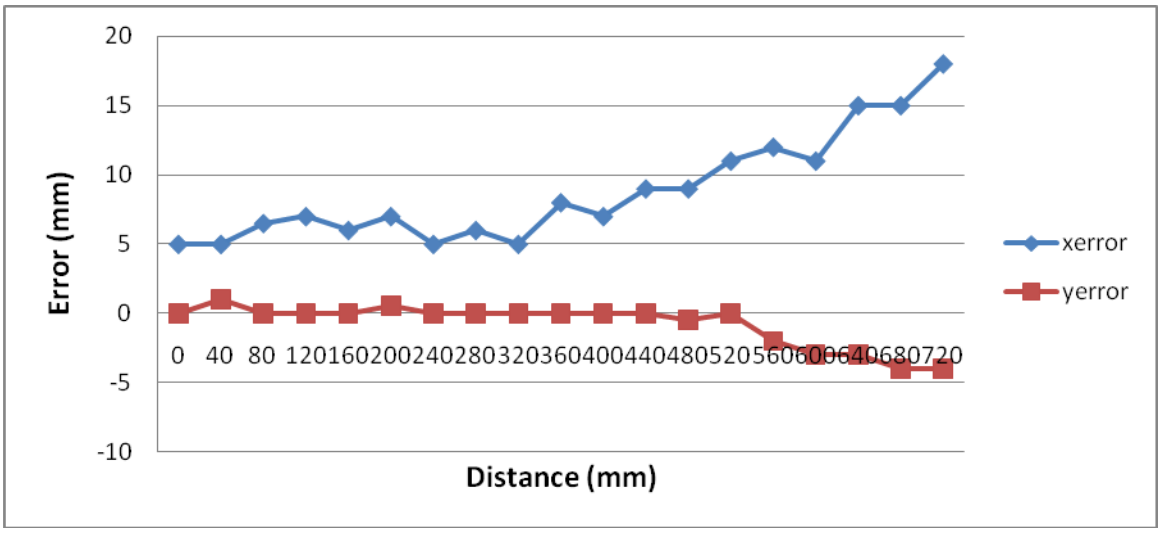

Figure 5.7: Plot of error in '*x*' and '*y*' world coordinates when the target is at a distance of about 208 cm from the projector lens.

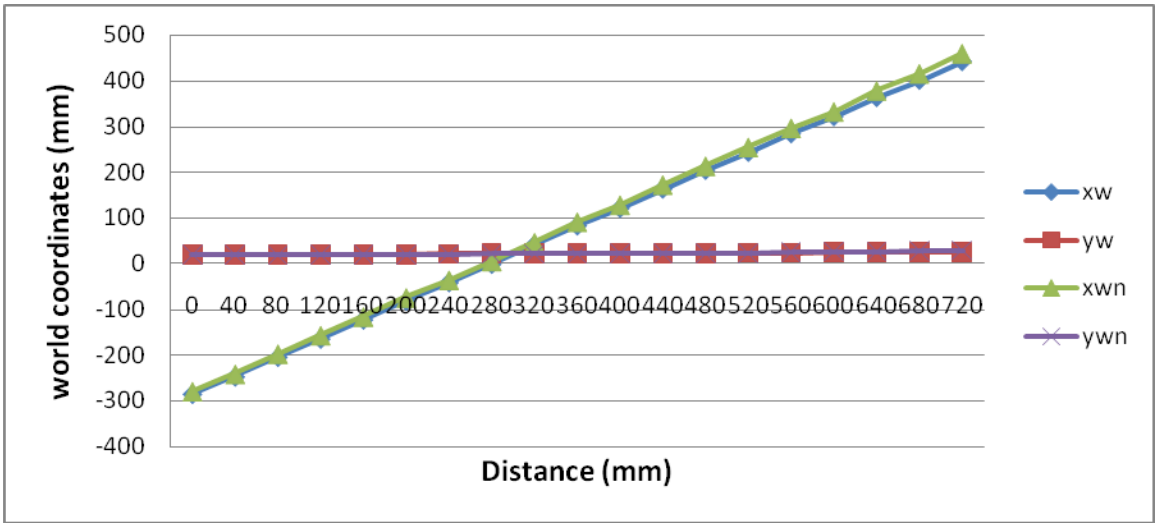

Figure 5.8: Plot of original and new '*x*' and '*y*' world coordinates.

Table 5.5: Error in 'x' and 'y' world coordinates when the target is at a distance of about 228 cm from the projector lens.

	d	x_w	y_w	z_w	x_p	y_t	x_{error}	y_{error}	x_{wn}	y_{wn}
0	225.6	-299.972	16.0909	864.9099	1579	-860	7.5	1	-292.472	15.0909
40	224.5	-259.013	14.8389	871.0232	1506	-863	7	0.5	-252.013	14.3389
80	223.5	-218.788	15.4483	877.8728	1434	-863	9	1	-209.788	14.4483
120	222.3	-175.094	15.9486	868.138	1363	-865	6	0.5	-169.094	15.4486
160	221.7	-132.917	14.6303	875.0929	1286	-868	7	0	-125.917	14.6303
200	220.5	-93.1524	15.2959	881.2223	1213	-868	8	0	-85.1524	15.2959
240	219.7	-52.7739	15.8452	871.7213	1145	-869	7	0	-45.7739	15.8452
280	219.2	-5.6142	16.2947	878.9027	1057	-870	9	0	3.3858	16.2947
320	218.4	32.46	16.8927	869.1655	991	-871	7	0	39.46	16.8927
360	217.9	74.3256	17.5078	875.5202	911	-871	9	0	83.3256	17.5078
400	217.4	117.3834	17.9361	866.6258	833	-873	7	-0.5	124.3834	18.4361
440	216.9	159.4067	18.5526	872.9649	751	-873	10	-0.5	169.4067	19.0526
480	216.7	200.1728	19.0286	863.8406	675	-875	10	-1	210.1728	20.0286
520	216.4	242.3432	19.6463	870.1646	591	-875	12	-1	254.3432	20.6463
560	215.9	284.7467	20.2671	876.5257	507	-875	15	-1	299.7467	21.2671
600	215.6	324.9627	20.7341	867.391	427	-877	15	-2	339.9627	22.7341
640	215.4	365.6614	21.4094	873.47	345	-877	18	-2	383.6614	23.4094
680	215.1	405.4209	21.8689	864.3785	263	-879	19	-3	424.4209	24.8689

The standard deviation in 'x' and 'y' world coordinates are found to be 10.8 mm and 1.13 mm respectively.

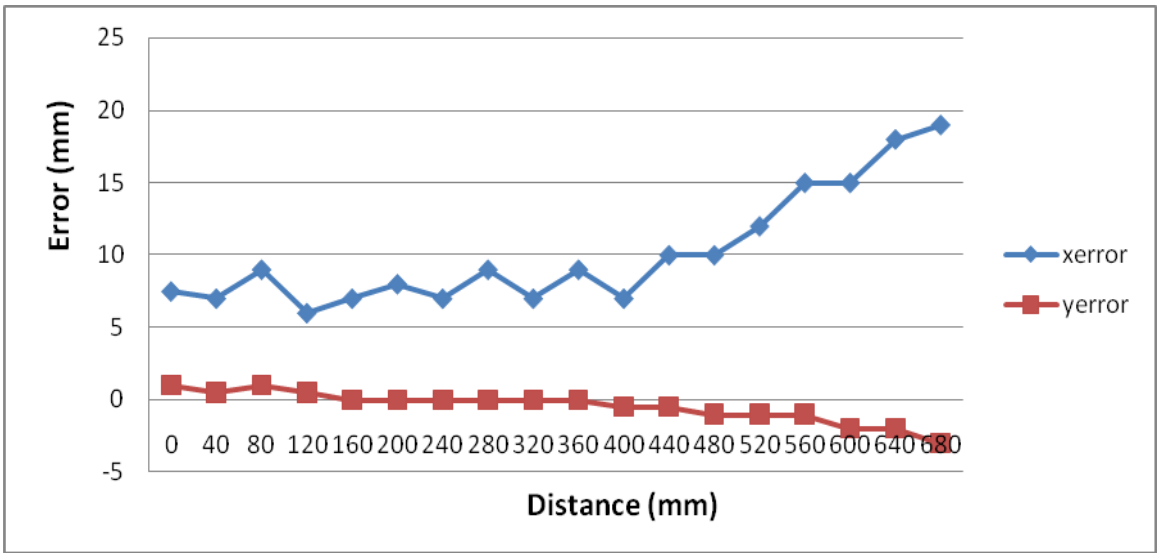

Figure 5.9: Plot of error in '*x*' and '*y*' world coordinates when the target is at a distance of about 228 cm from the projector.

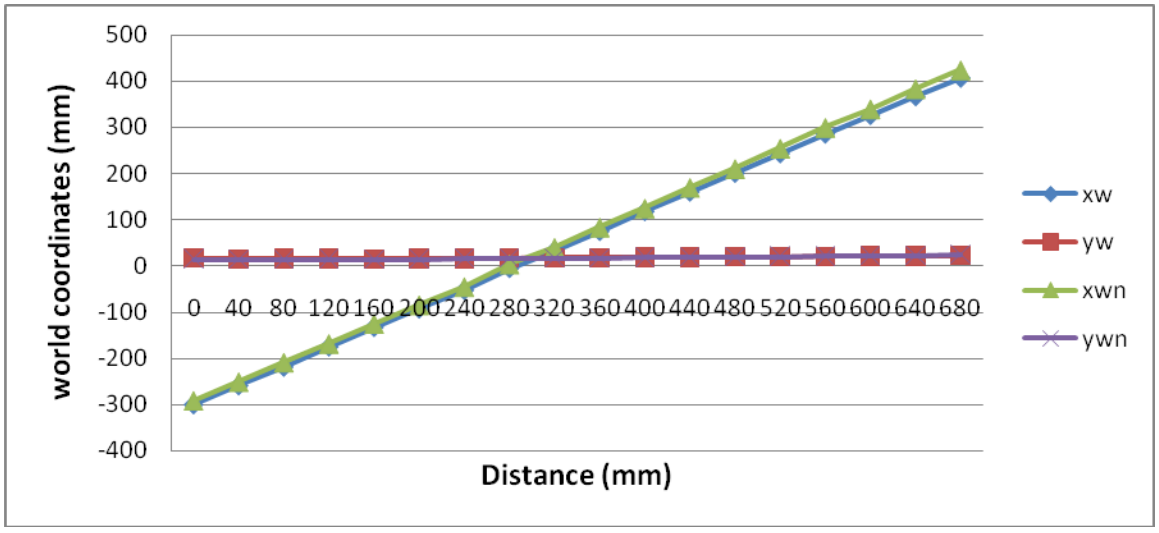

Figure 5.10: Plot of original and new '*x*' and '*y*' world coordinates.

Table 5.6: Error in 'x' and 'y' world coordinates when the target is at a distance of about 246 cm from the projector lens.

	d	x_w	y_w	z_w	x_p	y_t	x_{error}	y_{error}	x_{wn}	y_{wn}
0	244.9	-297.9	10.5	1093.3	1456	-845	6	0	-291.9	10.5
40	243.8	-256.4	11.3	1100.4	1388	-844	7	1	-249.4	10.3
80	243.1	-212.7	10	1107.6	1316	-847	7.5	0	-205.2	10
120	242.1	-171.7	10.8	1115.3	1249	-846	8	1	-163.7	9.8
160	240.8	-129.4	11.5	1102.2	1186	-848	7	0.5	-122.4	11
200	239.5	-83.6	12.1	1091.4	1116	-850	6	0	-77.6	12.1
240	239.5	-43	11	1116.6	1041	-850	10	0	-33	11
280	237.7	-0.4	11.7	1104.4	974	-852	9	0	8.6	11.7
320	237.7	43.1	12.4	1112.5	900	-851	10	0	53.1	12.4
360	237.5	86.1	13.1	1099.7	830	-853	9	0	95.1	13.1
400	236.7	128.7	13.9	1106.8	756	-852	10.5	0	139.2	13.9
440	236.2	170.2	14.6	1094.8	686	-854	12	0	182.2	14.6
480	236.2	214	15.3	1102.9	608	-854	12	0	226	15.3
520	235.7	255.8	16	1090.2	536	-855	12	-0.5	267.8	16.5
560	235.7	298.6	16.8	1097.2	458	-855	15	-1	313.6	17.8
600	235.5	338.9	17.4	1085.3	385	-857	15	-3	353.9	20.4
640	235.5	384	18.1	1092.6	302	-856	20	-4	404	22.1

The standard deviation in 'x' and 'y' world coordinates are found to be 10.9 mm and 1.3 mm respectively.

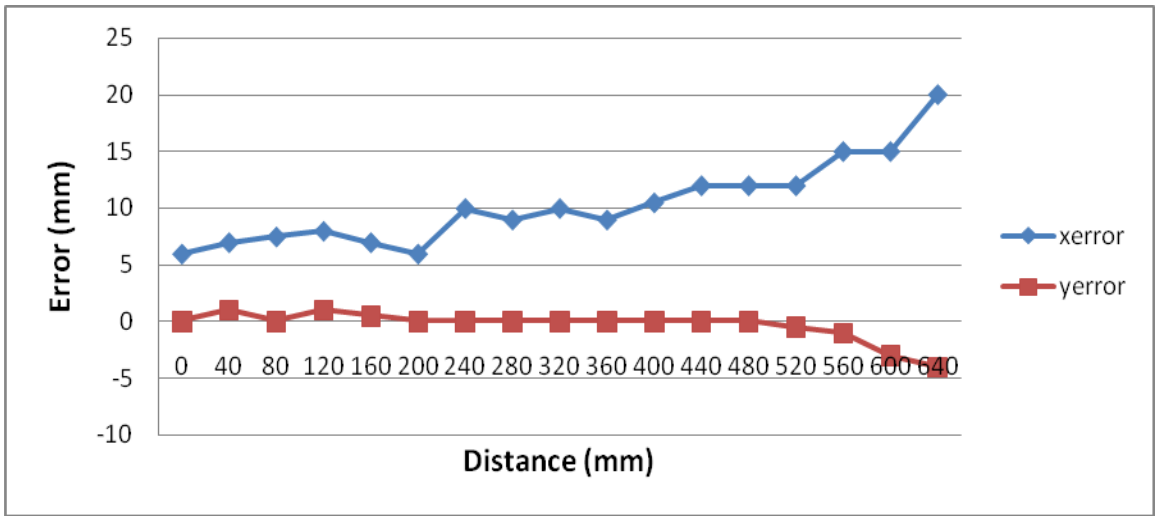

Figure 5.11: Plot of error in '*x*' and '*y*' world coordinates when the target is at a distance of about 246 cm from the projector lens.

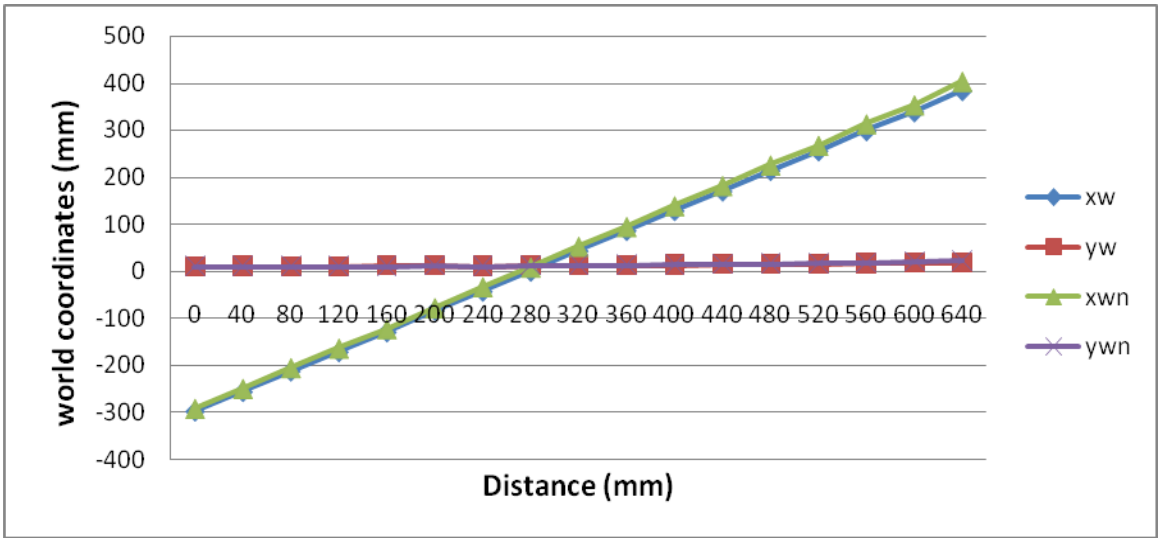

Figure 5.12: Plot of original and new '*x*' and '*y*' world coordinates.

As seen from the data provided in Table 5.1 to Table 5.6 and Figure 5.1 to Figure 5.12, the error in the '*x*' and '*y*' world coordinates of the target are very small and the predicted world coordinates of the target follows the original world coordinates with a very small difference. The standard deviation of the '*x*' world coordinate ranges between 3.2 mm to 10.9 mm and '*y*' world coordinate ranges from 0.52 mm to 1.7 mm. Therefore it can be said that the target is tracked quite accurately when it is moved from right to left.

5.3 Experiment 2.2 results

The Table 5.7 and Figure 5.13 and Figure 5.14 shows the result of the experiment when the target is moved backwards in '*z*' direction.

Table 5.7 Error in '*x*' and '*y*' world coordinates when the target in moved backwards in the '*z*' direction

	d	x_w	y_w	z_w	x_p	y_t	x_{error}	y_{error}	x_{wn}	y_{wn}
0	141	71.3233	39.2436	18.4544	1342	-984	2	0	73.3233	39.2436
10	151.9	71.6939	36.5878	117.1423	1267	-965	2	0	73.6939	36.5878
20	160	71.4265	34.4103	218.3991	1201	-946	3	0	74.4265	34.4103
30	170.2	73.732	32.7774	320.1604	1136	-928	4	1	77.732	31.7774
40	179.7	72.6755	28.1297	435.9011	1080	-917	5	0	77.6755	28.1297
50	195	71.3745	23.9622	548.7093	1033	-906	6	-1	77.3745	24.9622
60	200.5	71.0924	20.6868	665.8011	988	-894	8	0	79.0924	20.6868
70	210.5	69.4624	18.589	785.9785	949	-881	9	0	78.4624	18.589
80	220	65.213	17.7795	874.1772	929	-870	5	0	70.213	17.7795
90	230	68.3	14.9	1006.7	885	-860	9	0	77.3	14.9
100	240	66.5	13.8	1115.7	859	-849	10	0	76.5	13.8

The standard deviation in '*x*' and '*y*' world coordinates are found to be 3.26 mm and 0.58 mm respectively.

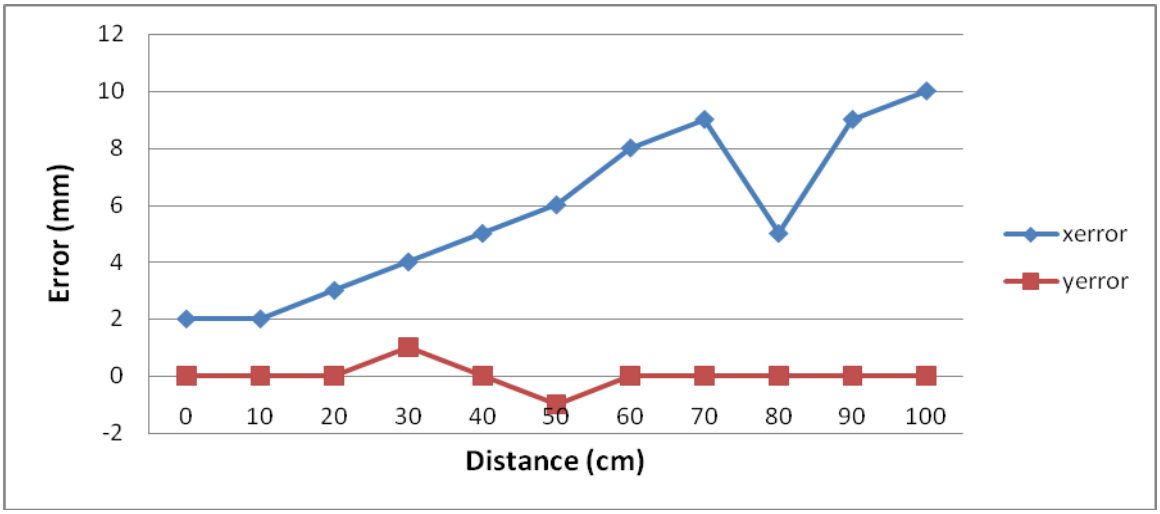

Figure 5.13: Plot of error in '*x*' and '*y*' world coordinates when the target is moved in the '*z*' direction.

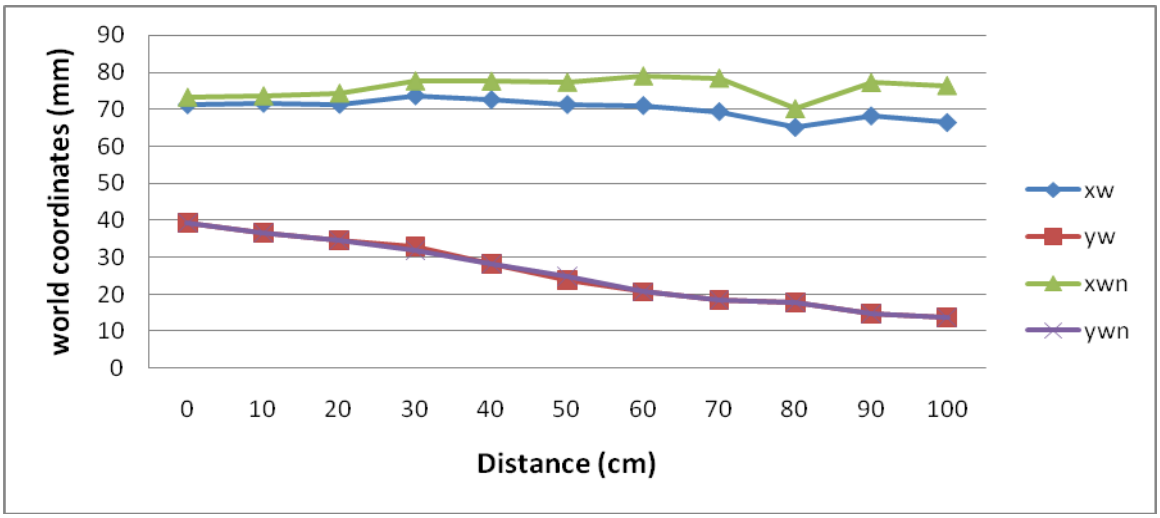

Figure 5.14: Plot of original and new '*x*' and '*y*' world coordinates.

From the data provided in Table 5.7 and the Figure 5.13 and Figure 5.14, it can be clearly seen the error in '*x*' increases as the distance between the target and the projector lens increases. In this experiment it was found that when the distance between the target and the projector lens was about 240 cm, the predicted '*x*' and '*y*' world coordinates of the target was slightly outside the circular target. Therefore it can be inferred from this experiment that, when the target is at a distance more that 250 cm from the projector lens, the error percentage in '*x*' world coordinate of the target increases.

Chapter 6 Conclusion and Future Work

We present a technique for tracking a human iris using two cameras, a pan-tilt control unit and a projector. A new calibration method is used for calibrating the pan-tilt control unit. Experiments were performed to test the system functionality and experimental error measurements are performed to estimate the error introduced by the system.

6.1 Conclusion

The two cameras are initially calibrated using a calibration grid consisting of 18 rings. The calibration gives the camera coordinates and the world coordinates of the calibration grid rings. The pan-tilt unit is also calibrated with the same calibration grid in the same position. A new method of calibrating the pan-tilt unit is used. In this method, the pan-tilt unit is first moved in a direction such that the projector projects a white spot on one of the calibration grid rings. The pan and tilt values for that ring is noted and similarly all the pan-tilt values for all different rings are noted and saved in the calibration program. The calibrated cameras are then used to determine the world coordinate of the projected spot. The method of Least Square Approximation is used to find the '*m*' coefficients of both the cameras and the pan-tilt unit. A Ring Filter is used to correlate the camera image in real time. Peaks are found in the correlated image using peak to side lobe ratio. In this way the target is detected and the camera coordinates of the target are found. The camera coordinates along with the '*m*' coefficients of the cameras are used to find the world coordinates of the target. The world coordinates of the target along with the '*m*'

coefficients of the pan-tilt unit are used to find the pan-tilt coordinates. The pan-tilt coordinates are modeled as a pinhole lens. The pan-tilt unit is moved using the pan-tilt coordinates and the projector points the white spot on the target.

According to the experiments conducted, the error percentage of the system is small when the target is not more than 240 cm from the projector lens. As the distance of the target from the cameras and the projector increases, the accuracy of tracking decreases.

6.2 Future Work

The system can be further improved by improvements in the filter used to correlate the camera image in order to detect the target and get the camera coordinates. In the current method, the radius of the positive target ring and the radius of the negative suppression ring of the ring filter has to be changed as the target moves backwards in order to detect the target perfectly and avoid any false detection. This also increases the processing time considerably.

Wide angle lenses can be used to increase the range of tracking.

The present system can be made more robust by making it an automated system so that experiments can be conducted on an actual human iris and a video of the whole tracking process can be recorded.

A laser set up can be used in place of the projector. In the present system we use separate programs to run the cameras and the projector and this slows up the whole process. By mounting a laser on the pan-tilt unit the tracking process can be made faster.

A digital camera can be employed in order to view the target detection and tracking directly on it. This helps to zoom the target spot and capture the image of the target and later analyze it.

In the present system we found that the error in the '*x*' world coordinate was higher than the error in the '*y*' world coordinate of the target. This error in the '*x*' world coordinate can be decreased by employing a 3 camera setup.

Appendix

A. Automated Calibration of Photogrammetric Grid

Initially the target image (grid) is captured by both the cameras. Down sampling of the image is done to speed up the process. The camera image is then binarized at some threshold level and segmented into a set of blobs. Minimum and maximum blob height/width is used to isolate the blobs from the surroundings.

Figure A, Figure B and Figure C show the actual image, binarized image and image after blob detection from both the cameras.

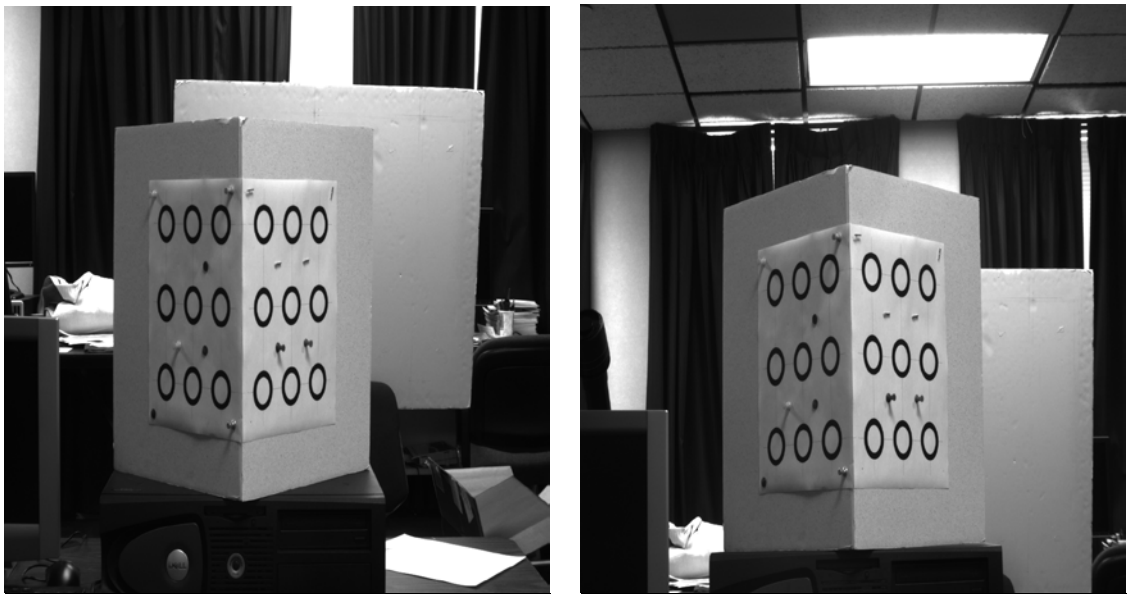

Figure A: Camera A (Top Camera) and Camera B (Bottom Camera) actual image.

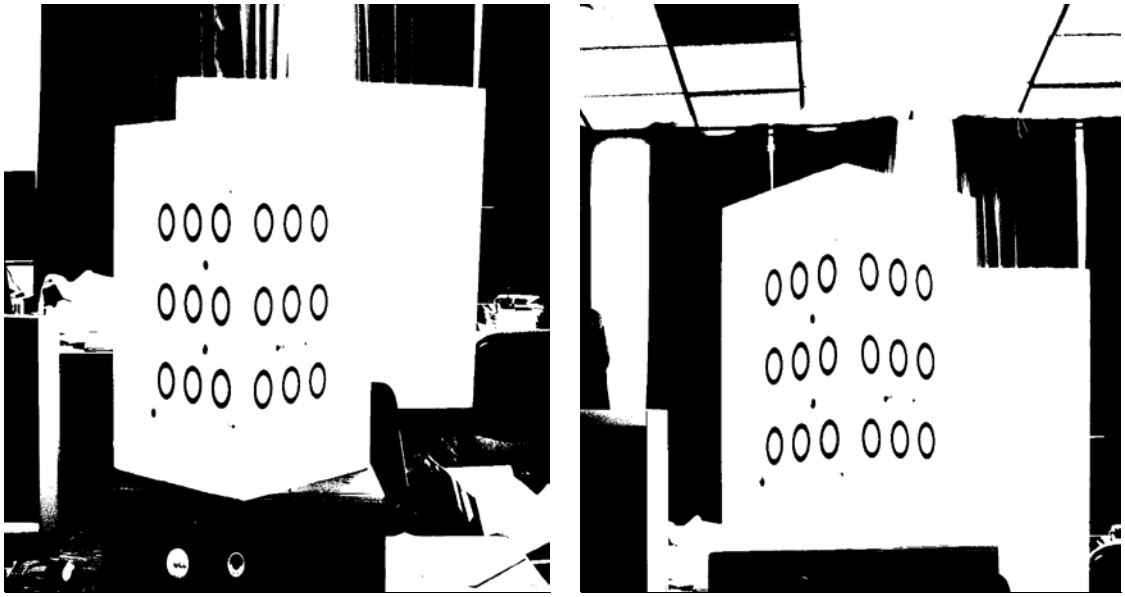

Figure B: Camera A (Top Camera) and Camera B (Bottom Camera) Binarized image.

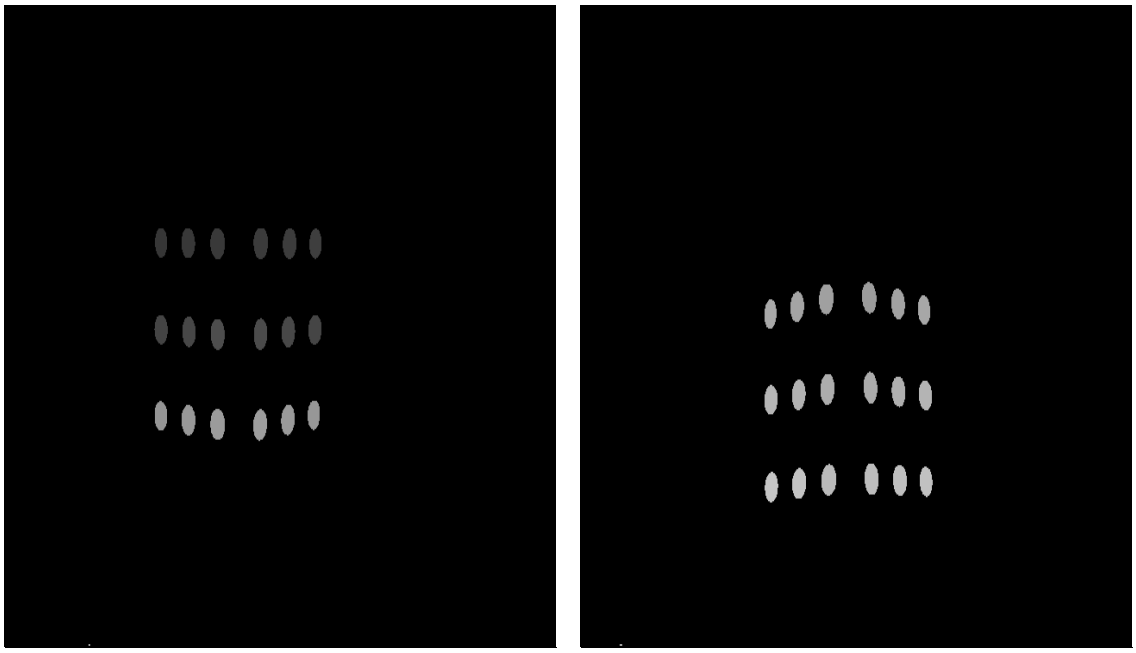

Figure C: Camera A (Top Camera) and Camera A (Bottom Camera) Blob detected Image.

The blobs shown in the Figure C can be processed to determine if their shape is consistent with the target feature shape such as a circle, ring, ellipse and elliptical ring. Given that they are, the dimensions of the shape can be measured along with the centroid of the shape. It is assumed that the exact number of features *N* is known so the *N* best fit shapes are assumed to be the correct ones.

References

[1] A. Vander Lugt, “Signal detection by complex spatial filtering,”IEEE Trans. Inf. Theory 10, 139–145 (1964).

[2] C. F. Hester and D. Casasent, “Multivariant technique for multiclass pattern recognition,” Appl. Opt. 19, 1758–1761 (1980).

[3] A. Mahalanobis, B. V. K. Vijaya Kumar, and D. Casasent, “Minimum average correlation energy filters,” Appl. Opt. 26, 3633–3630 (1987).

[4] P. Re´fre´gier, “Optimal trade-off filters for noise robustness, sharpness of the correlation peak, and Horner efficiency,” Opt. Lett. 16, 829–831 (1991)

[5] A. Mahalanobis, B. V. K. Vijaya Kumar, S. R. F. Sims, and J. F. Epperson, “Unconstrained correlation filters,” Appl. Opt. 33, 3751–3759 (1994).

[6] L. G. Hassebrook, B. V. K. Vijaya Kumar, and L. Hostetler, “Linear phase coefficient composite filter banks for distortion invariant optical pattern recognition,” Opt. Eng., Vol.29, 1033(1990); DOI:10.1117/12.55699

[7] Wei Su and Laurence G. Hassebrook, “Pose and position tracking with super image vector inner products,” Applied Optics, Vol. 45, Issue 31, pp. 8083-8091

[8] Wei Su, Laurence G. Hassebrook and Siddarth Hariharan, “Facial Feature Tracking with the Super Image Vector Inner Product,” Automatic Target Recognition XVII. Edited by Sadjadi, Firooz A. Proceedings of the SPIE, Volume 6566, pp. 656618 (2007)

[9] M. Sanjeev Arulampalam, Simon Maskell, Neil Gordon, and Tim Clapp, “A Tutorial on Particle Filters for Online Nonlinear/Non-Gaussian Bayesian Tracking,” IEEE Trans. Signal Process., vol. 50, no. 2, pp. 174-188, Feb. 2002.

[10] A. Doucet, N. Freitas, and N. Gordon “Sequential Monte Carlo Methods in Practice,”. New York: Springer-Verlag, 2001.

[11] M. Isard and A. Black, “Contour tracking by stochastic propogation of conditional density,” in Proc. Eur. Conf. Comput. Vis., 1996, pp. 343–356.

[12] Dornaika, F.; Davoine, F., "On Appearance Based Face and Facial Action Tracking,"Circuits and Systems for Video Technology, IEEE Transactions on, vol.16, no.9, pp.1107-1124, Sept. 2006

[13] S. B. Gokturk, J. Y. Bouguet, and R. Grzeszczuk, “Adata-driven model for monocular face tracking,” in Proc. IEEE Int. Conf. Comput. Vis., 001, pp. 701–708.

[14] M.L.Cascia, S.Sclaroff, and V.Athitsos, “Fast, reliable head tracking under varying illumination: An approach based on registration of texture-mapped 3-D models,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 22. no. 4, pp. 322=336, Apr. 2000.

[15] Jorgen Ahlberg “An active model for facial feature tracking,” EURASIP J. Appl. Signal Process. vol. 2002, no. 6, pp. 566–571, Jun. 2002.

[16] S. Zhou, R. Chellappa, and B. Mogghaddam, “Adaptive visual tracking and recognition using particle filters,” in Proc. IEEE Int. Conf. Multimedia expo, 2003, pp. 349-352.

[17] Richard I. Hartley, 1994, “Euclidean reconstruction from uncalibrated views,” In ‘Applications of Invariance in Computer Vision’, Mundy, Zisserman and Forsyth (eds.), Lecture Notes in Computer Science, Vol. 825, Springer-Verlag, pp. 237-256

[18] Heyden, A. and Aström, K., 1996 “Euclidean Reconstruction from Constant Intrinsic Parameters,” IEEE CVPR, pp. 339-343

[19] Salvi. J., Armangue. X. and Batlle. J., "A comparative review of camera calibration methods with accuracy evaluation," Pattern Recognition, Vol. 35, pp. 1617-1635.

[20] Fabio Remondino and Clive Fraser, "Digital camera calibration methods: considerations and comparisons".

[21] M. Ito, "Robot vision modeling camera modelling and camera calibration," Adv. Robotics 5 (1991) 321–335

[22] R.K. Lenz, R.Y. Tsai, "Techniques for calibration of the scale factor and image center for high accuracy 3Dmachine vision metrology," IEEE Trans. Pattern Anal.Mach. Intell. 10 (1988) 713-720.

[23] M. Penna, "Camera calibration: a quick and easy way to detection the scale factor," IEEE Trans. Pattern Anal. Mach.Intell. 13 (1991) 1240–1245

[24] Y. Liu, T.S. Huang, O.D. Faugeras, "Determination of camera location from 2-D to 3-D line and point correspondences," IEEE Trans. Pattern Anal. Mach. Intell.12 (1990) 28–37

[25] C.C. Wang, "Extrinsic calibration of a vision sensor mounted on a robot," IEEE Int. J. Robotics Automat. 8(1992) 161–175.

[26] Hall, E.L., Tio, J.B.K., McPherson, C.A. and Sadjadi, F.A., 1982, "Measuring curved surfaces for robot vision," Comput. J. 15 (1982) 42-54.

[27] J. Batista, H. Araujo, A.T. de Almeida, "Iterative multistep explicit camera calibration," IEEE Int. J. Robotics Automat.15 (1999) 897–916

[28] Wei, G. and De Ma, S., 1994, "Implicit and explicit camera calibration: theory and experiments," IEEE Trans. Pattern Anal. Mach. Intell.16 (1994) 469-480.

[29] Trigg, B., 1998, "Autocalibration from planar scenes," ECCV 98, pp. 89-105.

[30] Zhang, Z., 2000, “A flexible new technique for camera calibration,” IEEE Trans. On PAMI, Vol. 22(11), pp. 1330-1334.

[31] Fryer, J. and Brown, D., 1986 “Lens distortion for close-range photogrammetry,” PE&RS, Vol. 52(1), pp.51-58

[32] O.D. Faugeras, G. Toscani, “The calibration problem for stereo,” Proceedings of the IEEE Computer Vision andPattern Recognition, 1986, pp. 15–20.

[33] Roger Y Tsai, “A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-shelf TV Cameras and Lenses,” IEEE Int. Journal Robotics and Automation, Vol. 3(4), pp. 323-344.

Vita

Vikas Chandra Mehta

Born: September 19th, 1985

Hyderabad, Andhra Pradesh, India

Education

B.S. in Electronics and Communications Engineering, Karshak Engineering College, JNTU, 2006

Professional

Research Assistant, University of Kentucky, Signal and Image processing Lab, 2007-2008

Awards

Kentucky Graduate Scholarship, University of Kentucky, 2006-2008

Research Assistantship, University of Kentucky, 2007-2008