1989

Close range three-dimensional position sensing using stereo vision and point-of-correspondence through mathematical morphology.

Mohamed T. Boraie
University of Windsor

NOTICE

AVIS

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylogra phiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Canada

CLOSE RANGE 3-D POSITION SENSING USING STEREO

VISION AND POINT-OF-CORRESPONDENCE

THROUGH MATHEMATICAL MORPHOLOGY

by

Mohamed T. Boraie

A Dissertation
submitted to the
Faculty of Graduate Studies and Research
through the Department of
Electrical Engineering in Partial Fulfillment
of the requirements for the Degree
of Doctor of Philosophy at
the University of Windsor

Windsor, Ontario, Canada

1989

ABSTRACT

In this thesis, a procedure for calibrating 2 CCD cameras for the purpose of close-range position sensing will be introduced. The procedure results in an over-determined set of linear algebraic equations that can be utilized in determining the coordinate position of some identifiable feature in the scene relative to any set of orthogonal axes. The proposed approach relaxes the geometrical constraint required by most of the available procedures. The performance will then be evaluated and both accuracy and repeatability will be examined.

For the purpose of determining the feature position in a multi-feature image, the point-of-correspondence problem will be discussed and a procedure for the point-of correspondence in stereo vision will be introduced. The procedure utilizes mathematical morphology as an image processing tool in finding edge-strength images and in determining the location of the point-of-correspondence.

Finally, introducing lens distortion parameters to the calibration process will be presented and the effect of each term will then be illustrated.

To my father

To my sons

A generation past

and one to come

ACKNOWLEDGEMENTS

I would like to express my sincere thanks and appreciation to my supervisor, Dr. M.A. Sid-Ahmed, for his invaluable advice, guidance and constant encouragement throughout the progress of this thesis. No one could ask for a better advisor.

I extend my sincerest thanks and gratitude to my wife and my parents for their moral support, help and constant encouragement at all times.

Thanks are also due to the management and employees of Clay-Mill Technical Systems Inc. for their support and encouragement during the course of this research.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

Introduction

In recent years, the area of 3-D Machine Vision has been of great interest to both industrial and academic research institutions, with applications in numerous areas such as dimensional measurement, robot guidance, automatic assembly of mechanical parts and electronic components.

3-D machine vision can be broadly classified into two classifications, namely: active vision systems and passive vision systems. In active vision systems (Fig. 1.1), structured light (e.g., laser) highlights the points on the object to be measured.

Several approaches have been developed. A comprehensive review is given by Kak [1] and Jarvais [2] to approaches of retrieving position information using structured light, most of which are based on the triangulation method, with some specific geometrical setup.

Passive systems, on the other hand (Fig. (1.2)), utilize two or more cameras and natural features in the object to determine spatial position. The position information is retrieved from the disparity between the same features on the

1
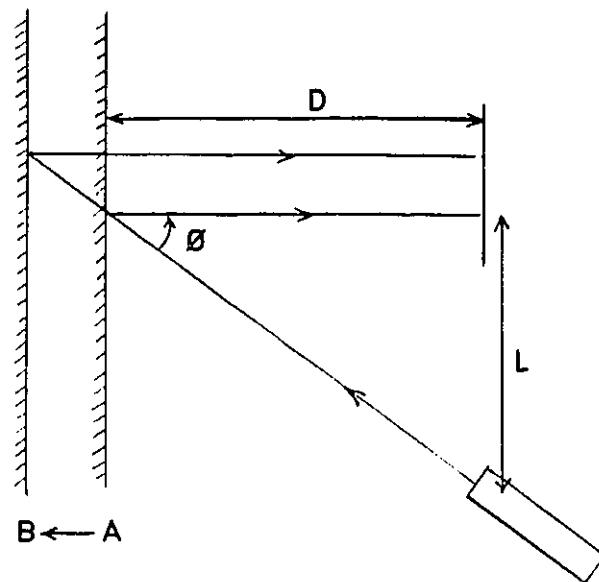
Fig. (1.1). An active vision system.
            If object moves from A to B, both
            D & L increase proportionally.
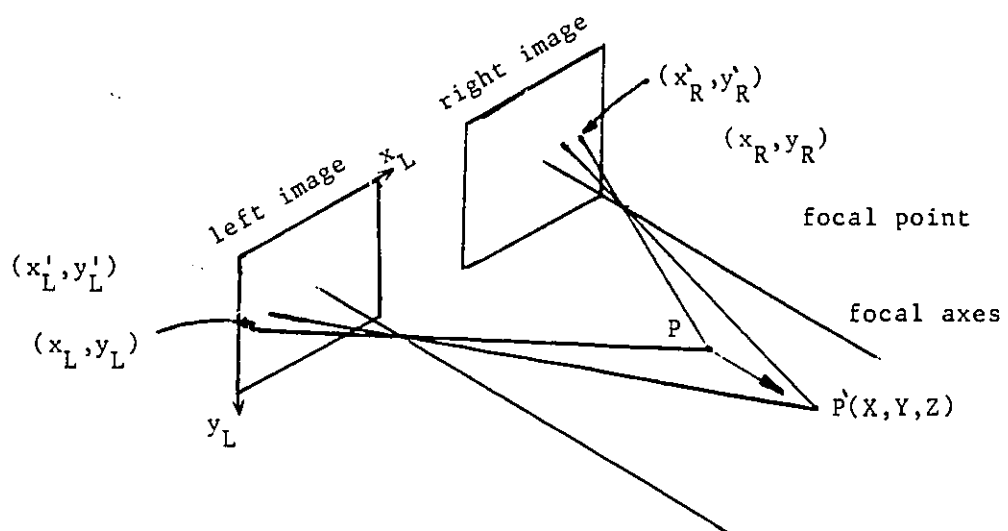
Fig. (1.2). Passive (stereo) vision system.

image planes of the cameras. Systems utilizing two cameras are often referred to as stereo vision. In stereo vision, the major problem is to find a set of matched features between the left and right cameras, known as the point-of-correspondence problem. The relation of the disparity between two matched feature points in the two image planes to a world coordinate system allows the computation of an object position in 3-D dimension. Figure (1.3) illustrates the disparity between the two imaging planes.

In the industrial environment, active systems are far more accepted and utilized than passive systems. This would mainly be due to the complexity of relating the cameras to each other and to the world coordinate system, and also due to the complexity of the point-of-correspondence.

In this thesis we will provide a new solution for the following:

1. Stereo vision calibration for close range position sensing.

2. Point-of-corresponding problem.

We will relax the geometrical constraints set on stereo vision cameras by most researchers in the field (e.g., [3-7]). The approach taken by most researchers can be summarized as:

1. The two cameras have to be co-planer.
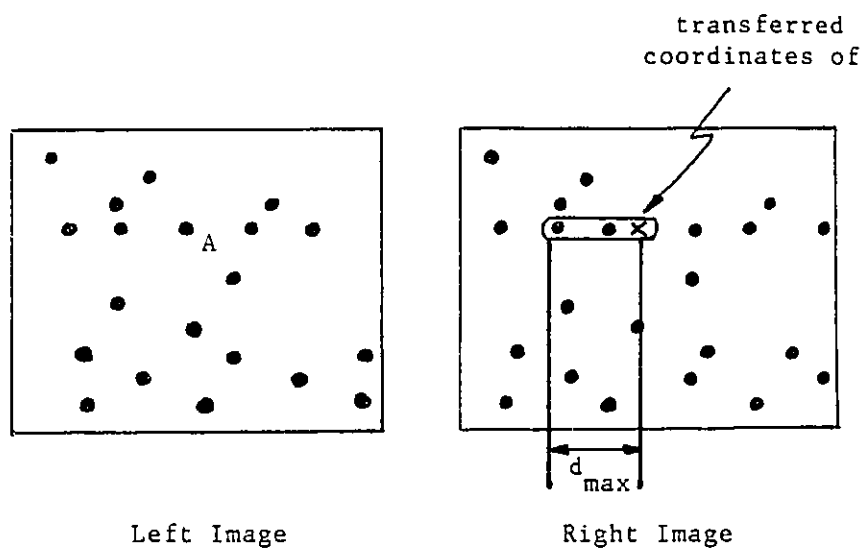
Fig. (1.3). Search space in parallel-axis-method.

2. A point in the left camera corresponds to a point in the right camera along the same horizontal line Fig. (1.4).

3. The neighborhood search for the corresponding point is a segment of a straight line defined by the depth range Fig. (1.4).

Relaxing the constraint of the co-planarity of the two imaging planes and their general positioning relative to each other complicates the problem to the extent that solutions provided in the literature are rendered useless. This situation of no specific geometrical constraints is more appropriate to most industrial problems and is considered in this thesis.

Examples pertinent to industrial problems are provided to illustrate the various solutions suggested.

1.1 3-D Stereo Vision:

Most of the techniques introduced in the literature by machine vision researchers had the same principal of operation. Specific geometrical constraints on the imaging setup are usually utilized, e.g. [3-7] and certain assumptions are made regarding the optical train such as the knowledge of the focal length. The effect of lens distortion in most cases was neglected.

Although these assumptions may be valid in some

applications, they cannot be a generalized approach for all applications where high accuracy measurements are required and the geometrical setup is subject to vibrations from the surrounding environment.

A typical example to this approach, is the method introduced recently by Kim and Aggrawal [3]. In this approach a geometrical setup, as shown in Fig. (1.4), was utilized where the two cameras are mounted on a platform with a fixed base-line and parallel focal axes. The horizontal scan line of both cameras are parallel to the base-line so that all disparities between the two imaging planes are horizontal with the same sign.

The main advantage of this approach, is the simplification of retrieving the 3-D information once a set of matched points in both cameras are found. It also simplifies the point-of-correspondence problem since the search for matches is limited over the same horizontal line and within a limited distance on that line proportional to the maximum allowable disparity (e.g., the maximum and the minimum distance the object can be from the camera reflects a maximum value for the disparity between two images, known as the maximum disparity).

The approach, however, has several drawbacks and disadvantages that will limit its use in industrial applications. First, the approach requires an elaborate amount of work to align the two cameras in order to achieve parallel
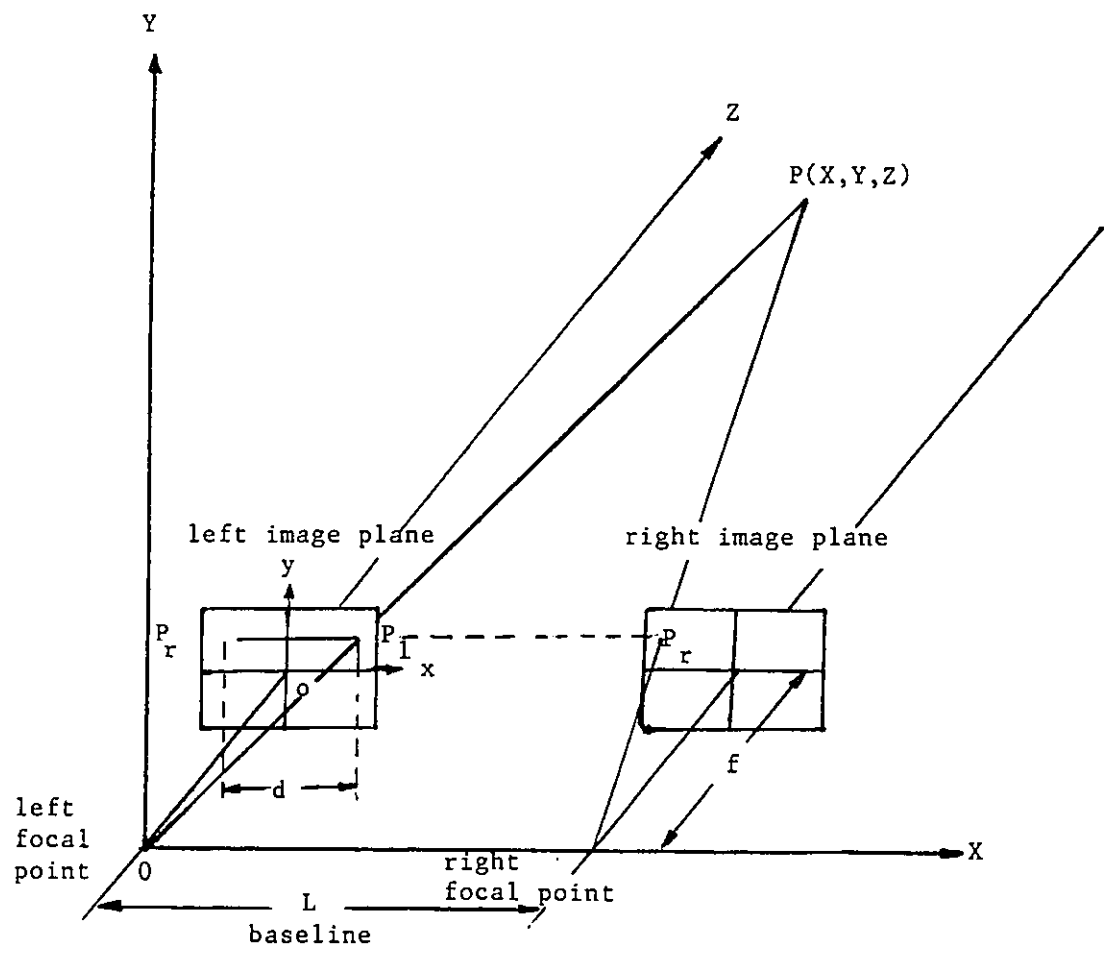
Fig. (1.4). Imaging geometry of parallel-axes method.

optical axes and parallel base-line to horizontal scan lines of the two cameras. Also, it requires highly skilled personnel to carry-out this task. Second, the approach assumes that the focal length of the two cameras are known, which if valid at initial setup makes replacing the lens or the use of close-up lenses rather complicated. Next, in this approach lens distortion is neglected, which (as will be demonstrated later) has a very noticeable effect on the accuracy of the measurements. Also , locating a corresponding point on the same horizontal line might not be valid due to the distortion (radial or tangential) which may dislocate the point from that line.

From the above discussion it is shown that although the approach simplifies the measurements it has major drawbacks.

It was not until recently that machine vision researchers considered camera calibration approaches that have been utilized by photogrammetrists for the purpose of obtaining quantitative information about an object via photographs. In the past, photogrammetrists have used metric cameras, comparators and plotters for their purpose. However the increasing use of computers introduced applications where the most desirable end product of the photogrammetric mapping is not a contour map, but a list of coordinates which define the spatial position of a finite but large number of discrete points. This change initiated the necessity of using non-metric cameras and therefore the development of calibration

techniques for the purpose of close-range photogrammetry and several calibration techniques have been introduced, e.g. , [8-21].

Camera calibration in the context of three dimensional machine vision is the process of determining the internal geometrical and optical characteristics of the camera and lens (intrinsic parameters) and/or the 3-D position and orientation of the camera frame relative to a certain world coordinate system (extrinsic parameters).

The techniques for camera calibration can be classified into three categories:

1. Techniques involving non-linear optimization.

2. Techniques involving only linear equations solving.

3. The two-plane method.

A typical example of techniques involving non-linear optimization is the approach introduced by Faig [10], where an analytical method for self calibration of non-metric close-range cameras was introduced. The approach uses a very elaborate model for imaging, where 17 unknowns are used as parameters to be optimized through a step by step iterative procedure. The reported results indicate high accuracy on an average basis. However, the large number of unknowns requires a considerable amount of CPU time and a good initial guess to ensure convergence. A professional format photo (Nikomat-FT 35 mm camera with a 50 mm Nikkor-4 lens) was used

for acquiring the photographs used for evaluation, which is a factor in the high accuracy reported compared to a CCD camera mounted with a relatively inexpensive lens. Sobel [13], introduced a calibration approach with the objective of automatic calibration of camera and robot for hand-eye project. The approach optimized 18 unknowns and utilized two optimization algorithms. One algorithm is used to provide fast convergence, the second is then applied to improve the results. The search was shown to be complicated due to the fact that a subset of the parameters are interdependent and that empirically it has been determined that there are multiple minimas of the cost function. Therefore, it is extremely important to obtain a good initial guess in order not to go astray. No detailed results were given to the measurement accuracy obtained by this approach.

Several number of approaches under the same category were introduced in the literature [8]-[13]. The general drawback of this category as was illustrated in the discussion of the previous two approaches can be summarized in the following points.

1. A good initial guess is required for the optimization parameters.
2. Computationally expensive, it is , therefore, not feasible to implement on small size computer suitable for industrial applications.

The second category techniques involve computing a perspective transformation matrix using linear equation solving [14], [15], [16]. A typical approach is the DLT introduced by Abdel-Aziz and Karara [14]. The approach relates the 3-D world coordinate to 2-D image coordinates through 3 x 4 perspective transformation matrix with the parameters of the matrix being the unknown parameters for the calibration. The approach utilizes the fact that although the equations characterizing the transformation from 3-D world coordinate to 2-D camera coordinate are non-linear, neglecting the lens distortion will result in a set of linear equations. The approach was implemented by Dainis and Juberts [16] to perform accurate measurement of robot trajectory. The accuracy reported was comparable to those obtained utilizing other approaches. However, several investigators showed that ignoring lens distortion has a great effect on the accuracy when performing 3-D measurements, e.g., Itoh et al. [17], Luh and Klassen [18]. The effect of lens distortion on the accuracy will be further investigated in this thesis. The effect of different terms of distortion on the accuracy of 3-D measurements will be illustrated.

The third category is the two-plane method developed by Marins et al. [20]. The approach utilized an empirically based formula for the transformation between image and object coordinates. Lens distortion was also neglected. Isaguirre

et al. [21] later generalized the two-plane technique for camera calibration purposes where a full scale non-linear optimization was used.

Recently, Tsai [19] introduced a "Two-stage" calibration technique. In his approach, the first stage determines the 3-D orientation, x and y positions by solving a set of over-determined system of linear equations. In the second stage, the focal length, lens distortion and the z position are computed. This is also done in two steps. First, to find an approximate value for the focal length f and z position by assuming no lens distortion, then solving a set of linear equations. The approximate values are then used as an initial guess to solve for the exact z,x,K , where K is the first coefficient in the radial distortion equation, using standard optimization scheme (such as Steepest Descent). Figure (1.5) shows the four steps transformation for Tsai's approach.

The two stage approach involves computing the exact value for the parameters by solving two sets of linear equations and then using non-linear optimization. It also assumes that only one term of radial distortion is sufficient and that tangential distortion can be ignored. In this thesis, a more simplified calibration approach will be introduced in which both types of lens distortion (tangential and radial) are considered.

(x ,y ,z) 3-D world coordinate

Step 1
Rig.d body transformation from (x ,y ,z) to (x,y,z)
Parameters to be calibrated: R,T

(x,y,z) 3-D camera coordinate system

Step 2
Perspective projection with pin hole geometry
Parameters to be calibrated: f

(X ,Y) Ideal undistorted image coordinate

Step 3
Radial lens distortion
Parameters to be calibrated: k ,k

(X ,Y) Distorted image coordinate

Step 4
TV scanning,Sampling,computer acquisition
Parameter to be calibrated: uncertainty factor s for image
X coordinate

(X ,Y) Computer image coordinate in frame memory

Fig. (1.5). Tsai's four steps transformation
from 3-D world coordinate to
computer coordinate.

1.2   Point-of-Correspondence

In stereo vision, the major problem is to find a set of matched features between the left and the right images, known as the point-of-correspondence problem. The studies addressing this problem can be classified into two classes.

One class of studies addresses the psychophysical and biological aspects of human visual system, Marr and Poggio [22-23] proposed a feature-point based model of aspects of human stereopsis. Grimson [14] implemented the algorithm and tested it on natural images and on the random dot stereograms introduced earlier by Julesz [25]. A number of additional psychophysical predictions of the model have been tested, and consequently several modifications and improvements have been introduced [26-28].

The second class of studies examines computational aspects in various application domains. Arnold [29] and Arnold and Binford [4] used an ad hoc voting scheme based on local context to match edges in aerial photographs. Barker [30] introduced an edge based approach and scan line correlation. An intensity correlation scheme was added later by Barker and Binford [31] to fill the gap between matched edges.

Recently, Kim et al. [3] introduced 16 zero-crossing patterns based on the local connectivity, Fig. (1.6). These patterns were used with a relaxation method based on continuity of local disparities for matching. The algorithm
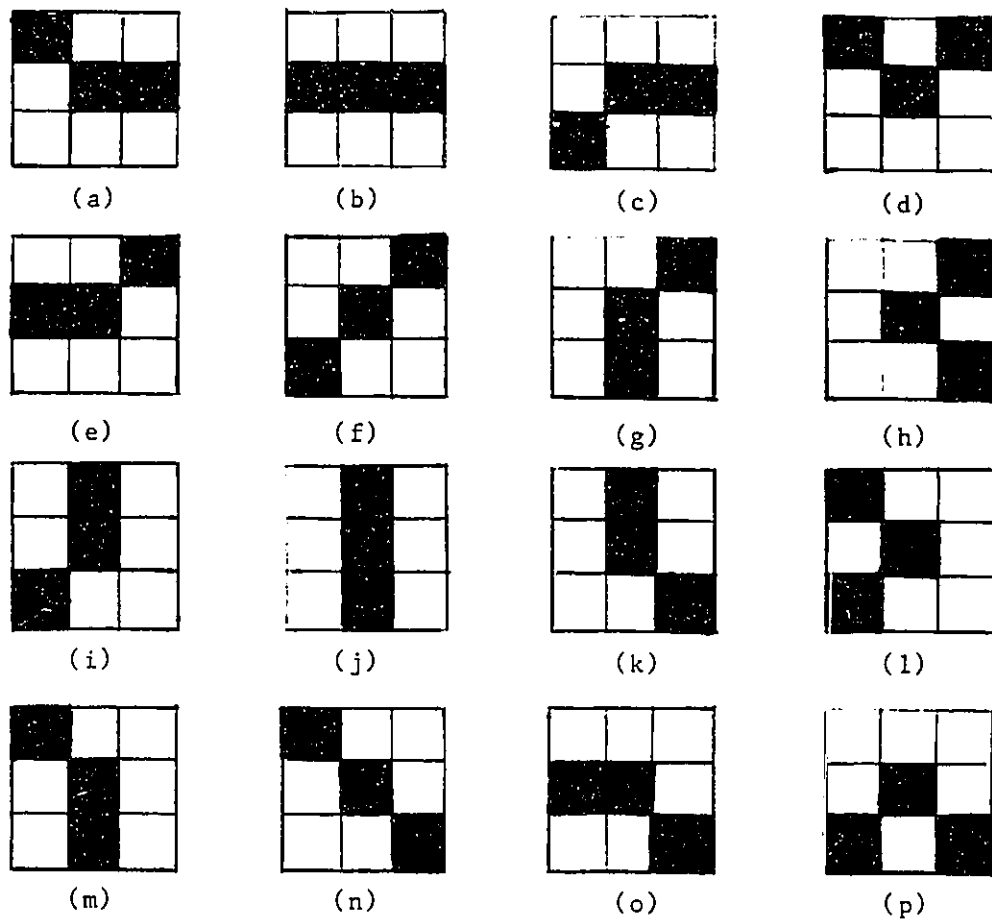
Fig. (1.6). Zero-crossing patterns.

was further simplified by constraining the imaging geometry. Similar to most of the techniques dealing with the point-of-correspondence, the approach utilized images from co-planar cameras. Therefore, if the two cameras are at an angle to each other none of the above techniques would have satisfactory performance. Since the zero-crossing connectivity or the edges do not have the same projection on the imaging planes. This problem is addressed in this thesis and a new approach for solving this problem will be introduced. The approach utilizes mathematical morphology operators and some of the recent work reported on mathematical morphology application in image analysis.

1.3 Mathematical Morphology

Mathematical morphology provides an approach to the processing of images which is based on the set theoretic concepts of shape.

It was formalized at the Ecole de Mines in Paris in the mid 1970's by Matheron [32] and later extended by Serra [33] and Sternberg [34-35]. It has grown to envelop a variety of applications and hardware. It has been widely used in industrial applications with computer architecture specialized to perform morphological operations, McCubbrey and Lougheed [36] proposed a raster pipeline processor for morphological image analysis. Crimmins and Brown [37] reviewed the special purpose theory, software and hardware

for automatic recognition that have been under development at ERIM (Environmental Research Institute of Michigan). It was also illustrated that shape recognition can be carried out with set theory operations by treating images and shapes as point sets in n-dimensional cartesian space.

Several machines which perform morphological operations are currently available such as Golay logic processor [38], and Diff3 [39]. 1987, Haralick et al. [40] indicated that although the techniques are being used in the industrial world, the basis and the theory of mathematical morphology are not covered in the literature. Henceforth, they provided a tutorial that reviewed both binary morphology and gray-scale morphology, covering the operations of dilation, erosion, opening, closing and their relations in an attempt to attract the attention of the academic researchers to this area and its importance as an image processing tool.

A limited number of applications for morphological operators has been reported in the literature. However, the interest is increasing and applications are being introduced. Peleg, Naor, Hartly and Avnir [41] introduced an approach for analysis and classification on multiple resolution texture. Crimmins and Brown [38] discussed an approach for automatic shape recognition. Lee, Haralick and Shappiro [42] introduced edge detection operators using mathematical morphology and compared the results to that obtained by applying the enhancement/thresholding edge detector [43] and

the cubic facet second derivative zero-crossing edge operator [44]. They reported that, as noise increases, the blur-minimum morphological edge operator has a comparable result to that obtained by the second-derivative zero-crossing edge detector and that both have much better performance than the rest of the operators tested.

In industrial applications involving objects or defects identification, mathematical morphology, when appropriately used, proves to be a useful tool for image processing that tends to simplify image data preserving their essential shape characteristics and eliminating irrelevancies. Since the identification of objects, object features and assembly defects correlate directly with shape, then mathematical morphology becomes the natural apparent approach to deal with industrial problems such as machine vision recognition and robot guidance. In this thesis mathematical morphology will be used for dealing with the matching process of the point-of-correspondence problem and also in obtaining the edge strength images.

1.4 Organization of the Thesis

The thesis is organized into 7 chapters. The first chapter, typically, is the introduction where some of the previous research work in related areas to this thesis is reviewed and the objectives of this thesis are outlined. The chapter is concluded with the outline of thesis organization.

Chapter 2 describes our proposed camera calibration procedure for the purpose of close range position sensing. The procedure relates camera coordinate systems to an orthogonal system through the calibration parameters. Using 2 CCD cameras calibrated to the same orthogonal axes system the 3-D position information of an object in view of both cameras will be retrieved.

Chapter 3 deals with performance evaluation of the proposed technique. Accuracy and repeatability of each of three axes is reported. A test procedure proposed by the Automated Vision Association is outlined and adhered to when the performance evaluation test performed.

Chapter 4 covers some of the fundamentals and mathematical concepts of mathematical morphology. Morphological operator, their residues and some of their applications to edge detection and automatic shape recognition are reviewed. Several illustrated examples of morphological applications to both binary and gray-scale images are given. This chapter provides a background for the definitions and operators of mathematical morphology to be used in chapter 5.

Chapter 5 deals with the point-of-correspondence problem in stereo images. A procedure for determining the point-of-correspondence is introduced. The procedure utilizes the calibration parameters and morphological operators. Several examples illustrating the procedure are presented.

Chapter 6 extends the calibration procedure to include lens distortion parameters. Both radial and tangential distortion are discussed and their effect on the accuracy of the measurements is evaluated.

Chapter 7 is a conclusion and discussion of this thesis.

CHAPTER II

CAMERA CALIBRATION

Abstract

In this chapter a procedure for calibrating 2 CCD cameras to a set of orthogonal axes for the purpose of close range position sensing will be introduced. The procedure involves solving an over-determined set of linear algebraic equations. The parameters obtained relate the camera coordinate system to the orthogonal system to which the camera was calibrated. These parameters are then utilized to determine the position of an object in view of both cameras.

2.1 Introduction

Several types of camera calibration techniques have been suggested in the literature. The mathematical formulation of calibration procedures have been introduced for non-metric cameras used by photogrammetrists for distance measurements from photographic data [8-15]. However, it is only recently that machine vision researchers became interested in

22

investigating some of the approaches developed for photogrammetry to be utilized for machine vision application [19].

In this chapter, a development of a camera calibration procedure will be introduced, based on the method of Direct Linear Transform (D.L.T.) introduced by Abdel-Aziz and Karara [14] with the objective of developing a passive position finding system using off-the-shelf cameras, lenses and computer hardware to be realized in an industrial application.

The calibration parameters obtained through the calibration procedure are to be used to determine the 3-Dimensional position of the object in the field of view of both cameras to the world coordinate system to which the cameras have been calibrated. The accuracy of the calibration procedure is evaluated by calculating the 3-dimension (3-D) information of the points used for the calibration.

The accuracy and repeatability of the system will be tested and discussed in the next chapter.

The procedure does not require knowledge of the focal length of the cameras, the distance between them or their relative angle of orientation.

2.2 Development of the Camera Calibration Equations

Camera calibration is defined as the process of determining the camera's extrinsic and intrinsic parameters. Extrinsic parameters are the 3-D relation of the camera frame in relation to a world coordinate system. Intrinsic parameters on the other hand are the internal camera parameters and optical train such as lens distortion coefficients. The following section will develop the equations dealing with the extrinsic parameters.

Let the position of a known target point be given in homogeneous coordinates $P = (X,Y,Z)$ measured from any set of coordinate axes, and its position in the imaging plane of the camera 1 be given by (x',y') as shown in Fig. (2.1). The 3-D matrix transformation, used commonly in computer graphics theory [45], can be utilized to relate (x',y') to (X,Y,Z) as

$$\begin{bmatrix} x'w_1 \\ y'w_1 \\ w_1 \end{bmatrix} = \begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 \\ \alpha_5 & \alpha_6 & \alpha_7 & \alpha_8 \\ \alpha_9 & \alpha_{10} & \alpha_{11} & \alpha_{12} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \tag{2.1}$$

where (x',y') is the camera imaging plane coordinate measured relative to the center of the plane.

The last parameter α_{12} in the transformation matrix is
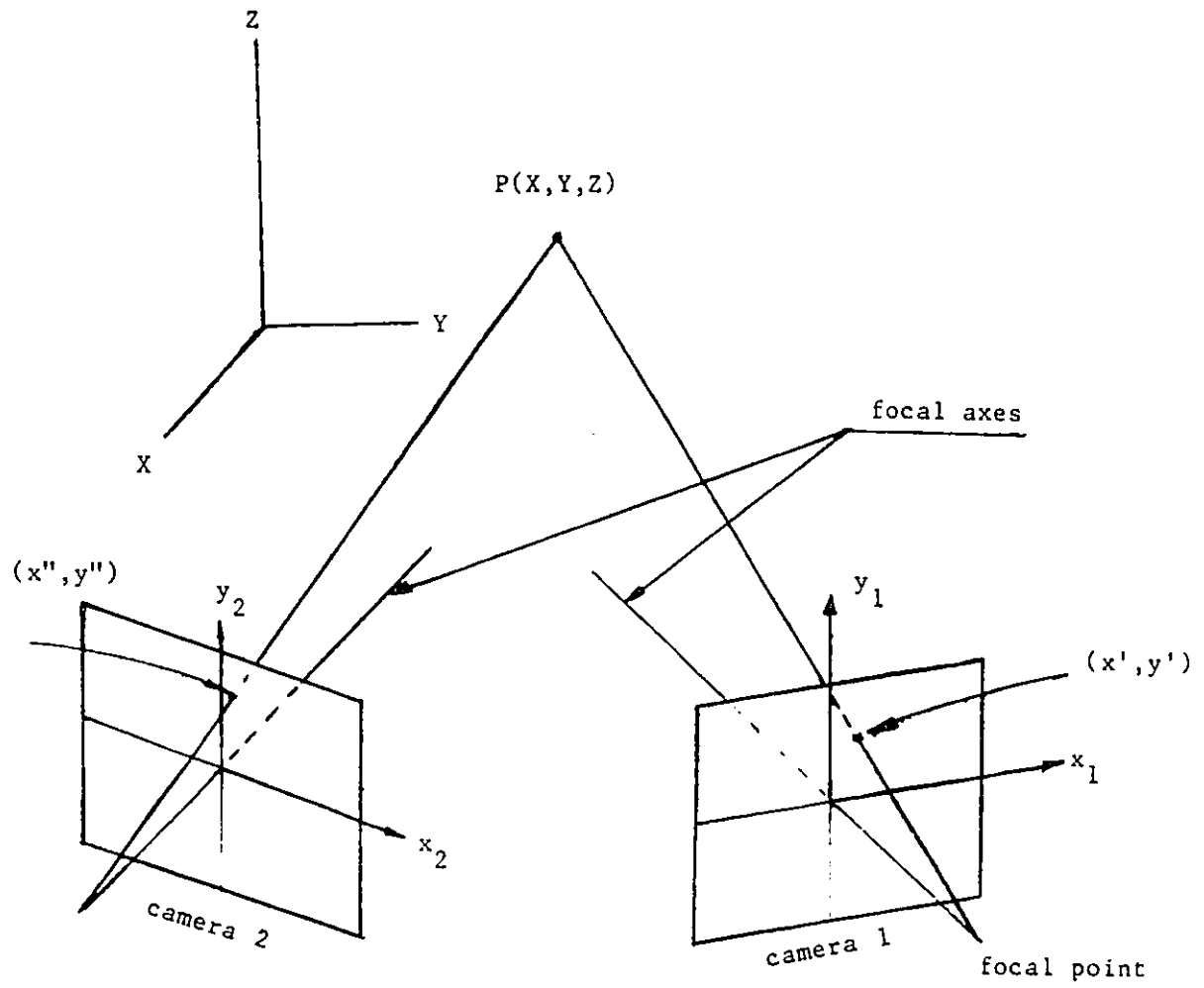
Fig. (2.1).   Projection of the point P(X,Y,Z).

set to 1 with no loss of generality since it can be considered in normalizing the 3 equations. Since, the parameter α_{12} which represents a scaling factor between the world coordinate and the imaging coordinate system [45] and, therefore, can never be equal to zero.

Now, rearranging equation (2.1) in the form

$$A\alpha = B \qquad (2.2)$$

we can write

$$\begin{bmatrix} X & Y & Z & 1 & 0 & 0 & 0 & 0 & -x'X & -x'Z \\ \\ 0 & 0 & 0 & 0 & X & Y & Z & 1 & -y'X & -y'Y \end{bmatrix} \alpha = \begin{bmatrix} x' \\ \\ y' \end{bmatrix} \qquad (2.3)$$

where

$$\alpha = \begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 \cdots \alpha_{11} \end{bmatrix}^T$$

The calibration in this form is to determine the values of the parameters α_i (i=1,...11) that satisfy these 2 equations. In order to solve the equation represented in (2.2) a minimum of six measured points (X_j, Y_j, Z_j) in the world coordinate relative to any desired axes system are required. Their corresponding position on the cameras imaging plane (x'_j, y'_j) j=1, N and N≥6 (usually more than 6 points are considered in order to compensate for small errors in the measurements) are obtained. Substituting these values into Eq. (2.3) results in the development of an over-determined set of linear algebraic equations in the form of

$$A \alpha = B \qquad (2.4)$$

The equation in (2.4) can be solved using the method of singular value decomposition [46] for a least-squares or by solving the set of linear equations given by

$$A^T A \alpha = A^T B \qquad (2.5)$$

which can be solved (using any of the available procedures, for example, Crouts method [47] (that was utilized in our implementation). Once, the parameters are obtained the camera coordinate system is related to the coordinate system through the transformation matrix of Eq. (2.1) and recalibration is not required as long as the camera's position in relation to the world coordinate system has not been changed.

2.3 Procedure for the Calibration and Data Collection

In the previous section the theoretical formulation of the equation for the calibration was established. In this section the implementation, the procedure followed and the results of the calibration will be discussed.

The accuracy of the calibration system in 3-D measurements depends on the accuracy of the target used for calibration and pre-processing algorithms.

A white target with 9 black discs with 5 mm equal spacings in both x and y direction was developed on a CAD system and a high precision plotter to plot the discs on mylar paper. A three axes table with 3-dial gage indicators mounted to the fixture for measurements of displacement (accuracy of dial gages is 0.0005 inch) with a resolution of 0.001 inch on the dial scale as shown in Fig. (2.2).

1. The image is first thresholded using the algorithm described in [48]. This basically yields a binary image from an 8 bit gray-level image (the image used have a resolution 256 gray-levels and is digitized to 492 x 510 pixels). This step is required to separate the objects (disks) from the white background. The thresholding selection algorithm is basically described by the iterative equation:

$$T_{\ell+1} = \frac{\sum\limits_{i=0}^{T_R} n_i\, i}{\sum\limits_{i=0}^{T_\ell} n_i} + \frac{\sum\limits_{i=T_{\ell+1}}^{255} n_i\, i}{\sum\limits_{i=T_{\ell+1}}^{255} n_i} \qquad (2.6)$$

The solution is obtained when $T_{\ell+1} = T_\ell$.

2. A border following algorithm utilizing the 8 points neighborhood described in [49] is then used to
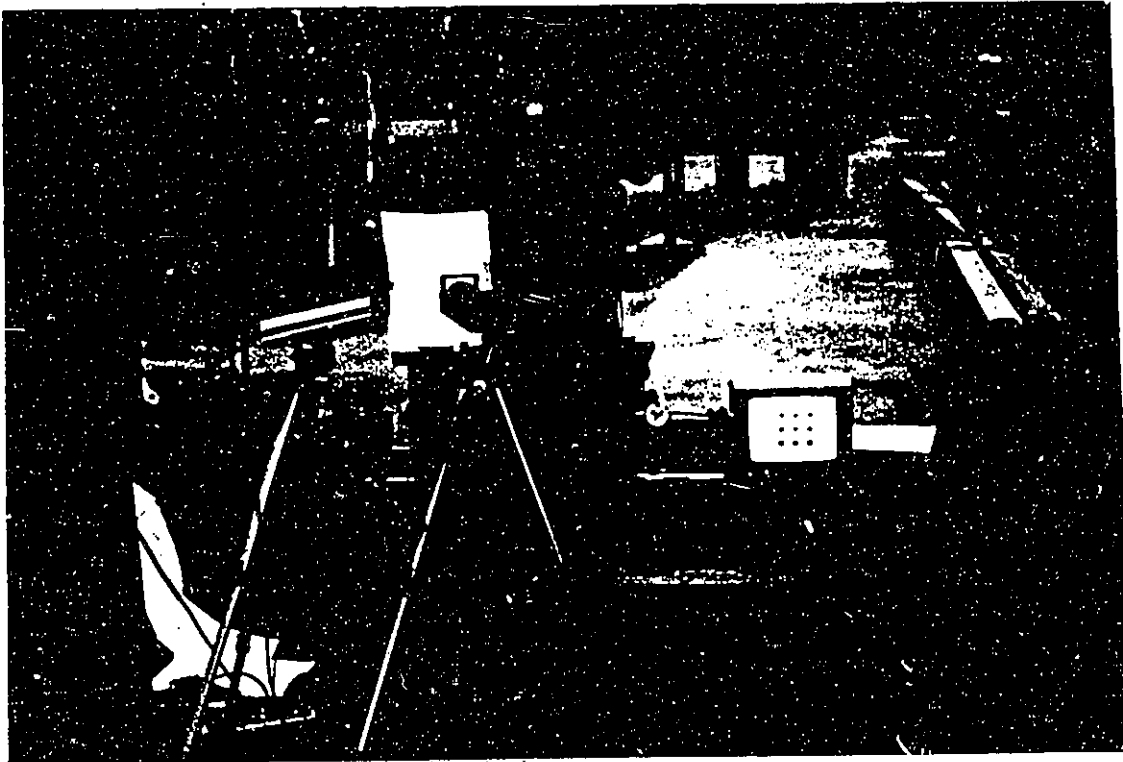
Fig. (2.2). Calibration and accuracy test set-up.

determine the border points of the disk.

3. From the border point the centroid location of the disk is obtained using the formula

$$c_x = \frac{\sum\limits_{i=1}^{N} x_i}{N}$$

$$c_y = \frac{\sum\limits_{i=1}^{N} y_i}{N} \qquad (2.7)$$

where N is the total number of border points.

4. The corresponding (X,Y,Z) location of the disk is given to the system.

5. Repeat the previous steps and construct a set of over-determined equation on the form

$$AX = B \qquad (2.8)$$

6. The set is then solved to determine the calibration parameters α's.

Figure (2.3) shows a flow-chart of the procedure, Fig. (2.4) shows typical values for calibration parameters obtained by following the above procedure and with 27 points.

Position target in
view of both
cameras

↓

Find centroid location
in each camera
coordinate system

(x'_i, y'_i), (x''_i, y''_i) $i = 1 \rightarrow 9$

↓

Enter the corresponding
X,Y,Z location of each disk
with respect to world
co-ordinate

↓

Construct a set of over-determined
linear equations as shown in
Equations 2.3 and 2.4

↓

Obtain the least square solution,
i.e., find α's
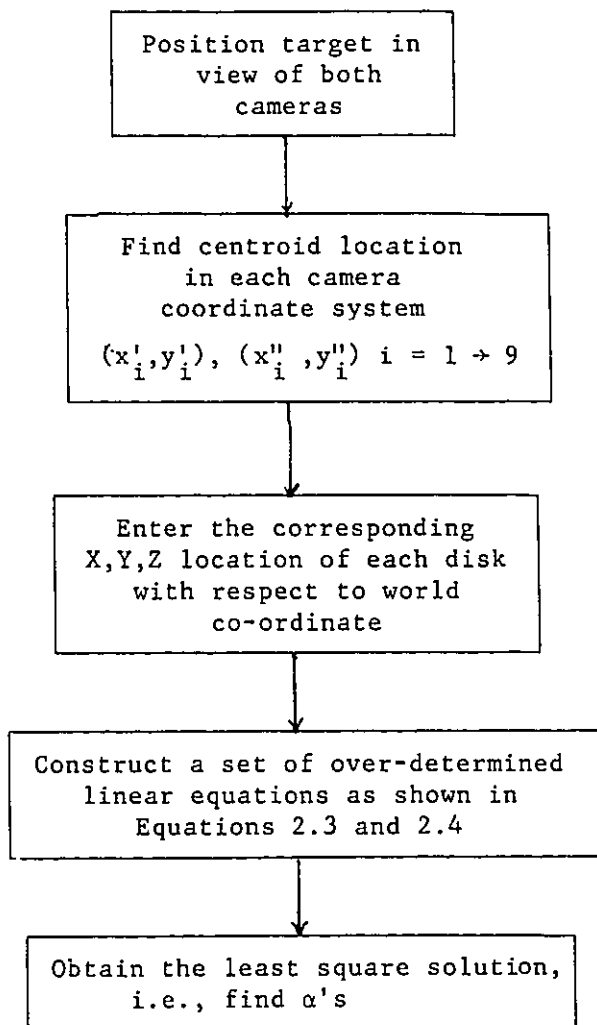
Fig. (2.3). Calibration procedure
flow chart

CAMERA 1   CAMERA 2

14.0091698401741400000000   13.8359876624018500000000
 -.3545243177730287000000   -.3802920905352070000000
-5.7236234274342670000000   5.3860936230904100000000
-31.6958465472026700000000   45.3257343549703900000000
 .0746646769533366900000   1.0099964542490380000000
18.1738469545501600000000   17.7548585541233700000000
-1.2880377150461280000000   -1.1408262534055380000000
-7.3244415747283100000000   6.7613431397991500000000
 .0000632694500682916700   .0006964065332755356000
 -.0006395128343810976000   -.0007957127392169839000
 -.0013349494728861380000   -.0014603486367017270000

Fig. (2.4). Shows a typical values for the calibration parameter
of Camera 1 and Camera 2.

2.4 3-Dimension Position Finding

Once the camera has been calibrated and the parameters $\alpha_i's;\ i=1,11$ are determined, the equation given in (2.1) can be re-arranged as

$$\begin{bmatrix} (\alpha_1-\alpha_2 x')(\alpha_2-\alpha_{10} x')(\alpha_3-\alpha_{11} x') \\ (\alpha_5-\alpha_9 y')(\alpha_6-\alpha_{10} y')(\alpha_7-\alpha_{11} y') \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} x'-\alpha_4 \\ y'-\alpha_8 \end{bmatrix} \qquad (2.9)$$

For a point on the target (X,Y,Z) the imaging (x',y') on the camera plane is measured and substituted in the above two equations yielding an under-determined set of equations (two equations in three unknowns).

In order to complete the above set of equations a second camera is required. Calibrating the second camera and obtaining its transformation matrix

$$\begin{bmatrix} x''w_2 \\ y''w_2 \\ w_2 \end{bmatrix} = \begin{bmatrix} \beta_1 & \beta_2 & \beta_3 & \beta_4 \\ \beta_5 & \beta_6 & \beta_7 & \beta_8 \\ \beta_9 & \beta_{10} & \beta_{11} & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \qquad (2.10)$$

Rearranging the equation as for camera 1 we obtain

$$
\begin{bmatrix}
(\beta_1-\beta_9 x'') (\beta_2-\beta_{10} x'') (\beta_2 -\beta_{11} x'') \\
\\
(\beta_5-\beta_9 y'') (\beta_6-\beta_{10} y'') (\beta_7 -\beta_{11} y'')
\end{bmatrix}
\begin{bmatrix}
X \\
Y \\
Z
\end{bmatrix}
=
\begin{bmatrix}
x''-\beta_4 \\
\\
y''-\beta_8
\end{bmatrix}
\quad (2.11)
$$

The four linear equations formulated in equations (2.9) and (2.11) represent an over-determined set of equations which can be solved in the least square sense (similar to the procedure followed in the calibration steps) to obtain the X,Y,Z of some unknown position of an identifiable feature on the target which provided the imaging points (x',y') and (x'',y'').

2.5 Calibration Results

Several calibration experiments using 18, 27 and 54 points were performed. The calibration parameters obtained were then used with the (x',y') and (x'',y'') that were utilized to determine these parameters to recalculate the position X,Y,Z of each of the calibration points and comparing the actual position to the calculated position of each point. Figures (2.5), (2.7) and (2.9) show the results of each case, whereas Figs. (2.6), (2.8), (2.10) show the maximum error between the original and the calculated values for each of the three examples for each of the X,Y,Z axes.

Viewing region in mm: 12.70 x 12.70 x 15.24
No. of Calibration points used: 18
Target: Dots on a flat surface

Original Position (mm)			Calculated Position (mm)		
Xo	Yo	Zo	X	Y	Z
-6.35000	6.35000	-7.62000	-6.39604	6.33245	-7.60733
.00000	6.35000	-7.62000	-.01630	6.35763	-7.56869
6.35000	6.35000	-7.62000	6.31797	6.35456	-7.58155
-6.35000	.00000	-7.62000	-6.33837	.00999	-7.55824
.00000	.00000	-7.62000	.11439	-.00152	-7.61407
6.35000	.00000	-7.62000	6.40791	-.01064	-7.67047
-6.35000	-6.35000	-7.62000	-6.38133	-6.34479	-7.61895
.00000	-6.35000	-7.62000	-.01385	-6.35730	-7.64143
6.35000	-6.35000	-7.62000	6.30679	-6.34175	-7.71953
-6.35000	6.35000	7.62000	-6.37424	6.33083	7.52602
.00000	6.35000	7.62000	-.03259	6.35241	7.59093
6.35000	6.35000	7.62000	6.33699	6.37270	7.61675
-6.35000	.00000	7.62000	-6.32270	.03923	7.62409
.00000	.00000	7.62000	.09150	-.03086	7.60071
6.35000	.00000	7.62000	6.37469	-.00858	7.66961
-6.35000	-6.35000	7.62000	-6.37336	-6.33632	7.62906
.00000	-6.35000	7.62000	.02986	-6.38231	7.64698
6.35000	-6.35000	7.62000	6.26870	-6.33574	7.67614
.00000	.00000	-2.54000	-.05000	.02000	-2.49000
.00000	.00000	-5.08000	-.02000	-.02000	-4.98000
.00000	.00000	-7.62000	-.03000	.00000	-7.54000
.00000	.00000	-10.16000	-.07000	-.04000	-10.09000
.00000	.00000	-12.70000	-.08000	-.02000	-12.70000
.00000	.00000	-15.24000	-.09000	-.02000	-15.15000
-2.54000	.00000	-10.16000	-2.56000	-.02000	-10.08000
-5.08000	.00000	-10.16000	-5.03000	-.02000	-10.23000
2.54000	.00000	-10.16000	2.46000	-.02000	-9.99000
.00000	.00000	-10.16000	-.07000	-.04000	-10.09000

Fig. (2.5). Original and calculated positions of the 18 calibration points.

|X - Xo|	|Y - Yo|	|Z - Zo|	|P - Po|
.04604	.01755	.01267	.05087
.01630	.00763	.05131	.05437
.03203	.00456	.03845	.05025
.01163	.00999	.06176	.06364
.11439	.00152	.00593	.11455
.05791	.01064	.05047	.07755
.03133	.00521	.00105	.03178
.01385	.00730	.02143	.02654
.04321	.00825	.00953	.10881
.02424	.01917	.09398	.09893
.03259	.00241	.02907	.04373
.01301	.02270	.00325	.02636
.02730	.03923	.00409	.04797
.09150	.03086	.01929	.09847
.02469	.00858	.04961	.05607
.02336	.01368	.00906	.02854
.02986	.03231	.02698	.05161
.08130	.01426	.05614	.09982
.05000	.02000	.05000	.07348
.02000	.02000	.10000	.10392
.03000	.00000	.08000	.08544
.07000	.04000	.07000	.10677
.08000	.02000	.00000	.08246
.09000	.02000	.09000	.12884
.02000	.02000	.08000	.08485
.05000	.02000	.07000	.08832
.08000	.02000	.17000	.18894
.07000	.04000	.07000	.10677

Average Position Error in X direction: .04552 mm
Average Position Error in Y direction: .01700 mm
Average Position Error in Z direction: .05050 mm
Average Overall Position Error   :   .07785 mm

Fig. (2.6). The error value between original and calculated positions.

Viewing Region in mm: 12.70 x 12.70 x 15.24
No. of Calibration points used: 27
Target: Dots on a Flat Surface

Original Position (mm)			Calculated Position (mm)		
Xo	Yo	Zo	X	Y	Z
-6.35000	6.35000	-7.62000	-6.40998	6.34506	-7.54862
.00000	6.35000	-7.62000	-.02844	6.33654	-7.62807
6.35000	6.35000	-7.62000	6.33139	6.36361	-7.61525
-6.35000	.00000	-7.62000	-6.29805	.03616	-7.61036
.00000	.00000	-7.62000	.07221	-.02702	-7.60373
6.35000	.00000	-7.62000	6.39212	.02348	-7.61017
-6.35000	-6.35000	-7.62000	-6.37089	-6.36120	-7.62981
.00000	-6.35000	-7.62000	.02972	-6.37142	-7.62091
6.35000	-6.35000	-7.62000	6.27744	-6.35584	-7.73181
-6.35000	6.35000	.00000	-6.39197	6.32484	-.02678
.00000	6.35000	.00000	-.02453	6.35031	.03500
6.35000	6.35000	.00000	6.34089	6.36901	.00096
-6.35000	.00000	.00000	-6.32024	.04209	.02234
.00000	.00000	.00000	.12372	-.02567	.01078
6.35000	.00000	.00000	6.37436	-.00768	-.04103
-6.35000	-6.35000	.00000	-6.39257	-6.32905	.04863
.00000	=6.35000	.00000	.01185	-6.37550	.01608
6.35000	-6.35000	.00000	6.28963	-6.33004	-.02812
-6.35000	6.35000	7.62000	-6.39054	6.31444	7.53973
.00000	6.35000	7.62000	-.01998	6.36034	7.55985
6.35000	6.35000	7.62000	6.33099	6.35786	7.66831
-6.35000	.00000	7.62000	-6.31756	.02785	7.62865
.00000	.00000	7.62000	.12942	-.01925	7.57740
6.35000	.00000	7.62000	6.36757	.00391	7.66101
-6.35000	-6.35000	7.62000	-6.41443	-6.32715	7.59525
.00000	-6.35000	7.62000	.02032	-6.37345	7.62217
6.35000	-6.35000	7.62000	6.28765	-6.35235	7.70863
.00000	.00000	-6.35000	.04000	.00000	-6.34000
.00000	.00000	-8.89000	.03000	.02000	-8.78000
.00000	.00000	-11.43000	.03000	.02000	-11.38000
.00000	.00000	-13.97000	.02000	.02000	-13.85000
.00000	.00000	-15.12000	.02000	.00000	-15.16000
-2.54000	.00000	.00000	2.49000	.02000	.70000
-5.08000	.00000	.00000	-4.99000	-.02000	.06000
-1.27000	.00000	.00000	-1.23000	.01000	.08000
-1.27000	.00000	2.54000	-1.22000	-.02000	2.49000
-1.27000	.00000	-5.08000	-1.24000	-.02000	-4.97000

Fig. (2.7). Original and calculated positions of the 27 calibration points.

|X - Xo|	|Y - Yo|	|Z - Zo|	|P - Po|
.05598	.00494	.07138	.09336
.02844	.01346	.00807	.03248
.01861	.01361	.00475	.02354
.05195	.03616	.00964	.06402
.07221	.02702	.01627	.07880
.04212	.02348	.00983	.04921
.02089	.01120	.00981	.02566
.02972	.02142	.00091	.03664
.07256	.00584	.11181	.13342
.04197	.02506	.02678	.05574
.02453	.00031	.03500	.04274
.00911	.01901	.00096·	.02110
.02976	.04209	.02234	.05618
.12372	.02567	.01078	.12681
.02436	.00768	.04103	.04834
.04257	.02095	.04863	.06794
.01185	.02550	.01608	.03239
.06037	.01996	.02812	.06952
.04054	.03556	.08027	.09670
.01998	.01034	.06015	.06422
.01901	.00786	.04831	.05251
.03244	.02785	.00865	.04362
.12942	.01925	.04260	.13760
.01757	.00391	.04101	.04479
.06443	.02285	.02475	.07271
.02032	.02345	.00217	.03111
.06235	.00235	.08863	.10839
.04000	.00000	.01000	.04123
.03000	.02000	.11000	.11576
.03000	.02000	.05000	.06164
.02000	.02000	.12000	.12329
.02000	.00000	.04000	.04472
.05000	.02000	.07000	.08832
.09000	.02000	.06000	.11000
.04000	.01000	.08000	.09000
.05000	.02000	.05000	.07348
.03000	.02000	.11000	.11576

Average Position Error in X direction: .04245 mm
Average Position Error in Y direction: .01748 mm
Average Position Error in Z direction: .04240 mm
Average Overall Position Error   :   .06956 mm

Fig. (2.8). The error value between original and calculated positions.

| Original Position (mm) | | | Calculated Position (mm) | | |
Xo	Yo	Zo	X	Y	Z
-6.35000	6.35000	-7.62000	-6.40023	6.31327	-7.63408
.00000	6.35000	-7.62000	-.04862	6.36990	-7.56809
6.35000	6.35000	-7.62000	6.35049	6.36278	-7.66938
-6.35000	.00000	-7.62000	-6.30659	.01290	-7.67483
.00000	.00000	-7.62000	.10814	-.01456	-7.51413
6.35000	.00000	-7.62000	6.39802	-.03187	-7.66397
-6.35000	-6.35000	-7.62000	-6.40194	-6.34134	-7.65920
.00000	-6.35000	-7.62000	.01984	-6.36173	-7.61522
6.35000	-6.35000	-7.62000	6.26843	-6.35752	-7.79435
-6.35000	6.35000	-5.08000	-6.36067	6.34542	-5.08696
.00000	6.35000	-5.08000	-.02196	6.34677	-4.99729
6.35000	6.35000	-5.08000	6.36479	6.36604	-5.08442
-6.35000	.00000	-5.08000	-6.31212	.04076	-5.02897
.00000	.00000	-5.08000	.10557	-.02700	-5.02468
6.35000	.00000	-5.08000	6.41168	-.01191	-5.06541
-6.35000	-6.35000	-5.08000	-6.36324	-6.33741	-5.08293
.00000	-6.35000	-5.08000	.01669	-6.35830	-5.11280
6.35000	-6.35000	-5.08000	6.30996	-6.31469	-5.08569
-6.35000	6.35000	-2.54000	-6.39158	6.32165	-2.56628
.00000	6.35000	-2.54000	-.02472	6.34813	-2.52770
6.35000	6.35000	-2.54000	6.30444	6.37186	-2.51781
-6.35000	.00000	-2.54000	-6.34399	.03320	-2.49532
.00000	.00000	-2.54000	.06119	-.00803	-2.47839
6.35000	.00000	-2.54000	6.35001	.01091	-2.48561
-6.35000	-6.35000	-2.54000	-6.39414	-6.35850	-2.54101
.00000	-6.35000	-2.54000	-.02790	-6.37911	-2.55043
6.35000	-6.35000	-2.54000	6.29637	-6.34133	-2.58055
-6.35000	6.35000	2.54000	-6.38103	6.32992	2.45171
.00000	6.35000	2.54000	-.03939	6.32743	2.52617
6.35000	6.35000	2.54000	6.33272	6.38162	2.60181
-6.35000	.00000	2.54000	-6.33336	.04436	2.54327
.00000	.00000	2.54000	.11504	-.02160	2.62808
6.35000	.00000	2.54000	6.34953	-.00791	2.56821
-6.35000	-6.35000	2.54000	-6.38620	-6.34039	2.53196
.00000	-6.35000	2.54000	.02505	-6.36005	2.58243
6.35000	-6.35000	2.54000	6.29309	-6.35439	2.50744
-6.35000	6.35000	5.08000	-6.38344	6.34048	5.03196
.00000	6.35000	5.08000	-.01139	6.35962	5.04471
6.35000	6.35000	5.08000	6.32131	6.37971	5.04212
-6.35000	.00000	5.08000	-6.29263	.03720	5.05658
.00000	.00000	5.08000	.11441	-.03390	5.06935
6.35000	.00000	5.08000	6.36529	.01180	5.11615
-6.35000	-6.35000	5.08000	-6.38863	-6.32546	5.13383
.00000	-6.35000	5.08000	.02385	-6.35666	5.03624
6.35000	-6.35000	5.08000	6.30695	-6.34457	5.07657
-6.35000	6.35000	7.62000	-6.38386	6.32191	7.59416
.00000	6.35000	7.62000	-.05355	6.33609	7.53559
6.35000	6.35000	7.62000	6.33701	6.35697	7.57271
-6.35000	.00000	7.62000	-6.29397	.03518	7.63239
.00000	.00000	7.62000	.09912	-.03548	7.66668
6.35000	.00000	7.62000	6.38022	.00525	7.65964
-6.35000	-6.35000	7.62000	-6.41725	-6.34635	7.62550
.00000	-6.35000	7.62000	.00985	-6.37119	7.64213
6.35000	-6.35000	7.62000	6.32343	-6.36471	7.62831
.00000	.00000	-2.54000	-.01000	.03000	-2.40000
.00000	.00000	-5.08000	-.01000	.02000	-4.99000
.00000	.00000	-7.62000	.00000	.04000	-7.58000
.00000	.00000	-10.16000	-.02000	.00000	-10.02000
.00000	.00000	-12.70000	-.02000	.00000	-12.64000
.00000	.00000	-15.24000	-.03000	.00000	-15.09000
-5.08000	.00000	-10.16000	-5.02000	-.03000	-10.17000
-2.54000	.00000	-10.16000	-2.55000	.00000	-10.06000
2.54000	.00000	-10.16000	2.49000	.00000	-10.04000
5.24000	.00000	-10.16000	4.49000	.04000	-9.98000

Fig. (2.9). Original and calculated positions of the 54
calibrated points.

	X - Xo			Y - Yo			Z - Zo			P - Po	
.05023	.03673	.01408	.06380
.04862	.01990	.05191	.07386
.00049	.01278	.04938	.05101
.04341	.01290	.05483	.07111
.10814	.01456	.10587	.15204
.04802	.03187	.04397	.07249
.05194	.00346	.03920	.06517
.01984	.01173	.00478	.02353
.08157	.00752	.17435	.19263
.01067	.00458	.00696	.01354
.02196	.00323	.08271	.08563
.01479	.01604	.00442	.02226
.03788	.04076	.05103	.07550
.10557	.02700	.05532	.12221
.06168	.01191	.01459	.06449
.01324	.01259	.00293	.01851
.01669	.00830	.03280	.03772
.04004	.03531	.00569	.05369
.04158	.02835	.02628	.05677
.02472	.00187	.01230	.02767
.04556	.02186	.02219	.05519
.00601	.03320	.04468	.05599
.06119	.00803	.06161	.08721
.00001	.01091	.05439	.05548
.04414	.00850	.00101	.04496
.02790	.02911	.01043	.04165
.05363	.00867	.04055	.06779
.03103	.02008	.08829	.09571
.03939	.02257	.01383	.04745
.01728	.03162	.06181	.07155
.01664	.04436	.00327	.04749
.11504	.02160	.08808	.14649
.00047	.00791	.02821	.02931
.03620	.00961	.00804	.03831
.02505	.01005	.04243	.05029
.05691	.00439	.03256	.06571
.03344	.00952	.04804	.05930
.01139	.00962	.03529	.03831
.02869	.02971	.03788	.05604
.05737	.03720	.02342	.07228
.11441	.03390	.01065	.11980
.01529	.01180	.03615	.04099
.03863	.02454	.05383	.07065
.02385	.00666	.04376	.05028
.04305	.00543	.00343	.04353
.03386	.02809	.02584	.05102
.05355	.01391	.08441	.10093
.01299	.00697	.04729	.04953
.05603	.03518	.01239	.06731
.09912	.03548	.04668	.11516
.03022	.00525	.03364	.05012
.06725	.00365	.00550	.06757
.00985	.02119	.02213	.03218
.02657	.01471	.00831	.03149
.01000	.03000	.14000	.14353
.01000	.02000	.09000	.09274
.00000	.04000	.04000	.05657
.02000	.00000	.14000	.14142
.02000	.00000	.06000	.06325
.03000	.00000	.15000	.15297
.06000	.03000	.01000	.06782
.01000	.00000	.10000	.10050
.05000	.00000	.12000	.13000
.75000	.04000	.18000	.77233

Average Position Error in X direction :   .04895 mm

Average Position Error in Y direction :   .01760 mm

Average Position Error in Z direction :   .04765 mm

Average Overall Position Error   :   .08097 mm

Fig. (2.10). The error value between original and calculated positions.

2.6 Discussion and Conclusion

A procedure for CCD camera calibration has been introduced. The approach relates the camera coordinates to the world coordinates through homogeneous transformation. By utilizing calibration procedures the geometrical constraints were relaxed. Retrieving the 3-D position information of an object within the field of view of both cameras is simplified to a procedure that involves solving only a set of four linear equations.

The approach does not require special cameras, lenses or computer hardware, it utilizes only off-the-shelf components which makes it suitable for most industrial applications. Relaxing geometrical setup constraints represents an important factor for industrial implementations since no special skills or equipment is required for system setup. Furthermore no initial guess is required for parameters value. But the procedure can be implemented for automatic calibration in robotic applications with little operator intervention.

CHAPTER III

PERFORMANCE EVALUATION

Abstract

The system performance in terms of accuracy and repeatability is evaluated. The evaluation procedure proposed by the Automated Vision Association* is adhered to as closely as possible. The test setup and procedure are illustrated. The results obtained for both accuracy and repeatability are discussed.

3.1 Introduction

Performance evaluation of machine vision 3-D metrology is concerned with three figures of merits, namely,

1. Accuracy, defined as the degree of conformance between a measurement of an observed quantity and a recognized standard or specification which indicates the true value of the quantity.

---

*AVA is a specialty association of Robotic Industries Assoc.

2. Repeatability, the degree to which repeated measurements of the same quantity vary about their mean.

3. Cycle time: the amount of time required to perform a stated task. The elapsed time between transmission of the "initiate" signal to the system under test and receipt of the corresponding "Done" signal.

Almost every article discussing 3-D machine vision reports one or more of these figures of merits to evaluate the performance of the proposed technique. However, these figures can be ambiguous or misinterpreted or a cause of incorrect evaluation of the technique, since no specific standard was utilized to obtain such figures.

The Automated Vision Association* (AVA), in recognition of the problem and in response to an industrial need developed the "A15.05 performance evaluation standard", aiming to create a common understanding of machine vision functionality. Instead of developing a standard that would test the performance of every facet of machine vision, tests were defined for very specific aspects of machine vision.

For the purpose of evaluating the performance of the proposed approach, A15.05 "Measurement of feature position in 3-D space - Single point test" will be adhered to as much as possible and any deviation will be indicated.

---

*AVA is located at 900 Victors Way, Ann Arbor, MI. 48106

3.2 Single Point Test

The equipment required to conduct the test is a machine vision system (to be tested), a target with associated light and translating table. The test controller is used to initiate the vision system, generate a random position for movements and generate the evaluation report.

In our test, the system was initiated manually from the keyboard and the positions were randomly selected to cover the range measurements. The vision system computer was used to store the data file, to be used later to generate the reports shown in Fig. (3.3) - Fig. (3.11).

The test target plate specified by the standard is a flat photographic quality glass coated via an optographic process with dense target shape. This target was replaced by a black disk of a radius 5 mm plotted with a high accuracy plotter (CALCOMP 1043) on a thick mylar. A 3-axes translation table with high precision anti-backlash ball-screw used as motion actuators, to move the target in 3 directions. A set of dial gages were mounted to indicate the actual movement value (dial gage accuracy of 0.0127 mm). An overview of the experimental setup is shown in Fig. (2.2) showing the vision system, the 3 axes table, the test target and the dial gages. Figure (3.1) shows a close-up of the table with the cameras and the dial gages.
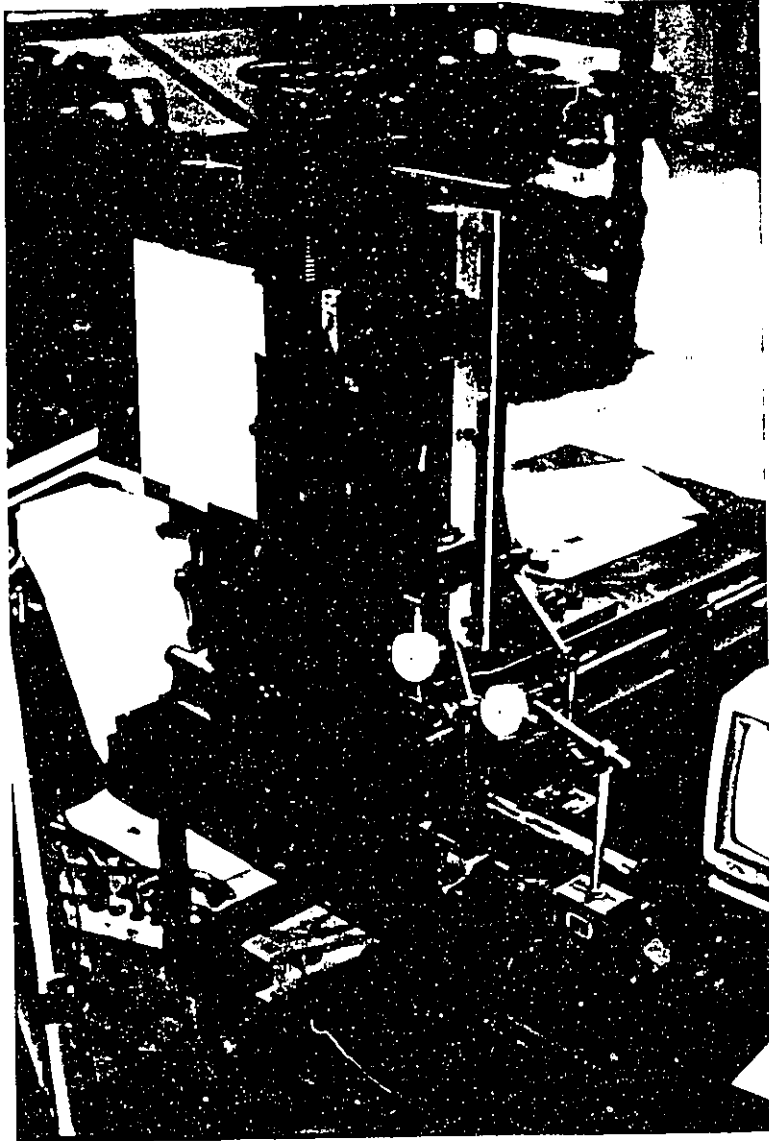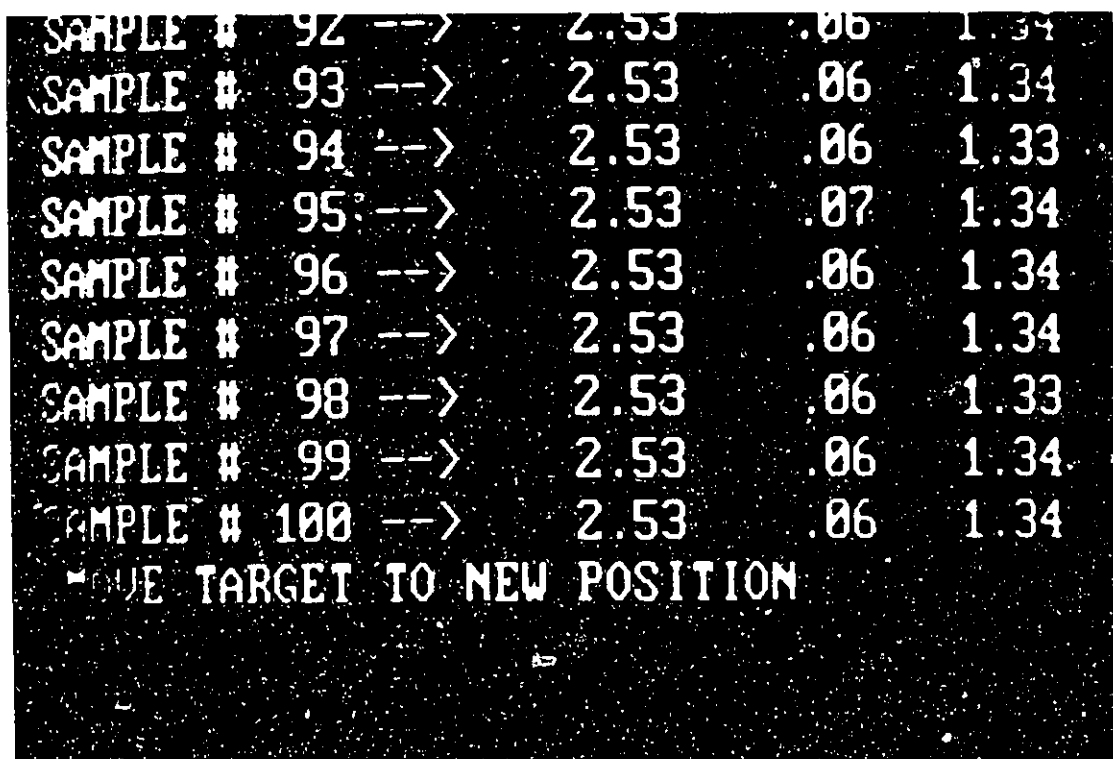
Fig. (3.1). A close-up to the test set-up
showing the 3-axis table and the
dial indicators.

```
SAMPLE #  92 --->      2.53      .06    1.34
SAMPLE #  93 --->      2.53      .06    1.34
SAMPLE #  94 --->      2.53      .06    1.33
SAMPLE #  95 --->      2.53      .07    1.34
SAMPLE #  96 --->      2.53      .06    1.34
SAMPLE #  97 --->      2.53      .06    1.34
SAMPLE #  98 --->      2.53      .06    1.33
SAMPLE #  99 --->      2.53      .06    1.34
SAMPLE # 100 --->      2.53      .06    1.34
MOVE TARGET TO NEW POSITION
```

Fig. (3.2).   The X,Y,Z position reported by the vision system
              when a test of 100 continuous readings were performed
              to illustrate the system's repeatability.

3.3 Test Procedure

The test was conducted with the following procedure:

a) The vision system was initiated to start taking measurements, display the results to the monitor and write the results to a file. Fig. (3.2) shows a sample of the 100 readings taken for one of the measurement points, indicating the sample number and the X,Y,Z position.

b) The following measurements sequence is iterated 90 times:

   i) The target is moved to a randomly selected position.

   ii) Once the target is in position the vision system is initiated taking 10 measurements.

The range of movements covered 12.7*12.7*12.7 mm (the limitation of the dial gage used, that had a maximum stroke of 12.7 mm).

3.4 Test Results

The test results are reported for each individual axis as shown in Fig. (3.3) through Fig. (3.11). Three values per axis were obtained. The first value is the standard deviation of the 10 measurements about their mean for each of the 90 points, shown in Fig. (3.3), Fig. (3.6) and Fig. (3.9) for the X,Y and Z axis respectively.
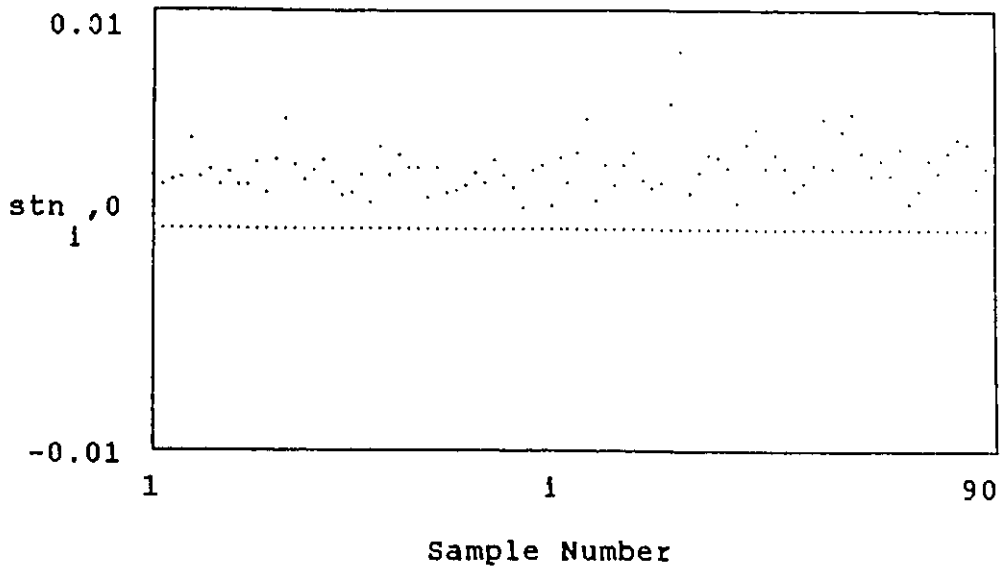
X-axis Standard deviation :

Fig. (3.3).   Standard deviation of the X-axis
samples about their mean value.

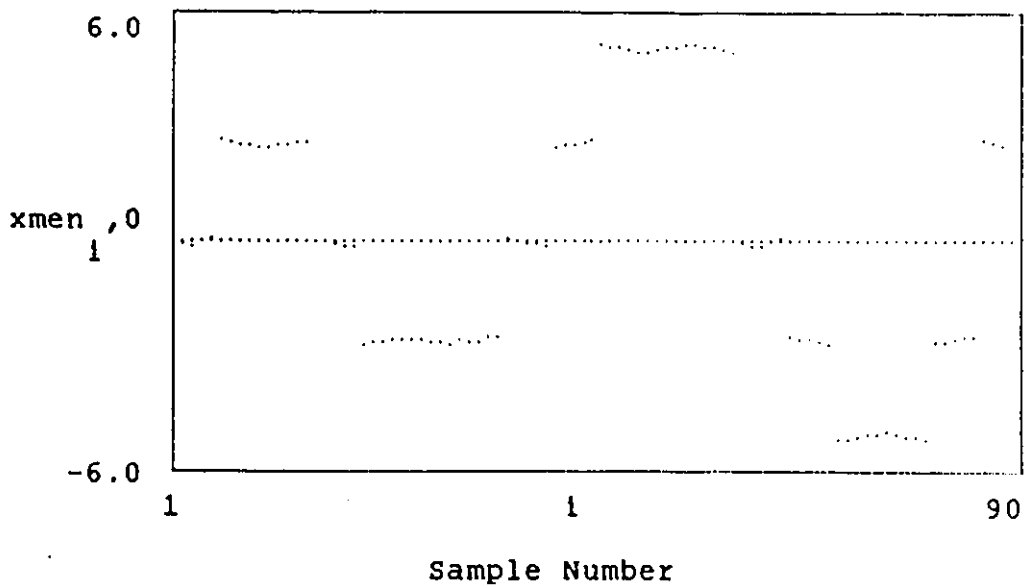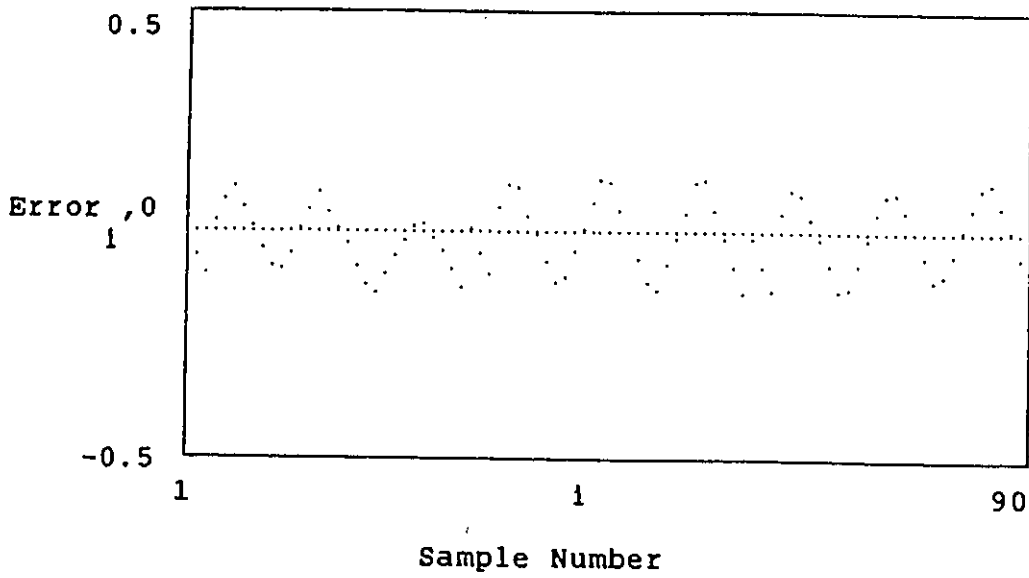X-axis Mean value for the samples position :

Fig. (3.4).   The mean value of the test point
for the X-axis.

minimum value in the error envelope: emin = -0.13 mm
maximum value in the error envelope: emax = 0.12 mm

Fig. (3.5). The error between the mean value of the
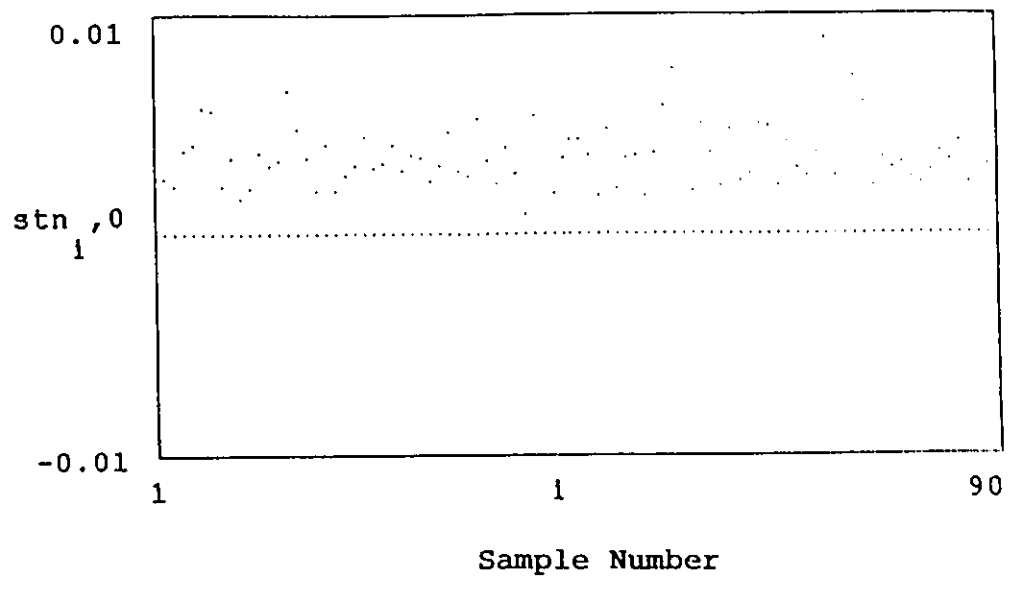   X-axis and the actual value

Fig. (3.6). Y-axis standard deviation.
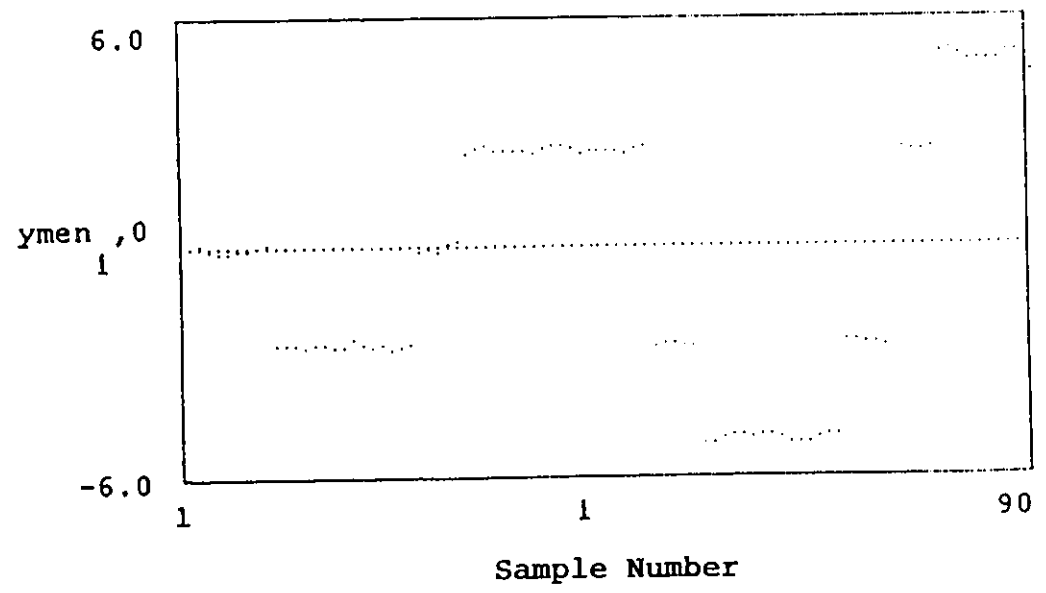
Fig. (3.7). Y-axis mean value for the samples position.

Error analysis for Y-axis :
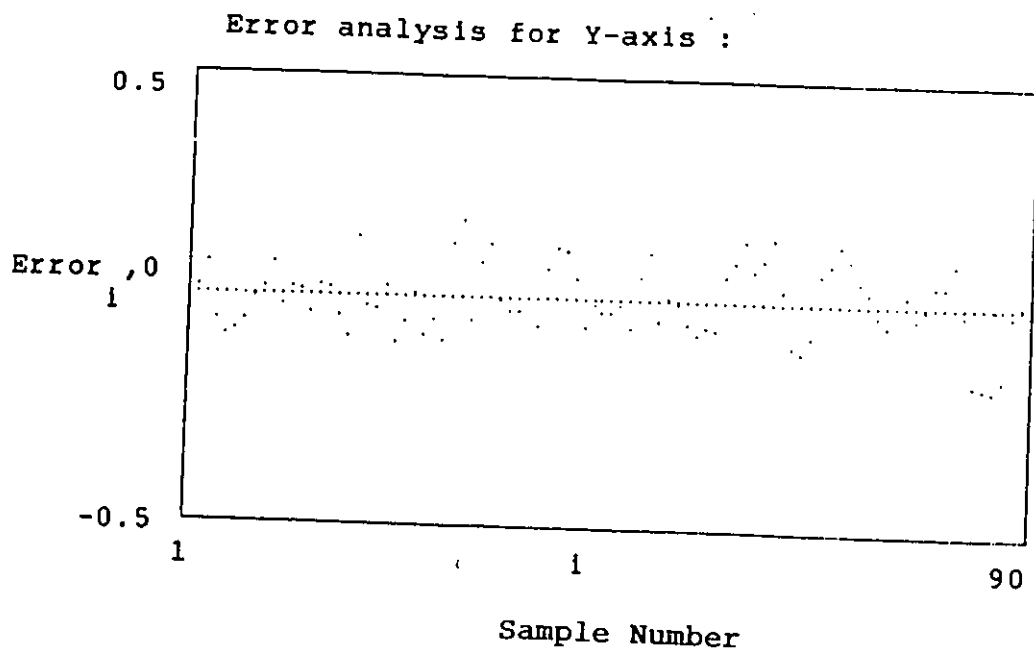
minimum value in the error envelope:
        emin = -0.18   mm
maximum value in the error envelope:
        emax = 0.172   mm
repeatability envelope:
        repy = 0.176   mm

Fig. (3.8).   Error analysis for Y-axis

Z-axis Standard deviation :
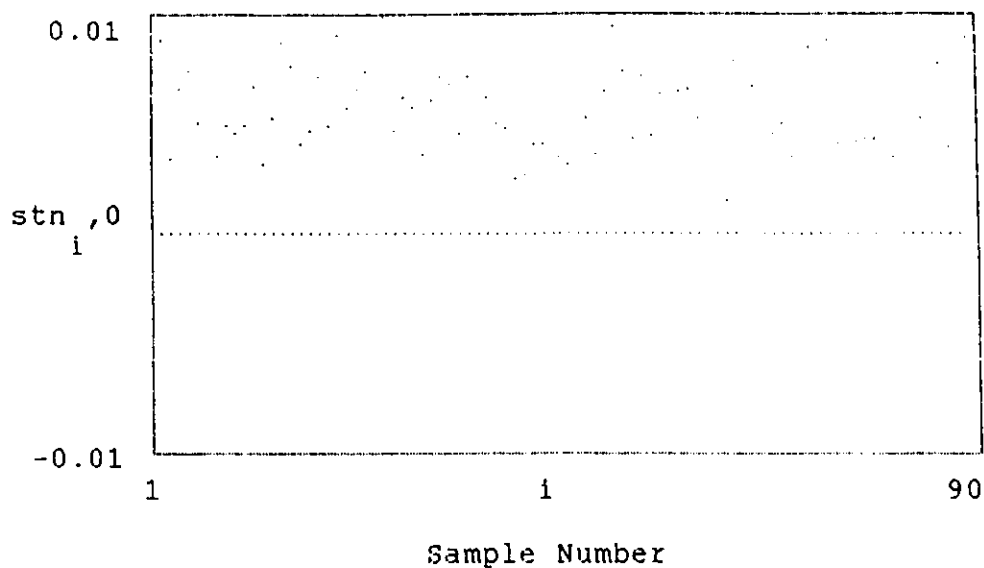
Fig. (3.9). Z-axis standard deviation.

Z-axis Mean value for the samples position :
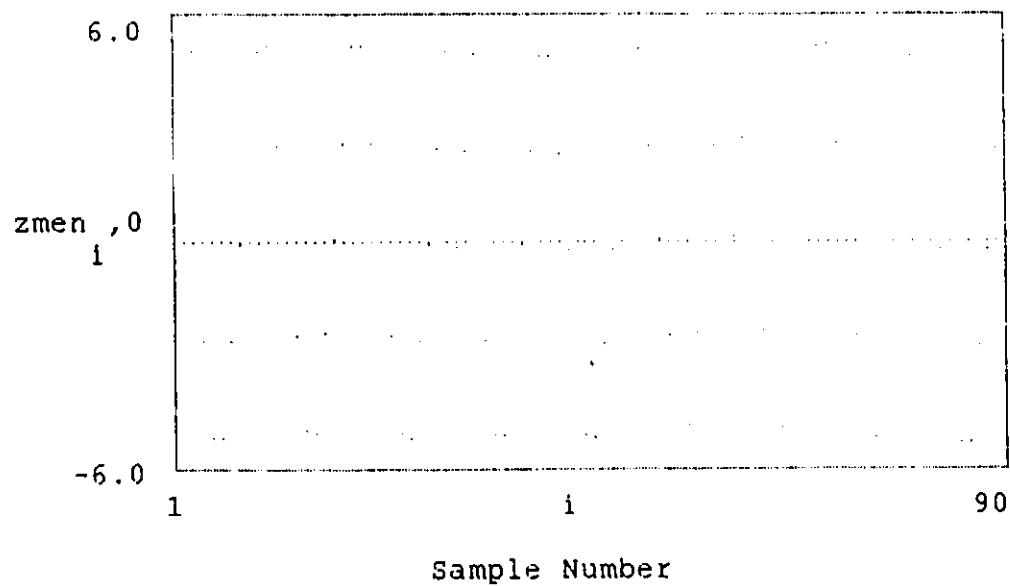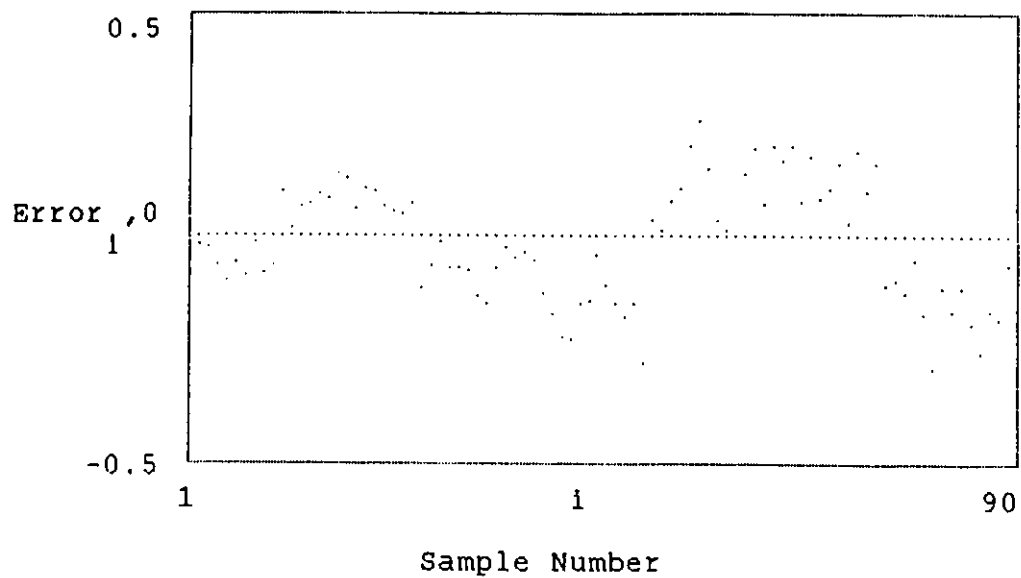

Fig. (3.10). Z-axis mean value for the samples
position

Error analysis for Z-axis :

minimum value in error envelope:
    emin = -0.29   mm
maximum value in the error envelope:
    emax = 0.264   mm
repeatbility envelope:
    repz = 0.277   mm

Fig. (3.11). Error analysis for Z-axis

The second value is the mean of the 10 readings for each of the 90 points, Fig. (3.4), Fig. (3.7) and Fig. (3.10). The last value is the error value between the actual and the mean of 10 measurements of each point, defined as:

$$\text{Error } (i) = x^c(i) - \bar{x}(i) \qquad i = 1,2,\ldots,90$$

where

$$\bar{x}(i) = \frac{1}{N} \sum_{n=1}^{N} x_n \qquad N = 10$$

and $x^c(i)$ is the actual x-axis value for the ith position.

Figures (3.5), (3.8) and (3.11) show the obtained results for each of the X,Y,Z axis, respectively.

3.5 Discussion and Conclusion

The test results obtained represent a measure for evaluating the performance of the proposed system.

The standard deviation of the 10 measurements about their mean represents a measure for evaluating system repeatability for the same position. The error between the mean and the actual, on the other hand, represents a measure for the accuracy of the system. Therefore, from the results illustrated by Figs. (3.5), (3.8) and (3.11), it shows that the system accuracy for each of the axis is:

X-axis accuracy   =   0.13 mm

Y-axis accuracy   = 0.18 mm

Z-axis accuracy   = 0.29 mm

over a range of measurements of 12.7 x 12.7 x 12.7 mm, which represents an accuracy of 2.2% of the field of view.

Since the depth information (i.e., Z-axis position) is retrieved through disparity in the X-axis, the accuracy of the X-axis is shown to be affected by the Z-axis position as can be noticed by the error in the X-axis, Fig. (3.5) and the Z-axis position, Fig. (3.10).

The system repeatability is shown in Figs. (3.3), (3.6) and (3.9) and also through the 100 measurements of Fig. (3.2), which indicates that the system's repeatability for all axes are within an envelop of 0.01 mm.

In conclusion, the performance of the proposed system is suitable for close range high accuracy measurements. The evaluation was performed in accordance to a standard test to eliminate any ambiguity about the obtained results.

CHAPTER IV

MATHEMATICAL MORPHOLOGY

Abstract

In this chapter, mathematical morphology operators and some of their applications pertaining to this thesis will be reviewed. The objective of this discussion is to pave the road to the next chapter where mathematical morphology is utilized in the procedure dealing with the point-of-correspondence problem. Erosion, dilation and their residues and application to edge detection and automatic shape recognition is therefore reviewed.

4.1 Introduction

Mathematical morphology is an approach for image processing, which is based on set theory and shapes. It has developed a great popularity in the industrial world and is currently utilized in numerous applications. The objective of discussing mathematical morphology in this chapter is to provide an overview of the morphological operator. Examples of their applications are to be employed in the

proposed procedure for determining the point-of-correspondence in stereo vision system.

The language of mathematical morphology is that of set theory where sets represent shapes which are manifested on binary or gray-scale images. In addition to the usual set theory operations of union, complements etc., two operators are used by mathematical morphology, namely, erosion and dilations. The morphological operator works with two images, the original data image to be analyzed and a structure element. Each structuring element has a shape which can be thought of as a parameter to the operation.

4.2 Morphological Operators

The two operators utilized in mathematical morphology application are erosion and dilation (often referred to as closing and opening). Morphological operator operates with two images (or sets) the original image to be analyzed and structure element utilized to perform the analysis with the shape of the structure element determining the form of analysis performed.

4.2.1 Erosion Operation

The erosion of binary image A by structure element B is denoted by

$$A \ominus B \qquad\qquad (4.1)$$

and defined by

$$A \ominus B = \{x \in E^N \quad x+b \in A \quad \text{for every } b \in B\} \qquad (4.2)$$

This definition considers A and B to be two sets in Euclidean-N space. Then the erosion of set A by set B is the set of all elements x for which $x+b \in A$ for every $b \in B$. To illustrate this definition further, let us consider the following example.

Let A be a binary set given as defined:

$$A = \{(1,0),(1,1),(1,2),(1,3),(2,1),(3,1),(4,1)\}$$
$$= \{x_1, x_2, x_3, x_4, x_5, x_6, x_7\}$$

and let B be a set given by

$$B = \{(0,0), (0,1)\}$$
$$= \{b_1, b_2\}$$

Now applying the definition given by Eq. (4.2) it can be easily be shown that

$$A \ominus B = \{x_1, x_2, x_3\}$$

since $x_1 + b_1 = (1,0) + (0,0) = (1,0) \in A$
$$x_1 + b_2 = (1,0) + (0,1) = (1,1) \in A$$

therefore $x_1 \in A \ominus B$.
Similarly x_2, x_3.

However, for the case of x_4

$$x_4 + b_1 = (1,3) + (0,0) = (1,3) \in A$$

$$x_4 + b_2 = (1,3) + (0,1) = (1,4) \notin A$$

and therefore $x_4 \notin A \ominus B$.

4.2.2 Dilation Operation

The operation of dilation is the morphological dual to the erosion operation. It operates on two sets in N-Space, A and B. The dilation of A by B is denoted by

$$A \oplus B \tag{4.3}$$

and defined by

$$A \oplus B = \{c \in E^N \mid c=a+b \text{ for some } a \in A \text{ and } b \in B].\tag{4.4}$$

To illustrate this definition, let A be a set defined as

$$A = \{(0,1), (1,1), (2,1), (2,2), (3,0)\}$$

and let B be a set defined as

$$B = \{(0,0), (0,1\}$$

Then applying the dilation definition given by Eq. (4.4) can be easily shown that

$$A \oplus B = \{(0,1), (1,1), (2,1), (2,2), (3,0),$$
$$(0,2), (1,2), (2,2), (2,3), (3,1)\}$$

This can be illustrated by considering a, $(0,1)$

$$c_1 = a_1 + b_1 = (0,1)$$
$$c_2 = a_1 + b_2 = (0,2)$$

Thus for a point in a set such as a_1 $(0,1)$ dilated with the set B the resulting set includes both c_1 and c_2.

The above two definitions were presented by Matheron [32] and Serra [33] for binary images. These definitions were extended later by Sternberg [34]-[35] to gray-scale morphology.

4.3 Gray-scale Morphology

The erosion of gray-scale image f by a structuring element b is denoted by e and defined by

$$e(r,c) = \min_{i,j} (f(r+i,c+j) - b(i,j)) \qquad (4.5)$$

where the minimum is taken over all (i,j) in the domain of b.

Dilation of gray-scale can also be defined as

$$d(r,c) = \max_{i,j} (f(r-i,c-j) + b(i,j)) \qquad (4.6)$$

where the maximum is taken over all (i,j) in the domain of b

and the domain of d is the domain of f dilated by the domain of b .

To illustrate these two definitions, let f be a gray-scale image as shown in Fig. (4.1). B is a structure element of a rod shape j with radius 1, i.e.,

B = {(0,-1), (0,1), (-1,0), (1,0)}

and the center of the rod is (0,0). Also, let the height of the rod be zero (this assumption is to obtain the results that will be used later with the edge detection examples) shown in Fig. (4.2).

Now applying the definition of the erosion e(r,c) to point (2,2).

$$e(2,2) = \min_{i,j} \{ \ f(2,1) - b(0,-1), \ (f(2,3) - b(0,1)),$$

$$f(1,2) - b(-1,0), \ f(3,2) - b(1,0)\}$$

From Figures (4.1) and (4.2)

$$e(2,2) = \min_{i,j} \{0,0,0,0\}$$

$$= 0$$

Now consider the point (3,3) for the erosion. Applying the same definition then

$$e(3,3) = \min_{i,j} \{(f(3,2) - b(0,-1), \ (f(3,4) - b(0,1)),$$

$$(f \ (2,3) - b(-1,0)), \ (f(4,3) - b(1,0))\}$$

	1	2	3	4	5	6	7	8
1	0	0	0	0	0	0	0	0
2	0	100	100	100	100	100	100	0
3	0	100	100	100	100	100	100	0
4	0	100	100	150	150	100	100	0
5	0	100	100	150	150	100	100	0
6	0	100	100	100	100	100	100	0
7	0	100	100	100	100	100	100	0
8	0	0	0	0	0	0	0	0

"I"

Fig. (4.1). A gray-scale image used for illustration of morphological operations.
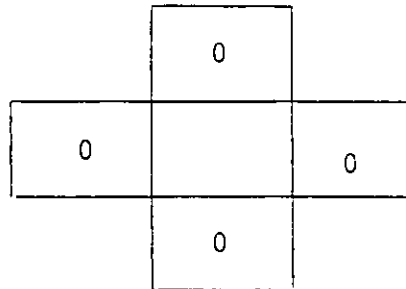
"B"

Fig. (4.2). A structure element used for illustration of morphological operations.

$$= \min_{i,j} \{0,\ 100,\ 0,\ 100\}$$

$$= 0$$

Applying the same definition over all values for r, c the results shown in Fig. (4.3) are obtained.

Similarly, the definition of the dilation can be illustrated by using the same structure element B and the same image f given in Figs. (4.2) and (4.1) respectively.

Applying the dilation definitions of Eq. (4.6) to point (2,2)

$$d(2,2) = \max_{i,j} \{(f(2,3) + b(0,-1),\ (f(2,1) + b(0,1)$$

$$(f(3,2) + b(-1,0)),\ (f(1,2) + b(1,0))\}$$

$$= \max_{i,j} \{100,\ 0,\ 100,\ 0\}$$

$$= 100$$

Applying the definition to the image results shown in Fig. (4.4) can be obtained.

Algorithms for mathematical morphology combines sequences of dilation, erosions and their residues. The two algorithms relevant to this thesis are the edge detection operation [42] and the automatic shape recognition [37].

0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	100	100	100	100	0	0
0	0	100	100	100	100	0	0
0	0	100	100	100	100	0	0
0	0	100	100	100	100	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

$$f \ominus B$$

Fig. (4.3). The result of $f \ominus B$ for f of Fig. (4.1) and B of Fig. (4.2).

0	100	100	100	100	100	100	0
100	100	100	100	100	100	100	100
100	100	100	150	150	100	100	100
100	100	150	150	150	150	100	100
100	100	150	150	150	150	100	100
100	100	100	150	150	100	100	100
100	100	100	100	100	100	100	100
0	100	100	100	100	100	100	0

$$"f \oplus B"$$

Fig. (4.4). The result of applying the dilated operation to the image of Fig. (4.1) with the structure element of Fig. (4.2).

4.4 Morphological Edge Operator

A simple method of performing gray-scale edge detection in a morphology based vision system is to take the difference between an image and its erosion by a small rod shaped structuring element.

Lee, Harlick and Shappiro [37] showed that a rod of radius 1 with flat top, the erosion of image f can be carried out by the rule

$$e(r,c) = \min_{i,j} \{f(r+i, c+j) - b(i,j)\}$$

and if the element height is zero, it leads to

$$e(r,c) = \min_{i,j} \{f(r+i, c+j)\}$$

then the erosion residue edge detector G_e can be defined as

$$
\begin{aligned}
G_e &= f(r,c) - e(r,c) \\
&= f(r,c) - \min_{i,j} f(r+i, c+j) \\
&= \max_{i,j} \{f(r,c) - f(i,j)\}
\end{aligned}
\qquad (4.7)
$$

which can be illustrated by subtracting image f of the erosion example Fig. (4.1) from the result of the erosion operation, Fig. (4.3). The result of this subtraction is as shown in Fig. (4.5), which represents the edge strength image for the image given by f. Figure (4.7) shows an example of edges obtained by applying the above procedure to the image

0	0	0	0	0	0	0	0
0	100	100	100	100	100	100	0
0	100	0	0	0	0	100	0
0	100	0	50	50	0	100	0
0	100	0	50	50	0	100	0
0	100	0	0	0	0	100	0
0	100	100	100	100	100	100	0
0	0	0	0	0	0	0	0

"$f - f \ominus B$"

Fig. (4.5). Edge-strength image resulting from
subtracting the original image from
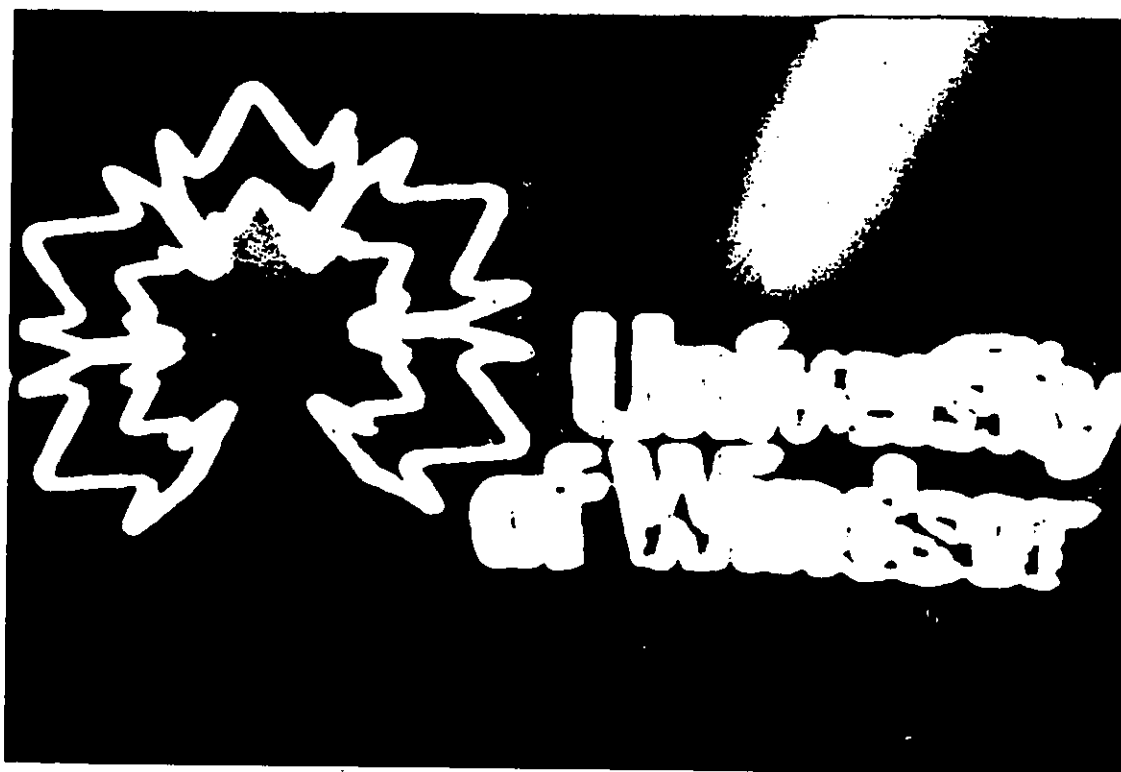its eroded image.

Fig. (4.6). Original Image.

Fig. (4.7). Edge-strength image obtained by applying erosion residue
edge operator

Dilation residue operators for edge detection can also be developed. Dilating the original image with a small rod structure element and taking the difference between the dilated image and the original will result in an edge strength image. This can be illustrated by subtracting the result of the dilation example Fig. (4.4) from the original image Fig. (4.1). The results are shown in Fig. (4.8).

Similar to the erosion residue edge detection operator, a dilation residue edge detection operator can be defined as

$$G_d(r,c) = \max_{i,j} \{f(i,j) - f(r,c)\}. \tag{4.8}$$

Figure (4.9) shows an example of the result of edge detection using the dilation residue operator.

The application of both approaches were discussed and evaluated in [42]. It was concluded that both approaches are either sensitive to noise or cannot detect ideal step edges. Lee et al. [42] then introduced the blur-minimum edge operator, which is noise insensitive. Also, on ideal step edges it produces results, which has non-zero edge strength on edge pixels.

The blur-minimum operator is defined by

$$I_{edge\text{-}strength} = \min\{I_1\text{-}erosion\ (I_1),\ dilation\ (I_1)\text{-}I_1\} \tag{4.9}$$

0	100	100	100	100	100	100	0
100	0	0	0	0	0	0	100
100	0	0	50	50	0	0	100
100	0	50	0	0	50	0	100
100	0	50	0	0	50	0	100
100	0	0	50	50	0	0	100
100	0	0	0	0	0	0	100
0	100	100	100	100	100	100	0

"f $\oplus$ B - f"

Fig. (4.8). Edge-strength image resulting from subtracting the dialated image from the original image.

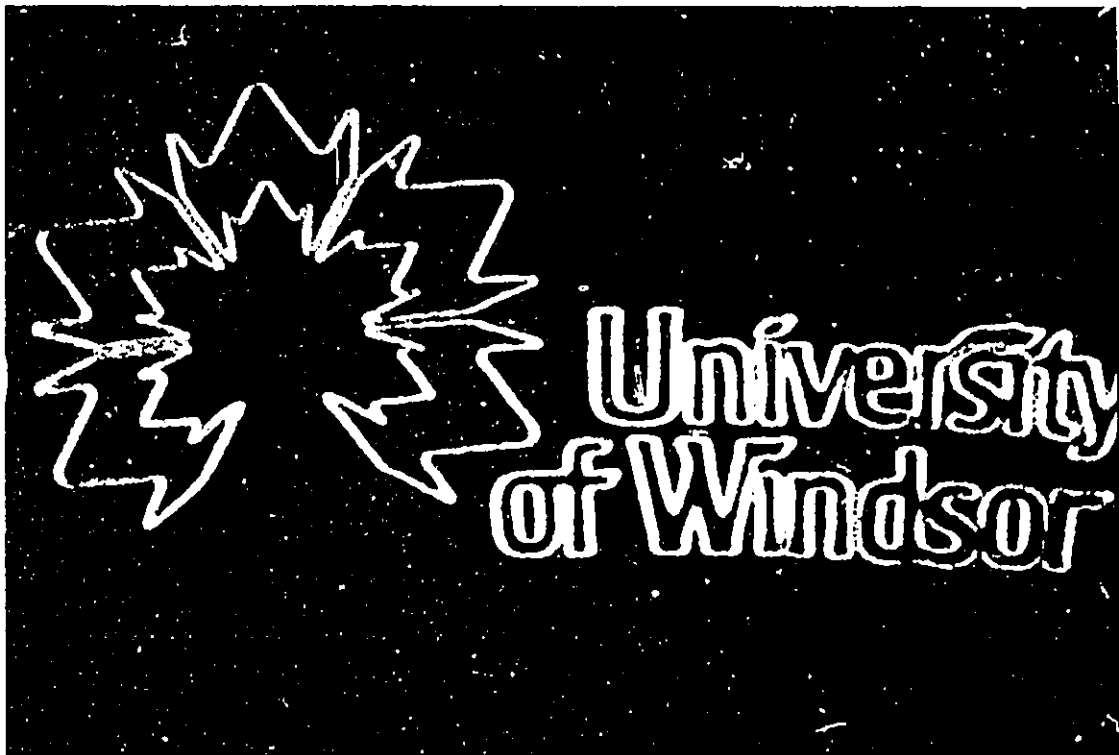Fig. (4.9) Edge-strength image obtained by applying dilation residue edge operator.

Fig. (4.10). Edge-strength image obtained by applying the blur-minimum operator.

where I_1=blur $\{I_{input}\}$, and blur $\{I_{input}\}$ is the input image with a blurring operation. Figure (4.10) shows the result of applying the blur-minimum operator to the image of Fig. (4.6).

4.5 Automatic Shape Recognition

The problem of shape or pattern recognition is one that is very closely related to the operation of erosion once the problem is converted to that of detecting the occurrences of specific sets within an image. Crimmins and Brown [37] introduced a theorem for automatic shape recognition, utilizing mathematical morphology and suggested a generic pattern recognition computer. The theorem states that the shape A in window W occurs in the image I at and only at the following locations:

$$[I \ominus A] \cap [I \ominus (W - A)]$$

where $W - A = W \cap \overline{A}$.

The sense of the theorem is, since $I \ominus A$ (the erosion of I with element A) is the locus of all points where A fits inside I, and $I \ominus (W-A)$ is the locus of all points where the background of A in window W fits inside the background of I, then the intersection provides exactly those places where the shape A with its required background occurs in the image I.

With this brief review of mathematical morphology we have paved the road to the next chapter, where both edge detection and automatic shape recognition will be utilized in the procedure for determining the point-of-correspondence.

CHAPTER V

POINT OF CORRESPONDENCE IN STEREO VISION

Abstract

A procedure for determining point-of-correspondence in stereo images for close range position sensing is introduced. The two images are not constrained to a specific geometrical relationship. The calibration technique introduced earlier is adopted to determine the transformation parameters between the world coordinate system and the camera planes.

The blur-minimum edge operator reviewed in the previous chapter is utilized to obtain the edge strength image for both images. A structure element centered about the selected point is extracted from one camera image and mapped to the second camera plane. A search neighborhood is calculated and erosion over that neighborhood with the mapped structure element is performed to determine the corresponding point.

5.1 Introduction

The point-of-correspondence problem is defined as the problem of finding matched features between the two images in

stereo vision system. These matches can then be used in finding depth information from the disparity between the two images. Research and theories formulation have been under development for several years. Studies addressing this problem have taken two main approaches. The first approach addresses the psychophysical and biological aspect of human visual system [22-28]. The second approach examines computational aspects in various application domains [29-31]. Both approaches, however, have a common assumption. That is the two cameras are co-planar (i.e., both cameras have parallel optical axes), however, the image of the object on both imaging planes are the same with the only difference being the disparity (translation) due to the distance between the two cameras. Although this assumption may be valid in some applications of stereo vision (e.g., aerial photography), for close range position sensing making, such an assumption would lead to the following two disadvantages.

1. The common field of view between the two cameras is not utilizing the full image plane, as shown in Fig. (5.1).

2. The setup of the two cameras to be aligned, require a highly qualified personnel and is time consuming task.

Furthermore, it is normally required that the two cameras be

Focal axes

Common field of view

Field of Depth

Focal Length f
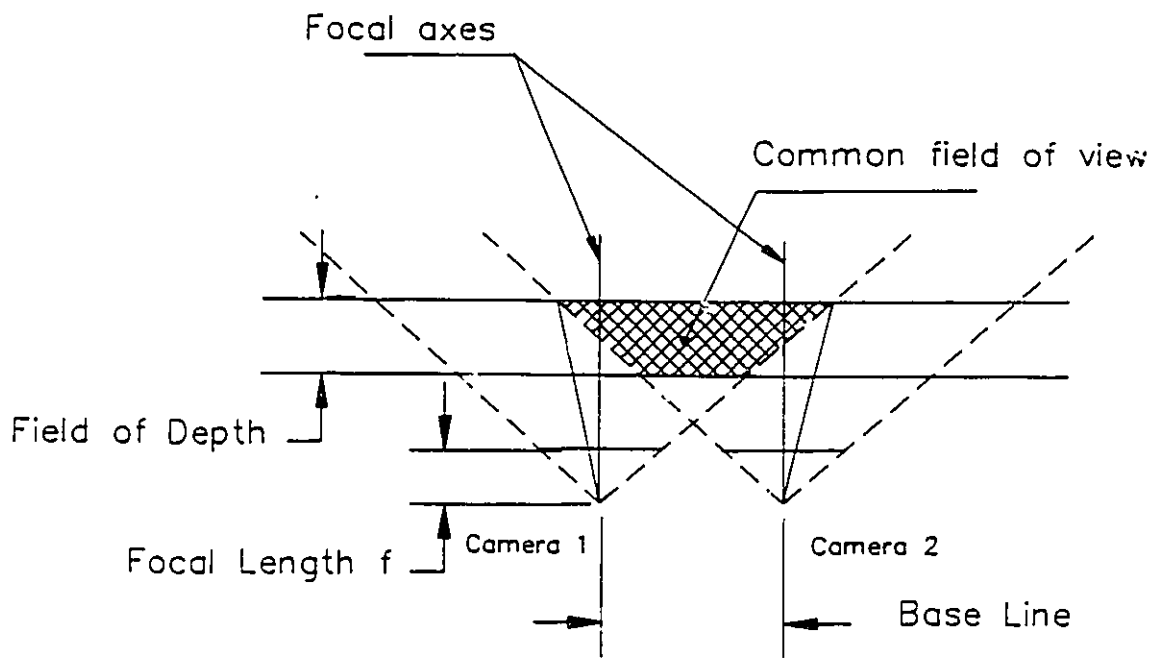
Camera 1

Camera 2

Base Line

Fig. (5.1). Common field of view of parallel axes approach.

at some angle to obtain a closer view of some features   for high accuracy measurements.

The alternative is to have no specific geometrical constraint and position the two cameras to have the maximum field of view as shown in Fig. (5.2)

However, this introduces a new set of problems that requires a different approach. Such problems include:

1. The object viewed by one camera will have a different projection on the other imaging plane due to the different viewing angles.

2. Points projected on one horizontal line may not be projected on the same horizontal line in the second imaging plane which is an assumption utilized by most of the techniques represented in the literature.

In this chapter a new approach for determining the point-of - correspondence in stereo images with no geometrical constraint using mathematical morphology will be introduced, several examples will then be given to illustrate the results obtained applying the proposed procedure.

In this approach the cameras are first calibrated to determine the extrinsic parameters of each camera (using the calibration procedure introduced in Chapter 2); i.e., the transformation from 3-D world coordinate to camera 2-D coordinate system.
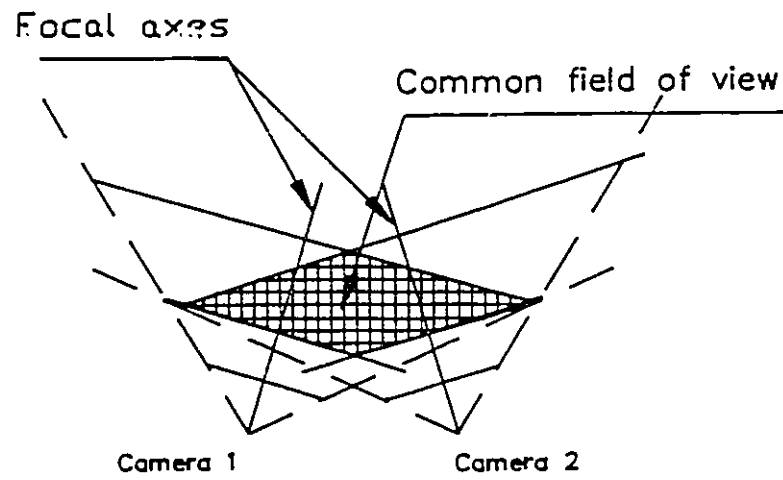
Fig. (5.2). Shows the common field of view of converged
two cameras.

The transformation matrices for the left camera image plane to the world coordinate are obtained. Using the blur-minimum edge operator introduced by Lee, Haralick and Shappiro]42], edges in both images are obtained. The point-of-interest is then selected and with the knowledge of the depth of field a window in the other image is calculated within which the point corresponding to the selected point should be found. A structure element is extracted from the first image, with the center of the element being the selected point, and mapped through a transformation in the second image plane to compensate for any distortion in the shape due to the different viewing angles. The image in the second image is eroded with the transformed structure element over the defined range to obtain the corresponding point. The advantages of this approach can be summarized as follows:

1. Morphological operators are not computationally complex and could be efficiently applied on a machine vision system that has special support for gray-scale morphology.

2. No geometrical constraints are required, which makes it suitable implementation in high accuracy close range measurements for industrial applications.

5.2 General Steps for Determining the Point-of-Correspondence

The proposed procedure for determining the point-of-correspondence can be summarized by the following steps:

1. Pre-processing performed on both images:

   a) Edge-strength detection using the blur-minimum operator.

   b) Automatic thresholding.

2. Select a candidate point in the first image.

3. Find the neighborhood in which the corresponding point can be found in the second image.

4. Extract a structure element from image 1 with the candidate point being the center of that element.

5. Using the calibration parameters, transfer the extracted element from camera 1 to camera 2 plane.

6. Use the transformed element and, over the neighborhood range determined at step 2, perform an erosion operation to find the corresponding point.

In the following section each of the above steps will be discussed in detail.

5.3 Pre-processing

The first step in the procedures is to perform a two-stage pre-processing operation. Stage one is to obtain edge-

strength image. In our implementation the blur-minimum edge operator was utilized. Since edge detection was covered in an earlier chapter, only the second stage involving thresholding will be discussed in this section.

Automatic Thresholding:

An iterative thresholding selection technique due to Trussel [48] based on the maximum separation between classes is utilized to obtain a binary image of the edge strength image. The iterative equation is given by:

$$t_{\ell+1} = \frac{\sum_{i=0}^{t_\ell} n(i) i}{\sum_{i=0}^{t_\ell} n(i)} + \frac{\sum_{i=t_{\ell+1}}^{N} n(i) i}{\sum_{i=t_{\ell+1}}^{N} n(i)} \qquad (5.1)$$

where

$n(i)$ = sum of gray-levels having the value i in the image

t_ℓ = threshold level at ℓ^{th} iteration.

N = maximum gray-level in the image for an 8-bit digitizer ($N=255$).

t_0 = $N/2$

convergence is reached when $t_\ell = t_{\ell+1}$

In applications where objects and background could be separated through thresholding the original image, the morphological edge strength operator need not be used.

5.4 Defining the Search Neighborhood

Narrowing down the search area in the second image for a match to the selected point in the first serves two purposes. First it reduces the possibility for false matches. Second it reduces the time required for match finding if the search is to be performed over the entire image.

It is typical of an optical system to have a specific depth of field within which objects will have to be positioned in order to perform accurate measurement. It is also true that this depth of field will vary from one application to the other. However, in a specific application that depth of field is considered constant and can be defined by Z_{upper} and Z_{lower}, where Z_{lower} is the closest allowable distance for the object to the optical system and Z_{upper} is the farthest distance the object can be located from the optical system as shown in Fig. (5.1) and Fig. (5.2).

Once the system has been calibrated, the point with coordinate (x', y') on the imaging plane of the right camera can be related to the world coordinate by the following:

$$
\begin{bmatrix} \alpha_4 - x' \\ \alpha_8 - y' \end{bmatrix} = \begin{bmatrix} \alpha_9 x' - \alpha_1 & \alpha_{10} x' - \alpha_2 & \alpha_{11} x' - \alpha_3 \\ \alpha_9 y' - \alpha_5 & \alpha_{10} y' - \alpha_6 & \alpha_{11} y' - \alpha_7 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \qquad (5.2)
$$

With α's and x', y' known, the equation (5.2) represents an under-determined set of equations (i.e., 2 equations in 3 unknowns with the unknowns being X, Y, Z).

However, since a prior knowledge of the depth of field was assumed and the values for Z_{upper} and Z_{lower} are known, Eq. (5.2) can be solved by substituting the value of Z to obtain X_{upper}, Y_{upper}, then by substituting the value of Z_{lower} to obtain X_{lower}, Y_{lower}.

The significance of the two points (X_l, Y_l, Z_l) and (X_u, Y_u, Z_u), where "l", "u" stands for lower and upper respectively, can be better understood by the following definition. For a point in 3-D space to have a projected coordinate (x', y') on the 2-D image plane and to satisfy equation (5.2), the point has to lie on the line joining the two points (X_u, Y_u, Z_u) and (X_l, Y_l, Z_l), assuming no lens distortion. Fig. (5.3) shows an illustration of the above definition.

Once the line $(X_u, Y_u, Z_u), (X_l, Y_l, Z_l)$ is defined, it can be related to the second camera through

$$
y'' = \frac{\beta_8 + \beta_5 X + \beta_6 Y + \beta_7 Z}{\beta_9 X + \beta_{10} Y + \beta_{11} Z + 1} \qquad (5.3a)
$$

$$x'' = \frac{\beta_4 + \beta_1 X + \beta_2 Y + \beta_3 Z}{\beta_9 X + \beta_{10} Y + \beta_{11} Z + 1} \qquad (5.3b)$$

Equation (5.3) is to be solved twice; once by substituting (X_u, Y_u, Z_u) to obtain (x'', y'') and a second time by substituting (X_1, Y_1, Z_1) to obtain (x''_1, Y''_1). The line joining (x''_u, Y''_u) and (x''_1, Y''_1) is the 2-D projection of the line (X_u, Y_u, Z_u) and (X_1, Y_1, Z_1) on the image plane of camera 2 and since the point with the projected coordinate (x', y') on the camera 1 plane lies on that line, then its projection on the camera 2 plane must lie on the line (x''_u, Y''_u) and (x''_1, Y''_1). Thus the line joining (x''_u, Y''_u) and (x''_1, Y''_1) defines the search neighborhood for the corresponding point to (x', y'). The procedure can be summarized in the following 3 steps:

1. Scan the image of camera 1 to determine an edge point (x', y').

2. From (x', y'), Z_u, Z_1 and α's solve the solution given in (5.2), obtain (X_u, Y_u) and (X_1, Y_1).

3. From (X_u, Y_u, Z_u) and (X_1, Y_1, Z_1) and β's of camera 2, and utilizing Eq. (5.3) obtain (x''_u, Y''_u) and (x''_1, Y''_1), which defines the search neighborhood.

5.5 Structure Element

Since the object is viewed by the two cameras from different viewing positions, approaches based on local

connectivity of the zero-crossing and matching corresponding to points on bases of the similarity of the patterns such as the approach suggested by Kim et al. [3] can not be utilized and do not prove useful in this case.

Therefore, a new approach had to be developed to deal with this problem. In this approach, since the object is projected on the imaging plane of camera 1 and the system has been calibrated, i.e., the parameters of the homogeneous transformation matrices are known, a good estimate of the 3-D object coordinate can be obtained. This 3-D coordinate through the second transformation matrix can be re-projected on the imaging plane of camera 2. This can be accomplished by assuming that a section (structure element) of the image plane of camera 1 with the point of interest at the center is extracted. Now let $(x'_{i,j}, y'_{i,j})$, where $i=1,2,\ldots,N$ and $j=1,2,\ldots,N$ and N is the Kernel of the element, be the coordinates of the pixels of the structure element on camera plane 1. Each pixel (x',y') is related to the world coordinate system through the homogeneous transformation of Eq. (2.1) which can be re-written in the form:

$$
\begin{bmatrix}
\alpha_4 - x' - (\alpha_{11}x' - \alpha_3)Z \\
\alpha_8 - y' - (\alpha_{11}y' - \alpha_7)Z
\end{bmatrix}
=
\begin{bmatrix}
\alpha_9 x' - \alpha_1 & \alpha_{10}x' - \alpha_2 \\
\alpha_9 y' - \alpha_5 & \alpha_{10}y' - \alpha_6
\end{bmatrix}
\begin{bmatrix}
X \\
Y
\end{bmatrix}
\qquad (5.4)
$$

It is known from Computer Graphics theory [45] that the viewer distance from the object does not affect the projection angle, however it affects the projected size of the object on the projection plane. A principal that will be utilized is the procedure of producing the structure element.

The procedure for producing the structure element can be summarized in the following two steps:

1. For each point of the structure element with coordinate (x',y') extract from camera 1 plane, find the X,Y coordinate in 3-D space by solving Eq. (5.4), and substituting Z_{upper} for Z.

2. Using the X,Y,Z_{upper}, calculate for each point in the element its coordinate (x'',y'') in camera 2 plane using Eq. (5.3).

Applying the Computer Graphics principal, the distance Z does not affect the viewing projection but affects the size of the object projected (i.e., the closer the object to the projection plane the larger the size of the projected image and vice versa). Therefore, we can assume that the object is at Z_{upper} and obtain the X,Y corresponding to each point of the element. Utilizing X,Y,Z_{upper} each point is then re-projected on the image plane of camera 2 using the transformation matrix of camera 2. This would result in obtaining a structure element that has the same viewing angle as the object on camera 2 and if not equal in size, it will be smaller due to the substitution of Z_{upper} for the Z
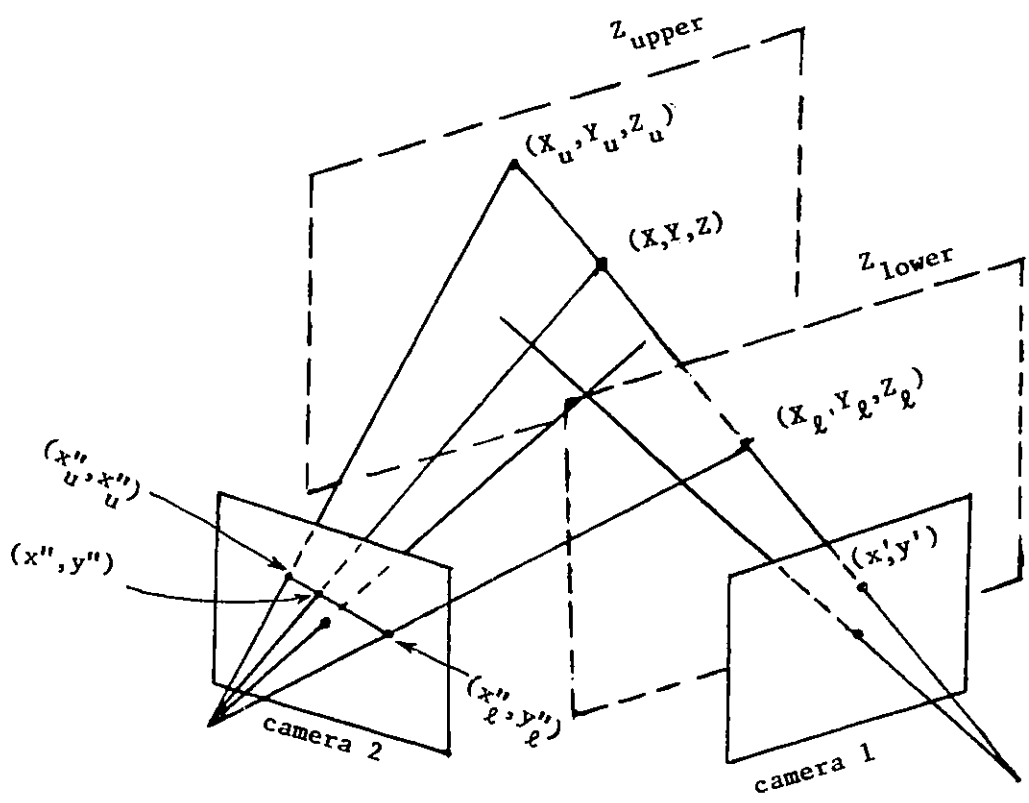
Fig. (5.3). The projection of the line (X_ℓ, Y_ℓ, Z_ℓ), (X_u, Y_u, Z_u) on camera 2 plane (x''_ℓ, y''_ℓ), (x''_u, y''_u).

value. This down sizing of the object is desired for the erosion operation as will be further explained in the next section.

5.6 Locating the point-of-correspondence

Treating images and shapes as point sets in n-dimensional cartesian space (n=2 for binary images), the problem of shape or pattern recognition is converted to the problem of detecting the occurrences of specific set within the search field.   This, as was shown earlier, is closely related to the operation of erosion discussed in Chapter 4.

The Crimmins and Brown theory [37], introduced earlier, stated that "The shape A in window W occurs in the image I at and only at the following locations:

$$[I \ominus A] \cap [\bar{I} \ominus (W-A)] \tag{5.5}$$

where $W-A = W \cap \bar{A}$.

The reason for the complement section in Crimmins and Brown's theory is to avoid the false recognition of objects of a similar shape yet smaller in size.   For our purpose the structure element was obtained with the objective of "if not equal, it is smaller" in size yet maintains the same zero-crossing connectivity.   Since the search for correspondence is performed over a limited neighborhood, only the erosion of the structure element "A" and the image of camera 2 "I":

I $\ominus$ A

over the domain of the neighborhood, results in the locus of the point where A fits inside I.

The locus of the structure element was selected to be the point of interest in camera 1 image, the result of the erosion is the corresponding point in camera 2 plane.

Since the structure element extracted from image 1 could contain spurious edges not necessarily present in the equivalent feature in image 2, plus the mapped feature within the structure element could be slightly distorted from the feature sought in image 2, the erosion is carried out as follows:

     1.   The structure element is scanned over the neighborhood search, i.e., at least the center of the structure element should be inside the neighborhood.

     2.   Over every region covered, the number of mismatch points between the structure element and region is counted.

     3.   The location in image 2, at the center of the neighborhood support of the structuring element, at which the minimum number of mismatch points is located provides the point-of-correspondence to the candidate point in image 1.

It should be clear that the above procedure is related to the

erosion definition given earlier.

Example 1

To illustrate the procedure and its accuracy in determining the point-of-correspondence, the setup shown in Fig. (5.4) is devised.

A robot arm picks up a calibration target from a pre-defined location and places it in view of the two cameras at a plane horizontal to the work area. The calibration target is a flat dark surface with nine white discs. The centroids of the discs in the two images obtained from the left and right cameras are calculated by first thresholding, followed by border following each disc. The robot arm then places the target at two more different horizontal planes. The reference axes used are shown in Fig. (5.5). The origin is at the center disc in the middle plane. The twenty-seven world coordinate points and their corresponding values in the two images are used to calculate the α's and β's.

The z location of the plane closest to the work area is z_{upper} and the z location of the plane closest to the cameras is z_{lower}.

For the setup used $z_u = -50$ mm, $z_l = +50$ mm, and the viewing area is ≈ 150 mm x 150 mm. The stand-off distance of the two cameras is about 150 cm.

Once the calibration is completed the robot places the target back in its original location, and a number of objects of varying heights are placed in view of the two
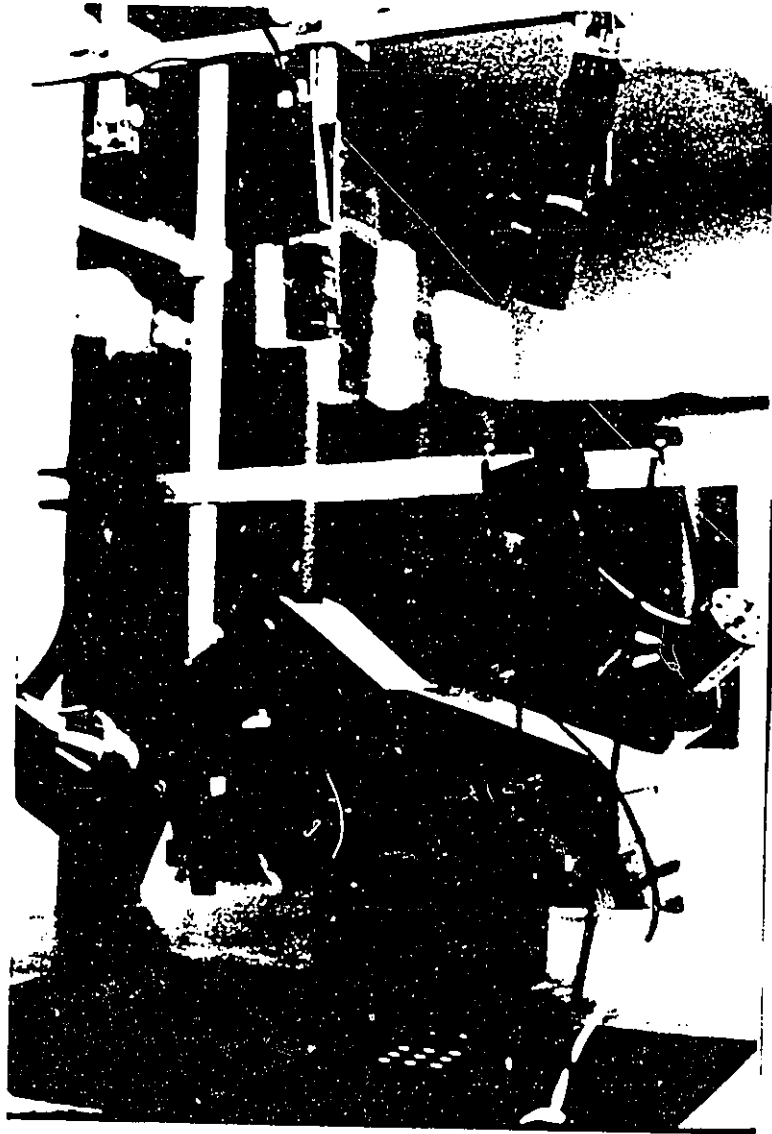
Fig. (5.4). Setup for calibration and for testing
the point-of-correspondence procedure.
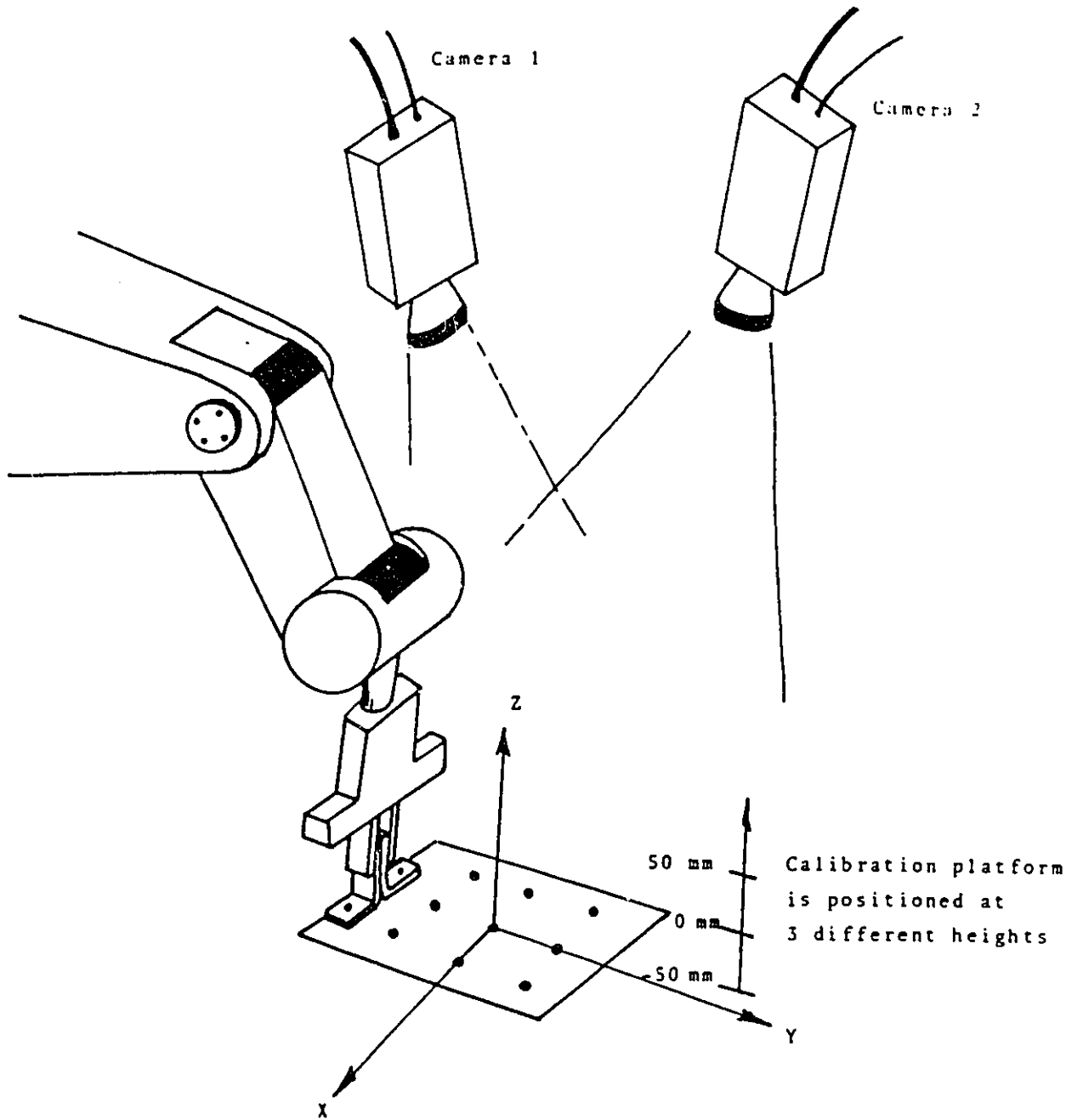
Fig. (5.5).   Calibration set-up.

cameras (see (5.6)). Figures (5.7) and (5.8) show the images as viewed by the left and right cameras, respectively. Figures (5.9) and (5.10) show the edge-strength of both images. Nine points are selected at the edges of the edge-strength image obtained from the left camera, and are used in the proposed procedure to determine their corresponding points in the right camera. A structure element of size 15x15 pixels (the images are of size 512x512 pixels), is used and centered about each of the 9 candidate points in turn. Figure (5.11) shows the structure element centered about one of the candidate points. Figure (5.12) shows the mapped image of the structure element to the right camera. Figure (5.13) shows the calculated region for the neighborhood search. Figure (5.14) shows the candidate points and Fig. (5.15) shows the corresponding points obtained through the mathematical morphology procedure.

From the corresponding points we can calculate the position of each point in space relative to the world coordinate axes using the over-determined set of linear equations.

$$
\begin{bmatrix}
\alpha_1 - \alpha_9 x' & \alpha_2 - \alpha_{10} x' & \alpha_3 - \alpha_{11} x' \\
\alpha_5 - \alpha_9 y' & \alpha_6 - \alpha_{10} y' & \alpha_7 - \alpha_{11} y' \\
\beta_1 - \beta_9 x'' & \beta_2 - \beta_{10} x'' & \beta_3 - \beta_{11} x'' \\
\beta_5 - \beta_9 y'' & \beta_6 - \beta_{10} y'' & \beta_7 - \beta_{11} y''
\end{bmatrix}
\begin{bmatrix}
X \\
Y \\
Z
\end{bmatrix}
=
\begin{bmatrix}
x' - \alpha_4 \\
y' - \alpha_8 \\
x'' - \beta_4 \\
y'' - \beta_8
\end{bmatrix}
\qquad (5.6)
$$
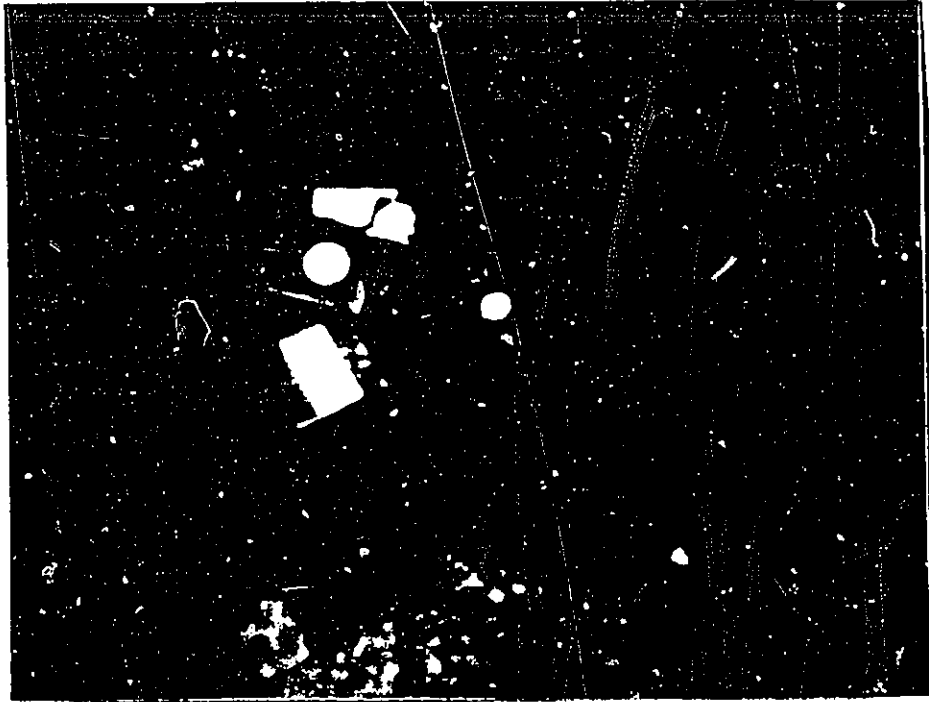
Fig. (5.6).   Objects of varying heights used to test
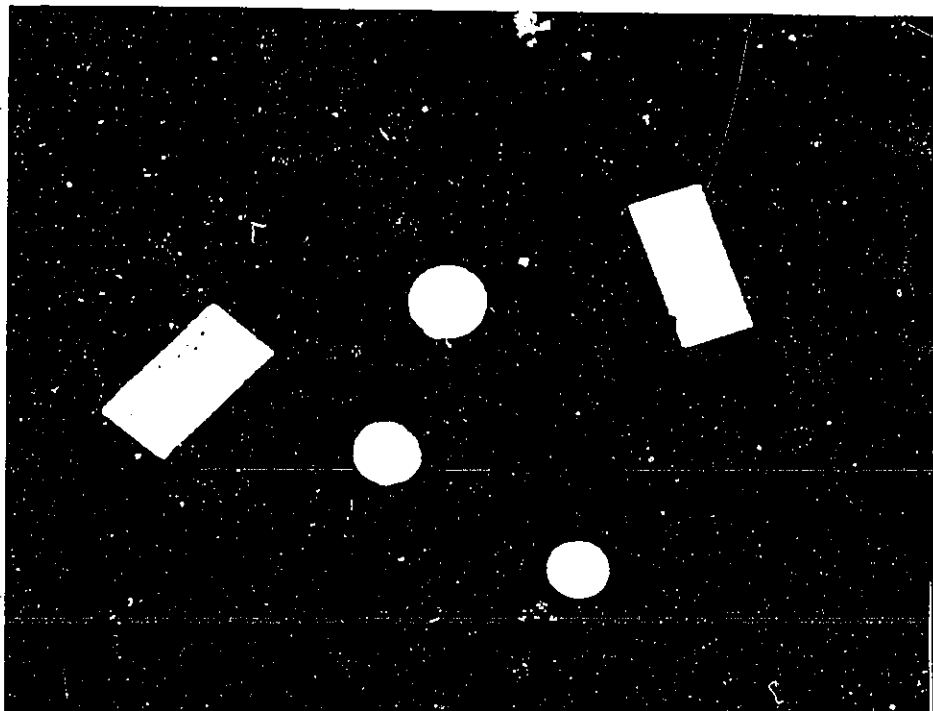the accuracy of the approach.
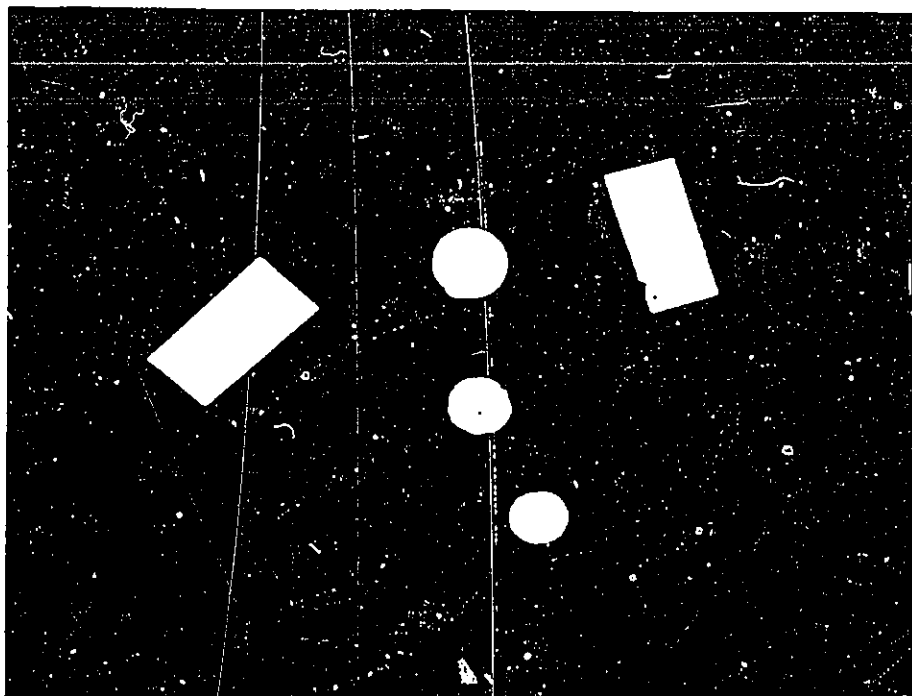

Fig. (5.7).   View from camera 1.
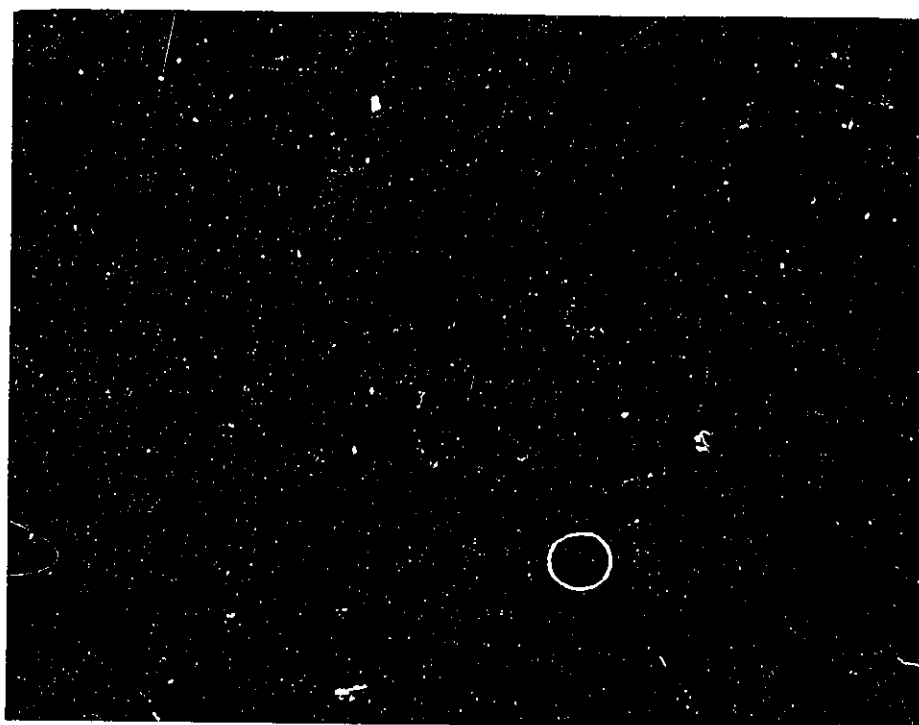
Fig. (5.8).   View from camera 2.

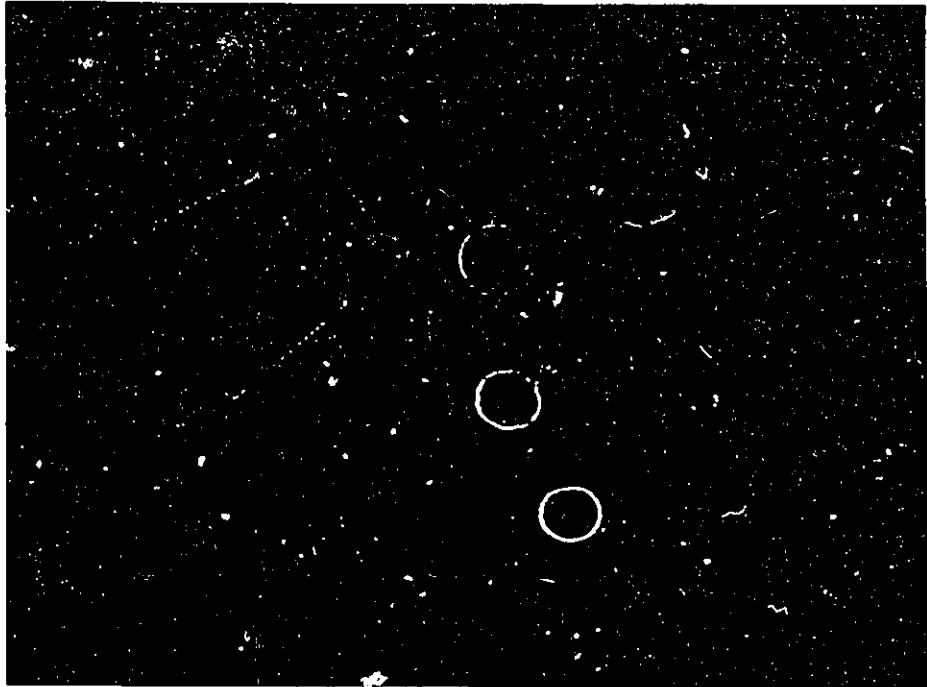Fig. (5.9).   Edge strength of image 1.

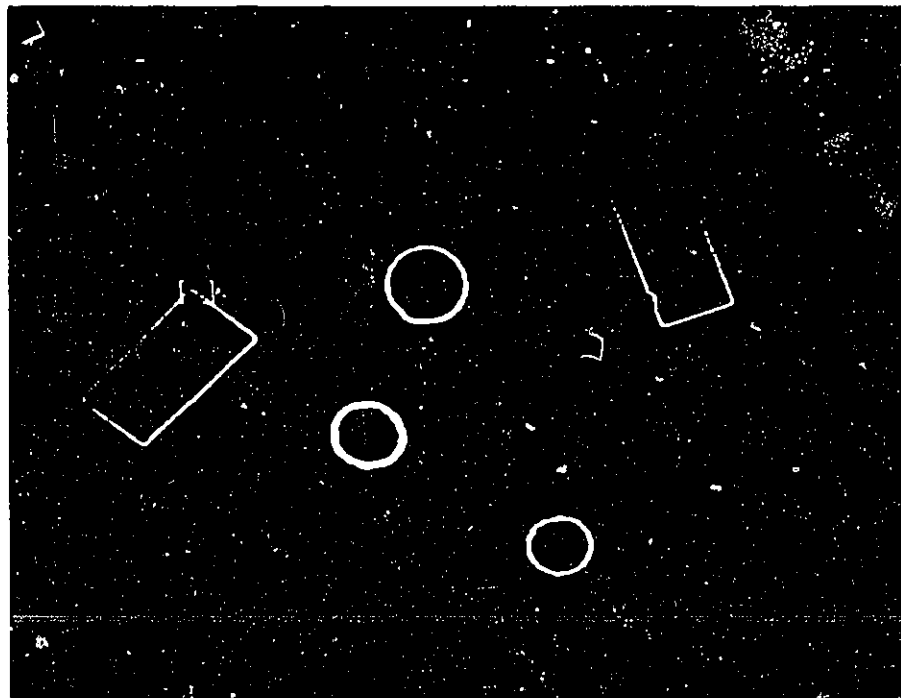Fig. (5.10). Edge strength of image 2.

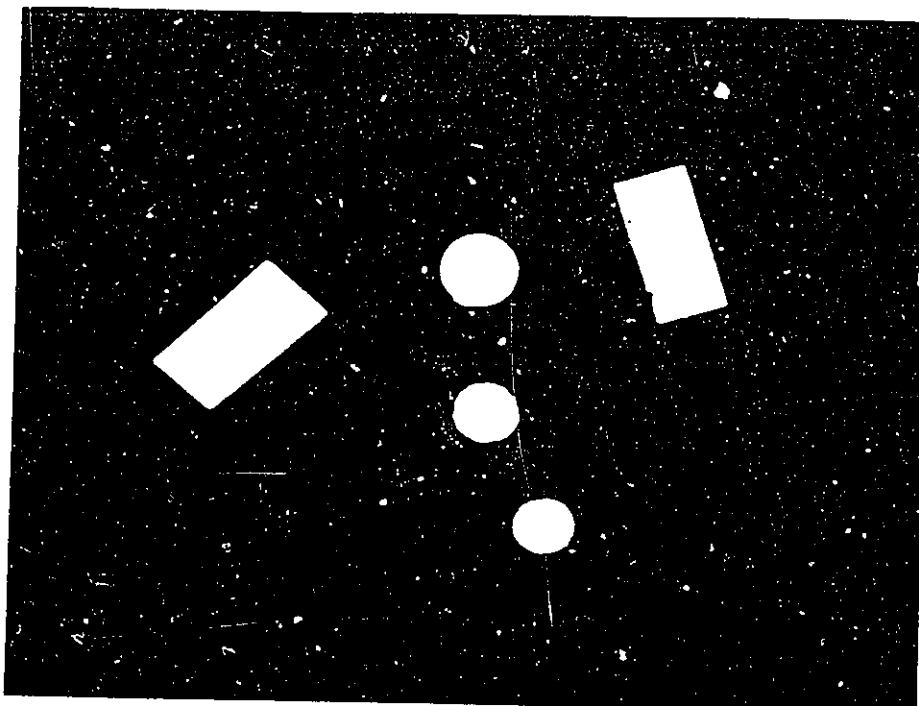Fig. (5.11). Forming the structure element about candidate point in image 1.

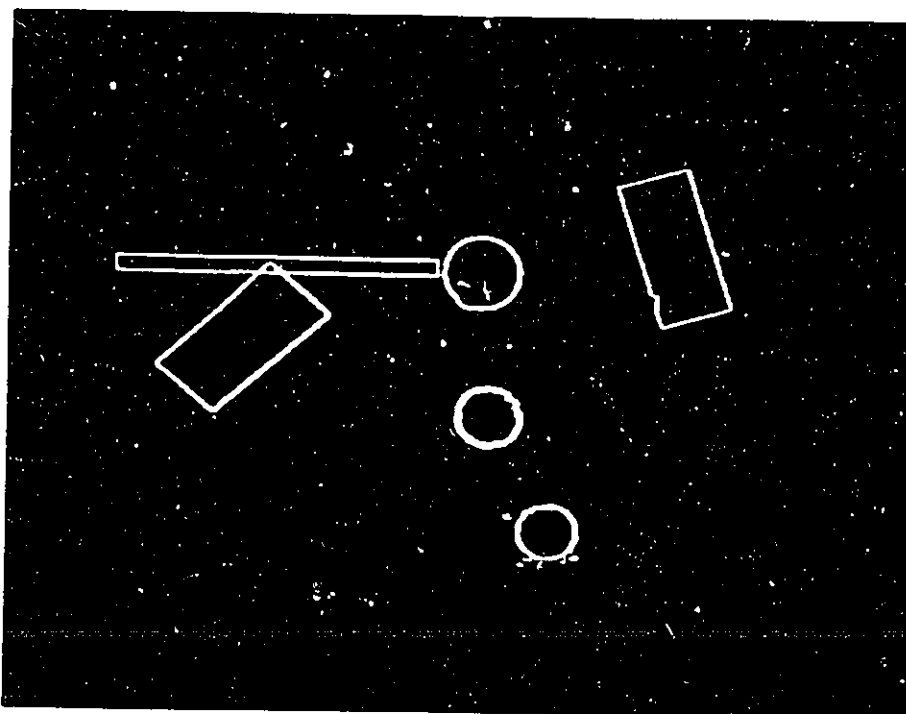Fig. (5.12). Mapped structure element in image 2.

Fig. (5.13). The search neighborhood.

The calculated Z-coordinate of each point is given in the following table:

Table 5.1

The Calculated Z-Coordinate

Point	Z
1	1.5 mm
2	1.34 mm
3	24. mm
4	24.48 mm
5	-4.67 mm
6	-4.44 mm
7	-24.06 mm
8	-23.22 mm
9	-25.90 mm

Since points 1 and 2 are on the same top surface of one of the objects, and so is 3 and 4, 5 and 6, 7 and 8 (Figs.(5.14 and 5.15)), it should be clear that the procedure for the point-of - correspondence was successful in obtaining the right solution.

The height of each object was measured accurately with a vernier having a digital readout. The height of each object is provided in Table (5.2) along with the results as obtained from Table (5.1).
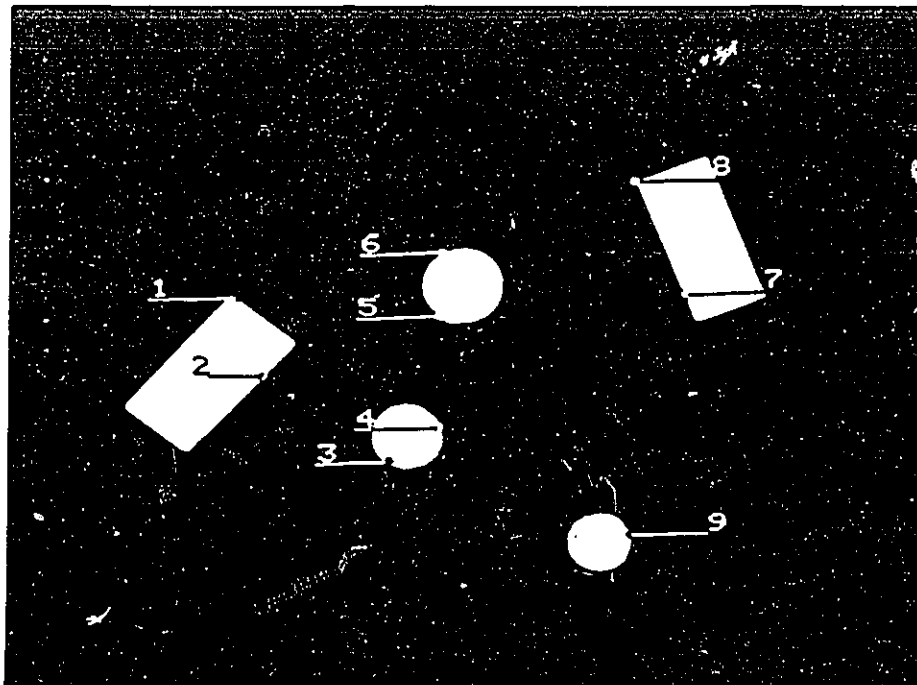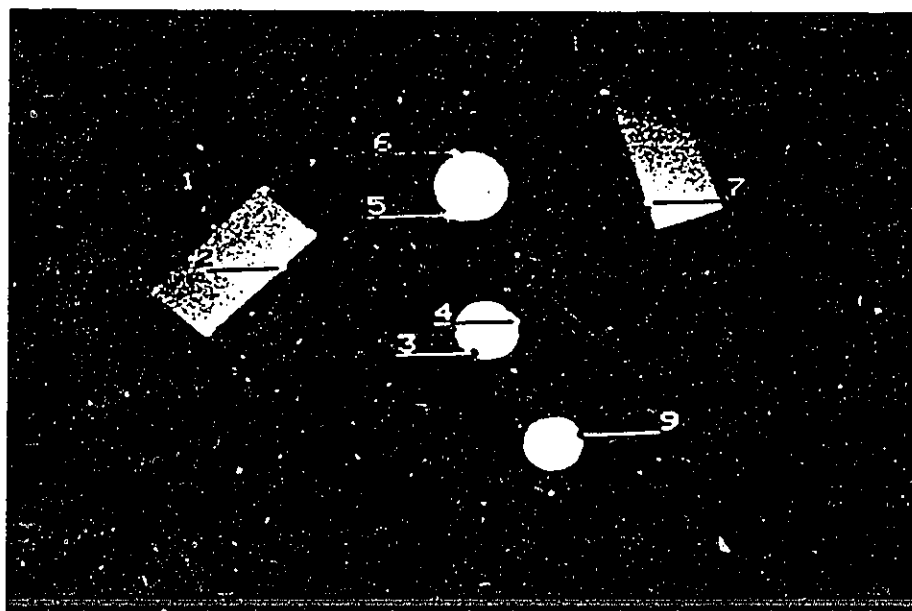
Fig. (5.14). Candidate points in image 1.

Fig. (5.15). Corresponding points in image 2.

Table 5.2

Actual Vs. Measured Z-Coordinate

Object labelled with the candidate points	Actual Height (mm)	Measured Height (Z + 50) mm
1 and 2	50.44	51.55/51.34
3 and 4	75.07	74. /74.48
5 and 6	44.18	45.33/45.56
7 and 8	25.19	25.94/26.77
9	25.06	24.1

(Note that the objects are placed at the work space which is
- 50 mm. from the defined origin).

Example 2

The same objects used in example 1 are utilized in this example to demonstrate the usage of a larger structure element. The images from both cameras are first threshold. The centroids of the white patches in the thresholded image of Camera 1 are calculated from the coordinates of their borders obtained from the border following algorithm. A structure element with its center being at the centroid of one of the targets in the edge strength image obtained from camera 1 is extracted (Fig. 5.16). The structure element should be large enough to encompass edges from the object. The structure element selected was 59x59. The purpose here is to determine the corresponding centroids in image 2.

The structure element is then mapped to image 2 (Fig. 5.17), and erosion of the neighborhood region (Fig. 5.18) with the mapped structure element follows. The above steps

are then repeated for all the centroids in image 1. Figure (5.19) shows the centroids derived through thresholding, and border following, and Fig. (5.20) shows the corresponding thresholds obtained from the mapping and morphological procedure.

To validate the results obtained, the centroids of the white patches in image 2 are obtained from the result of thresholding and border following, and compared to the results obtained from the morphological procedure. Table (5.3) provide the results thus obtained.

The results are almost identical (within an error of 1 pixel) testifying to the strength of the method.

Table 5.3

Centroids of the (x",y") Obtained From Border Vs. Morphological Procedure

L A B E L	Centroids (left camera)		Corresponding centroids (right camera). Morphological procedure		Centroid (right camera)calculated from the borders	
	x'	y'	x"	y"	x"	y"
1	132	-123	121	-130	121	-129
2	- 21	- 92	- 2	-107	- 2	-107
3	-174	- 16	-146	- 52	-146	- 52
4	- 59	31	2	9	2	0
5	58	125	39	105	38	105

Table 4 shows the calculated Z obtained through equation (5.6) using the corresponding centroids (x',y') and
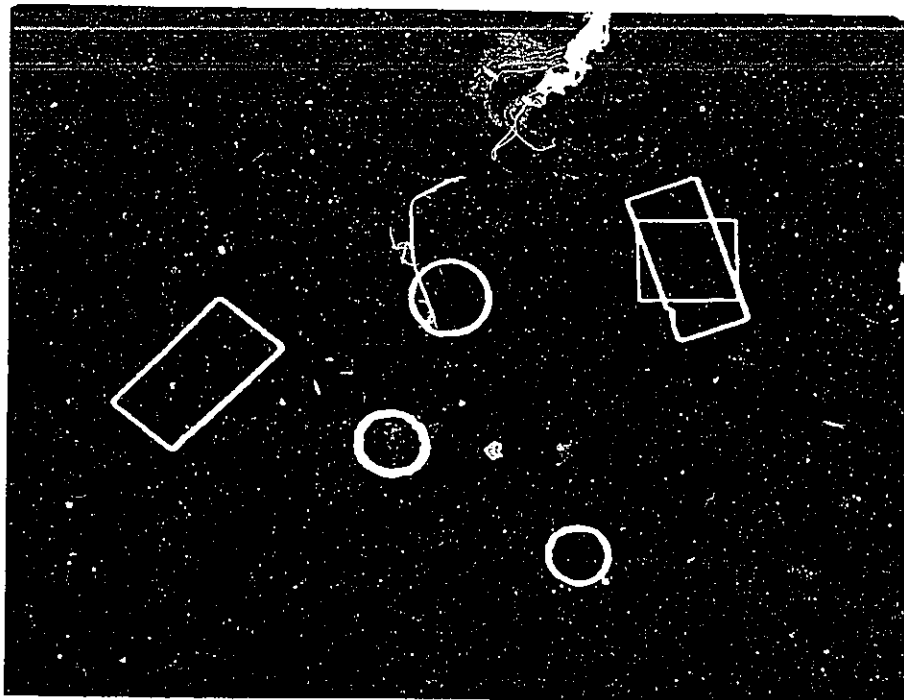
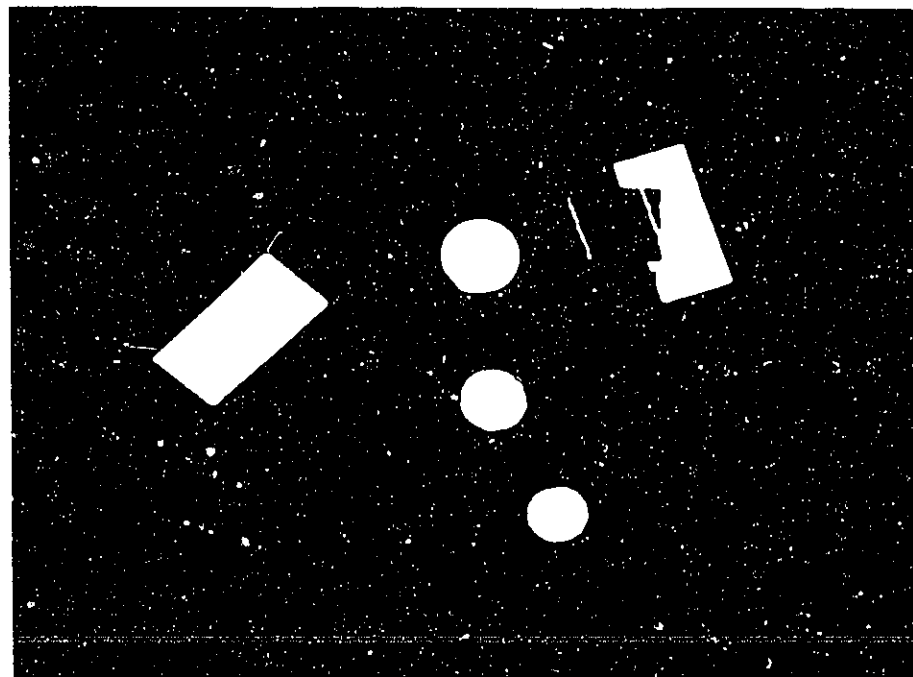Fig. (5.16). Selecting a structure element (60x60

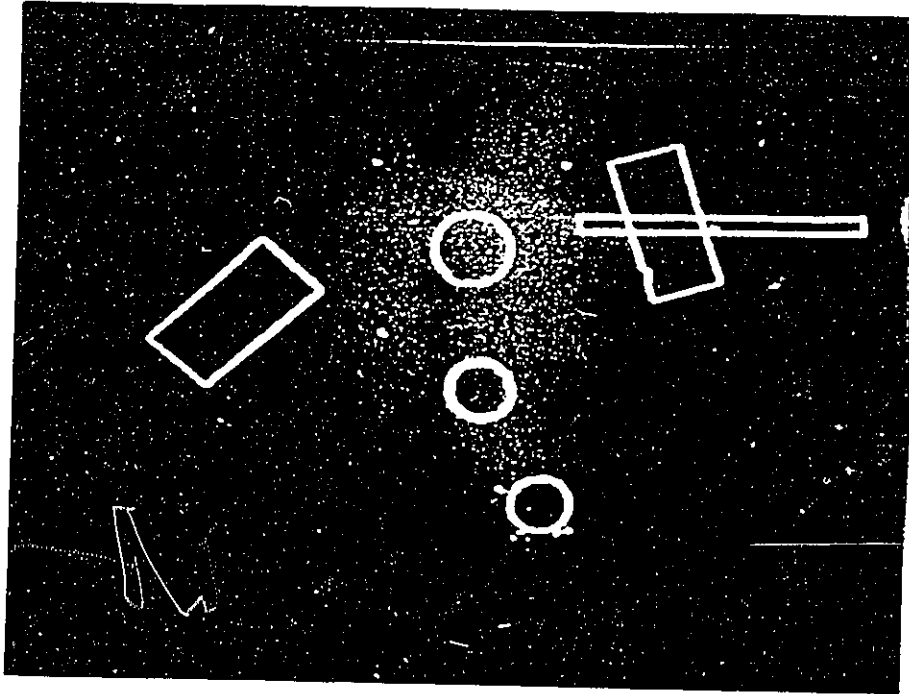Fig. (5.17). The mapped structure element.

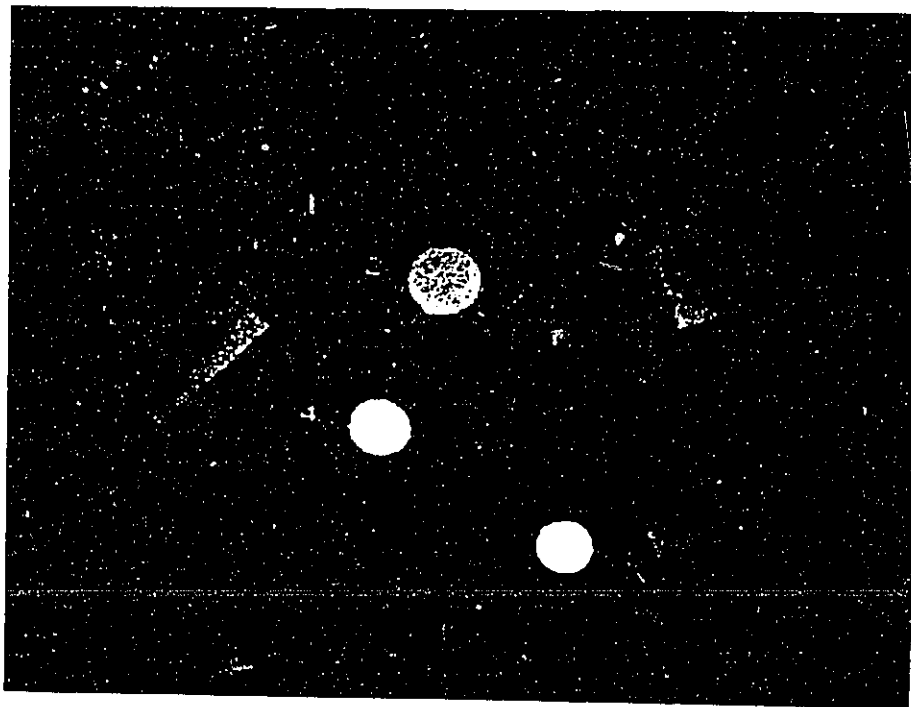Fig. (5.18).   The search neighborhood.

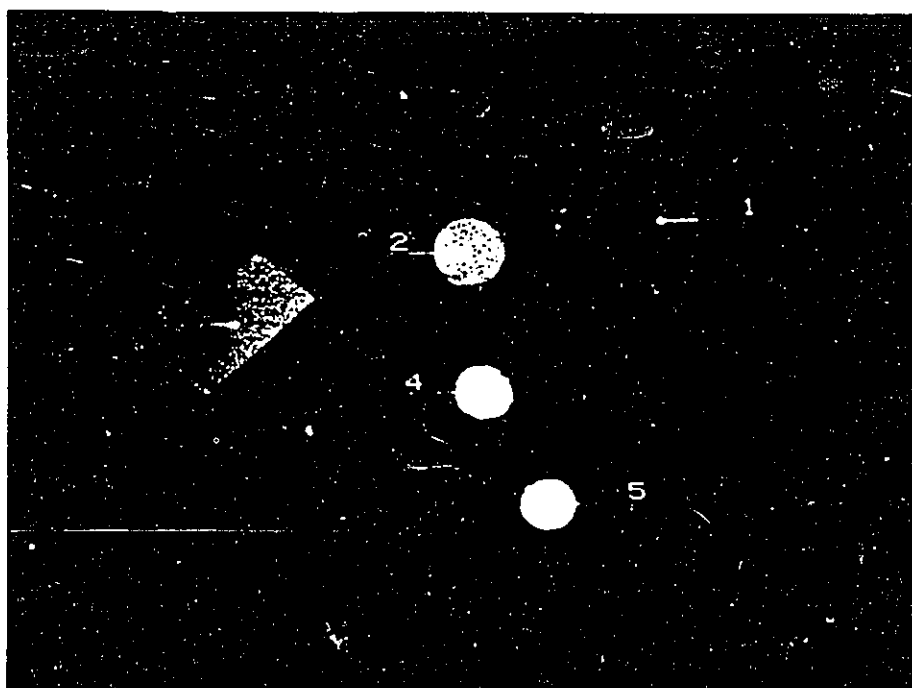

Fig. (5.19).   Centroid in image 1.

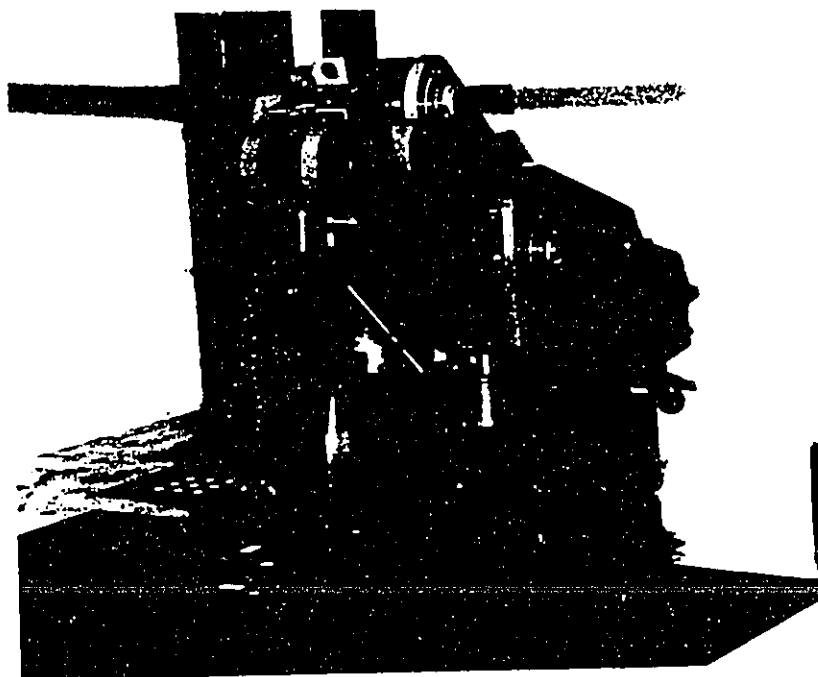Fig. (5.20). Corresponding centroids in image 2.

Fig. (5.21). Robot picking up one of the objects

(x",y") from the first two data columns of Table (5.3), and
the calculated heights versus the measured heights of each
object.

Table 5.4

The Calculated Z-Position for the
Values of Table 5.3

Points labelled	Z mm	(Z+50 mm)	Measured Height (mm)
1	-24.09	25.91	25.19
2	- 5.05	44.95	44.18
3	1.57	51.57	50.44
4	25.10	75.10	75.07
5	-26.06	23.94	25.06

The error in the calculated Z is not mainly due to the
point-of-correspondence calculation as can be seen from Table
3, but rather mainly due to other factors such as lens
distortion, error in robot-hand positioning and
repeatability, . . . , etc. Nonetheless, the maximum error
is less than 1.5 mm for a calibration range of 100 mm and a
stand-off distance of ≈1.5 m.

The values of the world coordinates (X,Y,Z) of each
centroid are provided to the controller of the robot through
a port in the computer, and the robot is programmed to pick-
up each object and place it aside. Figure (5.21) shows the
robot handling one of the objects.

Example 3

In this test example a spring on a base is provided in view of both cameras (see Fig. 5.23). Figures (5.24 to (5.30) are self explained. It should be obvious to the reader that point-of-correspondence procedure was successful for all the candidate points.

Example 4

To leave no further doubts a last example is presented (Fig. 5.3) which is basically three objects of different shapes and sizes. Figures (5.32) to (5.40) are self explained. Again it should be obvious to the reader of the success of the point-of-correspondence procedure.
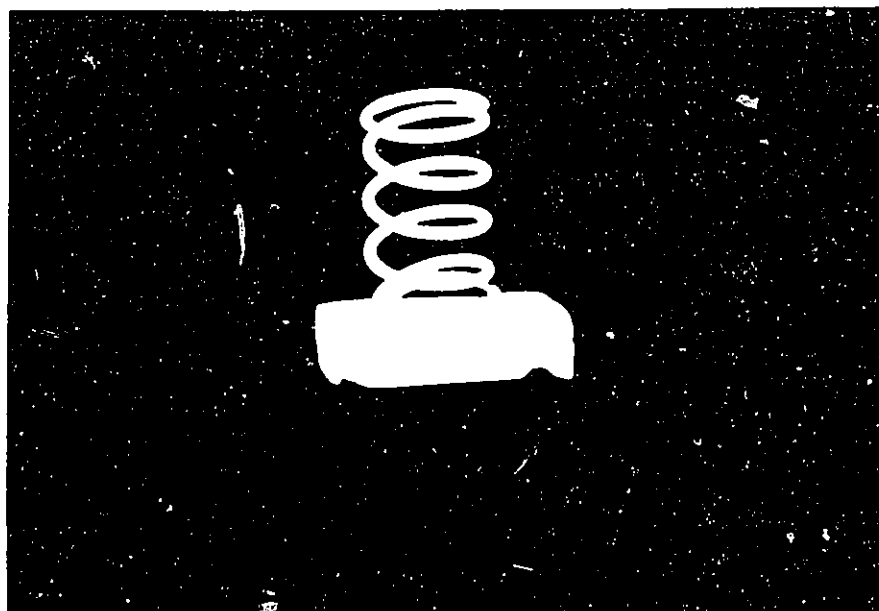
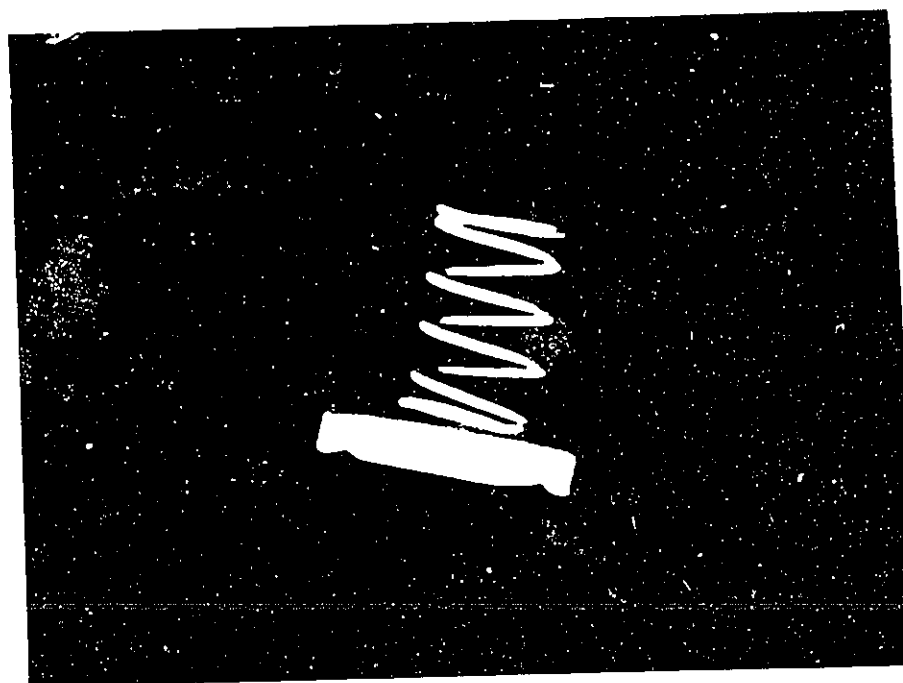Fig. (5.22). A spring on a base.

Fig. (5.23). Image from camera 1.
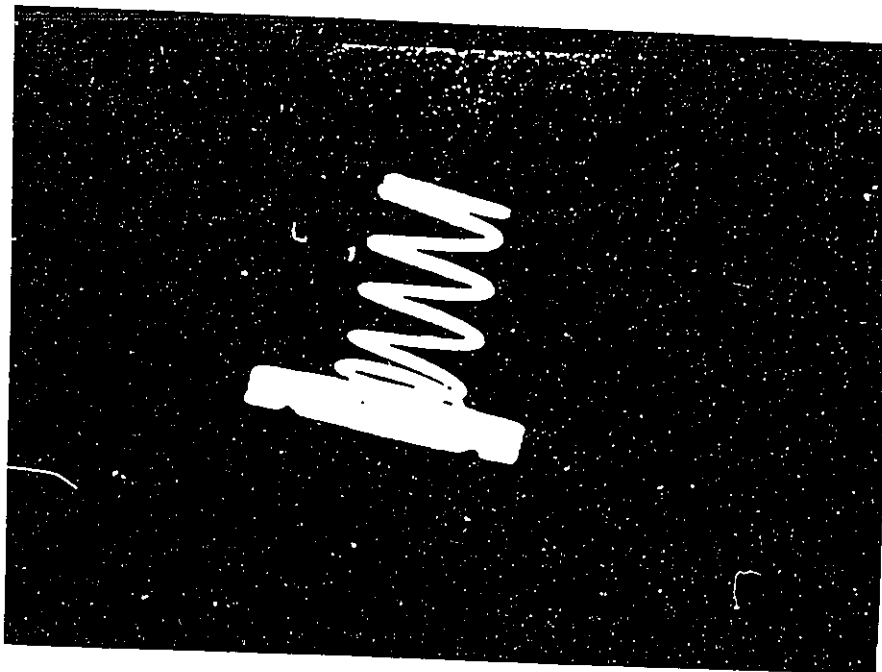
Fig. (5.24). Image from camera 2.

Fig. (5.25). A structure element surrounding a
candidate point.

Fig. (5.26). Mapped structure element.
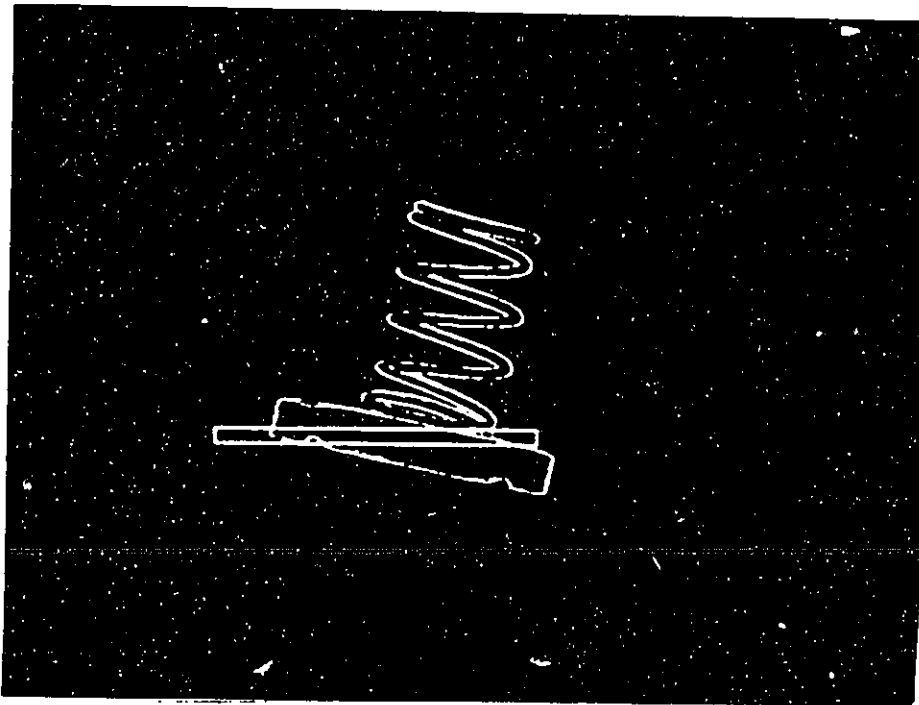
Fig. (5.27). The search neighborhood.
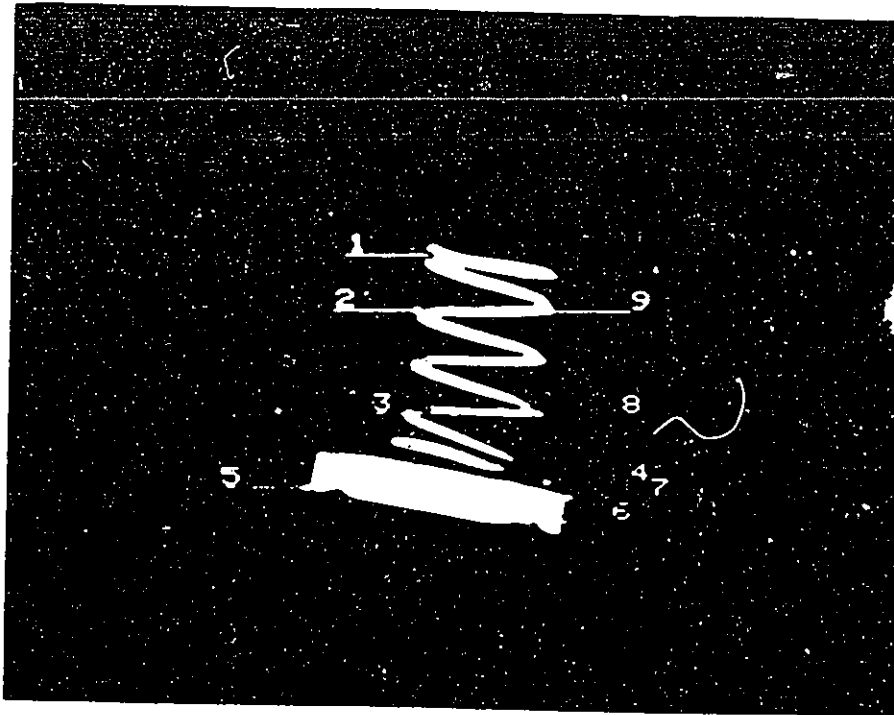
Fig. (5.28). Candidate points in image 1.
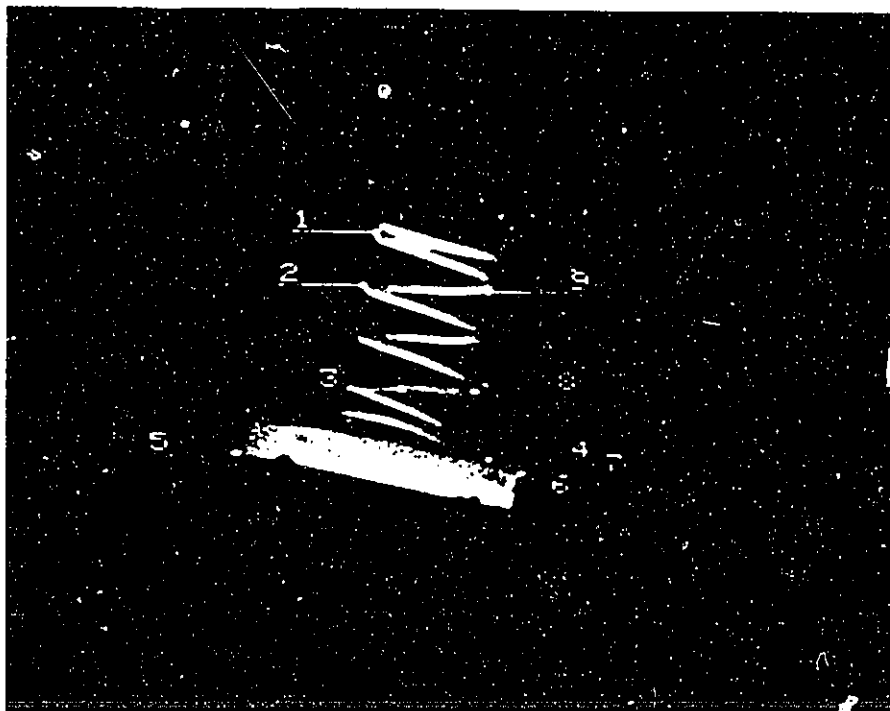
Fig. (5.29). Corresponding points in imate 2 obtained
through mapping and mathematical
morphology.

Fig. (5.30)   Image from camera 1.
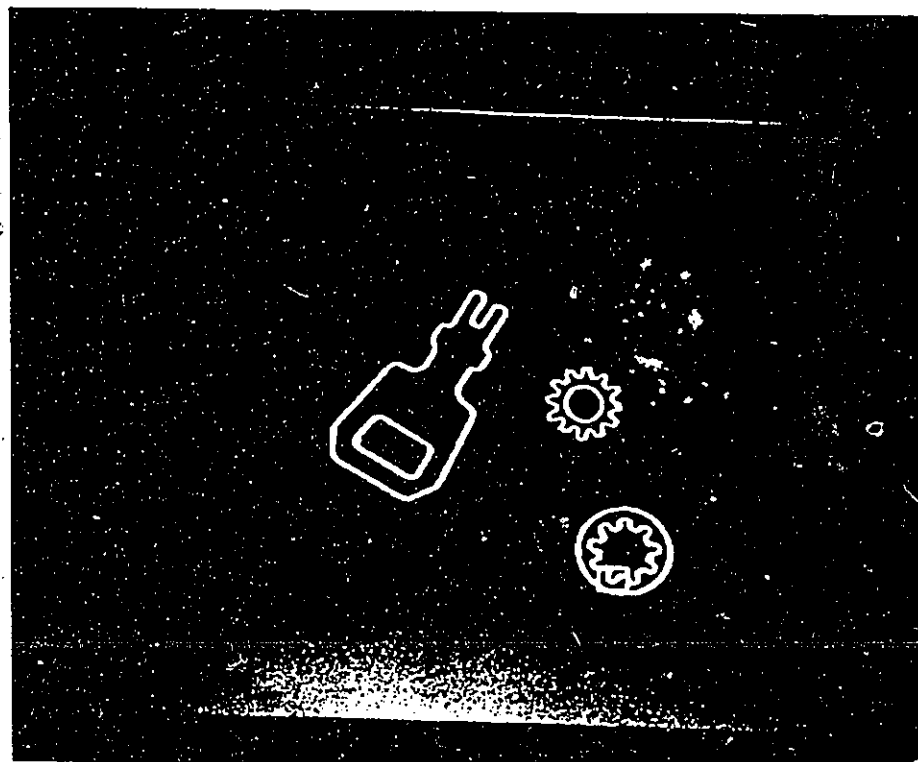
Fig. (5.31).   Edget strength and structure element
about a candidate point.
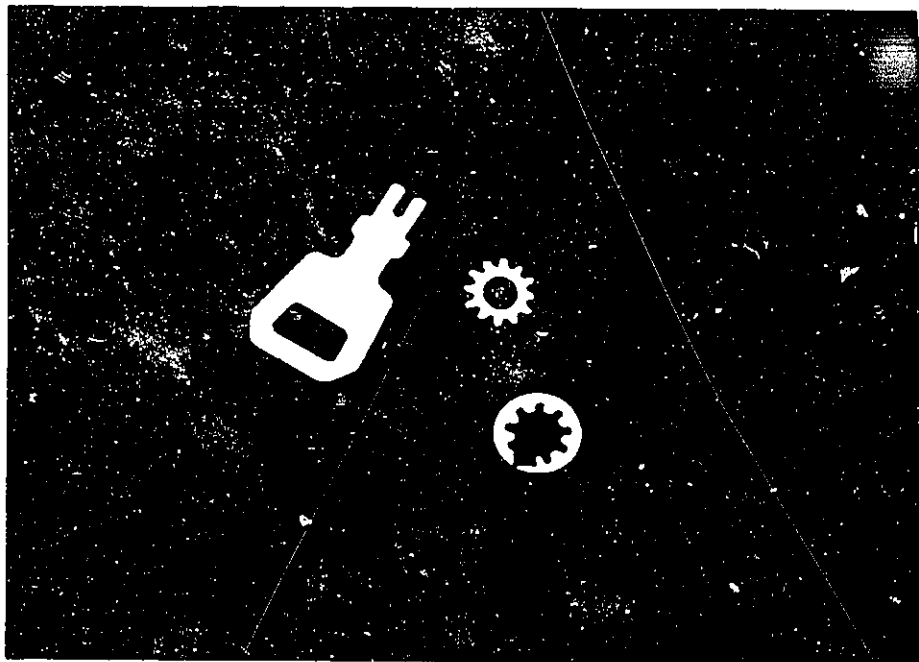
Fig. (5.32). Mapped structure element on image 2.
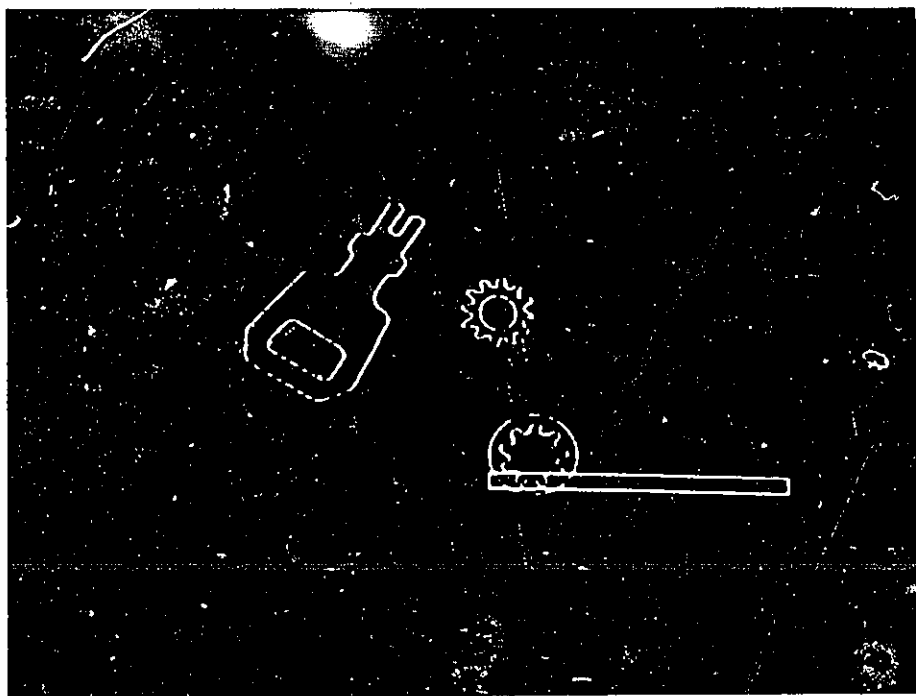
Fig. (5.33). The search neighborhood.
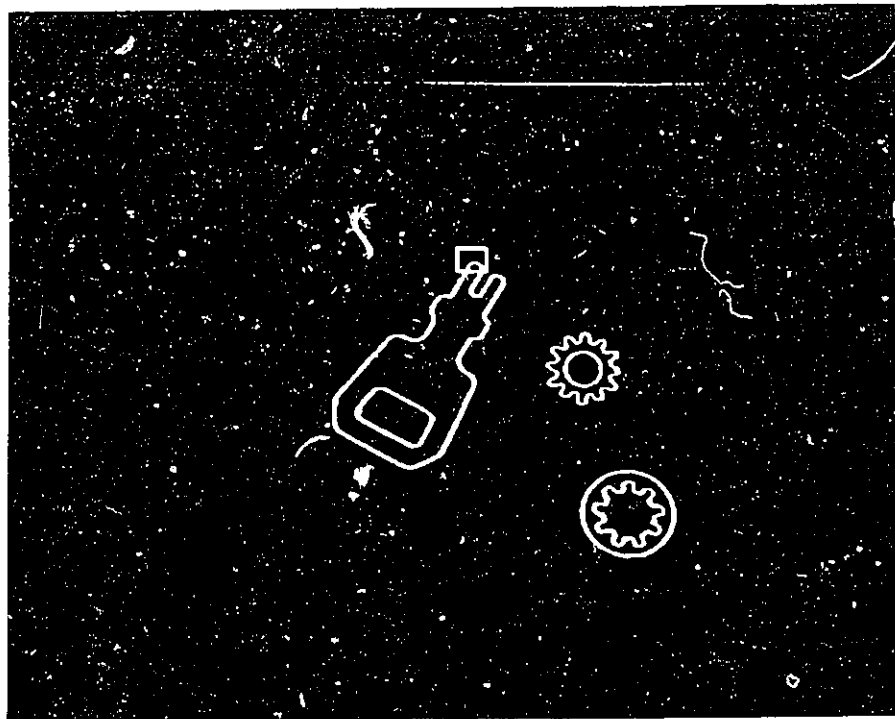
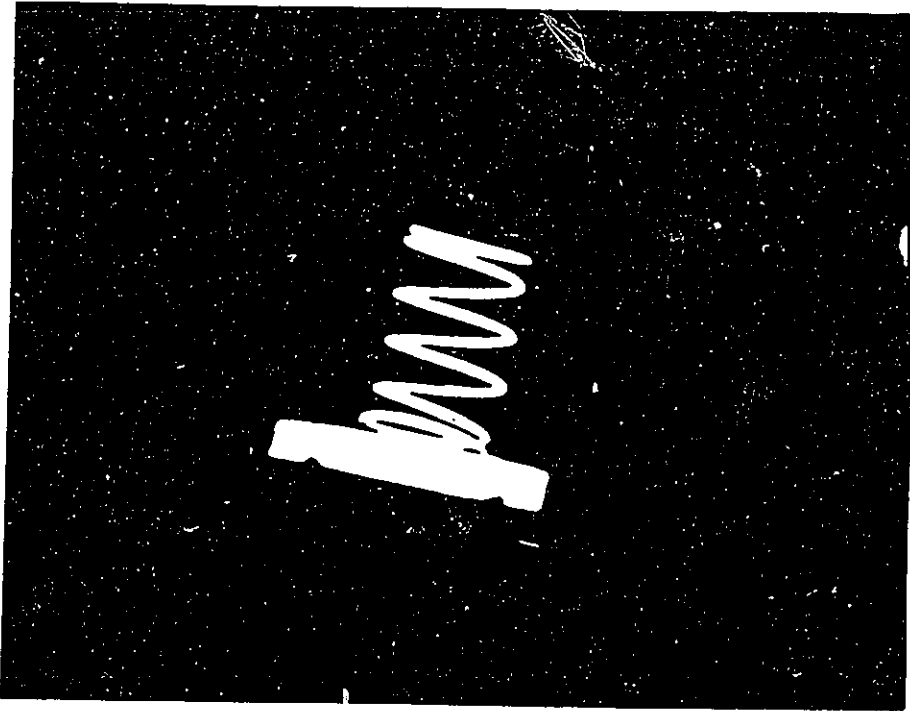Fig. (5.34). Structure element about a candidate point.
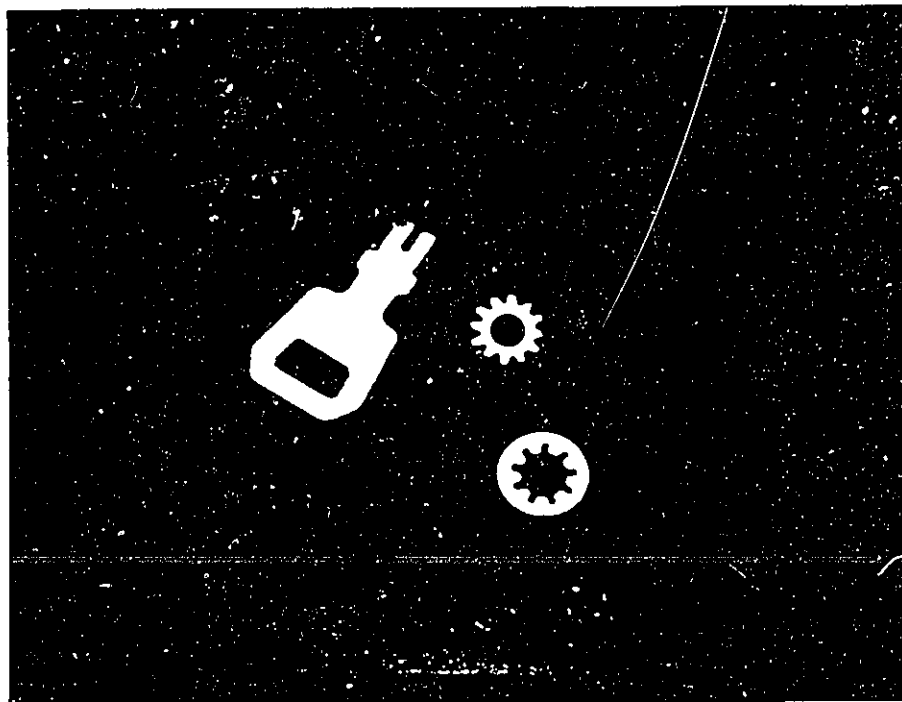
Fig. (5.35). Mapped structure element on image 2.

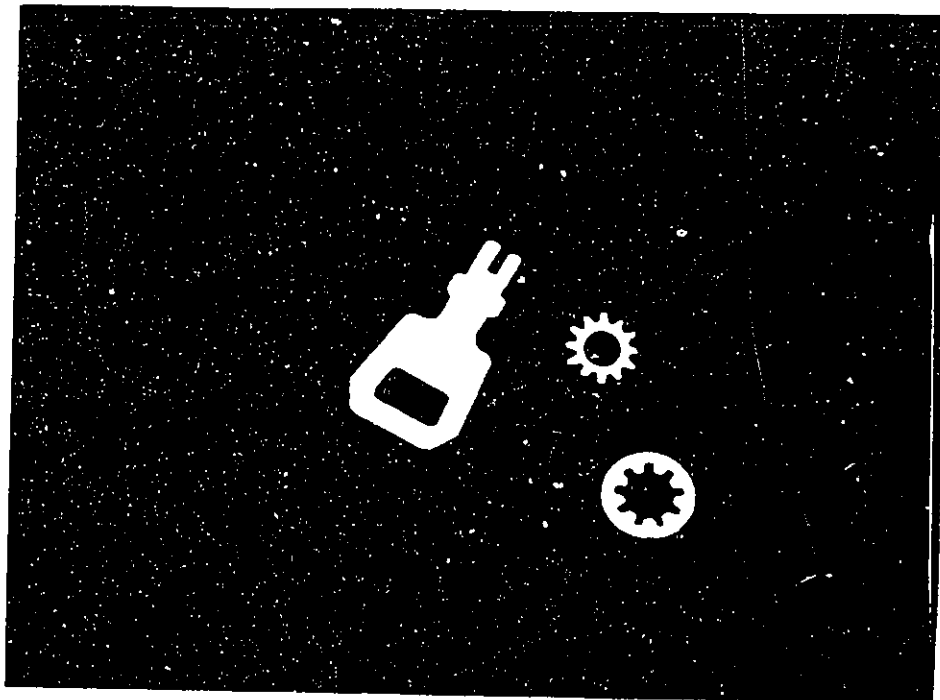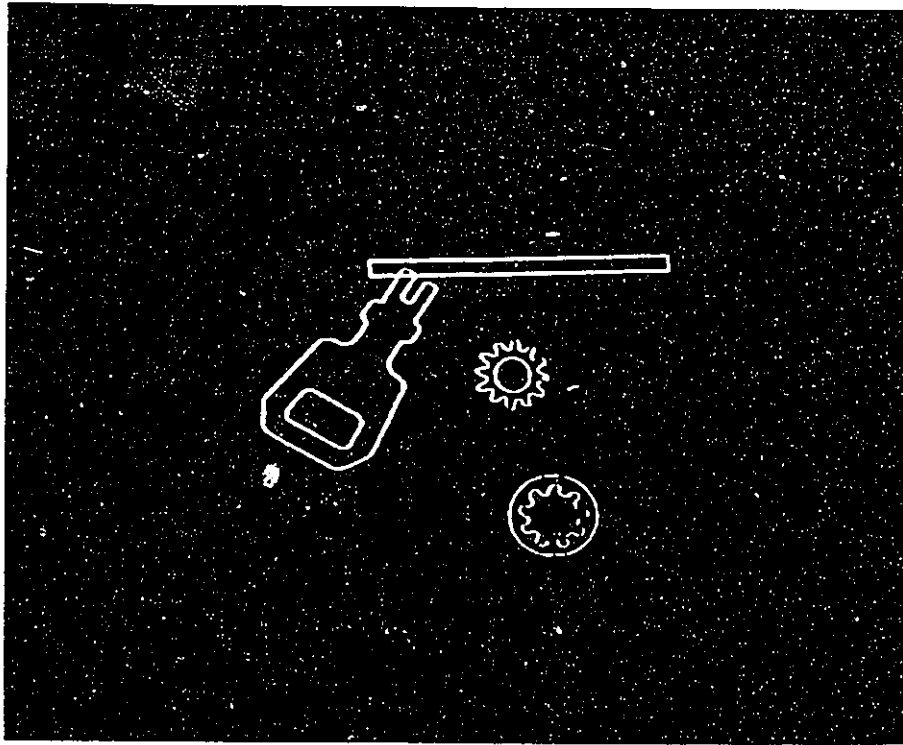Fig. (5.36). The search neighborhood.
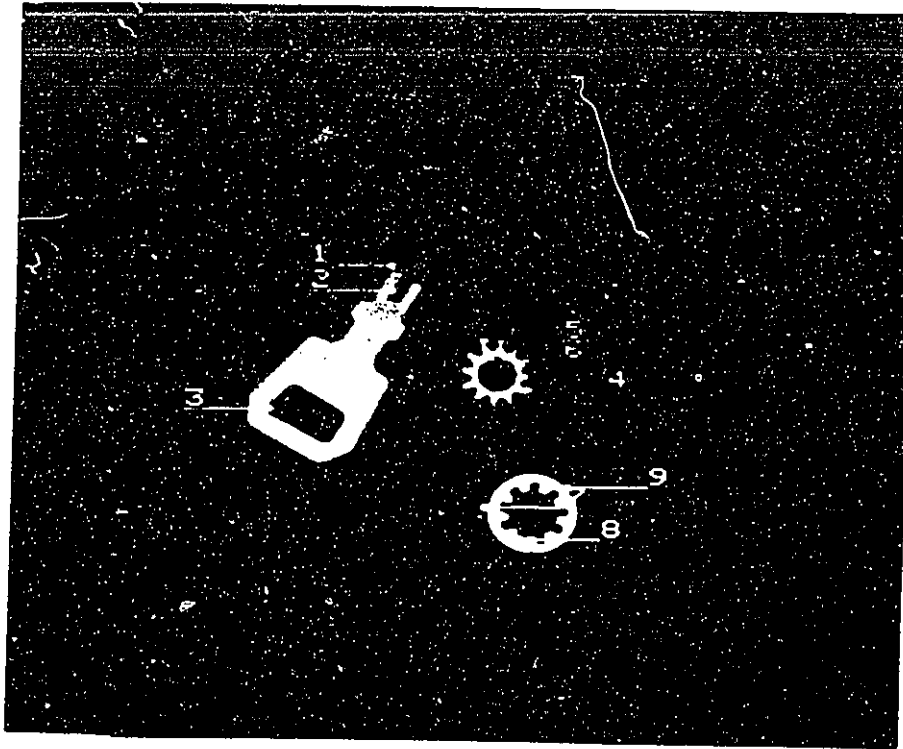
Fig. (5.37). Candidate points.
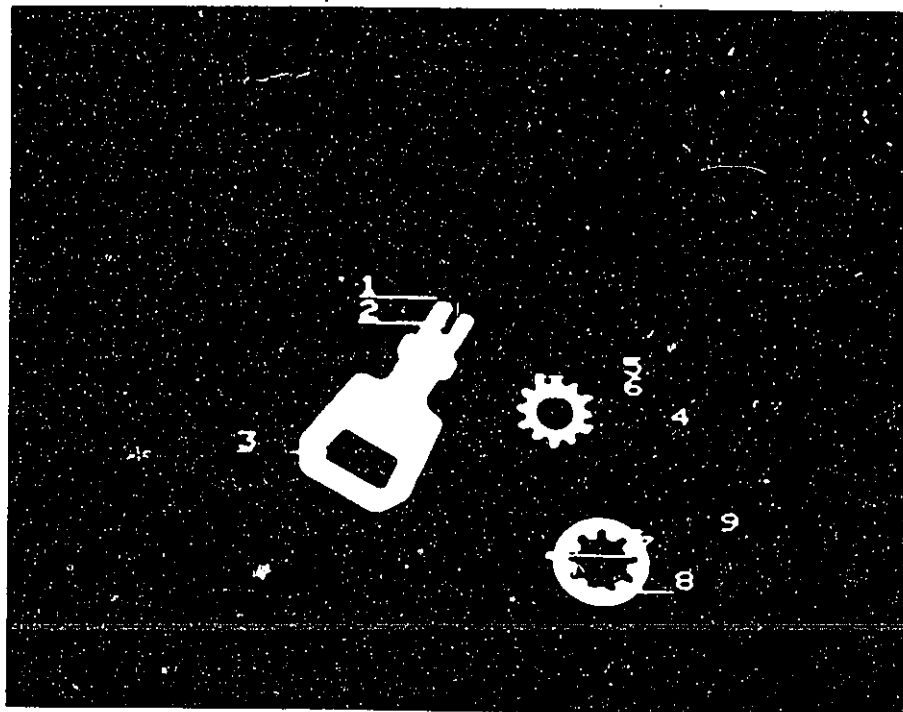
Fig. (5.38). Corresponding points.

5.7   Discussion and Conclusion

A   new   procedure   for   determining   the   point-of-correspondence was introduced. The procedure was then tested on   several   examples   covering various features and possible applications.   A test   setup including   a robot   arm and the vision system   was utilized   to locate   object position using the proposed procedure.   The accuracy   of the   procedure was evaluated and   shown to   be within   1.5 mm   for a calibration range of 100 mm and a stand-off distance of about 1.5 m.

The procedure utilized mathematical morphology operators for   finding   edges   and   then   for   determining the point-of-correspondence   in   the   second   image.   The   procedure was implemented   using   off-the-shelf   hardware   (IBM-AT,   Matrox frame buffer and two Pulnix cameras),   however,   hardware that supports morphological   operations is currently available and if utilized, the procedure   can be   implemented for real-time applications.

CHAPTER VI

CAMERA CALIBRATION WITH LENS DISTORTION

Abstract

Extending the camera calibration technique to include lens distortion is considered in this chapter. The lens distortion parameters representing radial and tangential distortion will be introduced to the camera calibration process. The effect of each parameter on the reduction of a specified error criterion will be utilized to illustrate the effect of these parameters. The initial value for solving the non-linear set of equations will be estimated by assuming no lens distortion and calibrating the cameras. The values obtained will then be used as initial guess with the lens distortion parameters estimated to be zero.

6.1 Introduction

Lens distortion is defined as the failure of a lens to image a straight line in object space as a straight line in image space and to maintain the same metric. Normally, the

resultant displacement of the image from that produced by a true projective transformation is referred to as distortion and is characterized by two distinct components.

One that is radially symmetric about the principal point and one that is symmetric along a line directed through the principal point. Terminology adopted by the photogrammetrist to describe these two components is radial and tangential distortion, respectively. Since distortion has to affect the position of the image point in the image plane, t, therefore represents an important factor that cannot be ignored in measurements when high accuracy is required.

6.2 Radial Distortion

Radial lens distortion is an aberration which causes a ray of light directed at the front nodal point to emerge from the rear nodal point changed in direction.

Radial lens distortion is illustrated in Fig. (6.1), where the ray directed from 0 along the optical axis passes through the lens undeviated in direction. The ray directed from A toward the front nodal point N should leave the rear nodal point N' undeviated in direction along the dashed line N'a'. The distortion, however, causes the emergent ray to travel along N'a. This places the image point too far out along a radial line from the axial point O. An outward displacement is considered a positive radial distortion. The direction of the ray BN is changed by distortion in such a
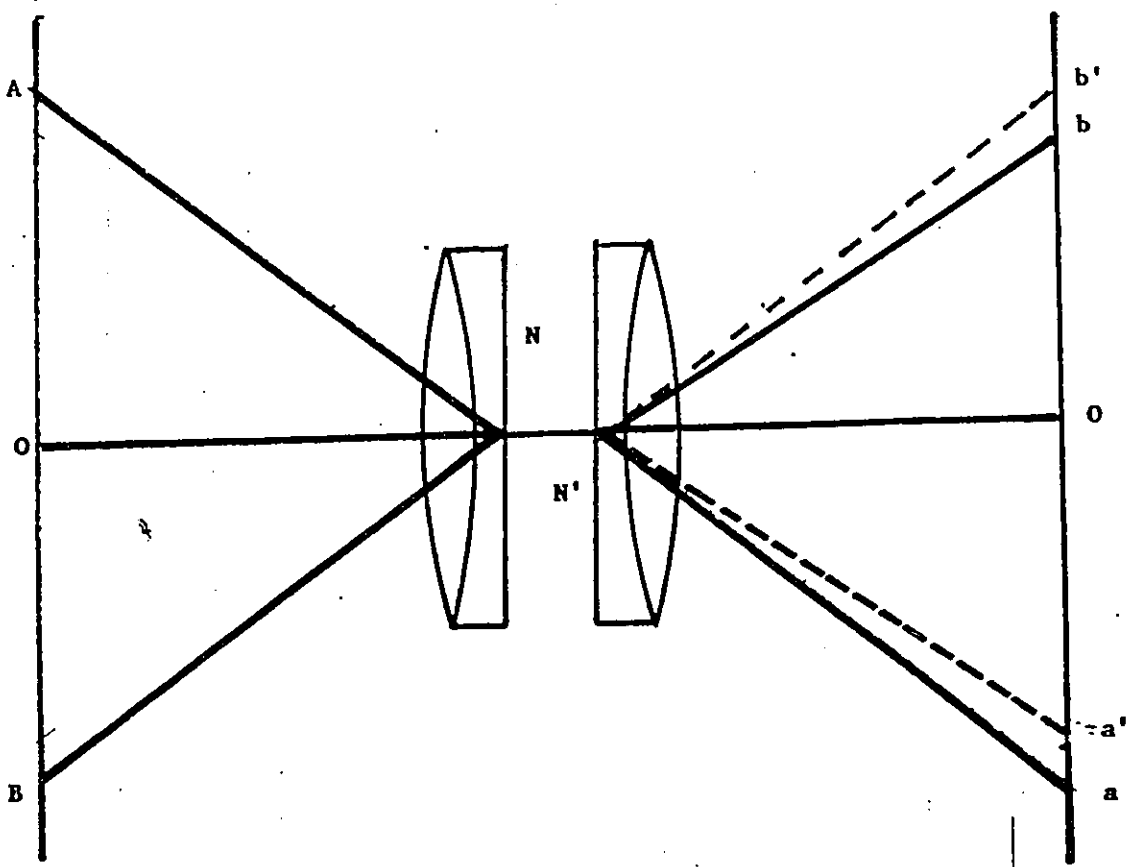
Fig. (6.1). Radial lens distortion.

manner as to cause the image point b to lie too far along a radial line toward the axial point   O.   This is a negative radial distortion.

Radial lens distortion for photogrammetric application has always been given in one of three ways.   One way is by representing the lens distortion by a curve in which the horizontal axis is the radial distance from the principal point   r   and the vertical axis representing the radial distortion value δ_r. However, since δ_r is a function of the radial distance "r" out from the principal point, this value must first be computed from

$$r = (x^2 + y^2)^{1/2}$$

in which x and y are the coordinates of the point with respect to the principal point.

A second approach to tabulate the values of the radial distortion δ_r as a function of the radial distance "r", the incremental value of "r" being chosen according to the shape of the curve.   An interpolation routine, based on the value of r for each point is then utilized to determine the corresponding value for δ_r .

A third method for presenting radial lens distortion is in the form of an odd-power polynomial of the form

$$\delta_r = K_o r + K_1 r^3 + K_2 r^5 + \ldots \qquad (6.1)$$
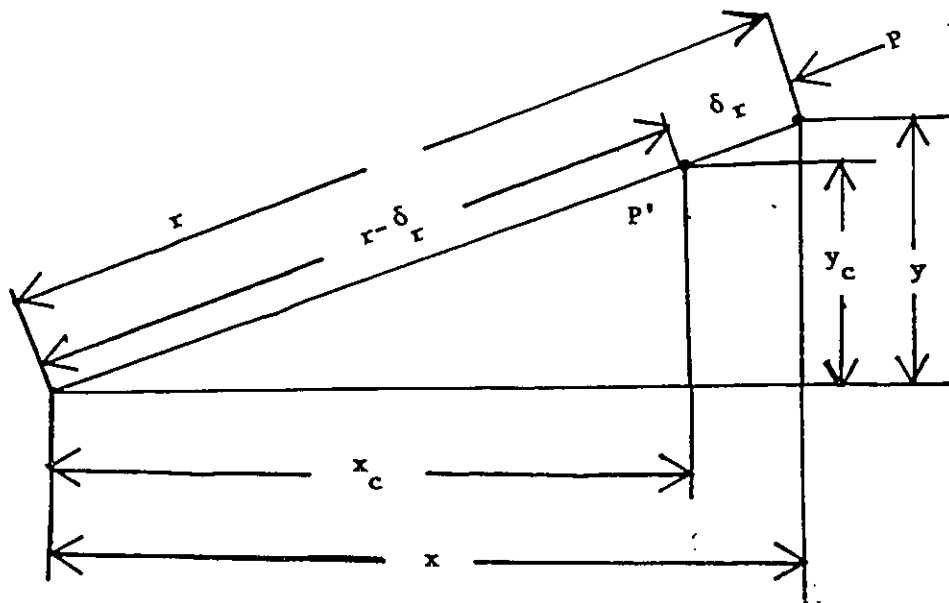
Fig. (6.2). Correction of coordinates for
radial distortion.

in which the K-coefficients are obtained by fitting the polynomial curve to the distortion data by the method of least squares. This form is the most convenient computer implementation.

After the value of the radial distortion of a point is determined, the co-ordinates of the point must be modified to eliminate the effect of this distortion. This can be accomplished by using the form

$$x_c = x(1 - \frac{\delta_r}{r})$$

$$y_c = y(1 - \frac{\delta_r}{r})$$

(6.2)

where x_c, y_c are the corrected values and x, y are the point location in relation to the principal point. The form given in Eq. (6.2) can be illustrated as shown in Fig. (6.2) where the point P' has been displaced outward (positive) to its image position P by radial distortion δ_r. The coordinates of P, uncorrected for line distortion are x and y. The corrected coordinates are x_c and y_c. Then by similar triangles,

$$\frac{x_c}{r-\delta_r} = \frac{x}{r} \quad \text{and} \quad \frac{y_c}{r-\delta_r} = \frac{y}{r}$$

(6.3)

Thus,

$$x_c = x(1 - \frac{\delta_r}{r})$$

$$Y_c = Y(1 - \frac{\delta_r}{r})$$

Then substituting Eq. (6.2) into (6.1) the following form is obtained:

$$x_c = x(1-(K_0 + K_1r^2 + K_2r^4 + \ldots))$$
$$Y_c = Y(1-(K_0 + K_1r^2 + K_2r^4 + \ldots))$$

$$(6.4)$$

The form that will be later used for the calibration procedure when extended to include lens distortion.

6.3 Tangential Distortion

The presence of tangential distortion was first noted by Pennington [50] when exploration of Multiplex models displayed unusual patterns, resulting in inaccuracies not attributable to radial distortion. Tangential distortion is the displacement of the image of a point normal to a radial line unlike radial distortion which is not a lens aberration but a consequence of imperfect fabrication and assembly of lens so that the centers of curvature of the individual element are not made to fall on a straight line. That results in a radial and tangential distortion of the image.

A commonly used model for correcting lens distorting for both radial and tangential is that developed by D. Brown [51].

$$\Delta x_j = \bar{x}_j [K_1 r^2 + K_2 r^4 + K_3 r^6 + \ldots]$$

$$+ [P_1 (r^2 + 2\bar{x}_j^2) + 2p_2 \bar{x}_j \bar{y}_j] \qquad (6.5)$$

$$\Delta y_j = \bar{y}_j [K_1 r^2 + K_2 r^4 + K r^6 + \ldots] + 2p_1 \bar{x}_j \bar{y}_j + p_2 (r^2 + 2 y_j^2)]$$

where Δx_j and Δy_j are corrections for geometric lens distortion present in the coordinates x and y of image point J and $r = (x_j^2 + y_j^2)^{\frac{1}{2}}$. The coefficients $K_1, K_2, K_3,$ P_1, P_2, P_3 are the parameters to be determined through the calibration process. This model accounts for both symmetric radial distortion and asymmetric tangential distortion. The terms which include the coefficient K_1, K_2 and K_3 represent the symmetric radial distortion which is the same as that of Eq. (6.1), and the terms which include P_1, P_2 and P_3 represent the asymmetric tangential distortion.

In the next section, the calibration procedure introduced in Chapter 2 will be extended to include line distortion parameters. The effect of different terms of distortion parameters will be illustrated through examples.

6.4 Camera Calibration with Lens Distortion

The prospective transformation matrix of Eq. (2.1) can be rewritten substituting

$$x_c = x_d + \Delta x$$

$$y_c = x_d + \Delta y$$

where x_d and y_d are the distorted coordinates of the point and x_c, y_c are the corrected coordinates of the same point.

and y is the distortion value represented by the model of Eq. (6.5). Therefore for camera 1, Eq. (2.1) can be rewritten as

$$
\begin{bmatrix}
(x_d + \Delta x)w \\
(y_d + \Delta y)w \\
w
\end{bmatrix}
=
\begin{bmatrix}
\alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 \\
\alpha_5 & \alpha_6 & \alpha_7 & \alpha_8 \\
\alpha_9 & \alpha_{10} & \alpha_{11} & 1
\end{bmatrix}
\begin{bmatrix}
X \\
Y \\
Z \\
1
\end{bmatrix}
\tag{6.6}
$$

Substituting Eq. (6.5) into Eq. (6.6) the following two equations can be obtained:

$$f_1[\bar{\alpha}] = [x_d[1+\alpha_{12}r^2+\alpha_{13}r^4+\alpha_{14}r^6] + \alpha_{15}[r^2+2(x^2) + {}_{16}2x_dy_d]$$

$$[\alpha_9X+\alpha_{10}Y+\alpha_{11}Z+1.0] - [\alpha_1X+\alpha_2Y+\alpha_3Z+\alpha_4] = 0$$

$$f_2[\bar{\alpha}] = [y_d(1+\alpha_{12}Y^2+\alpha_{13}r^4+\alpha_{14}r^6) + \alpha_{15}(2x_dy_d) \qquad (6.7)$$

$$+ \alpha_{16}[r^2+2y_d^2][\alpha_9X+\alpha_{10}Y+\alpha_{11}Z+1.0]$$

$$- [\alpha_5X+\alpha_6Y+\alpha_7Z+\alpha_8] = 0$$

where $\alpha_1, \alpha_2, \ldots, \alpha_{11}$ are the prospective transformation parameters

$$\alpha_{12} = K_1 \qquad \alpha_{15} = P_1$$
$$\alpha_{13} = K_2$$
$$\alpha_{14} = K_3 \qquad \alpha_{16} = P_2$$

of the lens distortion model.

Now, by taking a number of measurements (X,Y,Z) and the corresponding image points (x,y) a set of an over-determined non-linear equations can be formed.

$$\phi = f_i(\bar{\alpha})$$

where α is a 2-dimensional vector of unknowns, α where $j=1,\ldots,16$. It is customary in solving this form of equations to define an error criterion on the form,

$$\phi = \sum_{i=1}^{K} [f(\alpha)] \tag{6.8}$$

where K=2 in our application.

One of the most critical issues in optimizing such functions is the selection of good initial guess for the parameters that will insure a rapid and optimal convergence. Also the choice of an appropriate algorithm for solving a set of non-linear equations represents an important factor.

To deal with the initial guess issue, the system is first calibrated assuming no lens distortion, i.e., α_{12} to α_{16} are assigned zero value (this reduces the problem to the set of linear equations of Chapter 2) and then substituting these values as initial guess to the non-linear Eq. (6.8) for ϕ.

The algorithm developed by Marquardt [52] for the purpose of least-squares estimation of non-linear parameters was chosen as a suitable technique for this application. In this procedure the first derivative of $f_1(\alpha)$ and $f_2(\alpha)$ is required. The procedure is then implemented as follows:

1. Form the first derivative matrix on the form

$$\frac{\partial f_1(\bar{\alpha})}{\partial \alpha_1} \quad \frac{\partial f_1(\bar{\alpha})}{\partial \alpha_2} \quad \cdots \cdots \quad \frac{\partial f_1(\bar{\alpha})}{\partial \alpha_{16}}$$

$$\tag{6.9}$$

$$\frac{\partial f_2(\bar{\alpha})}{\partial \alpha_1} \quad \frac{\partial f_2(\bar{\alpha})}{\partial \alpha_2} \quad \cdots \cdots \quad \frac{\partial f_2(\bar{\alpha})}{\partial \alpha_{16}}$$

2. A P-matrix is formed by substituting the set of measured values for (X,Y,Z) and (x',y') in the first

derivative Eq. (6.9).

3. A set of linear equation in α , where $\Delta\alpha = [\Delta\alpha_1 \Delta\alpha_2 \ldots \Delta\alpha_{16}]$, is constructed in the form

$$[P^T P + \lambda I]\Delta\bar{\alpha}^{(r)} = -P^T f \tag{6.10}$$

where f is the vector obtained by substituting the measured value for (X,Y,Z) and (x',y') in

$$\begin{bmatrix} f_1(\bar{\alpha}) \\ f_2(\bar{\alpha}) \end{bmatrix} \tag{6.11}$$

where r is the iteration number and $\Delta\bar{\alpha}$ is defined as:

$$\Delta\bar{\alpha}^{(r)} = \bar{\alpha}^{(r+1)} - \bar{\alpha}^{(r)} \tag{6.12}$$

where $\alpha^{-(o)}$ is the initial guess value for the parameters as discussed earlier, through calibration with the assumption of no lens distortion, then utilizing the obtained parameters as initial guess for α_1 to α_{11} and that the lens distortion parameters α_{12} to α_{16} are initially estimated to be zero.

The value of λ is initially selected at 0.01, and modified in further iterations to ensure movements in the direction towards the minimum of the error criterion

$$\phi^{(r)} = \sum(f_1^2[\bar{\alpha}^{(r)}] + f_2^2[\bar{\alpha}^{(r)}]) \tag{6.13}$$

where Σ is considered over the measured points. Convergence is obtained when

$$\frac{|\alpha_j^{(r)}|}{10^{-3} + |\alpha_j^{(r)}|} \leq 10^{-5}, \quad \forall j$$

In all the experiments performed, the error dropped substantially in the first 6 iterations.

6.5 Calibration Examples

To demonstrate the effect of introducing various lines distortion terms in the camera calibration operations, data are collected as explained in Chapter 2. The data is then utilized to solve for each of the following assumptions:

1. No line distortion (parameters obtained are later used for assumptions 2-4).

2. One term of lens distortion.

3. One term of lens distortion plus tangential distortion.

4. Three terms of radial distortion plus tangential distortion.

The algorithm was executed for cases 2 to 4 until the convergence condition of Eq. (6.14) was met.

For comparison purposes between the effect of each of the above assumptions the error criterion of equation (6.14) is used. Two experiments were performed. The first involved 27 points and the second involves 63. The results obtained for both experiments are as follows:

Experiment 1: Number of data points collected = 27

      Assumption 1: error ϕ = 0.006885

      Assumption 2: error ϕ = 0.006278

      Assumption 3: error ϕ = 0.002591

      Assumption 4: error ϕ = 0.002204

Experiment 2: Number of data points collected = 63

      Assumption 1: error ϕ = 7.56 x 10

      Assumption 2: error ϕ = 7.50 x 10

      Assumption 3: error ϕ = 6.074 x 10

      Assumption 4: error ϕ = 5.62 x 10

From these results it is clear that assumption 2, in which only one term of the radial distortion is considered, had noticeable effect on reducing the error (approximately 9%) below the assumption of no lens distortion. In the third assumption, in which the tangential distortion terms was included, the error value was reduced by 62% from the first assumption. A further drop in error by 67% from the initial case was achieved by including three terms of radial distortion plus tangential distortion. The above results show that tangential distortion is of important

consideration in error reduction, contrary to Tsai's generalized claim that tangential distortion is of no effect and can be ignored. Also higher terms of radial distortion shows a considerable effect on the error reduction, and when including these terms no numerical instability was encountered, contrary to a statement given by Tsai in his paper [19].

6.6 Conclusion and Discussion

Lens distortion, both radial and tangential, was discussed and included in the camera calibration procedure. The proposed calibration procedure of Chapter 2 was extended to include lens distortion. This extension changed the nature of the set of equations from linear algebraic equations to a set of non-linear algebraic equations which constituted a different approach for solving the calibration parameters. A procedure for automatic estimation of the initial values for the optimization parameters, by assuming no lens distortion, which is an assumption that was tested in previous chapters and showed to obtain reasonable results, was used.

The effect of different terms distortion on reducing the error criterion was evaluated by introducing different terms to the optimization procedure. Through this evaluation it was found that in our application tangential distortion has a considerable effect and that the statement

made by Tsai [19] in his paper cannot be generalized. This is mainly due to the fact that since tangential distortion is, in fact, due to decentralization of lens elements during assembling, or due to the decentralization of different elements in the optical train, such as the case when close-up lenses are mounted on the optical system, then tangential distortion effect may vary from one set-up to the next and from one lens manufacturer to another and, therefore, its effect may have a great contribution to the accuracy of the measurements. Including tangential distortion and higher terms of radial distortion does not represent a numerical stability problem once a suitable algorithm such as the one used in this thesis is utilized.

CHAPTER VII

SUMMARY AND CONCLUSIONS

The principle objectives of the research work described in this thesis are to develop solutions to:

1) the calibration of randomly aligned stereo vision system of close-range sensing

2) point-of-correspondence problem for the general case of co-planar camera.

With these objectives in mind, approaches for stereo vision systems were investigated. It was concluded that most of the available approaches assume geometrical constraints on the imaging setup. This leads to the investigation of camera calibration techniques adopted by photogrammetrists for photogrammetric applications. An approach based on homogeneous coordinate transformation was developed to relate the camera 2-D coordinate system to the 3-D world coordinate system. With this approach the geometrical constraints were relaxed thus simplifying the setup procedure and reducing setup time. The approach involves solving a set of over-

131

determined linear equations. Retrieving 3-D position information of objects within the common field of view was then illustrated. Performance evaluation of the proposed procedure was carried out utilizing a standard testing procedure to determine both accuracy and repeatability of measurements reported by the system. The results obtained show that the proposed procedure has an accuracy of 0.1 mm over a range of measurements of 12.7*12.7*12.7 mm and a repeatability envelope of 0.01 mm.

The procedure was first implemented to deal with isolated features in the image (i.e., a single object is in view of both cameras). Therefore, finding matched points between left and right cameras can be done by simply locating the centroid of the object or the corner depending on the specific target. However, once several objects are within the viewing range, finding matched features between the left and right images introduces a new problem that requires the development of a procedure that deals with this point-of-correspondence problem.

A new procedure for determining the point-of-correspondence was developed. The approach utilizes mathematical morphology operators to obtain edge-strength images and then to identify matches between the two images.

The search for matches was performed over a limited search neighborhood rather than the entire image, thereby reducing the search time and increasing the accuracy of

match findings.

The lens distortion in optical train was then discussed and a proposed calibration procedure was expanded to include lens distortion parameters, both radial and tangential. The effect of the different terms of the lens distortion model on the accuracy of the results was then illustrated by error reduction to the cost function. The results obtained show that both radial and tangential distortion parameters have a significant effect on the accuracy of the measurements.

Further research based on the work presented in this thesis can be carried out. Some suggestions follow:

1. Development of an expert system for coordinate measurement and object recognition based on 3-D information.

2. Extending the proposed system to multi-camera viewing.

3. 3-D measurement using a single camera mounted on a translation table to provide several views of the object.

4. Optimum selection for the structure element in the morphological scheme.

REFERENCES

[1]   A.C. Kak, "Depth Perception for Robots," Purdue Univ., Lafayette, IN, TR-EE83-44, 1983.

[2]   R.A. Jarvais, "A Perspective on Range Finding Techniques for Computer Vision," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. PAMI-5, No. 2, March 1983, pp. 122-129.

[3]   Y.C. Kim and J.K. Aggrawal, "Positioning Three-Dimensional Objects Using Stereo Images," IEEE Journal of Robotics and Automation, Vol. RA-3, No. 4, August 1987, pp. 361-373.

[4]   R.D. Arnold and T.O. Binford," Geometrical Constraints in Stereo Vision," Proc. SPIE, Vol. 238, San Diego CA, 1980, pp. 281-293.

[5]   K.M. Mutch, "Determining Objects Translation Information Using Stereoscopic Motion," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. PAMI-8, No. 6, Nov. 1986, pp. 750-755.

[6]   M.D. Levine, D.A. O'Handley and G.M. Yagi, "Computer Determination of Depth Maps," Computer Graphics and Image Processing, Vol., 1973, pp. 134-150.

[7]   Y. Yakimovsky and R. Cunningham, "A System for Extracting Three-dimensional Measurements from a Stereo Pair of TV Cameras," Computer Graphics Image Processing, Vol. 7, Jan. 1978, pp. 195-210.

[8]   D.C. Brown, "Close-range Camera Calibration, Photogrammetric Eng., Vol. 37, 1971, pp. 855-866.

[9]   W. Faig, "Calibration of Close-range Photogrammetry Systems: Mathematical Formulation," Photogrammetric Eng. Remote Sensing, Vol. 41, 1975, pp. 1479-1486.

[10] A. Okomoto, "Orientation and Construction of Models, Part I: The Orientation Problem in Close-range Photogrammetry," Photogrammetric Eng. Remote Sensing, Vol. 47, 1981, pp. 1437-1454.

[11] A. Okomoto, "The Model Construction Problem Using the Collinearity Condition, "Photogrammetric Eng. Remote Sensing, Vol. 50, 1984, pp. 705, 711.

[12] K.W. Wong, "Mathematical Formulation and Digital Analysis in Close Range Photogrammetry," Photogrammetric Eng. Remote Sensing, Vol. 41, 1975, pp. 1342-1372.

[13] I. Sobel, "On Calibrating Computer Controlled Cameras for Perceiving 3-D Scenes," Artificial Intell., Vol. 5, 1974, pp. 185-189.

[14] Y.I. Abdel-Aziz and H.M. Karara, "Direct Linear Transformation into Object Space Coordinates in Close-range Photogrammetry," Univ. of Illinois, at Urbana, Champaign, Urbana, 1971, pp. 1-18.

[15] Y.I. Abdel-Aziz and H.M. Karara, "Photogrammetric Potential of Non-Metric Cameras," Civil Eng. Studies, Photogrammetry Series 36, Urbana, The Univ. of Illinois, 1984.

[16] A. Dainis and M. Juberts, "Accurate Remote Measurement of Robot Trajectory Motion," in Proc. Int. Conf. Robotics and Automation, 1985, pp. 92-99.

[17] H. Itok, A.M. Yauchi and Z. Ozawa, "Distance Measuring Method Using Only Simple Vision Constructed for Moving Robots," Proc. 5th Int. Conf. on Pattern Recognition, Montreal, P.Q., Canada, Vol. 1, 1984, p. 192.

[18] J.Y. Luh and J.A. Klassen, "A Three-Dimensional Vision by Off-shelf System with Multi-cameras," IEEE Trans. Pattern Anal. Machine Intell., Vol. PAMI-7 Jan. 1985, pp. 185-189.

[19] R.Y. Tsai, "A Versatile Camera Calibration Technique for High-accuracy 3-D Machine Vision Metrology Using Off-the-Shelf TV Camera and Lenses," IEEE Journal of Robotics and Automation, Vol. RA-3, No. 4, Aug. 1987, pp. 323-344.

[20] H.A. Martins, J.R. Birk and R.B. Kelley, "Camera Models Based on Data from Two Calibration Planes," Computer Graphics Image Processing, Vol. 17, 1981, pp. 173-180.

[21] A. Isaguirre, P. Pu and J. Summers, "A New Development in Camera Calibration: Calibrating a Pair of Mobile Cameras," Proc. Int. Conf. Robotics and Automation, 1985, pp. 74-79.

[22] D. Marr and T. Poggio, "Cooperative Computation of Stereo," Science, Vol. 194, Sept. 1976, pp. 283-287.

[23] D. Marr and T. Poggio, "A Theory of Human Stereo Vision," Proc. Roy. Soc. London, Vol. B204, 1979, pp. 301-328.

[24] W.E.L. Grimson, "A Computer Implementation of a Theory of Human Stereo Vision," Phil. Trans. Roy. Soc. London, Vol. B292, 1981, pp. 217-253.

[25] B. Julesz, "Binocular Depth Perception of Computer Generated Patterns," Bell Sys. Tech. J., Vol. 39 1960, pp. 1125-1162.

[26] J.E.W. Mayhew and J.P. Frisby, "Physiophysical and Computational Studies Towards a Theory of Human Stereopsis,"

[27] P. Nowforth, J.E.W. Mayhew and J.P. Frisby, "Vergence Eye Movement Made in Response to Spatial-frequency Filtered Random dot Stereograms," Perception, Vol. 10, 1981, pp. 299-304.

[28] J. Mayhew, "The Interpretation of Stereo-disparity Information: The Computation of Surface Orientation and Depth," Perception, Vol. 11, 1982, pp. 387-403.

[29] R.D. Arnold, "Local Contact in Matching Edges for Stereo Vision," Proc. Image Understanding Workshop Cambridge, MA., May 1978.

[30] H.H. Baker, "Edge Based Stereo Correlation," Proc. Image Understanding Workshop, College Park, MD, April 1980, pp. 168-175.

[31] H.H. Baker and Binford, "Depth from Edge and Intensity Based Stereo," Proc. 7th Int. Joint Conf. Artificial Intelligence, Vancouver, B.C., Canada, 1981, 631-636.

[32] G. Matheron, "Random Sets and Integral," Geometry, New York, Wiley, 1975.

[33] J. Serra, "Image Analysis and Mathematical Morphology," London Academic, 1982.

[34] S.R. Sternberg, "Languages and Architectures of
       Parallel Image Processing," Proc. Conf. Pattern
       Recognition in Practice, L.N. Kanal and E.S.
       Gelsema, Eds., Amsterdam, The Netherlands,
       North-Holland, 1980.

[35] S.R. Sternberg, "Biomedical Image Processing," IEEE
       Compact. Mag., Vol. 16, Jan. 1988.

[36] D.L. McCubbery and R.M. Lougheed, "Morphological Image
       Analysis Using a Roster Pipeline Processor,"
       IEEE Comput. Soc. Workshop Computer Architecture
       for Pattern Analysis and Image Database
       Management, Miami Beach, FL., Nov. 1985,
       pp. 444-452.

[37] T.R. Crimmins and W.M. Brown, "Image Algebra and
       Automatic Shape Recognition," IEEE Trans. on
       Aerospace and Electronics Systems, Vol. AES-21,
       No. 1, Jan. 1985, pp. 60-69.

[38] M.J.E. Golay, "Hexagonal parallel pattern transforma-
       tion," IEEE Trans. Comput., Vol. C18, 1969,
       pp. 733-740.

[39] D. Graham and P.E. Norgren," The Diff3 Analyzer:
       A parallel/Serial Golay image Processor,"
       Real Time Image Processing, M. Onoe, K. Preston
       and A. Rosenfeld, Eds., London: Plenum, 1980
       pp. 163-182.

[40] R.M. Haralick, S.R. Sternberg and X. Zhuang," Image
       Analysis Using Mathematical Morphology," IEEE
       Trans. on Pattern Analysis and Machine
       Intelligence, Vol. PAMI-9, No. 4, July 1987,
       pp. 532-550.

[41] S. Peleg, J. Naor, R. Harley, and D. Avnir, "Multiple
       Resolution Texture Analysis and Classification,"
       IEEE Trans. Pattern Anal. Machine Intell., Vol.
       PAMI-6, 1984, pp. 518-523.

[42] J.S. Lee, R.M. Haralick and G.V. Shapiro, "Morphologic
       Edge Detecting," IEEE Journal of Robotics and
       Automation, Vol. RA-3, No. 2, April 1987,
       pp. 142-156.

[43] I.E. Abdou and W.K. Pratt, "Quantitative Design and
       Evaluation of Enhancement/Thresholding Edge
       Detectors," Proc. IEEE, Vol. 67, May 1979.

[44] R.M. Haralick, "Digital Step Edges from Zero Crossing of Second Directional Derivatives," IEEE Trans. Pattern Anal. Machine Intell., Vol. PAMI-6, Jan. 1984, pp. 58-68.

[45] W.M. Newman and R.F. Sproull, Principles of Interactive Computer Graphics, McGraw-Hill, Computer Science Series, Second ed.

[46] G.H. Golub and C.F. Van Loan, Matrix Computations, Baltimore: John Hopkins University Press, 1983.

[47] L.W. Johnson and R.D. Riess, Numerical Analysis, Addison-Wesley, 1982.

[48] H.J. Trussel, "Common Picture Thresholding Using an Herative Selection Method," IEEE Trans. on SMC, Vol. SMC. 9, 1979, p. 311.

[49] A. Chottera and M. Shrider, "Feature Extraction of Manufactured Parts in the Presence of Spurious Surface Reflections," Can. Elect. Eng. J., Vol. 7, No. 4, 1982, pp. 29-32.

[50] J.T. Pennington, "Tangential Distortion and its Effect on Photogrammetric Extension of Control," Photogrammetric Engineering, Vol. 13, No. 1, 1947, pp. 135-142.

[51] D.C. Brown, "Decentering Distortion of Lenses," Photogrammetric Engineering, Vol. 32, No. 3, 1966, pp. 444-462.

[52] D.W. Marquardt, "An Algorithm for Least-squares Estimation of Non-linear Parameters," J. Soc. Indust. Appl. Math., Vol. 11, No. 2, June 1963, pp. 431-441.

VITA AUCTORIS

Mohamed T. Boraie

1958   Born on February, Cairo, Egypt.

1980   Received B.A.Sc. degree in Electrical Engineering
          from University of Assiut, Assiut, Egypt.

1984   Received M.A.Sc. degree in Electrical Engineering
          from The University of Windsor, Windsor, Ontario,
          Canada.

1989   Candidate for Ph.D. degree in Electrical Engineering
          at the University of Windsor, Windsor, Ontario,
          Canada.

Publications

Journals

M.A. Sid-Ahmed, M. Boraie and W. Loy, "Photogrammetric Aerotriangulation Using Matrix CCD Cameras for Close-Range Position Sensing." Computers in Industry, 1989, pp. 307-311.

Sid-Ahmed and M. Boraie, "Multiple Camera Calibration for 3-D Machine Vision Metrology." IEEE Instrumentation and Measurement. (Accepted).

M. Boraie and M.A. Sid-Ahmed, "Points-of-Correspondence in Stereo Images with no Geometrical Constraints Using Mathematical Morphology." IEEE Trans. Robotics and Automation. (In progress).

Conferences

M.A. Sid-Ahmed and M.T. Boraie, "Passive Range Sensing and Points-of-Correspondence using Mathematical Morphology." Canadian Conference on Electrical and Computer Engineering, Sept. 17-20, 1989, Montreal, Canada.

M. Boraie and M.A. Sid-Ahmed, "Points-of-Correspondence in Stereo Vision." 20th Annual Conference on Modelling and Simulation, May 4-5, 1989.

M.A. Sid-Ahmed and M. Boraie, "Calibration of Stereo Vision for 3-D Close-Range Sensing." The 31st Midwest Symposium on Circuits and Systems, August 9-12, 1988, St. Louis, Mo.