ANALYTIC MODELING FOR THREE DIMENSIONAL OBJECTS USING COMPUTER VISION

النمذجة التحليلية للاجسام ثلاثية الابعاد بأستخدام الرؤية الآلية

Yehia M. Enab
Clemson University, U.S.A.

Fayez W. Zaki
Mutah University, Jordan

A.I. Abd El-Fattah
Mansoura University, Egypt.

S.H. El-Konyaly
Mansoura University, Egypt.

خلاصة - ان الهدف الاساسي من هذا البحث هو ايجاد نموذج بسيط ودقيق للاجسام ثلاثية الابعاد وذلك كي نتمكن من تصميم مسار آمن لحركة الذراع الآلية في حقل الانتاج .

وقد اختير هذا النموذج على الشكل البيضاوي لما له من ميزات عديدة منها تقليل الحجم الضائع الى أدنى قيمة اذا قورن بالنماذج الاخرى المستخدمة مثل النموذج الكروي او صندوق متوازي المستطيلات ، هذا بجانب أن النموذج البيضاوي يمكننا من التعبير عن أي جسم بمعادلة ذات عدد ثابت من المعاملات (٩ معاملات) هي معاملات معادلة الجسم البيضاوي ، والفكرة الاساسية في البحث هي استخدام معلومات الرؤية الآلية لايجاد الاحداثيات الكارتيزية لمجموعة من النقط الموجودة على سطح الجسم وذلك باستخدام طرق " ايجاد المدى" هذه الاحداثيات تغذى بعد ذلك الى برنامج للحاسوب الذي يقوم بدوره في حساب معاملات المعادلة التحليلية بطريقة" متوسط المربعات الاصغر " بحيث يكون الخطأ أقل ما يمكن .

ABSTRACT: In this paper a method for analytic modeling of three dimensional objects is introduced. The basic idea is to make use of the information obtained using computer vision to determine the Cartesian Coordinates of a set of points existing on the surface of the object, this occurs through the use of range finding methods. These Coordinates are fed to a computer program which computes the coefficients of the analytic equation such that the error in object modeling is minimum, this occurs by using the least mean square estimation for determining the coefficients of the fitting equation. The ellipsoid is the primitive used to model the objects. This method enables us to represent any object by a constant number of coefficients(9 coefficients) which is the number of coefficients in ellipsoid equation. Also, the waste volume will be reduced compared to the waste volume resulted when modeling the object with other models such as spheres or bounding boxes. Our primary goal is to provide a simple and accurate model for the objects in order to be able to design a collision free path for the robot arm while moving in its working space.

1-INTRODUCTION

Geometric description of three dimensional(3-D) objects is a very important task in different fields such as solid modeling, computer-aided design, computer-aided manufacturing, and robotics. There exist different schemes for representing 3-D objects such as spatial occupancy of cells in an octree, Boolean combination of solids topological relationship of vertices, edges and faces in a boundary graph, and analytic solid modeling. In this paper, object modeling techniques suitable for designing a collision free-path for a robot arm while moving in its work space area are presented. Morvac [1] solved the find-path problem in two dimensions. He bounded all objects (obstacles) and the moving object by circles. Then the grown obstacles were all perpendicular cylinders and the problem could be projected back into two dimensions where rotations could be ignored. This method misses all paths that require rotational maneuvering [2]. Luh [3] proposed a scheme using a 'pillar Model' of unexpected obstacles which is constructed from the images by stereo cameras, and selecting a collision-avoidance path from prestructured paths. Using stereo cameras, however, often involves convolution which is computationally time consuming. To ensure the safety requirements the descriptive model of the obstacle should always enclose the physical object entirely. Since the description of the obstacle must be processed through computer, then it is desirable to avoid any complicated model. Recently, Lub and Klaasen [4] suggested a simple model which satisfies these conditions in which the physical obstacle is encloeed by a bounding box. When the Cartesian Coordinates are defined to be parallel with the edges of the box, it can be described by the lower(minimum value) and the upper (maximum value) bounds of the obstacle on the three axes. Three cameras arranged such that their axes are parallel to the Coordinate axes, are used to obtain these informations. Because the detailed description of the obstacle is not considered, the bounding box model causes a large amount of space wasted. The 'wire frame' model [5], in which the objects are represented by a collection of vertices connected by straight line edges, is used for many applications in path planning tasks.

In this paper the 'Ellipsoid Model' is introduced as a method for representing the objects. This model uses an ellipsoid as a primitive to represent 3-D objects and an ellips to represent 2-D objects. Section 2, discusses how to use least square estimation method [6] to derive the fitting equation in the Cartesian Coordinates of a limited number of points existing on the object surface. Section 3, discusses the role of computer vision to obtain the coordinates of these points which may be carried out using shape from shading techniques such as photometric stereo method [7], or stereo vision method [8]. Section 4, explains how to design a collision free path for the end effector of the robot arm using the ellipsoid model. Section 5, presents experimental work and practical considerations. The paper ends with conclusions and suggestions for further work.

2-MATHEMATICAL REPRESENTAION OF OBJECTS

Assume that the Cartesian Coordinates of N points existing on the surface of the object are given, these points are distributed across all the object surface (for example the coordinates of vertices of a convex polyhedron). It is required to obtain a simple analytic equation exist on or inside this simple surface, i.e, the required surface must envelop the object such that the waste space which exists inside the surface and at the same time does not belong to the object is minimum. This equation will be referred to as fitting equation.

a) Fitting Equation

Usually objects are described by the composition of 'PRIMITIVES'. A typical geometric model includes primitives such as a point, line, plane, ellipsoid, parallelopiped, sphere, cylinder, and cone. To ensure the safety requirements, the descriptive model of the object should always enclose the physical object entirely. Since the description of the obstacle must be processed through the computer, it is desirable to avoid any complicated model.

In this paper, the ellipsoid is chosen as an envelope for the object. The general equation of the ellipsoid using Cartesian Coordinates (x,y,z) is

$$a_1x^2 + a_2y^2 + a_3z^2 + a_4xy + a_5xz + a_6yz + a_7x + a_8y + a_9z + a_{10} = 0$$

where (a_i , i= 1, 2, ...,10) are the equation coefficients. Dividing both sides by $(-a_{10})$, then the above equation can be expressed in matrix notation form as:

$$\underline{C}^T \cdot \underline{X} = 1 \qquad (1)$$

where;

$$\underline{C} = [\; c_1 \;\; c_2 \;\; c_3 \;\; c_4 \;\; c_5 \;\; c_6 \;\; c_7 \;\; c_8 \;\; c_9 \;]^T \qquad (2)$$

$$c_i = a_i/(-a_{10}) \;, \qquad i= 1, 2, \ldots\ldots 9 \qquad (3)$$

and

$$\underline{X} = [\; x^2 \;\; y^2 \;\; z^2 \;\; xy \;\; xz \;\; yz \;\; x \;\; y \;\; z \;]^T \qquad (4)$$

$\underline{C}$ is called coefficients matrix and $\underline{X}$ is called coordinates matrix. It is easy to show that Eq.(1) represents all the points (x,y,z) existing on the surface of the Ellipsoid while the relation $\underline{C}^T . \underline{X} < 1$ represents the points inside it, and the relation $\underline{C}^T . \underline{X} > 1$ represents the points outside the Ellipsoid surface.

Given that the set of points (x_i, y_i, z_i); i = 1, 2,....N existing on the object surface, the problem becomes how to calculate the coefficients matrix $\underline{C}$ such that the error in representing the object by the Ellipsoid becomes minimum, or in other words the lost volume is minimum. There are variety of methods for solving this problem. The Least Mean Square (LMS) estimation method [6] is used here for this task. Substituting the points x_i, y_i, z_i into Eq.(1) yields:

$$\underline{C}^T . \underline{X}_i = 1 + e_i$$

where $\underline{X}_i$ is the coordinate matrix corresponding to the point x_i, y_i, z_i and e_i is the error due to approximating the object surface by the Ellipsoid surface. The goal is to choose $\underline{C}$ such that the mean square value of e_i is minimum across all the given points of the object surface. Therefore, the mean square error can be expressed as:

$$E = (1/N) \sum_{i=1}^{N} e_i^2$$

$$= (1/N) \sum_{i=1}^{N} (\underline{C}^T \underline{X} - 1)^2 \qquad (5)$$

Differentiating E with respect to $\underline{C}$ and equating the result to zero, yields

$$\underline{R} . \underline{C}^* = \underline{U}$$

where

$$\underline{R} = (1/N) \sum_{i=1}^{N} \underline{X}_i . \underline{X}_i^T$$

= Correlation Matrix

$$\underline{U} = (1/N) \sum_{i=1}^{N} \underline{X}_i$$

= Mean Matrix

and

$$\underline{C}^* = \underline{R}^{-1} \underline{U} \qquad (6)$$

= Optimum Coefficients Matrix

Note that $\underline{R}$ is always positive definite. In order to make sure that the whole object exists inside the Ellipsoid, Eq.(1) is modified as

$$\underline{C}^{*T} . \underline{X} = 1 + e^+_{max} \qquad (7)$$

where e^+_{max} is the maximum positive error occured when using $\underline{C}^*$ as a coefficient matrix, it can easily be obtained from the given information.

b) <u>Case of Complex Objects</u>

If the object under study is complex and represented by a single Ellipsoid , the waste volume will be large. In this case the complex object may be divided into smaller simple objects, and for each one of these minor objects a fitting equation is derived. For example, the T shape shown in Fig. 1, can be represented by two Ellipsoid, each one represents one of the two arms of the shape.

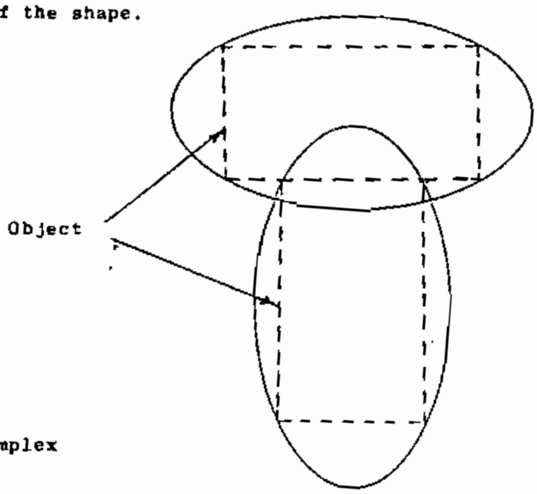

Fig. 1, Simplification of Complex Objects.

3- ROLE OF COMPUTER VISION

The importance of the computer vision is to provide a program in Cartesian Coordinates for a set of points that exist on the surface of the object. There exist different techniques for range finding using computer vision such as occlusion cues, texture gradient, focusing, image brightness, stereo disparity, camera motion, ultrasonic range finding, and laser techniques. For a good survey of these techniques see Jarvis[9].

Stereo disparity refers to the phenomenon by which the image of a 3-D object point shifts as the camera is moved laterally to the depth coordinate axis. For two such camera positions, simple geometry indicates that the image displacement (disparity) is inversely proportional to depth as measured from the camera. Processing the stereo image is known to take a considerable amount of time which creates a bottleneck and prevents the realisation of real time operation. To remedy this deficiency, the pixelwise processing is modified to become the patchwise processing. The left and the right cameras have rasers at a pixel rate of 1 MHz with an integration time of 0.25 ms [8].

The photometric stereo method [7] can produce a map for the surface orientation at each point of the object surface. In order to determine the depth of the point, the distance between at least one point on the object and the camera should be given and integration process is required to determine the locations of the other points.

4-COLLISION-AVOIDANCE PATH

In this section, design of collision-free path between the start and the goal positions of the robot is presented. Once the bounding Ellipsoid of the obstacle is constructed the collisionavoidance path can be determined as follows. In the conventional methods [3,10], the robot's end-effector is treated as a point at the origin of its coordinates and the robot's links is considered as sticks by expanding the bounding Ellipsoid by one-half of the arm-width of the robot.

In Fig. 2, assume the position S is the start position of the end-effector and the point G is its goal position, then it is easy to show that the shortest path between these two points is a straight line connecting them. Now, if this straight line intersects the extended Ellipsoid, say, at the two points A and B, then the end-effector should go around the Ellipsoid between these two points in order to avoid penetrating

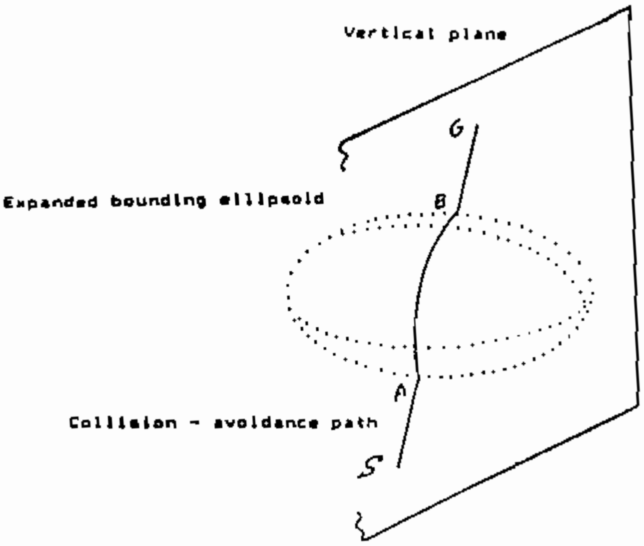

Fig. 2, Collision-Avoidance Path.

the obstacle. Geometrically a vertical plane can be constructed to pass through the straight line mentioned above, the intersection of this plane and the extended Ellipsoid will be an ellipse. Moreover, if a horizontal plane is used, another ellipse will result, the ellipse with the shortest length between the two points A and B is used. Finally the collision avoidance path consists of the line segment from S to A, followed by a portion of the ellipse between A and B and by the line segment from B to G. Fig. 3, shows a block diagram of the information flow of the overall scheme.

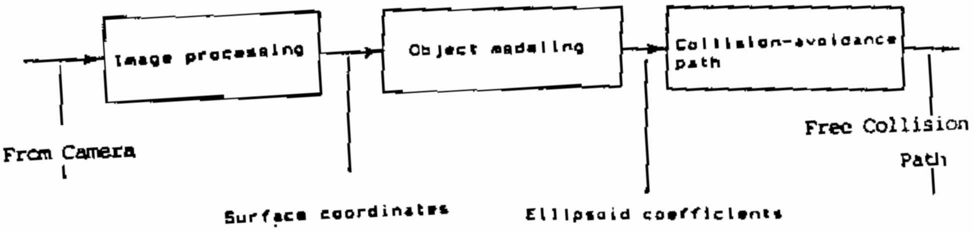

Fig. 3, System Block Diagram.

5-EXPERIMENTAL WORK

i- Practical Considerations

The vision system used in this work gives only two dimensional image of the objects. Therefore, two dimensional path planning is performed. Techniques which provide the third dimension (depth) from the 2-D images, such as shape from shading or stereo analysis, are not used here. These techniques require a long processing time which is not suitable for real time control of the robot arm. Laser range finder is one of the most suitable methods to evaluate the depth. However, it was not available for this work.

The vision system used is Autovision AV3 by Autnmax Incorporation. This system is able to take a picture of an object in its area of view and performs gray-and-connect analysis to determine 49 features for this object. These features are useful in different applications such as object recognition or inspection. For this work, only five features were used. These are:

1- OBJ-XCENT $\longleftrightarrow$ (X_c)
2- OBJ-YCENT $\longleftrightarrow$ (Y_c)
3- OBJ-MAJOR $\longleftrightarrow$ (a)
4- OBJ-MINOR $\longleftrightarrow$ (b)
5- OBJ-ANGLE $\longleftrightarrow$ (α)

These estimated features were used to formulate the ellipse model of the object as:

$$\underline{C}^{*T} \cdot \underline{X} = 1 + A$$

where $A = e^{+}_{max} + W + \Delta$

W = bandwidth of the robot arm

Δ = safety factor

The robot used is PUMA 650, by Unimation Incorporation. It has six joints and hence six degrees of freedom. The maximum bandwidth is about 90 mm, and the value of the safety factor Δ used was 10 mm.

Although, using OBJ-MAJOR and OBJ-MINOR do not guarantee that all the points of the object lay inside the ellipse. It was found that these features provide a suitable model for objects with less sharp edge variations.

ii-Calibrations

In order for the robot to acquire located parts by the vision system, there must exist a common cartesian frame of reference. This was accomplished by system calibration. First, the vision system is calibrated to establish an accurate two-dimensional coordinate system in the field of view. Second, the camera coordinate system must be related to the coordinate system of the robot. This was performed by determining a camera-to-robot coordinate transformation. A coordinate transformation is an operator which describes translation and rotation of a coordinate system from one location to another. Therefore, the camera-to-robot transformation describes the translation and rotation required to bring the camera coordinate system into coincidence with the robot coordinate system.

a) Vision System Calibration

The dimensions in the vision system are given in pixel units, wheras in the robot system they are given in millimeters. Therefore, the vision system must have its coordinates scaled in the same units as the robot coordinates system. The vision system is calibrated using the robot to measure a distance in the field of view. This was carried out as follows: A 'pointer' tool was held in the robot hand. A flat disk with a center hole was placed in the field of view. The robot is moved to place the pointer in the center hole of the disk. The robot location was recorded by typing HERE R1 on the terminal of the robot system, then the robot was moved out of the field view without disturbing the disk. The command PICTURE was typed on the terminal of the vision system, causing the vision system to determine the location of the disk in the field of view by recording OBJ-XCENT and OBJ-YCENT. If it is assumed that $R_1(x_{1r}, y_{1r})$ mm, and $V_1(x_{1v}, y_{1v})$ pixels, are the location of this point in the robot coordinates and in the vision coordinates respectivily. This procedure was repeated for a second location to determine $R_2(x_{2r}, y_{2r})$ and $V_2(x_{2v}, y_{2v})$. Now, with two robot locations and two corresponding disk locations in the field of view, the scale factor is defined as:

$$S = \frac{R_1B_2}{V_1V_2} = \sqrt{\frac{(x_{2r} - x_{1r})^2 + (y_{2r} - y_{1r})^2}{(x_{2v} - x_{1v})^2 + (y_{2v} - y_{1v})^2}}$$

The above procedure was repeated 10 times and an average scale factor of 1.9 mm/pixel was estimated.

b) Camera-to-Robot Calibration

Once the vision system has been calibrated, the transformation relating the field of view to the robot must be performed. In this experiment, the camera was mounted above the robot arm in its READY position with the axes chosen as: the x-axis of the vision system is parallel to the x-axis of the robot system and the y-axis of the vision system is anti-parallel to the y-axis of the robot system. Therefore, the relation between the two coordinate systems is pure translation (except the y-axis rotates by 180°). This linear transformation is expressed as:

$$x_r = C_1x_v + C_2$$

$$y_r = C_3y_v + C_4$$

where x_v and y_v are the coordinates in pixels, x_r and y_r are in millimeters, and C_1, C_2, C_3, and C_4 are constants. The above process was repeated and the coordinates of different points were determined. Using the values of these coordinates, the corresponding values of the constants C_1, C_2, C_3, and C_4 were calculated. In terms of average values of these constants, the linear transformation is given as:

$$x_r = 2.2\ x_v - 435$$

$$y_r = -1.9\ y_v - 769.3$$

iii - Program Execution

In the robot coordinates, the starting point of the robot hand (center of the gripper) was chosen at (500 , 500) mm and the goal point was at (-500 , 500) mm. That is the robot should move in a horizontal straight line if there is no obstacle on its way.

The object used was a rectangular hook of length 21 cm and width of 14 cm. This object was kept inbetween the start and the goal points as shown in Fig.4. The vision system was oriented to take a picture of the object through a program written in RAIL language. The results of this program is as follows:

```
OBJ-XCENT = 15T.18 pixels
OBJ-YCENT = 101.58 pixels
```

OBJ-MAJOR = 160.92 pixels
OBJ-MINOR = 125.83 pixels
OBJ-ANGLE = - 22.18°

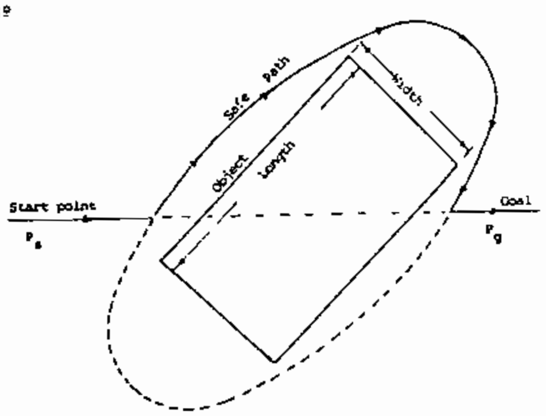

Fig. 4, Illustration of the Experimental Work.

These data were fed to another program (written in VAL II language) which controls the motion of the robot arm. The input to this program consists of the above five features and the output is a collection of points or locations, seperated by 1mm, through which the arm should pass. Analysis of the image features took about 0.5 sec. and the journey of the robot arm from the starting point to the goal point took about 51.61 sec.

6- CONCLUSIONS

A method for modeling 3-D objects using an Ellipsoid as a primitive is introduced. The least mean square estimation technique is used to calculate the coefficients of the Ellipsoid equation, such that the error in representing the object by the Ellipsoid becomes minimum. The use of Ellipsoid for modeling process has the following advantages:

1. It reduces the waste volume in representing objects more than the other models, such as bounding box or a sphere.
2. It makes the number of coefficients required to represent the volume occupied by the object always equal to 9, which is the number of the coefficients of the Ellipsoid equation. This number is very low compared to the other models, such as wire frame model.
3. Use of the analytic representation for the objects introduced in this paper enables us to express each object by only one constraint. This property will be very useful when looking to FIND-PATH problem as an optimisation problem.

The purpose of this paper was not to introduce an optimal path for the robot arm while it is moving, instead it introduced a simple and safe path. The determination of the optimal path using the proposed model may be done in a further work. Also, the study of another faster methods for calculating the coefficients of the fitting equation than that used in this paper may be a good extension in this direction.

7-REFERENCES

1- Morvac, H.P., "Obstacle Avoidance and Navigation in the Real World by a Seeing Robot Rover," Stanford Univ. Tech, Rep., AIM-340, Sep. 1980.
2-Brooks, R.A., "Solving the Find-Path Problem by Good Representation of Free Space," IEEE Trans. on System, Man, and Cybernetics, Vol. SMC-13, No.3, March/April 1983.

center hole of the disk. The robot location was recorded by typing HERE R1 on the terminal of the robot system, then the robot was moved out of the field view without disturbing the disk. The command PICTURE was typed on the terminal of the vision system, causing the vision system to determine the location of the disk in the field of view by recording OBJ-XCENT and OBJ-YCENT. If it is assumed that $R_1(x_{1r}, y_{1r})$ mm, and $V_1(x_{1v}, y_{1v})$ pixels, are the location of this point in the robot coordinates and in the vision coordinates respectivily. This procedure was repeated for a second location to determine $R_2(x_{2r}, y_{2r})$ and $V_2(x_{2v}, y_{2v})$. Now, with two robot locations and two corresponding disk locations in the field of view, the scale factor is defined as:

$$S = \frac{R_1R_2}{V_1V_2} = \sqrt{\frac{(x_{2r} - x_{1r})^2 + (y_{2r} - y_{1r})^2}{(x_{2v} - x_{1v})^2 + (y_{2v} - y_{1v})^2}}$$

The above procedure was repeated 10 times and an average scale factor of 1.9 mm/pixel was estimated.

b) Camera-to-Robot Calibration

Once the vision system has been calibrated, the transformation relating the field of view to the robot must be performed. In this experiment, the camera was mounted above the robot arm in its READY position with the axes chosen as: the x-axis of the vision system is parallel to the x-axis of the robot system and the y-axis of the vision system is anti-parallel to the y-axis of the robot system. Therefore, the relation between the two coordinate systems is pure translation (except the y-axis rotates by 180°). This linear transformation is expressed as:

$$x_r = C_1 x_v + C_2$$

$$y_r = C_3 y_v + C_4$$

where x_v and y_v are the coordinates in pixels, x_r and y_r are in millimeters, and C_1, C_2, C_3, and C_4 are constants. The above process was repeated and the coordinates of different points were determined. Using the values of these coordinates, the corresponding values of the constants C_1, C_2, C_3, and C_4 were calculated. In terms of average values of these constants, the linear transformation is given as:

$$x_r = 2.2\, x_v - 435$$

$$y_r = -1.9\, y_v - 769.3$$

iii - Program Execution

In the robot coordinates, the starting point of the robot hand (center of the gripper) was chosen at (500 , 500) mm and the goal point was at (-500 , 500) mm. That is the robot should move in a horizontal straight line if there is no obstacle on its way.

The object used was a rectangular book of length 21 cm and width of 14 cm. This object was kept inbetween the start and the goal points as shown in Fig.4. The vision system was oriented to take a picture of the object through a program written in RAIL language. The results of this program is as follows:

```
OBJ-XCENT = 157.18 pixels
OBJ-YCENT = 101.58 pixels
```

OBJ-MAJOR = 160.92 pixels
OBJ-MINOR = 125.83 pixels
OBJ-ANGLE = - 22.18°

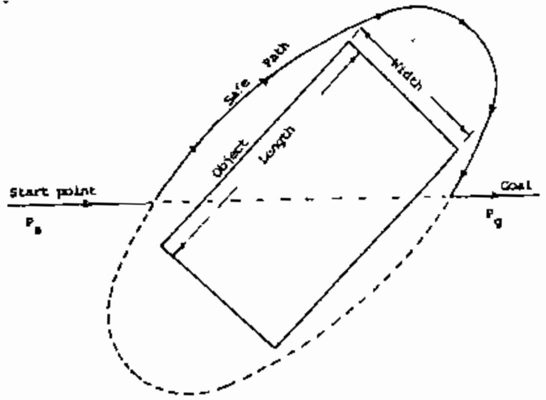

Fig. 4, Illustration of the Experimental Work.

These data were fed to another program (written in VAL II language) which controls the motion of the robot arm. The input to this program consists of the above five features and the output is a collection of points or locations, seperated by 1mm, through which the arm should pass. Analysis of the image features took about 0.5 sec. and the journey of the robot arm from the starting point to the goal point took about 51.61 sec.

6- CONCLUSIONS

A method for modeling 3-D objects using an Ellipsoid as a primitive is introduced. The least mean square estimation technique is used to calculate the coefficients of the Ellipsoid equation, such that the error in representing the object by the Ellipsoid becomes minimum. The use of Ellipsoid for modeling process has the following advantages:

1. It reduces the waste volume in representing objects more than the other models, such as bounding box or a sphere.
2. It makes the number of coefficients required to represent the volume occupied by the object always equal to 9, which is the number of the coefficients of the Ellipsoid equation. This number is very low compared to the other modele, such as wire frame model.
3. Use of the analytic representation for the objects introduced in this paper enables us to express each object by only one constraint. This property will be very useful when looking to FIND-PATH problem as an optimisation problem.

The purpose of this paper was not to introduce an optimal path for the robot arm while it is moving, instead it introduced a simple and safe path. The determination of the optimal path using the proposed model may be done in a further work. Also, the study of another faster methods for calculating the coefficients of the fitting equation than that used in this paper may be a good extension in this direction.

7-REFERENCES

1- Morvac, H.P., "Obstacle Avoidance and Navigation in the Real World by a Seeing Robot Rover," Stanford Univ. Tech. Rep., AIM-340, Sep. 1980.

2-Brooke, R.A., "Solving the Find-Path Problem by Good Representation of Free Space," IEEE Trans. on System, Man, and Cybernetics, Vol. SMC-13, No.3, March/April 1983.