MODELING AND CALIBRATION OF A STRUCTURED LIGHT SCANNER FOR 3-D ROBOT VISION

C. H. Chen and A. C. Kak

Robot Vision Lab
School of Electrical Engineering
Purdue University
West Lafayette, IN 47907

ABSTRACT

In this report we have used projectivity theory to model the process of structured light scanning for 3D robot vision. The projectivity formalism is used to derive a 4×3 transformation matrix that converts points in the image plane into their corresponding 3D world coordinates. Calibration of the scanner consists of computing the coefficient of this matrix by showing to the system a set of lines generated by suitable object edges. We end this paper by showing how the matrix can be used to convert image pixel locations into the world coordinates of the corresponding object points using two different scanning strategies.

1. INTRODUCTION

Structured light scanning is a rugged approach to range mapping a scene for 3D robot vision. In order to take full advantage of the flexibility for viewpoint selection made possible by a six-degree-of-freedom robot, we use a portable structured light unit that can be picked up by the robot when it wants to gather 3D vision data (Fig. 1). Within the constraints imposed by manipulator kinematics, the unit can then be oriented in any direction deemed desirable by the robot for the task at hand, and scanned either in a translational or a rotational mode for data collection.

A structured light unit consists basically of a light projector and a camera. The light projector throws a plane of light in the direction of the scene. The intersection of this plane with an object creates a stripe of illuminated points on the object surface, the stripe being recorded in the camera image plane. If the unit is properly calibrated, the world coordinates of the illuminated points can be calculated by using triangulation formulas, as has been done by Agin [agin82]. Agin used a 4×3 collineation matrix to write down a geometric relationship between the illuminated pixel coordinates and the world coordinates of the corresponding object points. The coefficients of this matrix are explicit functions of the camera and projector parameters. Calibration of the system implies determination of the coefficients of this matrix, which requires that the camera and the projector parameters be precisely known -- these parameters being positions and orientations of the camera and the projector, and the internal magnifications of the camera lens system. Because of the explicit dependence of the matrix coefficients on such parameters, Again had to first calibrate the robot joints so that the required positions could pinned down precisely, and then he had to individually calibrate the camera aim, camera scale and the projector aim.

In this report, we look at the calibration problem from a different point of view. The basic goal of structured-light calibration is to find a formula that converts the 2-D coordinates of a recorded pixel in the image plane to the world coordinates of the corresponding object point. Our position is that it should be possible to obtain this relationship for a structured light system without having to worry about such low-level details as the precise locations and aiming vectors for the camera and the projector. However, we do not believe that it is possible to do away with the requirement that the robot itself be mechanically calibrated before it can be used in conjunction with a structured light system. In fact, the accuracy of the methods to be proposed in this report will be no better than the absolute accuracy of the robot.

Note that the problem of deriving formulas that take us from 3-D world coordinates to 2-D image coordinates and vice versa also arises in straightforward camera imaging. As is well known [5], it is possible to write down a 3×4 homogeneous transformation matrix that for a given object point yields uniquely its corresponding image point; but, if we desire a transformation in the reverse direction, viz, from the image to the world, it is only possible to calculate the direction to the object point -- and not its location -- by using a similar matrix.

In Section 2, we will show that for the case of structured light imaging if we apply the theory of projectivity to relate the points in the light plane with the corresponding points in the image plane, it is indeed possible to derive a 4×3 homogeneous transformation matrix that is reversible. This implies that for each object point of *a priori* known location, we can uniquely determine its camera image plane coordinates; and for each image point we can uniquely determine the world coordinates of the corresponding object point.

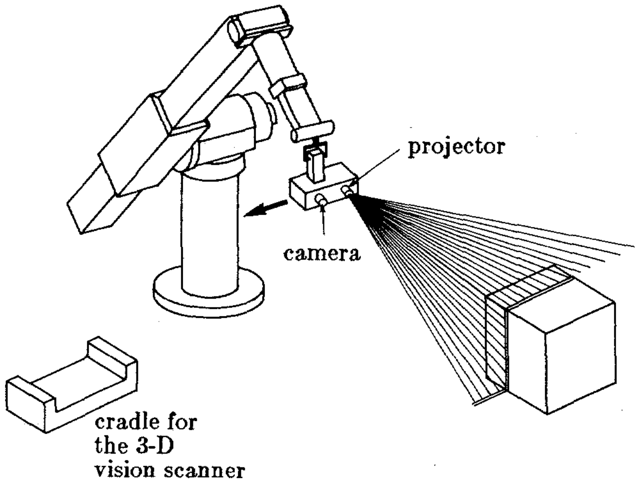


Fig. 1: Robot engaged in scanning a scene with a detachable structured-light scanner.

As we will show, the transformation matrix derived from the projectivity theory makes unnecessary the precise calculations of the locations of the camera and the projector and their aiming angles. Therefore, it is no longer critical that the robot joints be calibrated precisely, at least from the standpoint of enhancing the accuracy of range mapping.

We will also show that although from a purely theoretical standpoint only four object points at known locations are required for calibration -- meaning the computation of the elements of the transformation matrix -- the practical difficulty consisting of knowing where exactly the object points are located has caused us to seek other approaches. We will describe our procedure which consists of showing to the robot at least six lines generated by suitable object edges in the scene. In this procedure, it is not necessary to know the exact locations of the beginnings and the ends of the lines, as long as their relative separations are known. Section 3 presents a procedure for computing the optimum values of the calibration matrix when more than six lines are shown to the robot.

Once a structured light system is calibrated, the process of scanning for the purpose of range mapping a scene can take various forms. We will talk about two methods: rotational scanning and linear scanning. In Section 4, we will formulate coordinate transformations for both methods.

Finally, in Section 5, we will show some calibration results and compare our technique with the two-plane calibration method.

2. PROJECTIVE GEOMETRY

First, we will define the notation used in this report.

$X, U, P, A, \cdots$
: An italic upper case letter refers to a point which may be on a line, on a plane, or in 3D space. Usually, X, Y, Z are points in space, and U, V are points in the image plane.

$X_b, X_s, \cdots$
: An italic upper case letter with a subscript also refers to a point, but in this case the homogeneous coordinates of the point are also specified. The subscript denotes the coordinate frame in which the point is defined.

$\boldsymbol{X}_b, \boldsymbol{X}_s, \cdots$
: Bold italic upper case letters with subscripts are used to denote the regular coordinates of a point.

$\boldsymbol{r}, \boldsymbol{s}, \boldsymbol{t}, \cdots$
: A bold italic lower case letter is used to denote a line or a plane.

$F_{2s}, F_b, \cdots$
: Letter F with a subscript is used for representing a coordinate frame. The subscript 2 specifies a two dimensional coordinate frame.

$T_{cb}, \cdots$
: Letter T with a subscript represents a transformation from one coordinate system to another. The first letter of the subscript denotes the original coordinates system, while the second letter denotes the destination coordinate system.

One Dimensional Projectivity

On a plane, given a center of projection P and any two lines $\boldsymbol{s}$ and $\boldsymbol{r}$ not passing through P, as shown in Fig. 2, a one-dimensional projectivity is defined as follows: Let X be a point on line $\boldsymbol{s}$, its projective image X' on line $\boldsymbol{r}$ is the intersection of line PX with $\boldsymbol{r}$. Let A, B, C, D be any four

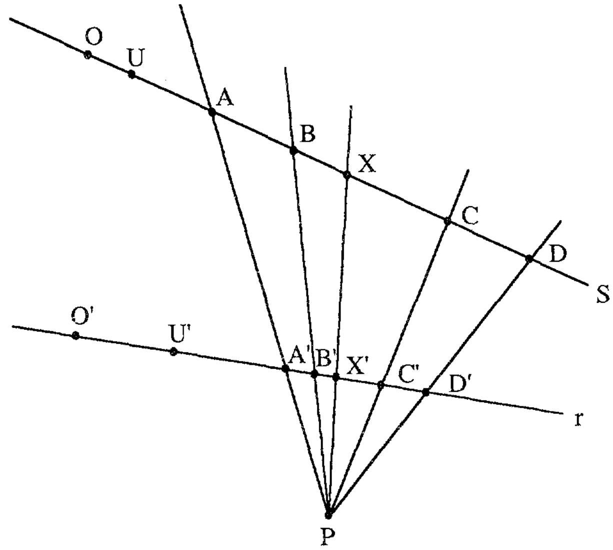


Fig. 2: One dimensional projectivity is illustrated here.

distinct points on line $\boldsymbol{s}$, the *cross ratio* of A, B with respect to C, D is

$$(A, B; C, D) = \frac{AC}{BC} \cdot \frac{BD}{AD}$$

Let A', B', C', D' on line $\boldsymbol{r}$ be, respectively, the image points of A, B, C, D under the projectivity shown. An important property that follows from projectivity is the invariance of the cross-ratio. This invariance can be expressed as

$$(A, B; C, D) = (A', B'; C', D') \tag{1-a}$$

or

$$\frac{AC}{BC} \frac{BD}{AD} = \frac{A'C'}{B'C'} \frac{B'D'}{A'D'} \tag{1-b}$$

With this relation established between $\boldsymbol{s}$ and $\boldsymbol{r}$, we can find the image point X' of X under this projectivity by substituting X for D, and X' for D':

$$(A, B; C, X) = (A', B'; C', X') \tag{2}$$

It is obvious that the two corresponding sets of triplets, $\{A, B, C\}$ and $\{A', B', C'\}$, completely describe the projectivity on $\boldsymbol{s}$ and $\boldsymbol{r}$ from the projection center P. One may raise the following questions at this point: Can we always find a projectivity on a plane which converts a set of points on one line to a set of points on another line ? Is this projectivity unique? Answers to these questions, which are crucial to the main theme of this paper, are provided by the following theorem [Ayres 67]:

> *The Fundamental Theorem of One Dimensional Projectivity*
>
> Given three distinct points on a line and another three points on a second line, there is one and only one projectivity which carries the first three points respectively into the second three points.

To illustrate the theorem, we first locate three points A, B, C at arbitrary places on a line $\boldsymbol{s}$ and another three points A', B', C' on a line $\boldsymbol{r}$ (Fig 3a). For finding the unique projectivity, we will fix the line $\boldsymbol{s}$ in the plane and move around the line $\boldsymbol{r}$ on the plane until the three lines AA', BB', CC' meet

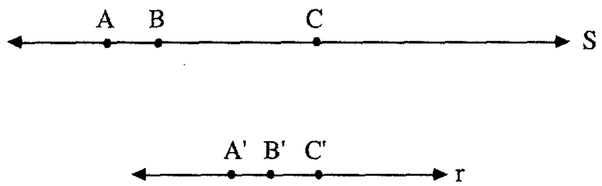


Fig. 3a: Three points defined on each of the two lines that will be used for demonstrating 1-D projectivity.

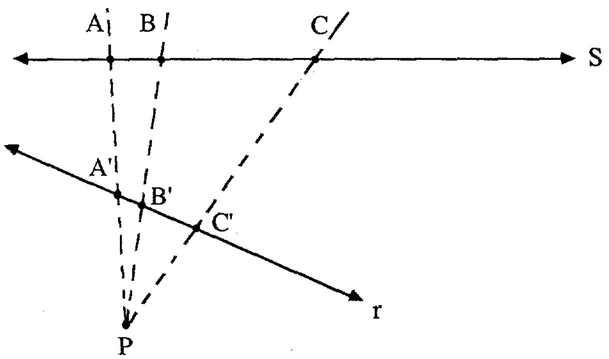


Fig. 3b: If we fix line s of (a) and move around line r shown there, there will exist only one projectivity for which AA', BB' and CC' will meet at a point.

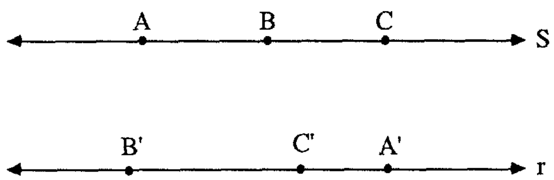


Fig. 4a: An example similar to that of Fig. 3a except that the order of the three points on line r is opposite to the order on line s.

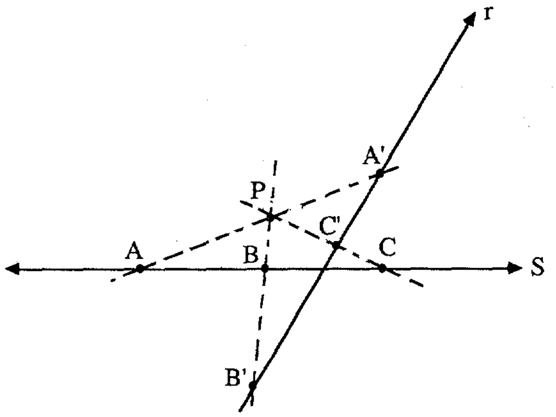


Fig. 4b: The unique projectivity that corresponds to the case shown in (a).

at one point (Fig 3b); this common point of intersection is the projection center of the projectivity. A more difficult case is shown in Fig. 4a, in which the corresponding points on lines s and r are ordered differently. The projectivity for this case is shown in Figure 4b.

For representing a point on a line, we need to define a coordinate system to express its position on the line. The familiar coordinate system on a line is established by selecting on the line a point O from which all measurements along the line are made, a unit of measure, and a sense of direction. Essentially, this consists of selecting a point O, called the origin, and U, called the unit point; to these two points we assign the coordinate values 0 and 1 respectively. The coordinate x of a point X on the line is then the directed distance of X from O. If on the other hand, a homogeneous coordinate system is desired, that can be done by assigning coordinates (0,1) to O, (1,1) to U, and (x_1, x_2) to any point X such that $x_1 / x_2 = x$. It is obvious that a point does not have a unique representation in a homogeneous coordinate system.

Let's say that we have chosen an origin O and a unit point U to define a coordinate system for a line s. Also, let O' and U' define a coordinate system for another line r. We do not require that the unit length OU on line s be equal to the unit length $O'U'$ on line r. We also do not require that the points O' and U' be the images of the points O and U under any projectivity. In fact, equation (2) is independent of the coordinate systems defined on either lines in the projectivity; this is a consequence of the following theorem that we present without proof:

Theorem of Cross Ratio

The cross ratio of any four points on a line is independent of the coordinate system established on the line.

Given a point x on, say, the line s, it is a simple matter to derive a formula for the corresponding point on line r. With respect to the coordinate system on line s, let the points A, B, C, X have coordinates a, b, c, x respectively. Similarly on line r, let the points A', B', C', X' have coordinates a', b', c', x' respectively. Then equation (2) can be rewritten as:

$$\frac{(a-c)(b-x)}{(b-c)(a-x)} = \frac{(a'-c')(b'-x')}{(b'-c')(a'-x')} \tag{3}$$

We now solve (2) for x' in terms of x. Setting $\frac{(a-c)}{(b-c)} = \alpha$ and $\frac{(a'-c')}{(b'-c')} = \beta$, we have

$$x' = \frac{a_{11}x + a_{12}}{a_{21}x + a_{22}}$$

where

$a_{11} = \alpha a' - \beta b'$, $a_{12} = ab'\beta - a'b\alpha$, $a_{21} = \alpha - \beta$, $a_{22} = a\beta - b\alpha$.

In terms of homogeneous coordinates, we have, by setting $x = x_1 / x_2$, and $x' = x'_1 / x'_2$,

$$\frac{x'_1}{x'_2} = \frac{a_{11}x_1 + a_{12}x_2}{a_{21}x_1 + a_{22}x_2}$$

or

$$\begin{cases} \rho x'_1 = a_{11}x_1 + a_{12}x_2 \\ \rho x'_2 = a_{21}x_1 + a_{22}x_2 \end{cases}, \rho \neq 0$$

In matrix form, we have

$$\rho \begin{bmatrix} x'_1 \\ x'_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{4}$$

Note that the existence of the free variable ρ. Since a point in homogeneous coordinates does not have a unique expression, that is, $x' = x_1 / x_2 = \rho x_1 / \rho x_2$. With the help of this free variable, we are ensured that regardless of the homogeneous coordinates chosen, the above expression for the projectivity solution will always satisfy equation (4). Also note that the roles of X and X' are exchangeable. We could consider X as the image of X', and we will get the same form of matrix equation as (3).

Two Dimensional Projectivity

We can establish a formalism for two dimensional projectivity in 3D space that is similar to the one dimensional projectivity in a plane. Let s and r be two planes in space and let there be a point P, which is neither on s nor on r, to be used as the center of projection (Fig. 5). For each point X on s, its image point X' on r is the intersection of line PX with plane r. It is obvious that the invariance of the cross-ratio is still valid for any four collinear points on s and their images point on r. Also, for any collinear points on s, their image points are also collinear. Extending the fundamental theorem of one dimensional projectivity, we have:

The Fundamental Theorem of Two Dimensional Projectivity

Given four distinct non-collinear points on a plane and another four distinct non-collinear points on the other plane, there is one and only one projectivity which carries the first four points respectively into the second four points.

A homogeneous coordinate system can also be established on a plane by a simple extension of what was done for line projectivity. Suppose we choose a point (0, 0, 1) as the origin in a plane and use two orthogonal unit points, (1, 0, 1) and (0, 1, 1), to lay out a coordinate frame in the plane. The homogeneous coordinates of any point in the plane are given by $(x_1 \, x_2 \, x_3)$ with $x_3 \neq 0$; $(x_1/x_3, \, x_2/x_3)$ are the regular coordinates of the point. Analogous to the derivation of equation (4), we can get a 3×3 conversion matrix which converts a point X on plane s to its image point X' on plane r, both points being expressed using homogeneous coordinates:

$$\rho \begin{bmatrix} x'_1 \\ x'_2 \\ x'_3 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{21} & a_{22} & a_{33} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

It is easy to verify that this equation preserves collinearity and invariance of the cross-ratio. Again, if we switch the roles of X and X', the above generic equation is still valid, i.e.,

$$\rho \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{21} & e_{22} & e_{33} \end{bmatrix} \cdot \begin{bmatrix} x'_1 \\ x'_2 \\ x'_3 \end{bmatrix} \tag{5}$$

A structured light scanner can be modeled by using 2D projectivity as follows. We use the camera-focus as the center of projection P, and treat the light stripe plane as plane s and the camera image plane as plane r. [This model is only valid under the condition that it be possible to use the pin-hole model for the camera (Fig. 6).] Although the coordinate system on the image plane can be arbitrary, a convenient definition consists of using the row index u and column index v of the digitized image as its two coordinates, and choosing the center of image plane as the origin. We will denote this coordinate frame on the image plane by F_{2c}. A point U in the image plane then has coordinates (u, v) or,

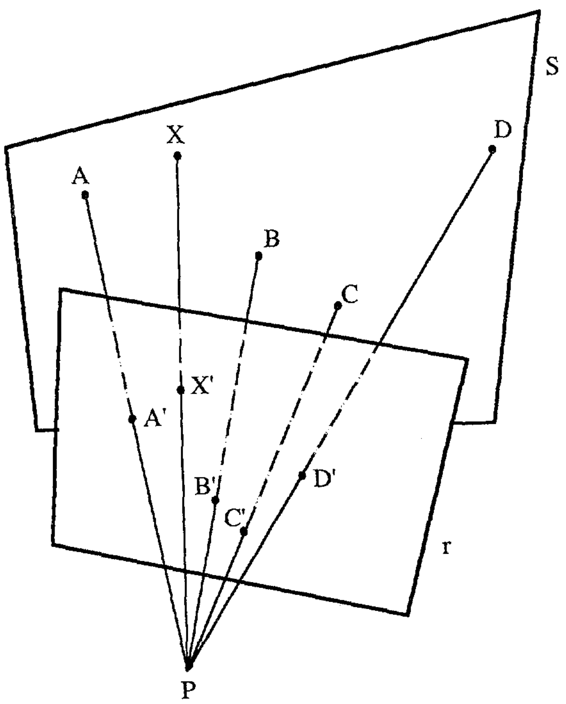


Fig. 5: Elements of two dimensional projectivity.

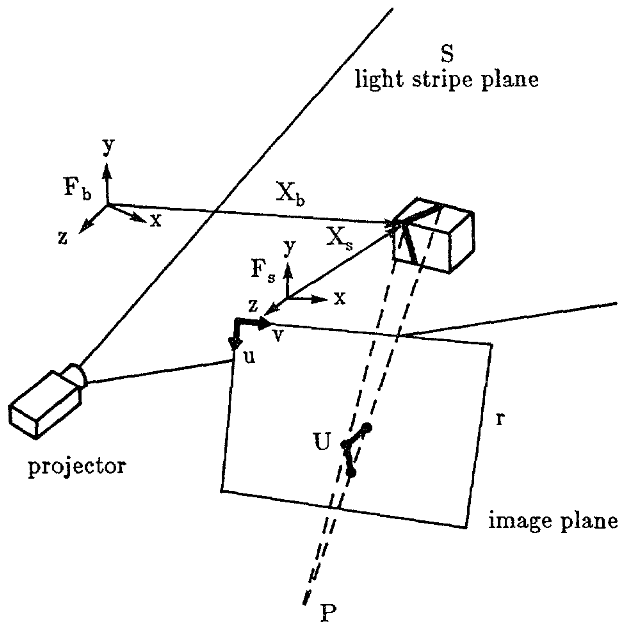


Fig. 6: This figure shows that the structured light imaging process can be fit precisely into 2-D projectivity. We can consider the light stripe plane as plane s and the camera image plane as plane r in drawing correspondence with Fig. 5. The camera focus center becomes the center of projection.

in a homogeneous coordinates system, $(u, v, 1)$ with respect to F_{2c}.

We also need to define a coordinate system on the light stripe plane. By virtue of the previous theorem, which says that the cross-ratios are independent of the choice of the coordinates system, we have considerable latitude in how we go about setting up this coordinate frame. We therefore choose one that can be easily related to the three dimensional base coordinate frame F_b for the robot. We will use x, y, z to represent the three orthogonal axes in F_b. Then a point X_b defined in the frame F_b will have homogeneous coordinates $w(x, y, z, 1) = (wx, wy, wz, w)$. Imagine a translation and a rotation that brings F_b to a coordinate frame F_s whose center is on the plane s and whose xy plane is aligned with the plane s. Since F_s is defined with respect to the base coordinate frame F_b, F_s contains all the information regarding the translation and rotation. Inheriting the coordinate system defined on the xy plane of the frame F_s, we can define a two dimensional coordinate frame F_{2s} on the plane s. Suppose a point X on plane s is assigned homogeneous coordinates (x_1, x_2, x_3) with respect to F_{2s}, where $x_3 \neq 0$. With respect to the frame F_s, which is three dimensional, the homogeneous coordinates of the same point are $(x_1, x_2, 0, x_3)$. The conversion of X from its two dimensional homogeneous coordinates in F_{2s} to its three dimensional homogeneous coordinates in F_s can then be written as

$$\begin{bmatrix} x_1 \\ x_2 \\ 0 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \tag{6}$$

Now let X_s be the homogeneous coordinates of X with respect to the frame F_s. We can convert X_s to the homogeneous coordinates representation X_b with respect to the base frame F_b by multiplying X_s with F_s, that is

$$X_b = F_s \cdot X_s \tag{7}$$

Here F_s is a 4×4 matrix.

Substituting $(u, v, 1)$ for (x'_1, x'_2, x'_3) in equation (5) and combining equations (6) and (7), we get a 4×3 conversion matrix T_{cb} that converts a point U in camera image plane to a light stripe point X_b in the robot base coordinate frame.

$$X_b = T_{cb} \cdot U$$

or

$$\rho \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \\ t_{41} & t_{42} & t_{43} \end{bmatrix} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \tag{8}$$

Note that we use subscript b to denote that X_b is in homogeneous coordinates with respect to the base coordinate frame F_b. Again, we use the free variable ρ to account for the non-uniqueness of homogeneous coordinate expressions.

3. SOLVING FOR THE CONVERSION MATRIX

We have shown that eq. (5) captures the general essence of two dimensional projectivity. For our particular case of transformations between the camera image plane and the light plane, the relationship represented by eq. (8) is however more suitable.

Obviously, the conversion matrix T_{cb} in eq. (8) depends upon both the positions and the orientations of the camera and the light plane projector. The purpose of calibration is to find this matrix without recourse to actually measuring these positions and orientations. Note that because of the free variable ρ in equation (8), we can set t_{43} in T_{cb} equal to 1 and the equation still holds. Our calibration is to determine the eleven unknown coefficients in T_{cb}.

We carry out our calibration by finding the 2-D projectivity that exists between the camera image plane and the light plane. By the fundamental theorem presented in Section 2.2, we can find this projectivity -- in principle at least -- by using four coplanar but non-collinear points in the light plane and their corresponding points in the image plane. By choosing four illuminated object points as calibration points, assuming that their 3-D coordinates and their cooresponding image coordinates can be measured correctly, we should be able to solve for the matrix T_{cb}. We will now show how one might set up equations for this purpose. Rewriting equation (8) as

$$\rho \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} T_1 \\ T_2 \\ T_3 \\ T_4 \end{bmatrix} \cdot U$$

and eliminating the free variable ρ, we have

$$\begin{aligned} x &= T_1 \cdot U \;/\; T_4 \cdot U \\ y &= T_2 \cdot U \;/\; T_4 \cdot U \\ z &= T_3 \cdot U \;/\; T_4 \cdot U \end{aligned} \tag{9}$$

or equivalently,

$$\begin{aligned} T_1 \cdot U - x\, T_4 \cdot U &= 0 \\ T_2 \cdot U - y\, T_4 \cdot U &= 0 \\ T_3 \cdot U - z\, T_4 \cdot U &= 0 \end{aligned} \tag{10}$$

Thus each calibration point produces a set of three linear equations in terms of the eleven coefficients of T_{cb}. Four calibration points would therefore lead to a set of twelve equations for the eleven unknowns. This number is one more than what we need. Since we could pick any eleven equations out of the twelve and get a solution for T_{cb}, we could ostensibly get different T_{cb}'s depending on the choice of the eleven equations; this would evidently be in contradiction to the uniqueness implied by the fundamental theorem of projectivity. However, we should note that the fundamental theorem requires the four calibration points to be coplanar. Therefore, the twelve 3-D coordinate values of the four points are not independent of one another, and, in fact, they obey the constraint of the co-plane equation:

$$\det [X_b^1 \, X_b^2 \, X_b^3 \, X_b^4] = 0$$

That is, one of the twelve coordinates is determined by the other eleven values. Since the above co-plane constraint is in fact implicit in equation (8), one of the twelve equations generated by the four calibration points is redundant. As a consequence, we can use any eleven equations and arrive at the same unique solution for T_{cb}.

A Procedure for Automatic Calibration

In practice, using four object points at *a priori* known locations for computing the matrix T_{cb} is beset with difficulties for the following reasons:

1) There are always some errors associated with the measurement of locations of the four calibration points in the robot base frame. On account of such errors, their coplanarity can not be completely guaranteed.
2) It is unrealistic to assume that the camera can be modeled perfectly by a pin-hole. A pin-hole model is of questionable validity, especially when zoom lenses are used. When the pin-hole approximation breaks down, there may be no unique center of projection.
3) Because of the non-zero thickness of the illumination stripe and other digitization aspects of camera imaging, there will always be some non-zero error associated with the location of the image point corresponding to an object point.

Since for these reasons T_{cb} can not be found exactly, our best hope is to estimate it by minimizing some error criterion in an over-determined system of linear equations. In other words, given more than 4 calibration points, we want to find the T_{cb} which best fits those calibration points in equation (10). This is equivalent to a linear least square problem, which can be solved simply by calling appropriate IMSL subroutines.

At this point the reader probably has the impression that, in order to calibrate a structured light system, one must first install in the robot work area a set of object points at *a apriori* known locations. However, that is not the case in practice. Since the robot is programmed to move the structured-light unit in discrete steps, it is possible that the light planes emitted from any of the allowed positions of the scanner will not illuminate the object points. One way to get around this difficulty is to use extended objects in the work area, the objects being of such a shape that at least four non-collinear points are illuminated by the light plane emitted from the projector. After the vision data is collected, the world coordinates of these object points are measured by moving to their locations the robot end-effector. Clearly, this method would only work if the mechanical calibration of the robot is accurate. This method is hard to automate. By automating a vision calibration procedure we mean the following: We want to place certain objects at strategic locations in the robot work area; then by simply having the robot record structured-light data on these objects at any time a calibration is desired, it should be possible for the associated computer to figure out the calibration parameters.

We will now propose a procedure that is easier to automate. A flat trapezoidal object is located permanently in the work area. The object is shaped in such a manner that no two edges of the top-surface are parallel to each other. The end coordinates of the top edges of these objects are known to the robot; therefore, one might say that the equations that define the lines corresponding to these edges are known. Consider one such line: Since a line can be defined as the intersection of two planes, it is described by the following two equations corresponding to the two planes.

$$\begin{cases} a_1 x + b_1 y + c_1 z = d_1 \\ a_2 x + b_2 y + c_2 z = d_2 \end{cases} \tag{11}$$

When the scanner projects a stripe intersecting this calibration line, it generates an illuminated point whose image coordinates are given by, say, U. While, of course, we can record the image coordinates of U, its world coordinates are unknown. In the procedure being described, we have no need for the world coordinates of the illuminated object point on the line. By substituting the right hand side of equations (9) for the x, y, z in (11), we have

$$\begin{cases} a_1 T_1 \cdot U + b_1 T_2 \cdot U + c_1 T_3 \cdot U = d_1 T_4 \cdot U \\ a_2 T_1 \cdot U + b_2 T_2 \cdot U + c_2 T_3 \cdot U = d_2 T_4 \cdot U \end{cases}$$

It shows that each calibration line is capable of producing a set of two equations in terms of the 11 coefficients of T_{cb}. Therefore, if we use at least six calibration lines, we will have a system of over-determined linear equation to estimate the conversion matrix. As we will describe below, it is not necessary to use six different calibration lines, although one could certainly do so.

In our current implementation of this procedure, we use only two distinct object edges, which are not parallel, for generating two calibration lines from any single viewpoint. By moving the structured light unit to different heights above the table, we can record the image coordinates of the same two edges for generating as many equations as we like. We will now describe a step-by-step description of the procedure. First note though that mounted in the robot work area is a flat object whose top surface is not parallel to the light plane of the scanner. After this initial setup, each time a calibration is carried out by the robot, it automatically carries out the following steps:

1) The robot moves the scanner to an initial position. The coordinate frame of the robot tool center is recorded.
2) The scanner makes projects a light plane onto the calibration block. This generates on the block a segment of the light stripe, whose two end points must lie on the two calibration lines respectively.
3) From the digitized image, record the image coordinates of the illuminated points corresponding to the two calibration lines. Substitute these image coordinates for U in the two line equations; this gives us four linear equations.
4) To acquire more calibration lines, use the robot to move the scanner by (d_x, d_y, d_z) to a new position. Now the line equations will become

$$\begin{cases} a_1 T_1 \cdot U + b_1 T_2 \cdot U + c_1 T_3 \cdot U = \\ \quad (d_1 - a_1 d_x - b_1 d_y - c_1 d_z) T_4 \cdot U \\ a_2 T_1 \cdot U + b_2 T_2 \cdot U + c_2 T_3 \cdot U = \\ \quad (d_2 - a_2 d_x - b_2 d_y - c_2 d_z) T_4 \cdot U \end{cases}$$

Go back to step 2).
5) Call the IMSL subroutine *llbqf* to find the best estimate of T_{cb}.

Note that the estimated conversion matrix is with respect to the scanner at the initial position only. We will remove this constraint in the next section.

$$\begin{cases} a_1 T_1 \cdot U + b_1 T_2 \cdot U + c_1 T_3 \cdot U = d_1 T_4 \cdot U \\ a_2 T_1 \cdot U + b_2 T_2 \cdot U + c_2 T_3 \cdot U = d_2 T_4 \cdot U \end{cases}$$

It shows that each calibration line is capable of producing a set of two equations in terms of the 11 coefficients of T_{cb}. Therefore, if we use at least six calibration lines, we will have a system of over-determined linear equation to estimate the conversion matrix. As we will describe below, it is not necessary to use six different calibration lines, although one could certainly do so.

In our current implementation of this procedure, we use only two distinct object edges, which are not parallel, for generating two calibration lines from any single viewpoint. By moving the structured light unit to different heights above

the table, we can record the image coordinates of the same two edges for generating as many equations as we like. We will now describe a step-by-step description of the procedure. First note though that mounted in the robot work area is a flat object whose top surface is not parallel to the light plane of the scanner. After this initial setup, each time a calibration is carried out by the robot, it automatically carries out the following steps:

1) The robot moves the scanner to an initial position. The coordinate frame of the robot tool center is recorded.
2) The scanner makes projects a light plane onto the calibration block. This generates on the block a segment of the light stripe, whose two end points must lie on the two calibration lines respectively.
3) From the digitized image, record the image coordinates of the illuminated points corresponding to the two calibration lines. Substitute these image coordinates for U in the two line equations; this gives us four linear equations.
4) To acquire more calibration lines, use the robot to move the scanner by (d_x, d_y, d_z) to a new position (Fig. 7). Now the line equations will become

$$\begin{cases} a_1 T_1 \cdot U + b_1 T_2 \cdot U + c_1 T_3 \cdot U = \\ \quad (d_1 - a_1 d_x - b_1 d_y - c_1 d_z) T_4 \cdot U \\ a_2 T_1 \cdot U + b_2 T_2 \cdot U + c_2 T_3 \cdot U = \\ \quad (d_2 - a_2 d_x - b_2 d_y - c_2 d_z) T_4 \cdot U \end{cases}$$

Go back to step 2).
5) Call the IMSL subroutine *llbqf* to find the best estimate of T_{cb}.

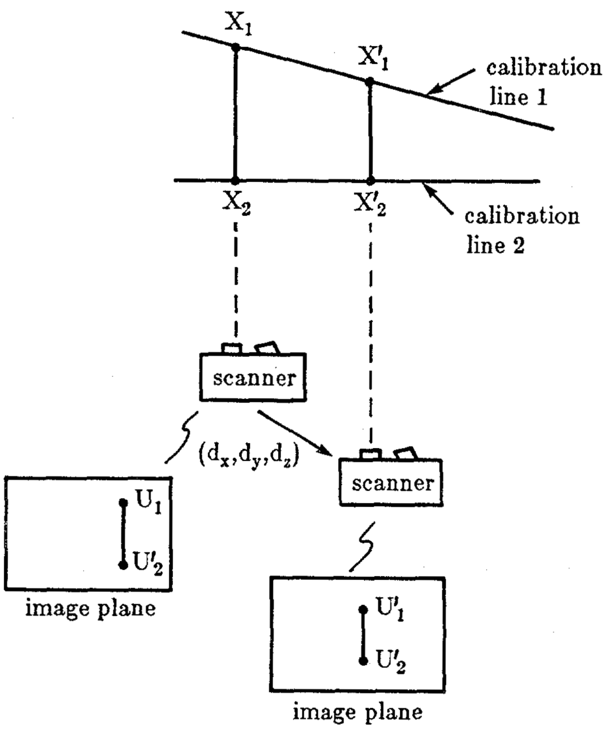


Fig. 7: To acquire more calibration lines, the robot moves the scanner by (d_x, d_y, d_z) to a new position and makes projection.

Note that the estimated conversion matrix is with respect to the scanner at the initial position only. We will remove this constraint in the next section.

4. LINEAR AND ROTATIONAL SCANNING

Formulation

If the range map of a scene is desired, the scene must be scanned in some manner with the structured-light unit. Linear scanning and rotational scanning are the two schemes used in our lab. In linear scanning, the orientation of the scanner is fixed, only its position is changed equally between successive light stripe projections, as shown in Figure 8. In

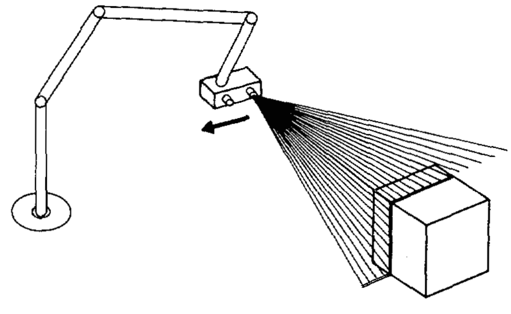


Fig. 8a: In linear scanning shown here, the orientation of the scanner is kept fixed while the scanner is translated along a line.

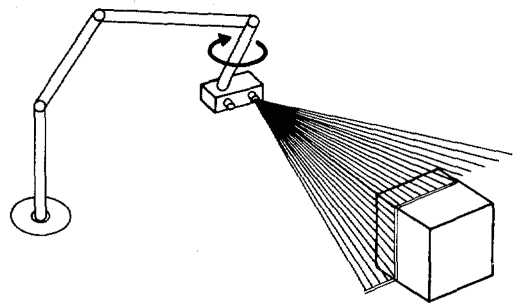


Fig. 8b: In rotational scanning, while holding the scanner at a fixed position the robot rotates the scanner in equal angular increments about the axis of the wrist joint.

rotational scanning, the robot holds the scanner at a fixed position, but rotates the scanner in equal angular increments about the axis of the wrist joint. The movement of the scanner is specified by the position and orientation of its end effector on which the tool-center is defined. Let us define the coordinate frame of the tool-center as F_t such that the z axis of F_t aligns with the axis of the robot's wrist joint. For the case of linear scanning, we will express the translational movement from projection to projection by $D = (d_x, d_y, d_z)$. This movement can be written as a translation transformation matrix:

$$H_d = \begin{bmatrix} 1 & 0 & 0 & d_x \\ 0 & 1 & 0 & d_y \\ 0 & 0 & 1 & d_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Similarly, for rotational scanning, if the angular increment between successive rotational positions of the scanner is δ, we can write down the following for a rotational transformation matrix

$$R_\delta = \begin{bmatrix} \cos\delta & -\sin\delta & 0 & 0 \\ \sin\delta & \cos\delta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The conversion matrix T_{cb}, as obtained from the calibration process, is defined in the base coordinate frame F_b with the scanner at a specific position. Let the tool-center coordinate frame used for calibration be F_{t_c}. When scanning a scene, the position and orientation of the scanner will differ from those used during calibration. Therefore, during scanning, the tool-center coordinate frame, as represented by F_t, will be different from F_{t_c}. As a result, the T_{cb} matrix obtained from calibration can not be plugged directly in equation (8) for the purpose of computing the range map of a scene.

To get over this problem, we can convert the matrix T_{cb} to a matrix T_{ct}, which is defined in the tool-center coordinate frame F_{t_c}. This is done by

$$T_{ct} = (F_{t_c})^{-1} \cdot T_{cb} \tag{12}$$

This relation is depicted in Fig. 9. Thus T_{ct} converts an image point U into the corresponding object point in homogeneous coordinates with respect to frame F_t. Let F_{t_0} be the tool-center coordinate frame at the beginning of a scan and let j denote the j^{th} projection in a scan. In linear scanning, we have

$$F_{t_j} = F_{t_0} \cdot (H_d)^j$$

Therefore, we get

$$X_b = F_{t_j} \cdot T_{ct} \cdot U$$
$$= F_{t_0} \cdot (H_d)^j \cdot T_{ct} \cdot U \tag{13}$$

Similarly, for rotational scanning, we have

$$X_b = F_{t_0} \cdot (R_\delta)^j \cdot T_{ct} \cdot U \tag{14}$$

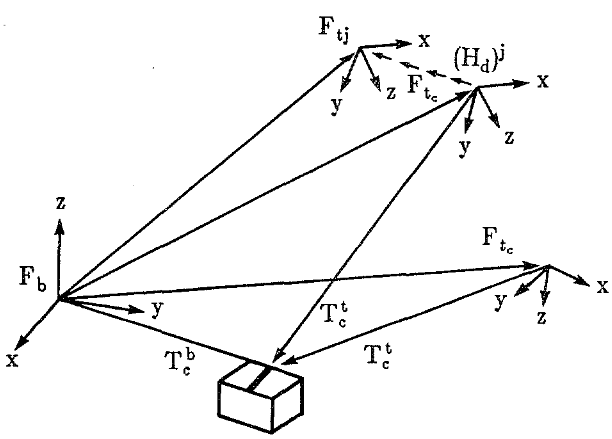


Fig. 9: Relation among coordinate frames for linear scanning.

Analysis of a Range Map

Equations (13) and (14) provide us with formulas for computing the range map of a scene. For each light stripe projection during scanning, we record the column index v of the sampled illuminated object point in each row of the camera image. By applying equation (13) or (14), for each row indexed by u we have the 3-D coordinates $[x(u), y(u), z(u)]$ of the object point. These 3-D coordinates are then collected into a range map.

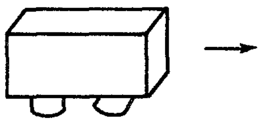
At this time, a few comments about the parametrization of the object surface are in order. Let row index of the scene range map be the same as the row index u of camera image plane; and let its column index be the index j associated with successive projections of the light stripes. Thus the range map can be expressed as $[x(u,j), y(u,j), z(u,j)]$. For example, if the camera image plane is of 480×512 resolution, and there are 80 projections in a scan, we will have a range map of size 480×80. Now consider the range map of a scene as the sampling of a visible surface, and assume that the surface is expressed as $f = [f_x, f_y, f_z]$. Its range map $f(u,j) = [f_x(u,j), f_y(u,j), f_z(u,j)]$ is the quantized parametrization of this visible surface. Note that the direction represented by the j index is directly related to the movement of the scanner from projection to projection. We want this "movement direction" to be perpendicular to the column direction of the camera image plane so that (u,j) will form an orthogonal parametrization of the surface. This can be important for later processing of the range map. For example, most 3-D edge detection operators are derived with the assumption of orthogonal parametrization.

5. EXPERIMENTAL RESULTS AND CONCLUSION

The structured light scanner used in our experiment consists of a Sony DC-37 CCD camera and an infrared projector. For conducting a calibration experiment, the calibration block is placed on the table and the scanner is moved to its initial position, which is about 20 inches above the table. A scan is then conducted along a line that is horizontal with respect to the work table; during the the calibration block is illuminated by three stripes. This process is repeated at four different heights, -- 20, 14, 8 and 2 inches -- above the work table, leading to range data on a total of 12 stripes. This data leads to 48 linear equations for the computation of the conversion matrix. The total time expended in the collection of calibration data is about a minute and the computer time for processing this information is about 3 seconds.

Although ultimately the evaluation of a calibration procedure must be carried out by determining the absolute accuracy of the system, for many purposes it is sufficient to compute the relative accuracy. By absolute accuracy we mean the precision with which the system locates a point with respect to the origin in the robot base coordinate system; and by relative accuracy we mean the precision with which the system makes a dimensional measurement of an object feature located in the robot area. In our experimental evaluation of the procedure described in this paper, we will only use relative accuracies. This is primarily owing to the fact that absolute accuracy tends to be a function of the accuracies of both the vision calibration and the robot arm calibration, meaning that a measurement of absolute accuracy may or may not tell us about the performance of a vision calibration technique.

After calibration, the relative accuracy of the procedure is evaluated by the computing the dimensions W and H of a block, like the one shown in Fig. 10. As expected, our exper-



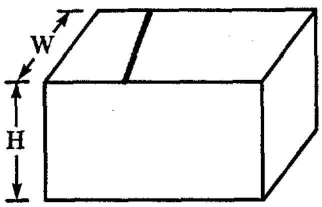


Fig. 10: The width and the height of the block are computed from the range data in order to test the relative accuracy of calibration.

Table 1. Relative accuracy test results

$W = 5.66$ *inch* $H = 6$ *inch*

d	8 *inch*	14 *inch*	20 *inch*
W_δ	--	<0.04 *inch*	<0.05 *inch*
H_δ	<0.03 *inch*	<0.14 *inch*	<0.30 *inch*

d: distance form the scanner to the block
W_δ : difference between the computed width and the real width
H_δ : difference between the computed height and the real height

imental results show that the accuracies with which these two measurements can be made depend upon the distance of the block from the structured-light unit and the orientation of the block with respect to the scan direction. For the results reported here, the long axis of the block was kept approximately parallel to the scan direction. The results are shown in Table 1.

The reader might note that we have not taken into account any nonlinear lens distortions in our development of the calibration procedure. We have seen that these distortions become important for object points that are far away from the camera lens, usually farther than two feet. Lens nonlinearities may be taken into account by a variety of techniques presented by Tsai [tsai86].

REFERENCES

[1] Gerald J. Agin and Peter T. Highnam "Movable Light-stripe Sensor for Obtaining Three-dimensional Coordinate Measurements" Proceedings of the SPIE Int. Tech. Symp., San Diego, CA, Aug 21-27, 1982. p 326-333

[2] Gerald J. Agin "Calibration and use of a Light Stripe Range Sensor Mounted on the Hand of a Robot" CMU-RI-TR-85-20 The Robotics Institute, Carnegie-Mellon University. 1985.

[3] Frank Ayres, Jr. "Projective Geometry" Schaum Publishing Co. 1967.

[4] Ballard, D. H., and Brown, C. M. "Computer Vision" Pretice-Hall 1982

[5] R. O. Duda and P. E. Hart "Pattern Recognition and Scene Analysis" New York, Wiley. p 379-423

[6] H. A. Martin, J. R. Birk and R. B. Kelley "Camera Models Based on Data from Two Calibration Planes" Computer Graphics and Image Processing, 17. p 173-180

[7] Roger Y. Tsai "An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision" IEEE Proceedings of Conf. on Computer Vision and Pattern Recognition 1986. p 364-374