Washington University in St. Louis

Washington University Open Scholarship

All Computer Science and Engineering
Research

Computer Science and Engineering

Report Number: WUCS-81-06

1981-10-01

Surface Geometry Acquisition using a Binary-Coded Structured Illumination Technique

Jeffrey L. Posdamer

Follow this and additional works at: https://openscholarship.wustl.edu/cse_research

Recommended Citation

Posdamer, Jeffrey L., "Surface Geometry Acquisition using a Binary-Coded Structured Illumination Technique" Report Number: WUCS-81-06 (1981). *All Computer Science and Engineering Research*. https://openscholarship.wustl.edu/cse_research/887

Department of Computer Science & Engineering - Washington University in St. Louis
Campus Box 1045 - St. Louis, MO - 63130 - ph: (314) 935-6160.

**Surface Geometry Acquisition using a Binary-Coded
Structured Illumination Technique**

Jeffrey L. Posdamer

WUCS-81-06

October 1981

**Department of Computer Science
Washington University
Campus Box 1045
One Brookings Drive
Saint Louis, MO 63130-4899**

***Research supported by the United States Air Force, School of Aerospace
Medicine, Contract F33615-78-D-0617.***

Submitted to Computers in Industry
10/1/81

Surface Geometry Acquisition Using a Binary-Coded
Structured Illumination Technique

by

Jeffrey L. Posdamer*

Medical Image Processing Group
Department of Computer Science
State University of New York at Buffalo

Research Supported by the United States Air Force,
School of Aerospace Medicine, Contract F33615-78-D-0617

Task #41

The U.S. Government may reproduce this paper for its purposes.

* Current Address: Department of Computer Science,
Washington University, St. Louis, MO 63130

0. Introduction

We live in a three-dimensional world. The observable domain of an (artificial or natural) object is delineated by the surface encompassing the object.

The ability to ascertain the surface geometry of solid objects is of increasing importance in manufacturing, materials handling and biomedical applications. While a significant amount of research has been performed on the problem of inferring geometric information from visual data [7], the direct, automatic acquisition of geometric information seems preferable.

Technologies for surface geometry acquisition include contact techniques and energy measurement (non-contact) techniques. Non-contact techniques may be further categorized as:

- (1) echometric
- (2) reflectometric; and
- (3) stereometric techniques.

Echometric systems make use of time-of-flight measurements of an energy pulse (sound, microwave, light) reflected from a surface to determine the distance to the surface. A collimated beam may yield more detailed surface structure. Reflectometric systems process one or more images of the object to determine, point-by-point, surface orientation and thus shape. Stereometric systems determine surface coordinates by triangulation of a single common distinguishable point between two separate surface sensors of the object.

1. Surface-mapping Mathematics

1.1 Triangulation

The standard mathematical technique is computational geometry for the representation of geometric data is homogeneous coordinates. After introducing the notation of homogeneous coordinates, the mathematics of triangulation will be reviewed.

1.1.2 Homogeneous Coordinates [4, 5, 6]

Homogeneous coordinates represent a data point in n-space with the use of an n+1-vector. Thus in three-space, a point is represented as:

$$\bar{P} = [hx \ hy \ hz \ h].$$

where hx , hy and hz are biliteral symbols, not products. The h coefficient may be viewed as a scale factor carried with each point. A projection from four-space to three-space, homogeneous normalization (H), provides for utilization of computational results.

$$H(\bar{P}) = \begin{cases} [\frac{hx}{h} \ \frac{hy}{h} \ \frac{hz}{h} \ 1] & h \neq 0 \\ [hx \ hy \ hz \ 0] & h = 0 \end{cases} \quad (1)$$

The condition $h=0$ is interpreted as a point, infinitely distant, reached by traversing the semi-line from the origin through the Cartesian point $[hx \ hy \ hz]$.

Homogeneous coordinates are used in computer graphics and computational geometry because they simplify the representation and computation of rigid body motion, scaling

$$\begin{bmatrix} T_{11} & T_{12} & T_{13} & 0 \\ T_{21} & T_{22} & T_{23} & 0 \\ T_{31} & T_{32} & T_{33} & 0 \\ T_{41} & T_{42} & T_{43} & \sigma \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1/D \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3)$$

Transformations

Perspective/Scaling

Thus, the transformation/perspective process yields a point in two-dimension

$$P_i = [hx_i^* \ hy_i^* \ 0 \ h_i^*] = [hx_i \ hy_i \ hz_i \ h_i] \begin{bmatrix} T_{11} & T_{13} & 0 & T_{14} \\ T_{21} & T_{22} & 0 & T_{24} \\ T_{31} & T_{32} & 0 & T_{34} \\ T_{41} & T_{42} & 0 & T_{44} \end{bmatrix} \quad (4)$$

where the product T matrix contains the positioning transformation as well as its perspective and scaling factors.

1.2 Perspective Inversion

Sutherland [5, 6] has presented an excellent discussion of the process of inverting the perspective projection so that three-space or model coordinates can be obtained from points whose images can be "distinguished" in two or more images. Two particular uses are of importance:

- (1) Image coordinates and T matrix are known; model space coordinates are unknown; and
- (2) Model coordinates and image coordinates are known; the T matrix is unknown.

be solved in some best-fit sense. Thus, any point which is distinguished for a pair of views satisfying case (1) assumptions may have its three-space coordinates computed. This is the essence of the stereometric or triangulation process.

Since the system of case (1) is overdetermined, an alternative to a best-fit solution is to use only the number of equations necessary for an exact solution. Consider the case in which a camera (defined by its T matrix) is used to determine an (x^*, y^*) for a point. A second device (defined by a transformation matrix L) which measures only one linear dimension (e.g. column) provides a single equation in coordinate u. Thus for the case of a camera and a columnar scanner we get:

$$\text{Camera } x^*: (T_{11}-T_{14}x^*)x+(T_{21}-T_{24}x^*)y+(T_{31}-T_{34}x^*)z+(T_{41}-T_{44}x^*)=0 \quad (6a)$$

$$\text{Camera } y^*: (T_{12}-T_{14}y^*)x+(T_{22}-T_{24}y^*)y+(T_{32}-T_{34}y^*)z+(T_{42}-T_{44}y^*)=0 \quad (6b)$$

$$\begin{array}{l} \text{Columnar} \\ \text{Sensor:} \end{array} \quad (L_{11}-L_{14}u)x+(L_{21}-L_{24}u)y+(L_{31}-L_{34}u)z+(L_{41}-L_{44}u)=0 \quad (6c)$$

Thus, with only the columnar equation known for the second sensor, the 3-space coordinates may be computed.

The solution of case (2) of the above is somewhat more complex. Given model points of known location and their images, the T matrix for the imaging system must be determined. Rewriting the pair of equations (6a, 6b) above

$$T_{11}x+T_{21}y+T_{31}z+T_{41}-T_{14}x x^*-T_{24}y x^*-T_{34}z x^*-T_{44}x^*=0 \quad (7a)$$

$$T_{12}x+T_{22}y+T_{32}z+T_{42}-T_{14}x y^*-T_{24}y y^*-T_{34}z y^*-T_{44}y^*=0 \quad (7b)$$

in which all T_{ij} are unknown. We have two equations in twelve unknowns. For six points whose model space and image coordinates are known, there are twelve equations in the

including stripes [3], line grids, random dots and regular dot arrays have been discussed. Many of these schemes encounter difficulty in matching the individual points between pairs of images. For arbitrary objects and surfaces there exists a second difficulty relating to the non-detection of points (i.e. illumination elements which do not appear in a sensed camera image), as for example out-of-scene surface elements, low-reflectance surface elements or imaging artifacts. These difficulties may be resolved if illuminated points carry an identification or code.

A simple example of such a scheme is the use of a single directed ray for illumination [10]. In this case a single bright spot illuminates a single surface sample point. Obviously, such an approach may suffer from excessive sampling time. For a 128×128 sample grid operating at conventional video rate (30 Hz), it would take more than eight seconds to acquire the more than 16000 frames of data. Clearly more efficient schemes of pattern illumination are needed.

Many schemes for encoding a light beam exist. Certain limitations in the imaging process limit the encoding technique. In order to make use of readily available imaging hardware, most specifically video, a typical frame rate of 30 Hz with a maximum of approximately 100 Hz, is necessary. Within each frame (image) any coding information may be thought of as being integrated over the image time. Thus no time-based encoding at frequencies greater than frame rate is usable. Another pragmatic limitation is that in many situations there is significant ambient illumination (thus noise). This requires the ability to distinguish

centric" coordination.

Clearly, in the case where a (relative) coordinate system is defined with respect to projector (or camera) coordinates, the L (or T) matrix is known directly. Because the camera can "see" fiducial points on an object which might not fall exactly on a projected ray, it is far easier to determine T with a known L than vice versa. Thus, a relative coordinate system based on the projector will be chosen for the center of coordinates. This yields an L matrix with only linear scaling and true perspective projection terms:

$$L = \begin{bmatrix} \sigma_u & 0 & 0 & 0 \\ 0 & \sigma_w & 0 & 0 \\ 0 & 0 & 0 & 1/\delta \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (8)$$

where σ_u , σ_w and δ are determined directly from the projector geometry.

The first variant of the system to be considered is the column-coded projector. In this version (currently operational) a column in the ray matrix is assigned a space code (u). The calibration problem to be solved is to calculate the plane equation coefficients of each plane determined by a shutter column and common focus point ($\bar{F}$) of the array. An absolute coordinate system will be used.

The reference object, a rectangular prism (block)(Fig. 1), is placed so that it is viewable by the camera and so that as many columns of rays as possible fall onto at least two

nonsingular. We thus have, for each space code (value of u), a set K of points, whose three-space coordinates are known.

$$K = \{K_u\} \equiv \{p \in R^3 \mid \sum_u (p) \approx 0\} \quad (11)$$

For any set of such points which cross two or more reference cube faces, plane equation coefficients may be calculated. For those sets failing to cross at least two reference object faces, a priori knowledge of shutter and projector geometry allow for interpolation to be used to determine $\sum_u$. Alternatively, one can use every imaged point (say N points altogether) to solve for the coefficients L_{11} , L_{21} , L_{31} , L_{41} , L_{14} , L_{24} , L_{34} of Eq. (6c) on the assumption that $L_{44} = 1$.

Whether we have coefficients σ_{u1} , σ_{u2} , σ_{u3} for each u plane or solve for the seven L_{ij} directly by least squares, we can solve for the scene intersection point (x,y,z) of an incident laser beam once we are given the corresponding values of x^* , y^* , and u .

An obvious variant of this system is to use two cross-mounted columnar shutters to allow for individual (as opposed to columnar) encoding of projected rays. In this situation, the T_{ij} matrix would again be determined by use of the reference block. Since each ray has a unique space-code which maps to its (u, w) coordinates, the determination of the L matrix is identical in method with determining T . For each ray projected onto a face of the reference object, a

a binary "space-code" allowing it to be distinguished from its neighbors [Fig. 2].

A ray projector system which takes advantage of a recently developed computer-controlled electro-optic shutter has been constructed. The shutter switches from transmission to occlusion of an individual laser ray at microsecond speeds. The laser/shutter/optics/microprocessor controller developed by Altschuler, Altschuler, Taboada and Frieder is described in [11].

Typically, the projector system is used as follows. An initial frame is projected in which no ray is blocked [Fig. 3(a)]. The camera views every beam/surface intersection dot that can be sensed. The image of each dot may vary with surface reflectivity, camera focus, the angle of the surface with both the beam and camera and ambient light conditions. By use of appropriate image processing and pattern recognition techniques, a record of each dot's (x^*, y^*) image coordinates is created. In each succeeding frame [Fig. 3(b), 3(c)] of the space-encoding sequence, the locations associated with the initial-frame dots are analyzed. If the location is "bright" it is encoding as a one in the appropriate bit position, otherwise a zero is inserted. Every detectable dot in the image, at the end of the encoding sequence, has an x^* , y^* and space-code u .

The list of (x_i^*, y_i^*, u_i) triples plus the T and L data determined at calibration time, serve as input to the computations used to generate the (x_i, y_i, z_i) data of the digitized, sampled surface. The result of the process is a set of points organized into subsets by binary space-code.

- (2) image dissector computer controlled video camera; and
- (3) processor to perform computations.

The projector array is created as a 128×128 array of beams. The shutter is a 128 column, columnar coded system under the control of a dedicated 8080-class processor. For experimental use a second, crossed columnar shutter may be mounted in the apparatus as a row shutter.

The image sensor is a computer-controlled Hamamatsu C-1000 digital camera system. The system is capable of operating at up to 1000 lines of resolution and 256 gray levels. It is equipped with a variety of lenses, extension tubes, etc.

The processor being used is an LSI-11/23 system with 192 KB memory, dual density floppy disks and RT-11. It serves as the camera controller as well as performing most computations.

In addition, various auxiliary computations have been performed on a Data General Eclipse S/200. Photographs in this paper were obtained from a COMTAL display on the Eclipse [Fig. 4-6].

References

1. M.D. Altschuler, B.R. Altschuler, and J. Taboada,
"Measuring Surfaces Space-Coded by a Laser-Projected
Dot Matrix", SPIE (Society of Photo-optical Instrumentation
Engineers), 182, 187-191, (1979).
2. B.R. Altschuler, J. Taboada, and M.D. Altschuler,
"A Laser Electro-optic System for 3-D Topographic Mensuration",
SPIE (Society of Photo-optical Instrumentation Engineers),
182, 192-195, (1979).
3. R.E. Herron, "Light Beam Profiler - Past, Present and Future",
SPIE (Society of Photo-optical Instrumentation Engineers),
283, (1981).
4. W. Newman and R. Sproull, Principles of Interactive Computer
Graphics, McGraw-Hill, 1973.
5. I.E. Sutherland, "Three-dimensional Input by Tablet",
Proc. IEEE, April, 1974.
6. Rogers and Adams, Mathematical Elements of Computer Graphics,
McGraw-Hill.
7. B. Horn, "Obtaining Shape from Shading Information" in the
Psychology of Computer Vision, P.H. Winston, Ed., McGraw-Hill,
1975.
8. M. Potmesil and H. Freeman, Computer Descriptions of Curved
Surface Objects from Multiple Projected-Pattern Images, TR
CRL-58, Electrical and Systems Engineering Department,
Rensselaer Polytechnic Institute, Troy, NY, June, 1978.

Fig. 1(a) One-inch reference cube, white light

Fig. 1(b) One-inch reference cube, Mask=1111111₂

Figure 2(b) Illumination of pattern enabled by left bit of space code. Two bit code.

```

o o o o
o o o o
o o o o
o o o o

```

Reference frame -
columns illuminated

Mask= 11

```

o o ● ●
o o ● ●
o o ● ●
o o ● ●

```

First illuminated frame

Mask=10

```

● o ● o
● o ● o
● o ● o
● o ● o

```

Second illuminated frame

Mask = 01

Fig. 3(a) Example of illumination patterns for space coding of 2 x 2 bit, 4 column, column-coded shutter. o is an illuminated spot; ● is an obscured beam.

Fig. 4(b) Mask = 1111111_2

Fig. 4(c) Mask = 1000000_2

Fig. 4(f) Mask=0000100₂

Fig. 4(g) Mask00000010₂

Fig. 5(a) Surface of tooth, White light

Fig. 5(b) Surface of tooth, Mask= 0001000₂

Fig. 6 (a) Cracked tip of turbine blade, White light

Fig. 6(b) Cracked tip of turbine blade, Mask=1111111₂
9(Note depth of field effects.)

Fig. 4(h) Mask=0000001₂

Fig. 4 (d) Mask=0100000₂

Fig. 4 (e) Mask=0010000₂

Fig. 4 Scanning sequence for 128 x 128 system

Each frame is labelled with the value of the column space-code mask. Consider the columns of the shutter to be indexed 0..127 from left to right. If the mask is ANDed with the columns binary representation and the result is non_zero, the column is illuminated.

Fig. 4(a) The curved sided pyramidal object in white light.

Figure 2(c) Illumination of pattern enabled by right bit of space code. Two bit code.

Figure 2(a) Illumination of reference pattern on object. Two bit space code.

9. J.L. Posdamer, Geometry of Orbiting Earth Satellites (GOES): Specular Geometry Study, Rome Air Development Center, in publication.
10. T. Kanade, H. Asada, "Non-contact Visual 3-D Ranging Devices", SPIE (Society of Photo-optical Instrumentation Engineers), 282, (1981).
11. M.D. Altschuler, J.L. Posdamer, G. Frieder, M.J. Manthey, B.R. Altschuler and J. Taboada, "A Medium-range Vision Aid for the Blind", 1980 International Conference on Cybernetics and Society, October 1980, pp. 1000-1002.
12. J.L. Posdamer, Spatial Sorting Techniques and Applications, in preparation.

3. Conclusion

The techniques described in association with a device such as a laser/shutter projection system will allow for high-speed geometry acquisition. Analysis of the mathematical and computational results from the engineering feasibility model have encouraged the construction of a real-time system in the immediate future. The use of direct acquisition of geometric representations as exposed to indirect techniques associated with vision/scene analysis systems provide obvious advantages in those applications in the precise geometry of the environment is the critical issue. Such applications include industrial and biomedical mensuration, robotics and aids for the visually handicapped.

Acknowledgement

This work has been done in conjunction with Drs. Bruce Altschuler and J. Taboada of the United States Air Force School of Medicine whose work on the laser/shutter system motivated this paper and Drs. M.D. Altschuler and Gideon Frieder of MIPG/SUNY at Buffalo who collaborated on this project.

1.6 Three-dimensional Geometric Modelling

The result of perspective inversion is a list of three-dimensional point coordinates organized by space code. Such a list constitutes a sampling of the three-dimensional surface. The generation of a usable, a posteriori, geometric model requires additional processing. The modelling schema which suggests itself in this circumstance is a faceted model. (For alternatives, see [15] or [16]).

To organize sample points into a faceted surface model, it is necessary to specify a connection or mesh topology. A natural approach is to utilize the grouping of points by space code since each code represents a planar slice of the object. Connections of points from adjacent space-code groups is similar (but clearly not identical) to the problem of triangulating a surface defined by sample points on parallel planar slices [12]. Another approach to generating a topology is the use of a divide-and-conquer approach based on octal-tree spatial sorting [14]. This technique allows for the merging of the results of geometry acquisition from multiple views/illuminations of the object needed to obtain full surface data. The generation of adequate surface models from sampled surface geometry is an open question.

2.0 Feasibility Study System

In order to investigate the problems associated with a laser-projector surface digitizer, an engineering model of the system has been implemented at the USAF School of Aerospace Medicine. It consists of

- (1) computer-controlled binary space-coded laser projector;

three-space coordinate is determined as in Eq. (10). In a manner identical with computing the T matrix, we thus write the equations of computing with the substitutions:

$$x \rightarrow u; y \rightarrow w; T_{ij} \rightarrow L_{ij}$$

Six or more non-coplanar points are thus sufficient to calculate the L matrix directly. This is in contrast to the technique used for the columnar shutter in which a collection of planes, each associated with a space-code, is determined. The L matrix now is equivalent to the second T matrix used to solve case (2) in Section 1.2.

1.5 Space-encoding of Projected Rays [1, 2]

For a system to triangulate an illuminated point, it must distinguish that point's image from all others. By means of interference techniques, a single collimated laser beam may, by two-stage interference, be converted into an array of projected rays. The rays do not differ sufficiently from one another to allow any visual distinguishing techniques. When such a pattern illuminates a surface, all visible, illuminated points will be detected, yielding a reference image.

We thus have an $n \times n$ array of projected light beams, almost everyone of which intersects the surface of interest and creates a bright "dot". Since all rays in the array are present in the initial reference image, the deletion of any projected ray will be detected by the absence of its dot in any succeeding image. By blocking a ray with a controllable shutter, a sequence of images can map a sequence of 1's and 0's for that ray. Thus, each position in space, illuminated by a ray is assigned

faces of the prism. The camera matrix T is determined from the image of the block and its known corner coordinates by the previously described method of Section 1.2. At this point the T matrix is known, as is the plane equation of each face of the reference cube.

Consider now the set of rays associated with a single column of the shutter. We wish to calculate $(\sum_u)$, defined by the shutter column with space code u and the focus point

$$\sum_u(\bar{p}) : \sigma_{u1}x + \sigma_{u2}y + \sigma_{u3}z + 1 = 0; \bar{p} = [x \ y \ z \ 1] \quad (9)$$

where the σ_{ui} are unknown.

For each ray of the column which intersects the surface of the cube at a point seen by the camera, an image space coordinate (x^*, y^*) is known. From the process used to determine the T matrix, the screen coordinates of the images of the visible corners of the reference cube are known. Thus, the region of the image associated with each cube face may be determined and, consequently, the face on which (x^*, y^*) we thus have:

$$(x^*) : (T_{11} - T_{14}x^*)x + (T_{21} - T_{24}x^*)y + (T_{31} - T_{34}x^*)z + (T_{41} - T_{44}x^*) = 0 \quad (10a)$$

$$(y^*) : (T_{12} - T_{14}y^*)x + (T_{22} - T_{24}y^*)y + (T_{32} - T_{34}y^*)z + (T_{42} - T_{44}y^*) = 0 \quad (10b)$$

$$(\text{face}) : \phi_{l1} x + \phi_{l2} y + \phi_{l3} z + 1 = 0 \quad (10c)$$

(where Eq. (10c) is the plane equation of face l of the reference cube).

Geometric considerations require that, if a cube face is visible to the camera, the system of equations is

the pattern from ambient "noise".

To summarize, the requirements for a structured illumination system, are:

- (1) Each pattern element must carry an identifying label or code;
- (2) The system must be able to operate with a frame rate somewhere in the range between 30 and 100 Hz;
- (3) The decoding scheme should be insensitive to ambient light noise;
- (4) The coding should be such that unimaged elements can be tagged as missing.

For the purposes of the next section, it will only be necessary to state that a system exists which projects an array of encoded rays, each a semi-line lying in a plane of known equation. The intersection of any such ray with the surface of an object creates an image which can be distinguished from any other.

1.4 Calibration

Calibration of this system consists of determining the T matrix for the camera and an L matrix (T 's projector equivalent) for the laser projector. An initial issue is the choice of a coordinate reference system. Two alternatives are suggested by differing applications. The more obvious system is one in which an "absolute" coordinate system is established in which both T and L will be determined. Typically, applications associated with object measurement would require such a system. A less obvious alternative would be to define a relative coordinate system defined with respect to either the camera or projector. Applications such as mobile robots would use "projector-

twelve unknowns. Since the equations are homogeneous, one unknown, (e.g., T_{44}) may be arbitrarily set. Thus for $T_{44}=1$, only eleven unknowns need be determined. Given six or more non-coplanar points whose model space coordinates are known, the imaging transformation for a specific camera/position, defined in terms of the model space, may be determined.

1.3 Distinguished Points

For stereometry or triangulation to be at all useful, it is necessary to distinguish points common to pairs of images.

Automatic surface geometry acquisition techniques for images of fully illuminated scenes are based on pattern recognition or image analysis approaches. This should be distinguished from image understanding or vision systems in which the desired result is the recognition of abstract three-dimensional objects in the scene as opposed to surface geometry. An area of interest is chosen in one image. Typically the choice might be made on the richness of texture or syntactic features of an area. Correlation techniques are then used to find a maximally matched area in the second scene. The success rate of the method will be degraded by substantial mismatches of the lighting, sensor orientation, or directional sensitivity between images. Additionally, images of the classes which are troublesome for humans may also be difficult for correlation techniques.

The alternative to full illumination of scenes is structured illumination of scenes. A wide variety of such patterns

In the former case, which is of primary concern to this discussion, the model space coordinates x_i, y_i, z_i are the unknowns to be determined from the image data and the known geometry of the imaging system; in the latter case the T matrix for a particular imaging system is to be found from the known spatial (model) and image coordinates of a small number of sample points. For both cases, the starting point is (from (4) above):

$$\begin{aligned}hx^* &= T_{11}hx + T_{21}hy + T_{31}hz + T_{41}h \\hy^* &= T_{12}hx + T_{22}hy + T_{32}hz + T_{42}h \\h^* &= T_{14}hx + T_{24}hy + T_{34}hz + T_{44}h\end{aligned}\tag{5}$$

(where the subscript on coordinates are understood).

For the case of image points not at infinity (the case for any imaged object), we have

$h^* \neq 0$ and can eliminate h^* in Eq. (5) by application of (1). Thus, $hx^* = h^* \times x^*$ and $hy^* = h^* \times y^*$

$$(T_{11}-T_{14}x^*)x+(T_{21}-T_{24}x^*)y+(T_{31}-T_{34}x^*)z+(T_{41}-T_{44}x^*)=0\tag{6a}$$

$$(T_{12}-T_{14}y^*)x+(T_{22}-T_{24}y^*)y+(T_{32}-T_{34}y^*)z+(T_{42}-T_{44}y^*)=0\tag{6b}$$

Now, in case (1) above (T matrix and image points known), there are three unknowns (x, y and z) in two equations. If a point on the model at $[x_i, y_i, z_i]$ can be identified in two views (with different viewpoints) of the object it is distinguished for those two views. Assuming both of these views satisfy case (1) assumptions, four equations (two per view) in three unknowns are generated. The overdetermined system can

and perspective projection.

The use of homogeneous coordinates to represent three-space geometry permits the definition of rigid body transformations, scaling and perspective projection as 4×4 matrices which operate (by post-multiplication) on $n \times 4$ matrix representations of n three-space points. Thus, the general transformation (T) on n points $\{\bar{P}_i = [hx_i \ hy_i \ hz_i \ h_i] \ i = 1 \dots n\}$; may be represented as:

$$\begin{array}{ccc}
 \begin{bmatrix} hx_1 & hy_1 & hz_1 & h_1 \\ & & \vdots & \\ hx_i & hy_i & hz_i & h_i \\ & & \vdots & \\ hx_n & hy_n & hz_n & h_n \end{bmatrix} & \begin{bmatrix} T_{11} & T_{12} & T_{13} & T_{14} \\ & \vdots & & \vdots \\ T_{21} & T_{22} & T_{23} & T_{24} \\ & \vdots & & \vdots \\ T_{31} & T_{32} & T_{33} & T_{34} \\ - & - & - & - \\ T_{41} & T_{42} & T_{43} & T_{44} \end{bmatrix} & = \begin{bmatrix} hx'_1 & hy'_1 & hz'_1 & h'_1 \\ & & \vdots & \\ hx'_i & hy'_i & hz'_i & h'_i \\ & & \vdots & \\ hx'_n & hy'_n & hz'_n & h'_n \end{bmatrix} \\
 n \times 4 & 4 \times 4 & n \times 4
 \end{array} \quad (2)$$

It should be noted that the T matrix may represent transformations in addition to those listed above. The upper left 3×3 sub-matrix contains the rotation terms, the lower left 1×3 is used for translation. The fourth column contains the perspective projection and/or scaling terms.

A specific case is the "standard" system with projection plane at $z=0$, focal point at $[0, 0, -1]$. This standard system simplifies the presentation of the transformation, projection computation by allowing for specification of the two steps in separate multiplied matrices, i.e.

A stereometric system must: Recognize each model measurement point of a scene in two (or more) images of the scene; measure and pair the image positions of each sample of the point; determine the relative geometry of the imaging devices; and, perform the actual triangulation to determine surface point coordinates.

The system described is a stereometric system developed at the USAF School of Aerospace Medicine, Dental Investigation Service based on a laser projector which illuminates the scene with a structured, binary-coded light pattern. A computer-controlled electronic shutter array assigns a binary space code to each ray in a matrix of laser generated beams. This code allows for the proper identification of image points in multiple images of a scene. The design of the shutter allows for calibration of shutter and camera geometry with only a single simple reference object. Point locations in the 3-D scene are computed automatically from two data items:

- (1) the location in a camera image of the light dot created by the intersection of a laser ray with a surface in the 3-D scene, and;
- (2) the space code associated with that dot.

Abstract

A wide variety of techniques have been used to automatically measure the surface geometry of three-dimensional objects. The computational and mathematical techniques associated with binary-coded structured-light system is developed. Analysis and testing of a feasibility system indicates that such a system is capable of directly measuring several thousand surface points in real-time.

A system which projects a square array of laser beams into a three-dimensional environment has been developed. By use of a computer-controlled electro-optic shutter, a binary code may be imposed on a beam or group of beams. The use of this structured, binary-coded illumination and a single camera are sufficient to determine the three-dimensional coordinates of any illuminated point.