A ROBOT VISION SYSTEM FOR DETERMINING 3-D COORDINATES OF OBJECT POINTS

by

Tea-Quin Kim

Thesis submitted to the Faculty of the Graduate School of
the New Jersey Institute of Technology in partial fulfillment of
the requirements for the degree of
Master of Science in Mechanical Engineering
1989

APPROVAL SHEET

Title of thesis: A robot vision system for determining 3-D coordinates of object points

Name of candidate: Tea-Quin Kim
Master of science in Mechanical Engineering, 1989

Abstract and thesis approved:

Dr. M. C. Leu
Professor
Date

Dr. R. Dave
Assistant Professor
Date

Dr. Y. Park
Assistant Professor
Date

VITA

Name: TEA-QUIN KIM

Permanent address:

Degree and date to be conferred: MASTER OF SCIENCE
IN MECHANICAL ENGINEERING
MAY, 1989

Date of birth:

Place of birth:

Collegiate Institutions Attended:

Name	Date	Degree	Date of Degree
Korea University, Seoul, Republic of Korea	Feb. 1981 to Feb. 1985	BSME	Feb. 1985
New Jersey Institute of Technology, Newark, NJ	Sep. 1986 to May. 1989	MSME	May, 1989

Major: MECHANICAL ENGINEERING

ABSTRACT

A ROBOT VISION SYSTEM FOR DETERMINING 3-D COORDINATES OF OBJECT POINTS (MAY 1989)

Tea-Quin Kim, M.S.M.E., New Jersey Institute of Technology

Thesis Advisor: Dr. M. C. Leu

A stereo disparity algorithm is implemented on an industrial robot vision system. One camera manipulated by an industrial robot is used to obtain stereo images of an object. After a pair of stereo images are processed, the points of interest are selected from each image. The similarities of three types of feature characters are used to assign initial weights of the matching probabilities to a set of point candidates. Then a relaxation method, which is based on probabilities of local connectivity, similarity of feature characters, and smoothness of matching, is used iteratively to improve the initial probabilities of image matching. The positions of object points are determined by image disparities based on a triangulation method. Experiments performed on an AdeptOne Robot with based vision system show that the algorithm works well.

Blank Page

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLE

ACKNOLEGEMENT

I thank JESUS for finishing this thesis without any trouble. I also thank my mother, my wife Hee-Kyung, and my family for their encouragement through this work. I sincerely appreciate Prof. M. C. Leu and Dr. Z. Ji for their helpful comments and suggestions. I also thank SEJONG FOUNDATION which partially supported this work.

Chapter 1

INTRODUCTION

The techniques in extracting depth information using a vision system can be classified into two broad categories: active and passive. In the active methods, light sources are controlled to illuminate objects at places of interest. Depth information at those places can then be measured by triangulation. In the passive techniques, natural lighting is used. Stereo disparity vision, the subject of this thesis, is an important passive method for extracting depth information.

In the stereo disparity vision, two images of the same object taken from two different locations are compared. The disparity (i.e. difference) between the two images is used to extract the range information. In principle, the two images can be taken arbitrarily, as long as the object of interest is included in both images and essential features can be preserved for reliable matches. To simplify the matching process, however, the two images are usually taken with the camera axes parallel to each other and two image planes coplanar (i.e. there is no relative displacement along the camera axis direction).

One essential step in the stereo disparity method is the matching process. That is, for a point in one image, find its correspondence in the other image so that the disparity can be computed. However, it is not a good practice to find corresponding points for all image pixels because it is costly and not all points can be matched with equal confidence. Usually feature points are selected for matching because of their distinguished geometric properties. One of the most important image features is "edge", which can be a physical edge or a boundary line of a uniform region. A number of approaches have been proposed for extracting features from two-dimensional image data and solving the correspondence problem. In general, edge points can be determined by applying an operator which calculates the first derivative and/or the second derivative of image intensity data [1]. Marr and Hildreth [2] proposed a "zero-crossing" method which first applies the Gaussian operator to the original image for smoothing, and then applies the Laplacian operator. For the matching of edge points, a number of different techniques have been suggested. Several of these techniques will be briefly reviewed.

1.1 Survey of image matching techniques

A number of techniques have been proposed for solving the corresponding problem. We can divide these techniques into four categories based on their feature extraction method (area-based and feature-based) and matching process method (psychophysical and biological aspects of the human visual system and mathematical aspects with probability). A brief review is provided below.

Area-based point of view

An area-based technique finds corresponding points on the basis of similarity between two corresponding areas in the left and right images. The corresponding area is the neighboring regions of the edge and is called a "window" (also called a "patch"). Many similarity functions have been proposed for specific applications[1]. The most popular one is the correlation of light intensity (i.e grey-level intensity) between the left and right windows. Barnea and Silverman[5] introduced SSDA(Sequential Similarity Detection Algorithm) for finding the best matching region with a difference function and iteration.

Feature-based point of view

A feature-based approach is performed by comparing physical attributes of object in the two images. A lot of different features have been proposed in different domains. Leu and Pherwani[6] used combined data of area, centroid, coordinate, perimeter, and principal axis as matching feature data. Roach and Aggarwal[7] proposed corners of polyhedral as features. Marr and Poggio[8] and Grimson[3] used the zero-crossings of the second derivatives of the intensity values in the image as the matching features. Kim and Aggarwal[11] used the zero-crossing pattern type as the matching feature.

Psychophysical and biological aspects of the human visual system

Marr and Poggio[8] proposed a model of the human stereo vision system on the basis of computational theory. They introduced a multi-channel hierarchical matching scheme which uses zero-crossing as the matching feature. Mayhew and Frisby[12] introduced figural continuity in human visual mechanism and improved the Marr and Poggio algorithm. Grimson[3] implemented and tested the Marr and Poggio model with continuity constraint.

Mathematical aspects with probability

Bannard and Thomson [4] presented a locally parallel model which utilizes the discreteness, similarity, and consistency properties to reduce the chance of matching error using a relaxation method.first introduced by Hummel and Zucker[9]. Shmuel[10] showed a new relaxation scheme which derived the coefficients more analytically than the existing schemes (i.e., Bannard and Thomson[4]).

1.2 Objectives

The objective of this thesis is to determine 3-D coordinates of object points using one single camera manipulated by an industrial robot. A stereo range-finding algorithm is implemented on an AdeptOne Robot with a PULNIX-CCD camera under general lighting. The camera positions can be changed by manipulating robot motion, since the camera is mounted on the robot arm. An epipolar-line based matching method is used in this implementation. First , edge points are detected with the AdeptVision XGS system. Then, epipolar lines are constructed to locate the edge points. For each edge point in one image, a set of points within a range along the same epipolar line on the other image are chosen as a candidate point set for that point. The similarities of three feature characters (i.e., connectivity pattern type, grey-level gradient, and average grey level of the neighbor region) are used to assign initial weights to that set of the matching candidates. The matching process uses a relaxation method based on probability of local connectivity, similarity of feature characters, and smoothness of probability matching. The use of the feature characters is reliable for reducing the possibility of false initial weighting. A practical program written in VAL-II is developed for experiments on the AdeptOne robot based vision system.

There are three possible error sources in stereo disparity vision: (1) image resolution,, (2) robot positioning inaccuracy, and (3) lens-to-image plane length error. The inaccuracy due to image resolution is found to be the main error source of the stereo vision system.

The thesis is organized as follows. In chapter 2, general stereo imaging geometry and its related transformation are expressed. Chapter 3 describes the matching procedure. The edge detection and feature characters, which are used to find the similarity of the edge points of the left and right images, are described. A relaxation method is adopted to find the best matching edge points. Chapter 4 describes the experimental set-up and procedures. Chapter 5 describes the experimental results on algorithm effectiveness and accuracy. Chapter 6 states the conclusion.

Chapter 2

STEREO IMAGE GEOMETRY

2.1 Coordinate frames

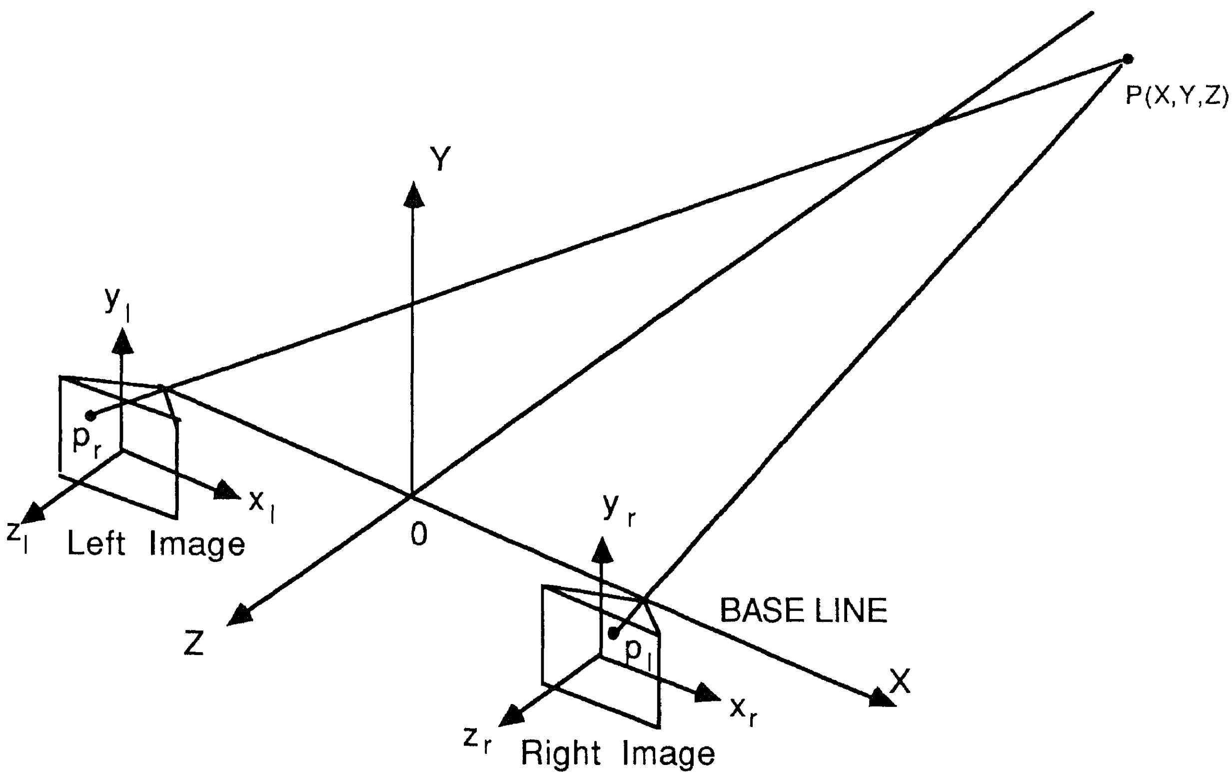


Figure 2-1. Parallel-axis stereo vision

In Fig. 2-1, a reference frame OXYZ is defined with its X-axis being the line passing through the two lens centers (also called the baseline), its origin being the mid-point between the lens centers, and its Z-axis parallel to the camera axis. The two images are labeled left image and right image, respectively. Each image plane is fixed with a coordinate frame which has its axes parallel to those of the reference coordinate frame. The origin of a coordinate frame fixed to an image is the intersecting point of the image plane and the camera axis. Thus, the image of an object point P(X,Y,Z) will have the coordinates of its right image as $(x_r,y_r,0)$ and the coordinates on its left image as $(x_l,y_l,0)$.

2.2 Epipolar lines and related transformation

Epipolar lines play an important role in our implementation of stereo disparity range-finding. From an object point and the two camera lens centers, one can construct a plane in passing the object point and the plane parallel to the line connecting the two lines centers. This plane is called the epipolar plane. The intersection of an epipolar plane with the two image planes forms the so-called epipolar line. Since the two image planes are coplanar in our arrangement, an epipolar line is exactly a line connecting the two corresponding image points, and it is parallel to the baseline. The coordinate frames fixed to the image planes will be referred as epipolar coordinate frames since their X-axes are parallel to the epipolar lines. On the other hand, it is more convenient to refer to an image point by its pixel location. Thus, the normal coordinate frame x_{ni}- y_{ni}(i= l and r) has its x- and y-axis parallel to its pixel arrays as shown in Fig. 2-2. The transformation between the normal coordinate frame and the epipolar coordinate frame (X-Y coordinate) is a rotation about the z-axis.

In order to determine the angles θ_l and θ_r, we need to perform another transformation between the normal coordinate frame and the world coordinate frame. As shown in Fig. 2-3, positions and orientations of x_{nl} - y_{nl} and x_{nr} - y_{nr} with respect to the X_w - Y_w frame are (x_1,y_1,ϕ_1) and (x_2,y_2,ϕ_2) , respectively. Since the epipolar lines are parallel to the line connecting the two image centers, the orientaion of all epipolar lines is the same as that of the line connecting the two image centers, which can be found as

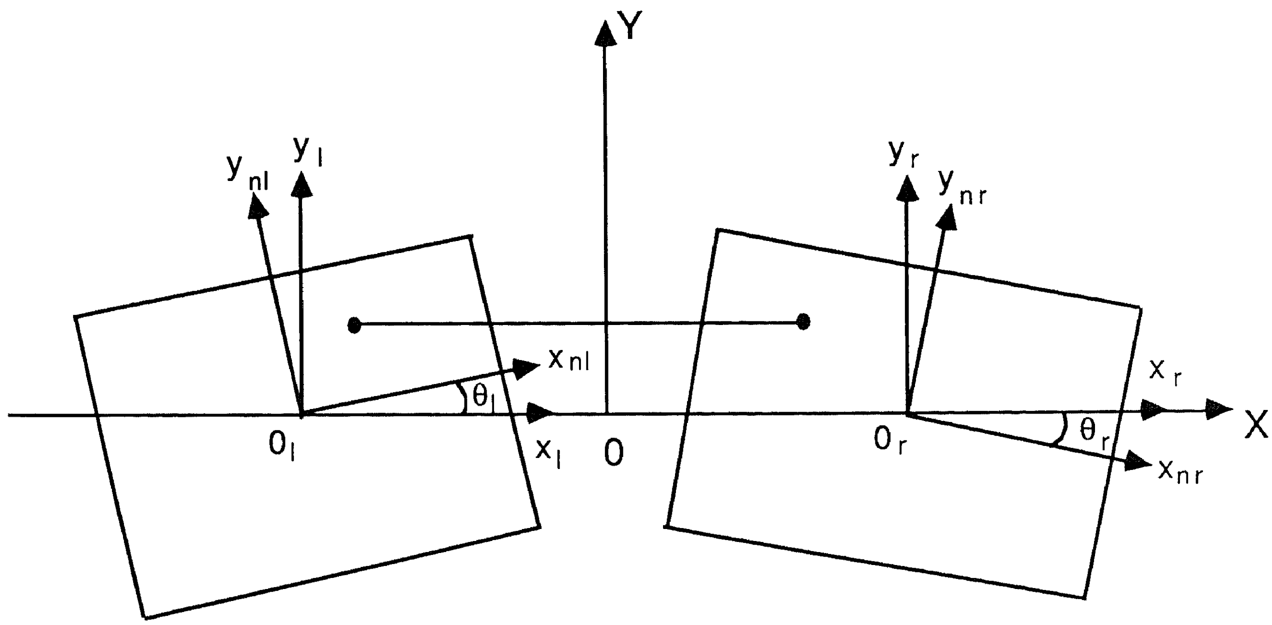


Figure 2-2. Normal coordinate frame and epipolar coordinate frame

$$\alpha = \text{atan2}\,(y_2 - y_1,\ x_2 - x_1)$$

Therefore

$$\theta_l = \phi_l - \alpha$$

$$\theta_r = \phi_2 - \alpha$$

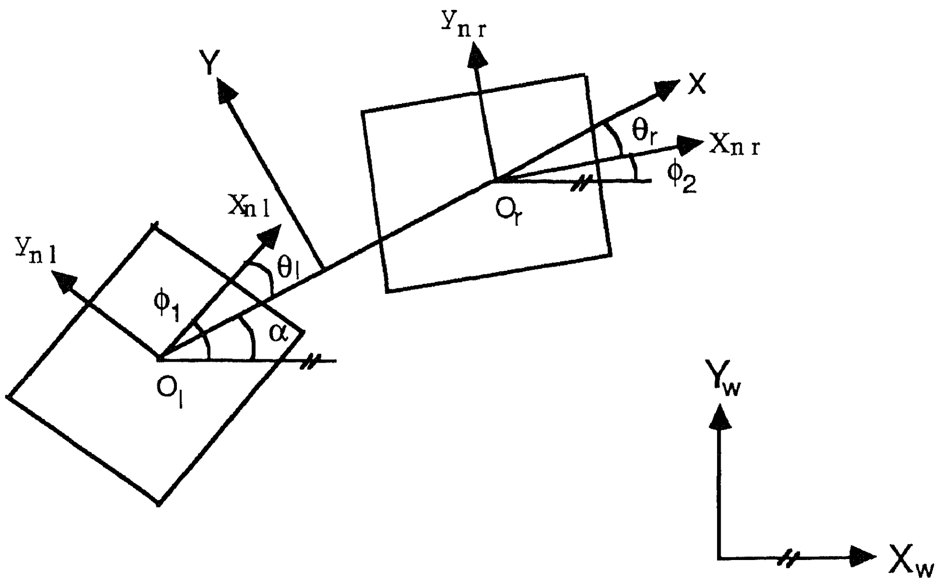


Figure 2-3 World coordinate frame and epipolar coordinate frame

Chapter 3

ALGORITHMS

In this chapter, we will discuss algorithms for matching and for determining positions.

3.1 Edge detection

Edge detection is done on the basis of gray-level variation in image data. A point is regarded as an edge point if its gray level varies considerably from that of the other points in the region having the point as the centers. A typical way to find edge points is to apply gradient operators that are sensitive to variations.

Edge detection in our implementation uses a Gaussian weighted convolution with w=5 pixels (see Fig. 3-1) to smooth out noise in the original image. Spurious edge points are removed through thresholding. The threshold value is selected based on the intensity gradient value which is 8.0 in this work.

2	2	2	2	2
2	10	10	10	2
2	10	15	10	2
2	10	10	10	2
2	2	2	2	2

Figure 3-1. Averaging-Gaussian filter

The AdeptVision XGS uses a built-in function VRULER to locate the edge points along a line. That is, edge points are found as the intersection of the VRULER and edges. Since an epipolar-line based matching is chosen in our implementation, epipolar lines are used to construct the ruler lines. Thus the locations of edge points, as well as their associated characters, are registered along the epipolar lines.

3.2 Feature characters

Edge points in a filtered image mark the locations of significant changes in the intensity data. We use the distance from an edge point to a predetermined start point on the epipolar line to represent the location of the edge point. In addition to their locations, four other attributes of edge points are recorded for evaluating matching candidates. The four attributes are: (1) contrast sign (i.e,whether the convolution value changes from positive to negative, or from negative to positive), (2) magnitude of intensity gradient, (3) connectivity pattern type (i.e. the relation of an edge point with edge points in its immediate upper and lower neighboring epipolar lines), (see Fig. 3-2c), and (4) average grey level (of a 5-by-5 pixel window see Fig. 3-3). Thus, the j^{th} edge point along the i^{th} epipolar line can be denoted as P[i,j](d,s,m,p,a) as shown in Fig. 3-2, where d represents the distance, s the contrast sign, m the gradient magnitude, p the pattern type, and a the average grey level.

Since there are usually more than one match candidate point for each edge point in the other image, the next step is to find the best match from the set of candidate points based on the degree of similarity of their attributes. Three similarity indices are used as discussed below.

(a) Intensity gradient similarity index, GI

An edge point has an gradient value with respect to its neighbor pixels. This value is used to determe the similarity of matchable edge points as follows:

$$GI = 1 / (1 + | G_l - G_r |)$$

where G_l and G_r are the gradients of grey level intensity at the two points to be matched.

(b) Connectivity pattern similarity index, PI

An edge is one element of an edge group. The pattern type of an edge point is defined as follows: see Fig. 3-2. The distance of an edge point from the start point position on the i^{th} epipolar line is d. Similarly, d' and d" represent distances of edge points on the $i\text{-}1^{th}$ and $i+1^{th}$ epipolar lines, respectively. If the distance differences d - d' and d - d" are below one pixel length, then the edge point on the i^{th} epipolar line is connected to the upper and/or lower edge points lying on the $i\text{-}1^{th}$ and $i+1^{th}$ epipolar lines. A detected edge point connected with both upper *and* lower edge points is represented by a type number from 1 to 9, as shown in Fig. 3-2c. A detected edge point connected with only upper edge point is represented by a type number from 10 to 12. A detected edge point connected with *only* lower edge point is represented by type number from 13 to type number 15. Type number 16 represents a point which is disconnected from both upper and lower edge points.

The difference between the pattern type of an edge point in the left image P_l and that of a candidate edge point in the right image P_r is defined by the following rule. First, the difference of the upper edge point in the right image is compared with the upper edge point in the left image. The possible value of difference is an integers from 0 (i.e., there is no locational difference between two edge points) to 3 (i.e., one edge point is connected, but the other edge point is not connected). Then, the locational difference is checked on the lower epipolar line. Combined together, the total difference of the pattern type between P_l and P_r ranges from 0 to 6. The following formula represents the similarity index, PI:

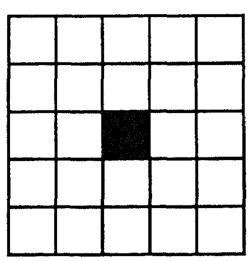
$$PI = 1 / (1 + | P_l - P_r |)$$

(c) average grey level similarity index, AI

Given an edge point in the left image and a point in the right image, the average grey level similarity index is used to identify the similarity of the regions containing the points. That region is a 5 pixel by 5 pixel "window"., which has the edge point at the center of the window (see Fig. 3-3). The average grey level of the window is the mean value of grey levels of the 25 pixels. The average grey level similarity index is calculated as follows:

$$AI = 1/ (1+ | A_l - A_r |)$$

where A_l and A_r are the average grey level intensities of the two points to be matched.



window-1

Figure 3-3 Average grey level of window

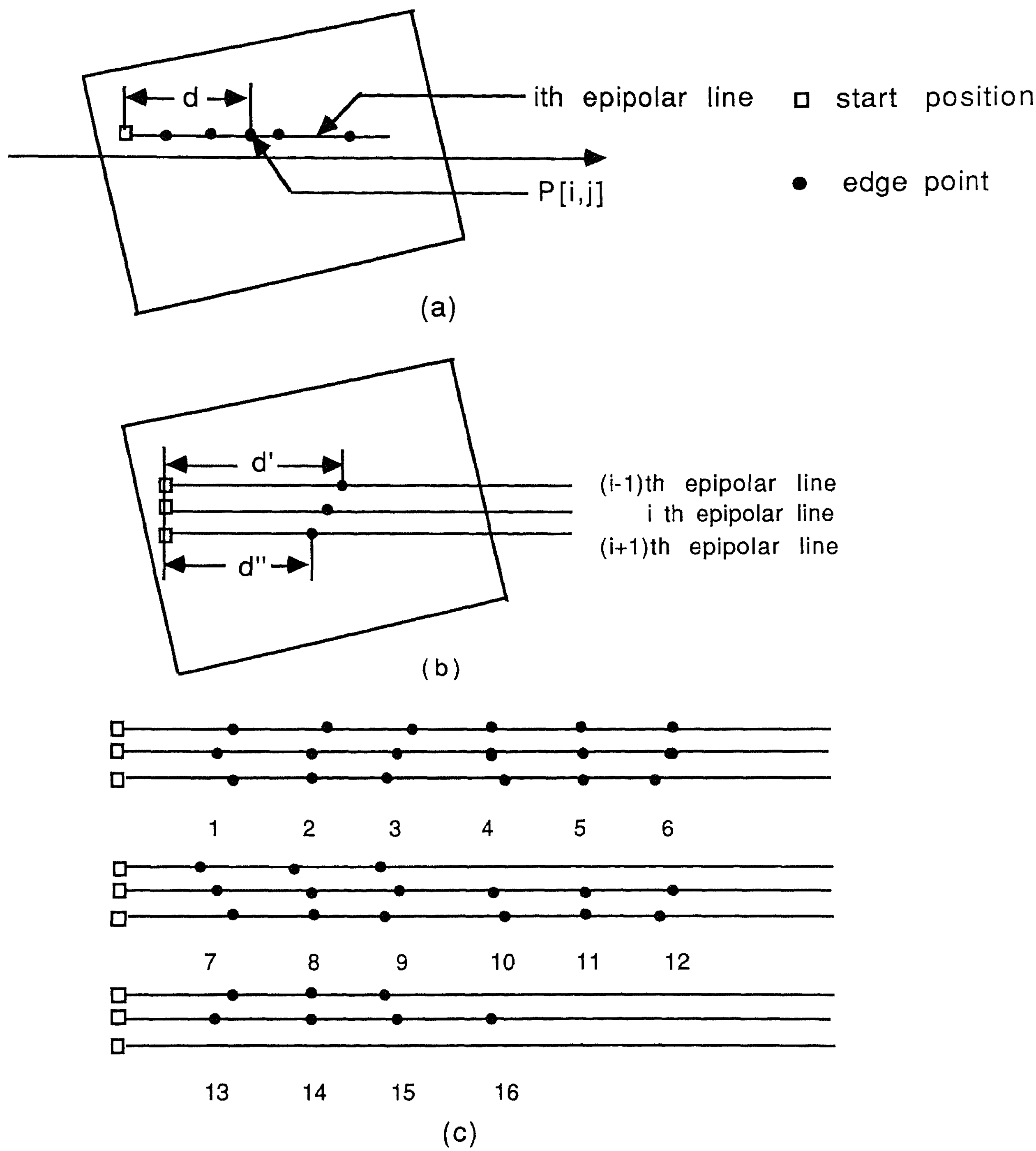


Figure 3-2. Representation of an edge point

For each of the above three similarity indices , an index will have value one if the compared feature characters are the same. The value will be less than one if the characters compared are different. The bigger the difference, the smaller the index value.

3.3 Initial probability

Let L: $\{l_1, l_2, ..., l_n\}$ denote the candidate point set, where n is the number of candidate points in the set and l_i is the label of the i^{th} point in the set. Once GI, PI, and AI are computed for l_i, its combined similarity index, w, can be computed as:

$$w(l_i) = \lambda_1 GI_i + \lambda_2 PI_i + \lambda_3 AI_i \quad (0 <= \lambda_i <= 1)$$

$$\Sigma\lambda_i = 1$$

The value of λ_i can be chosen depending on the relative importance of the similarity indices. The indices which vary randomly depending on the shape and/or surface characteristics of the object. Here we chose λ_i=1/3 (i =1...3). The value of $w(l_i)$ is between 0 and 1.

The point with the highest combined similarity index in the candidate set is most likely to be the correct match in many cases. However, it may be the case that the point with the maximum combined similarity index is still a valid match. A probability value is therefore needed for checking the consistency of the matches.

The initial probability that l_i is the match point can be calculated with the Bayes' formula[4]

$$P^0(l_i) = P(l_i/m)\ P_m \qquad (1)$$

where P_m is the matchable probability defined as

$$P_m = 1 - \max \{ w(l_i), i=1,... n \} \quad (2)$$

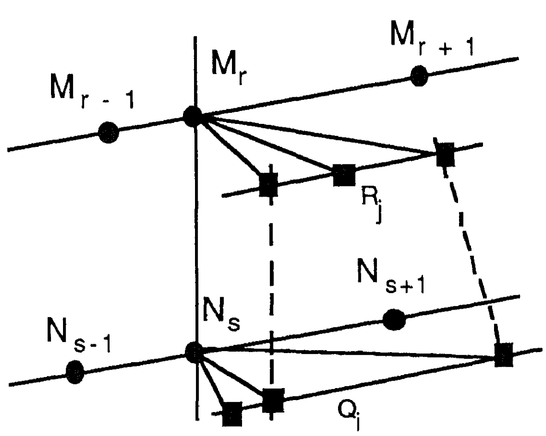
and $P(l_i/m)$ is the conditional probability for l_i to be the match point if the match exists, and is

$$P(l_i/m) = w(l_i) / \Sigma\, w(l_j) \quad (3)$$

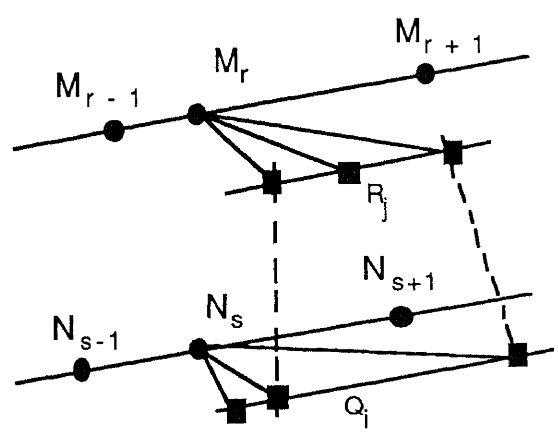
The initial probability, which depends only on the similarity of candidate points, gives us some information about which point in the candidate set is most likely to be the matching point. But there could be multiple matches or false match. To obtain more reliable matches, the initial probability should be improved by using a relaxation rule for checking consistency, as will be discussed next.

3.4 Relaxation rule

As the initial probability $P^0(l_i)$ determined in the previous section could produce false match if it is used alone, a relaxation rule is adopted to improve the initial probability.



Edge points M_r and N_s are connected
(a)



Edge points M_r and N_s are disconnected
(b)

Figure 3-4. Possibility of edge connectivity

As shown in Fig. 3-4, M_r denotes the edge point on the i^{th} epipolar line in the left image and R_j (j = 1,..., n) are its matching candidates on the same epipolar line in the right image. Similarly, N_s denotes the edge point on the $i+1^{th}$ epipolar line in the left image and Q_i (i=1,...,n) are its matching candidates on the same epipolar line in the right image. Given the probabilities of matching R_j to M_r, we want to update the probabilities of matching Q_i to N_s. There are two different cases, depending on whether M_r and N_s are connected or not. Fig. 3-4a shows the case where M_r and N_s are connected, and Fig 3-3b is the case where M_r and N_s are not connected. Among the candidate points Q_i, some are connected to the edge points in the upper epipolar line (as shown by the dotted lines) and others are not. The information on the connectivity between Q_i and R_j should have effect on whether the probability of matching M_r to N_s should be increased or decreased. Let us introduce the coefficients r_{ij} to represent such connectivity information. Thus, the value of r_{ij} reflects the degree of support Q_i can get, for it to match with N_s, from the probabilities of matching between R_j and M_r.

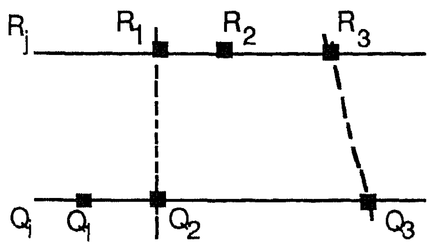
3.4.1 Probability coefficient

The probability coefficient can be defined in the following. Let the coefficients $r_{ij} = a_{ij}+b_{ij}$, where the constant a_{ij} is the combined similarity index w(li), and b_{ij} will be assigned with the following rules.

CASE I: Edge points M_r and N_s are connected

There are three possibilities in the edge connectivity in the matched points of the right image. First, the upper and lower edge points are connected, then we give value 1 to b_{ij} (e.g. R_1 to Q_2 in Fig. 3-5a). Second, the two edge points are disconnected, but one of them are connected to other edges in Q_i or R_j, then the value of b_{ij} is 0.2 (

e.g. R_1 to Q_1). Third, the two edges are disconnected and both are connected to other edges in Q_i or R_j, then value of b_{ij} is 0.0 (e.g. R_1 to Q_3). The table marked case I in Fig. 3-5b shows the values of b_{ij} for this case.



(a)possible matching diagram

b_{ij}	R_1	R_2	R_3
Q_1	0.2	0.2	0.2
Q_2	1	0.2	0.0
Q_3	0.0	0.2	1

CASE I

b_{ij}	R_1	R_2	R_3
Q_1	0.2	1	0.2
Q_2	0.0	0.2	0.2
Q_3	0.2	0.2	0.0

CASE II

(b) b_{ij} coefficient table

Figure 3-5. The coefficient of probability

CASE II: Edge points M_r and N_s are disconnected

For this case the values of b_{ij} are shown in the table of case II in Fig 3-5b. The largest matchable probability is when two edge points are both disconnected from their neighbor edge points., for which the value of b_{ij} is is 1 (e.g. R_2 to Q_1) . A second possibility is that the two edge points are disconnected but they are connected with edge points in the upper or lower epipolar line. Then the value of b_{ij} for this case is 0.2 (e.g. R_1 to Q_1). The worst matchable possibility is that two edge points are connected. This is the reverse of the connectivity of M_r and N_s. So its value of b_{ij} is 0.0 (e.g. R_1 to Q_1).

3.4.2 Updating equation

Once r_{ij} is determined for all (Q_i, R_j) pairs related to M_r and N_s, the initial probability of matching Q_i to N_s can now be updated iteratively as follows.

Let $P_i^k(N_s)$ be the probability of matching Q_i to N_s and $P_j^k(M_r)$ be the probability of matching R_j to M_r on the k^{th} iteration, then the total probability of both Q_i to N_s and R_j to M_r where Q_i and R_j represent the sets of candidate points in the right image which are to matched.can be represented by following equation:.

$$\gamma^k = [\, \underline{P}_i^k(N_s)] \, [\, \underline{r}_{ij}] \, [\, \underline{P}_j^k(M_r)] \qquad (4)$$

where $\underline{P}_i^k(N_s)$ and $\underline{P}_j^k(M_r)$ are the vector form of $P_i^k(N_s)$ and $P_j^k(M_r)$, $\underline{r}_{ij}$ is the matrix form of r_{ij}.

The updated probability of matching Q_i to N_s becomes.

$$P_i^{k+1}(N_s) = P_i^k(N_s) * \{\Sigma_j \, P_j^k(M_r) * r_{ij}\} / \gamma^k \qquad (5)$$

3.5 Matching region

Once the edge points are identified along with their feature characters in both images, a matching process is performed to identify each edge point in the left image with an edge point in the right image on the same epipolar line. This is done in two steps. First, the match candidates are picked up within the estimated matchable region. The matchable region along the epipolar line in the right image for a point in the left image is

$$\{d_r: \ d_l - d_i - w < d_r < d_l - d_i + w\}$$

where d_l and d_r are the distances of the edge points to their corresponding start points in the two images, and d_i distance between corresponding start points as shown in Fig. 3-6. For an edge point in the left image with a distance d_l, the search for the match edge points in right image is constrained to the distance range given by the above formula.

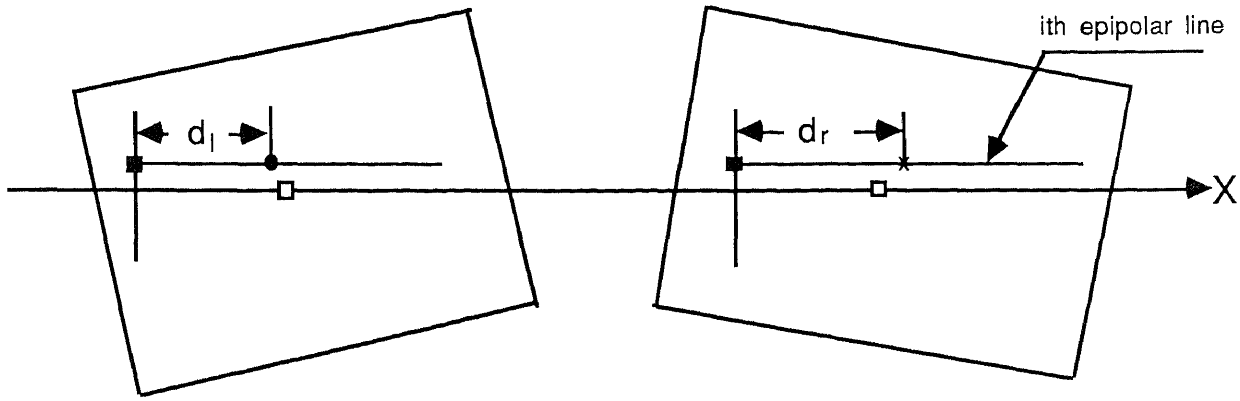


Figure 3-6. Matching region

3.6 Determination of the 3-D position

Once an image pair has been successfully matched for an object point, the three-dimensional coordinates of that point can now be calculated with a simple triangulation technique [6]. If the matched pair has their image coordinates as (x_r,y_r) and (x_l,y_l), respectively, then the three-dimensional coordinates of that point are (see Fig. 3-7)

$$x = \frac{D}{2}\frac{x_r+x_l}{x_r-x_l} \tag{6}$$

$$y = \frac{D}{2}\frac{y_r+y_l}{x_r-x_l} \tag{7}$$

$$z = Ds\frac{x_r+x_l}{x_r-x_l} \tag{8}$$

where D denotes the distance between the lens centers, and s is the distance between the lens plane and the image plane of the camera.

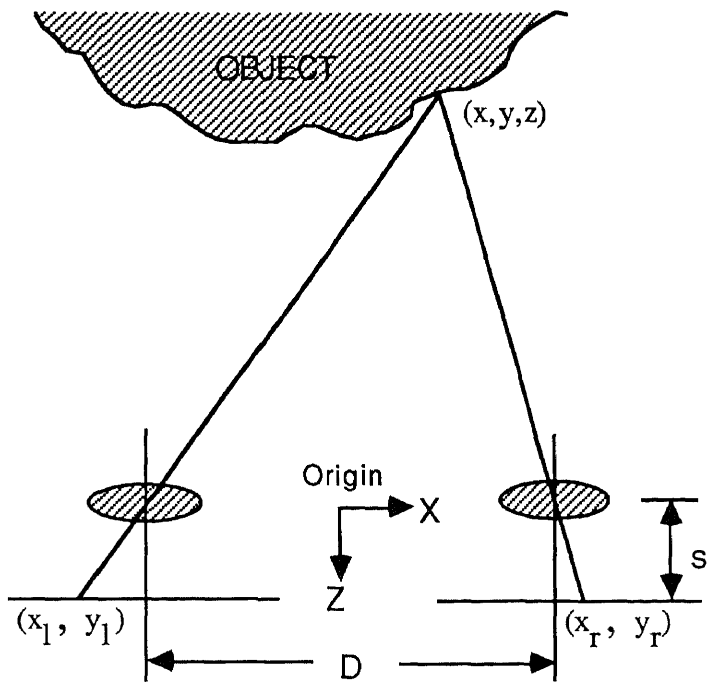


Figure 3-7. Position determination with triangulation method

3.7 Error analysis

There are some errors encountered by any computation of the distance range finding based on disparities.

The overall system accuracy can be determined quantitatively by analyzing the equation (8). From this equation, we can see that the sources of error are camera distance D, lens to image planes distance s, and the digital image quantization effect. The overall error E_z can be expressed as

$$E_z = \left| \Delta D \frac{\partial z}{\partial D} \right| + \left| \Delta s \frac{\partial z}{\partial s} \right| + \left| \Delta(x_r - x_l) \frac{\partial z}{\partial (x_r - x_l)} \right| \qquad (9)$$

Both the baseline distance and the image distance can be controlled with high accuracy by the AdeptOne robot. However, the image resolution is constrained by the finite number of pixels in the image plane of the solid-state camera, which may cause the resolution error $\Delta(x_r - x_l)$ fairly appreciable.

Chapter 4

EXPERIMENTAL SET-UP AND PROCEDURE

4.1 Experimental set-up

The experiment is performed with an AdeptOne robot with XGS vision system [13] as show in Fig. 4-1.

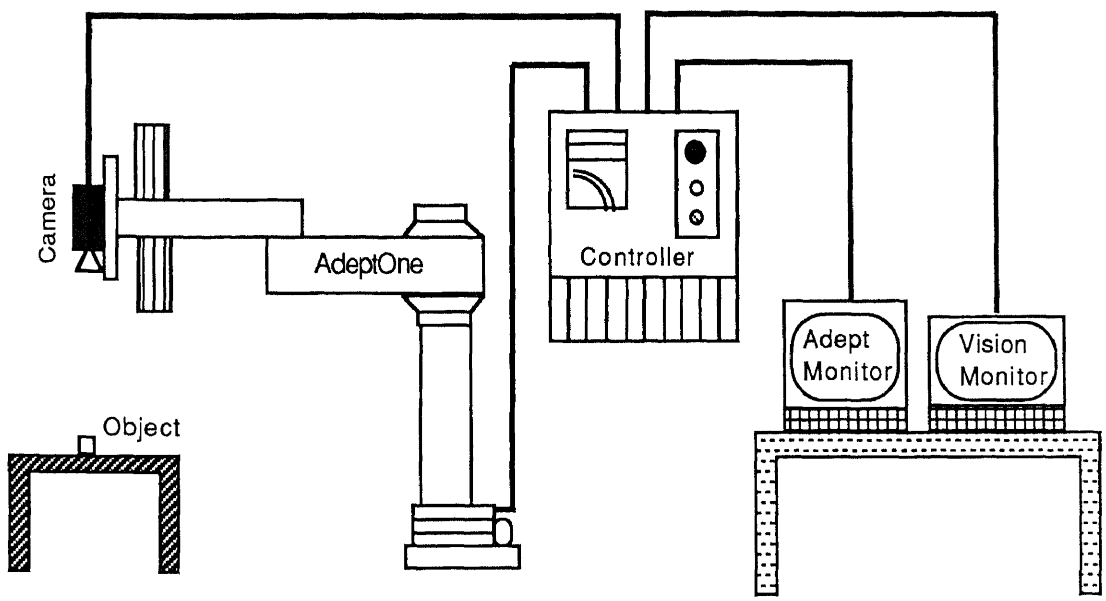


Figure 4-1. The Experimental Set-up

The AdeptOne is a four-degree-of-freedom SCARA type industrial robot. Starting from the base, the first two joints are rotational about their vertical joint axes, the third joint is translational , moving up and down along the vertical direction, and the last joint is rotational about a vertical axis which coincides with the translation joint. A PULIX CCD vision camera is mounted on the distal end of the forearm. The camera can only have two-degree-of-freedom horizontal movement. The CCD image plane in the PULNIX camera has 509x481 pixels. The images can be viewed through a vision monitor.

The AdeptVision XGS system is a grey level vision system with an intensity level from 0 to 127.

The geometry for camera viewpoints (two distinct locations) is depicted in Figure 4-2. The robot is manipulated so that the camera can view above the object considered. Then, the camera base line is chosen, which is shown in the figure. The angle between X and X_w is α . The camera location is used as the origin of the epipolar coordinate frame. The two viewpoints of the camera are located on the X axis, with their X coordinates being $\pm(D/2)$. A program in VAL-II is developed for automatically manipulating the camera.

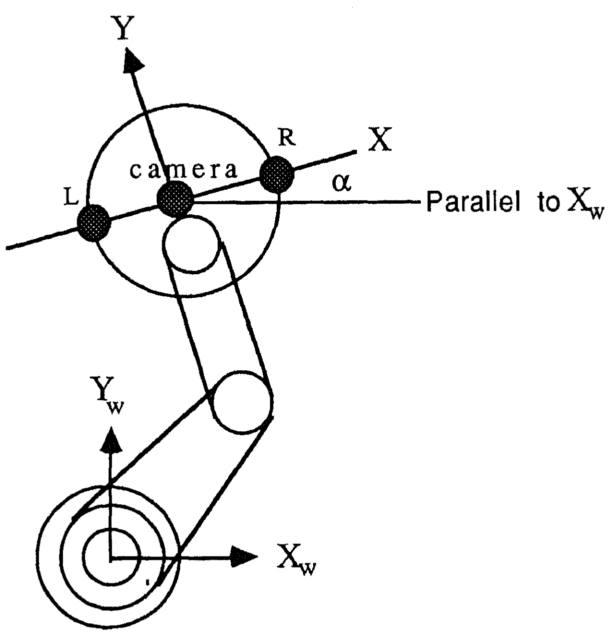


Figure 4-2. Camera Viewpoints

4.2 Experimental Procedure

4.2.1 Image processing

After the robot moves the camera to two specified locations above the object, the VAL II instruction VPICTURE is used to grab an image at each of the two locations for processing. The raw images are convolved with a 5x5 Gaussian Average operator for smoothing purpose by using VAL II instruction VCONVOLVE. The epipolar lines are then constructed with the VAL II system function VRULER on both images for extracting edge points information. VRULER is a linear edge point detection routine. Only edges that are nearly vertical to the ruler can be effectively detected. Since this feature is readily available in the vision system, this linear detector routine is used for simplicity, Edge detection is performed along an epipolar line. Threshold value of the edge detection is specified with VRULER parameter V.EDGE.STRENGTH. Fig. 4-3 shows the detected edges with various threshold values for a rectangular block.

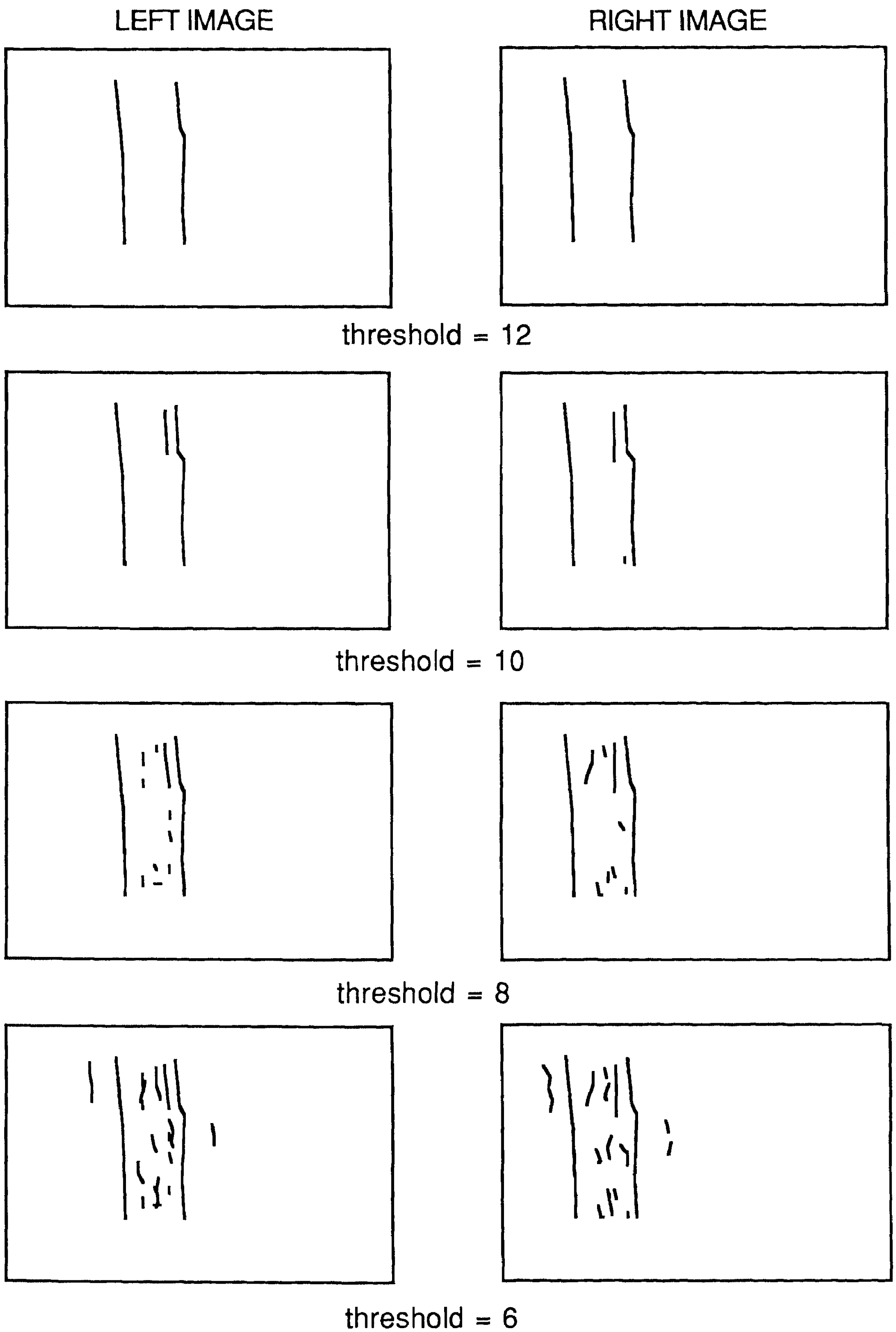


Figure 4-3. Detected edges

4.2.2 Matching

The matching starts with grouping the candidate points and calculating the combined similarity indices. With these combined similarity indices as initial probabilities, matching is improved through the updating procedure given in Eqs. (1-7) in chapter 3. Figure 4-4 shows the matched points of the above rectangular block with different numbers of iteration.

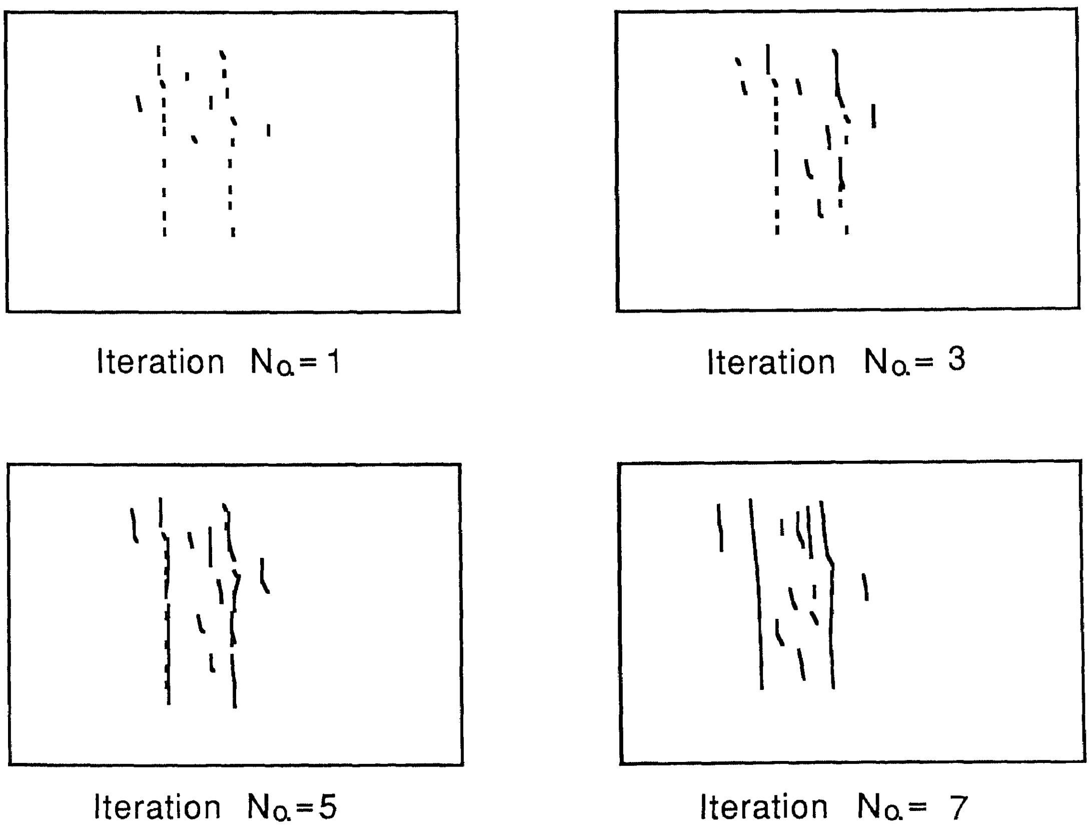


Figure 4-4. Matched points with different numbers of iterations

Chapter 5

Experimental results and discussion

Experiments were performed on a hub-pully(gear in an auto part) and a precision rectangular block. The hub-pully is used to test the algorithm effectiveness for complex objects. The precision rectangular block is used to check the accuracy of stereo-range finding algorithm implemented on the AdeptOne robotic vision system.

5.1 Algorithm effectiveness

The hub-pully is depicted in Fig. 5-1. In the left image 818 edges were detected and in the right image 815 edges were detected using the threshold value 8. The matched edge points were showed in Fig. 5-2. In the experiment, 86% of the points were appeared to be matched after the fifth iteration., and the iteration appeared converging. Among the matched points,, less than 1.65% (12 edge points out of 728) were found to be false matches.

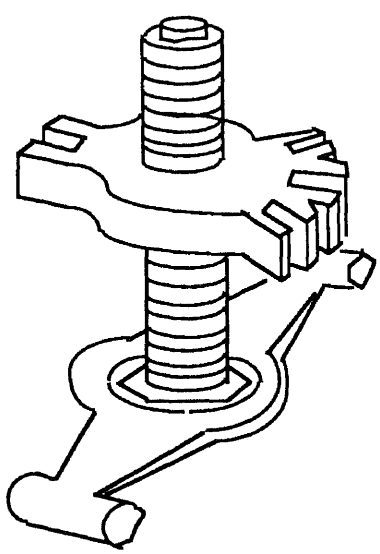
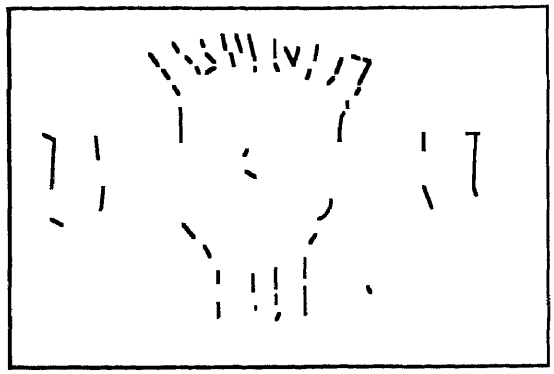
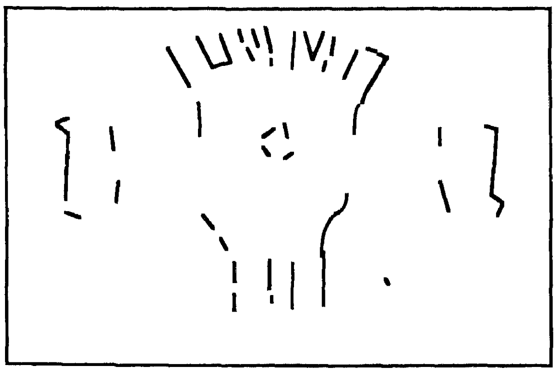


Figure 5-1. Hub-fully

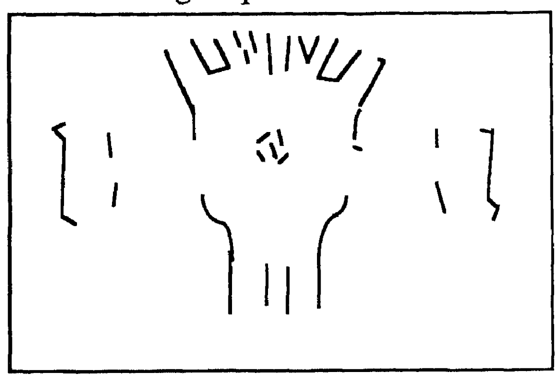
Iteration.N_0= 1
Edge points= 148



Iteration N_0= 3
Edge points= 532



Iteration N_0 = 5
edge points = 710



Iteration N_0 = 7
Edge points = 728

Figure 5-2. Matched edge points after iterations.

5.2 Accuracy

Relative accuracy

For all the matched points, Eqs. (6-8) in chapter 3 can now be used to calculate the corresponding object point coordinates in the epipolar coordinate system. Table-1 lists the experimental results for the above rectangular object. The two positions of the camera are separated by a distance 20mm, and the focal length of the camera is 25mm. Two sets of data are obtained for evaluating the relative accuracy of the vision system. The first set of data is obtained with the object located at a distance about 1000mm from the camera. The mean value of the Z-coordinates for this set of data is -1007.6 mm. The second set of data is obtained with the object lifted up 79.65mm from the the first location. The mean value of the Z-coordinates for the this set of data is -1086.3 mm. The relative error in the Z-coordinate measurement is thus less than 1mm (see Figure 5-3 and Table 5-1).

Absolute accuracy

The calculation of the coordinates is illustrated below using the left edge point on the fifth epipolar line as an example (see Fig. 5-3). The parameter data for this test point were

$s = 2.5$ cm.
$D = 2.0$ cm.
$(x_r - x_l) = 0.046$ cm.

Calculating the value of z coordinate using the above data gives $z = -108.32$ cm. Similar calculations were done for x and y coordinates of this point and for the 3-D coordinates of other points.
Computing the various partial derivatives using Eq. (9) gives

$$\frac{\partial z}{\partial s} = \frac{D}{x_r - x_l} = 43.0$$

$$\frac{\partial z}{\partial D} = \frac{s}{x_r - x_l} = 54.3$$

$$\frac{\partial z}{\partial (x_r - x_l)} = \frac{-Ds}{(x_r - x_l)} = 2362.9$$

Also, the upper-bound errors of the individual parameters were

Δs = 0.001 cm. (image distance inaccuracy)
ΔD = 0.001 cm. (base line distance inaccuracy)
$\Delta(x_r - x_l)$ = 0.00172 cm. (image resolution)

The predicted z-depth error bound using Eq. (9) is = 4.1 cm. The maximum measured z-depth error was 2.8 cm., which is within this error bound.

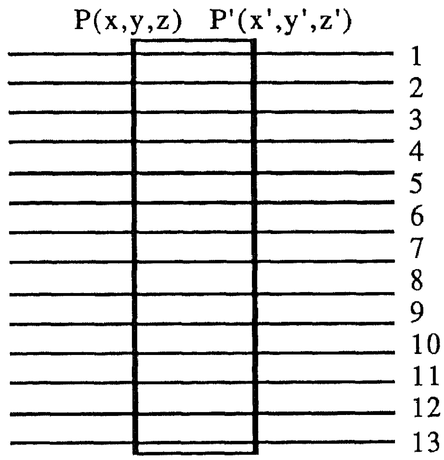


Figure 5-3 Rectangular block

k	x	y	z	x'	y'	z'	D
1	-55.3	54.7	-1009.3	-31.8	55.3	-1011.3	23.5
2	-55.4	49.4	-1010.4	-31.9	50.1	-1017.3	23.4
3	-55.5	44.1	-1013.8	-31.8	44.3	-1011.0	23.6
4	-55.2	38.5	-1007.5	-31.4	38.7	-1004.1	23.7
5	-55.1	33.1	-1008.1	-31.3	33.1	-998.9	23.7
6	-55.0	27.7	-1007.5	-31.5	28.1	-1009.8	23.5
7	-54.9	22.2	-1005.4	-31.3	22.6	-1007.6	23.6
8	-55.1	16.9	-1008.1	-31.2	17.2	-1004.2	23.8
9	-55.1	11.5	-1009.8	-31.1	11.6	-998.9	24.0
10	-55.1	6.1	-1008.7	-31.2	6.3	-1006.5	23.8
11	-55.1	0.7	-1009.8	-31.1	1.0	-1008.1	24.0
12	-55.1	-4.7	-1010.5	-30.8	-4.4	-999.2	24.3
13	-55.0	-10.2	-1012.3	-30.7	-9.8	-1001.2	24.3

(a)

k	x	y	z	x'	y'	z'	D
1	-56.5	57.3	-1087.8	-32.5	57.2	-1085.6	23.4
2	-56.7	51.7	-1092.4	-32.6	51.3	-1079.3	24.1
3	-56.9	45.8	-1092.4	-32.7	45.6	-1078.6	24.2
4	-56.5	39.6	-1082.5	-33.1	40.4	-1093.2	23.4
5	-56.6	33.8	-1083.2	-33.3	34.4	-1090.4	23.3
6	-57.6	28.3	-1097.2	-32.9	28.2	-1077.9	24.7
7	-56.9	22.1	-1078.3	-33.3	22.5	-1082.6	23.6
8	-56.9	16.3	-1081.9	-33.3	17.2	-1107.3	23.6
9	-57.6	10.6	-1089.1	-33.8	11.1	-1091.1	23.8
10	-57.8	4.7	-1087.2	-33.5	5.1	-1079.8	24.4
11	-57.5	1.1	-1083.2	-33.4	0.7	-1078.6	24.1
12	-57.2	-6.8	-1077.5	-34.3	-6.6	-1096.2	22.9
13	-57.1	-12.5	-1072.1	-34.3	-12.4	-1091.1	22.8

(b)

Table 5-1 The experimental data

Chapter 6

CONCLUSION

This thesis study involves implementation of a stereo disparity algorithm on a robotic vision system for determining 3-D coordinates of objects points. In the image matching process, feature-based characters (intensity gradient, and similarity of connectivity pattern) and area-based character (similarity of connectivity in the average grey-level intensity) are combined to determine the initial matching probabilities of candidate edge points. To reduce the probability of false matching, the initial matching probabilities are updated iteratively with a relaxation process. The algorithm has been implemented on the vision system of AdeptOne robot. The experimental results show that the combination of three characters provides good initial matching. The updating by the relaxation process is fast in convergence. False matching in the experiments is only 1.65%. Analysis of the experimental data shows that the inaccuracy in the determined coordinates is due primarily to the image resolution.

Bibliography

[1] Yoshiaki Shirai, *Three-Dimensional Computer Vision*, Springer - Verlag, 1987.

[2] D. Marr and E. Hildreth, "Theory of edge detection," Proc. Roy. Soc. London, vol. B207, pp. 187-217, 1980.

[3] W. E. L. Grimson, "Computational experiments with feature based stereo algorithm," IEEE Trans. Pattern Anal. Mach. Intell., PAMI-7, no.1, pp. 17-34, 1985.

[4] S. T. Barnard and W. B. Thompson, "Disparity analysis of images," IEEE Trans. Pattern Anal. Mach. Intell., vol. PAMI-2, no.4, 333-340, 1980.

[5] D. J. Barnea, and H. F. Silverman, "A class of algorithms for fast digital image registration," IEEE Trans on Computers, C-21, no.2, pp. 179-186, 1972.

[6] M. C. Leu and R. M. Pherwani, "A vision system for three-dimensional positional measurement based on stereo disparity," Tech. Report RARP-009, Dept. of Mech. & Ind. Engr., New Jersey Inst. of Tech., 1989.

[7] J. W. Roach and J. K. Aggarwal,"Computer tracking of objects moving in space," IEEE Trans. Pattern Anal. Mach. Intell., vol. PAMI-1, no.2, pp. 127-135, 1979.

[8] D. Marr and T. Poggio, "A computational theory of human stereo vision," Proc. Roy. Soc. London, vol. B204, pp.301-328, 1979.

[9] A. Rosenfeld, R. A. Hummel, and S. W. Zucker, "Scene labeling by relaxation operations," IEEE Trans. Syst., Man, and Cybern., vol. SMC-6, 420-433, 1976.

[10] Shmuel Peleg, "A new probabilistic Relaxation Scheme," IEEE Trans. Pattern Anal. Mach. Intell., PAMI- 2, no.4, pp. 362-369, 1980.

[11] Y. C. Kim and J. K. Aggarwal, "Positioning three dimensional objects using stereo images," IEEE Transactions on Robotics and Automation,J- RA3, no.4, pp. 361-373, 1987.

[12] J. E. W. Mayhew and J. P. Frisby, "Psychophysical and computational studies towards a theory of human stereopsis," Artificial Intell., vol. 16-17. pp. 349-385, 1981.

[13] Adept Vision Reference Guide, Adept Technology, Inc., Apr. 1987.