

DESCRIPTION

Method for Measuring Geometric Properties of Patterns
Displayed on A Curved Surface

5

Field of The Invention

This invention relates in general to 3-D machine vision methods for determining the locations of points on the surface of an object from one or more digital images

10 captured by one or more 2-D sensors such as a camera.

This invention is intended to be used in the adjustment of picture geometry for cathode-ray tube (CRT).

Background of The Invention

15 In CRT geometry adjustment, geometrical patterns are displayed on the CRT. Then, geometrical dimensions are measured from these patterns. The display characteristics of the CRT is adjusted based on the differences between these measurements and a set of desired dimensions. Conventional
20 methods of automated CRT adjustment have either assumed that the CRT under analysis is placed firmly on a palette where its pose is completely known, or measured the relative displacements of pattern features with respect to the bezel. The former requires a high precision placement mechanism
25 which is costly, while the latter requires the CRT screen to be flat so that simple mathematics can be used to correct for placement error. Related methods can be found in the U.S. Pat. Nos. 5,510,833 (Webb et al.) and 5,216,504 (Webb et al.).

In order to measure the geometrical properties of patterns on a curved CRT surface, 3-D coordinates measuring techniques are required. 3-D machine vision systems utilize a number of
5 range finding techniques to locate points on the surface of an object. Typically, these techniques use active light-stripe triangulation, time-of-flight or stereovision. The drawbacks of these techniques include the requirement of a projection of structured light onto the scene, an expensive
10 laser source or the matching of features between two images. The matching of features between two images is itself a difficult problem. Related methods are disclosed in the U.S. Pat. Nos. 4,412,121 (Kremers et al.); 4,099,880 (Kano) and 3,961,851 (Gerharz).

15

In a different strategy, the perspective information of a camera image can be used to deduce the pose of a 3-D object of known geometry. Once the object pose is determined, the surface equation of this object can be calculated and the
20 coordinate of any feature point on this surface can be determined. The additional requirement for using such a method is that the shape of the object must be known in advance.

25 The technique of pose determination using perspective information is disclosed in the U.S. Pat. No. 4,942,539 (McGee et al.). In this method, certain dimensions of the object must be known in advance. These dimensions are defined by a set of at least three non-collinear points on the object.

These points will be referred to as the control points. However, the approach used in this prior art is computationally intensive for a large set of control points. In this method, there is an extra stage of geometrical error
5 minimization before the calculation of pose parameters. In addition, the number of terms in the error function used in the minimization process increases quadratically with the number of control points on the object.

10 The problem to be solved is to determine the 3-D coordinates of feature points on a curved surface from as few as a single digital image generated by as few as a single 2-D sensor such as a camera. The solution to this problem requires the pose of the surface to be determined. In the prior art (U.S. Pat.
15 No. 4,942,539, issued to McGee et al.), the method of finding the pose of an object is via a search to arrive at geometric consistency of the control points. However, the resultant error function involves in the minimization process requires the comparison of distances for every pairs of the control
20 points. This leads to a quadratic complexity of the number of terms in the error function. As a result, the computational effort increases quadratically with the number of control points.

25 Disclosure of The Invention

Instead of minimizing the errors in the geometrical distances among the control points, the search for the optimal pose can be found by minimizing the fitting errors between the set of control points with the set of 3-D lines where these control

points are known to lie on. These 3-D lines are referred to as the projection lines whose exact definition will be presented in Preferred Embodiments. It is adequate here to point out that the projection line can be calculated from the location of a point in the image. The set of control points are first positioned in the 3-D space with ideal geometrical distances among them. The search for the optimal fitting is carried out in the rigid transformation space where the original geometry of the control points is maintained throughout the search. The optimal rigid transformation, consists of a rotation and a translation, is searched such that the fitting errors between the control points and the projection lines are minimized. An efficient computational solution is achieved by this approach and there is no need of an extra stage of pose parameters calculation. The number of computation steps in this method increases linearly with the number of control points involved. This method is more efficient than the prior art.

That is, the present invention is a method (hereafter referred to as "Method 1") for automatically determining the position and orientation of an object in 3-D space from one or more digital images captured by one or more 2-D sensors, the method comprising the steps of:

generating sensor calibration data relating the position and orientation of the 2-D sensor to the 3-D space;

generating reference data relating to at least three non-collinear control points of a reference model;

generating the as few as one digital image containing the at least three non-collinear control points of the object;

5 locating each of the control points in the as few as one digital image;

 computing at least three non-parallel 3-D lines as a function of the locations of the control points and the sensor calibration data, each of the lines passing through its respective control point on the object; and

10 determining the pose of the object by minimizing the fitting errors of the control points of the reference model to the 3-D lines.

The present invention is also directed to a method (hereafter referred to as "Method 2") for determining the locations of feature points on a surface of an object in 3-D space from one or more digital images captured by one or more 2-D sensors, the method comprising the steps of:

20 generating sensor calibration data relating the position and orientation of the 2-D sensor to the 3-D space;

 generating reference data relating to at least three non-collinear control points of a reference model;

25 generating the as few as one digital image containing the at least three non-collinear control points of the object;

 locating each of the control points in the as few as one digital image;

 computing at least three non-parallel 3-D lines as a function of the locations of the control points and the

sensor calibration data, each of the lines passing through its respective control point on the object;

determining the pose of the object by minimizing the fitting errors of the control points of the reference model
5 to the 3-D lines;

determining the equation for the surface of the object;

locating feature points in the as few as one digital image;

computing 3-D lines as a function of the locations of
10 the feature points and the sensor calibration data, each of the lines passing through its respective feature point on the object; and

determining the locations of the feature points by the intersections of their respective 3-D lines and the surface
15 of the object.

Other embodiments of the present invention are as follows:

i) In Method 1 or 2, the object is a rigid body.

20 ii) In Method 1 or 2, the 2-D sensor is a camera.

iii) In Method 1 or 2, said step of pose determination involves minimizing the sum of the squared distances from the control points of the reference model to the 3-D lines.

25

iv) In Method 2, said step of equation determination makes use of the pose information of the object.

Brief Description of The Drawings

Figure 1 shows a schematic diagram of the system used in the current invention.

Figure 2 shows the algorithm flowchart of the current
5 invention.

Figure 3 is a diagram of a control point, p , and its related projection line useful in visualising the mathematics utilised in the method.

10 Preferred Embodiments for Carrying out The Invention

An object to be measured is placed in front of the set of cameras. An image is taken from each camera. The control points on the object are detected in the images to generate the projection lines. Then, the optimal pose is determined to
15 fit a set of control points with ideal geometrical distances among them through this set of lines. To determine the coordinates of the feature points detected on the images, the projection lines related to these feature points are calculated. The intersections between these projection lines
20 and the surface of the object give the coordinates of these feature points.

The current invention applies to a system of a single camera or multiple cameras. In the following description, by
25 referring to a set of cameras, we also mean to include a system of a single camera.

Referring to Figure 1, a set of cameras 10 are positioned to view at different portions of an object 14. Object 14 is

meant to be a CRT, but the current method also applies generally to any other object. Images of CRT 14 are captured by cameras 10 and processed by a computer 11. The patterns displayed on CRT 14 are produced by a TV signal generator 13 which is under the instruction of computer 11.

In camera calibration, process 20 in Figure 2, object 14 is replaced by a calibration chart. All cameras are calibrated to obtain their 2-D-to-3-D image projection parameters using the DLT (Direct Linear Transformation) method, where all cameras are calibrated to a common world coordinate system. A reference to the DLT method is given in:

Y.I.Abdel-Aziz and H.M.Karara, "Direct linear transformation from comparator coordinates into object space coordinates in close-range photogrammetry", *Proc. Symposium on Close Range Photogrammetry*, Urbana, Illinois, pp.1-18, January 1971.

The DLT camera model is expressed as follow:

$$u = \frac{A_{11}x + A_{12}y + A_{13}z + A_{14}}{A_{31}x + A_{32}y + A_{33}z + 1} \quad (1)$$

$$v = \frac{A_{21}x + A_{22}y + A_{23}z + A_{24}}{A_{31}x + A_{32}y + A_{33}z + 1} \quad (2)$$

where (x,y,z) is a world coordinate, and (u,v) is its projected image point expressed in the image coordinate. The

calibration process is a process of estimating the values for the eleven parameters $A_{11}, A_{12}, A_{13}, \dots, A_{33}$.

Note that given an image point (u,v) , it is possible to
 5 determine the line that passes through the camera's focal point and the object point (x,y,z) . This line is referred to as the projection line. An example of such a line is 12 which passes through the corner A of object 14. Since the DLT equations (1) and (2) hold true for both the image point and
 10 the object point, the projection line can be calculated from these two equations as will be described in the next paragraph.

Given an image point (u,v) , the projection line, after a few
 15 steps of algebraic manipulations from equations (1) and (2), has the following parametric form in terms of parameter t ,

$$x = P_x t + Q_x \quad (3)$$

$$y = P_y t + Q_y \quad (4)$$

$$20 \quad z = P_z t + Q_z \quad (5)$$

where

$$P_x = \frac{(A_{13} - uA_{33})(vA_{32} - A_{22}) - (uA_{32} - A_{12})(A_{23} - vA_{33})}{(uA_{31} - A_{11})(vA_{32} - A_{22}) - (uA_{32} - A_{12})(vA_{31} - A_{21})}$$

$$Q_x = \frac{(A_{14} - u)(vA_{32} - A_{22}) - (uA_{32} - A_{12})(A_{24} - v)}{(uA_{31} - A_{11})(vA_{32} - A_{22}) - (uA_{32} - A_{12})(vA_{31} - A_{21})}$$

$$P_y = \frac{(A_{23} - vA_{33})(uA_{31} - A_{11}) - (vA_{31} - A_{21})(A_{13} - uA_{33})}{(uA_{31} - A_{11})(vA_{32} - A_{22}) - (uA_{32} - A_{12})(vA_{31} - A_{21})}$$

$$25 \quad Q_y = \frac{(A_{24} - v)(uA_{31} - A_{11}) - (vA_{31} - A_{21})(A_{14} - u)}{(uA_{31} - A_{11})(vA_{32} - A_{22}) - (uA_{32} - A_{12})(vA_{31} - A_{21})}$$

$$P_i = 1$$

$$Q_i = 0$$

In object pose determination, an object is placed in front of the set of cameras 10. This object must possess at least three control points, all of which must be non-collinear, where the geometrical relationships among them have been measured (in process 28 of Figure 2). An example of control points are the corner points of object 14 (A, B, C and D). These points are chosen because they can be physically measured and reliably detected in the image. An image of object 14 is taken from each camera 10. The image points corresponding to the control points are detected from these images. It does not matter from which camera's image these control points are found.

Once the projection lines are calculated, an iterative process, 30 in Figure 2, is invoked to find the optimal pose. The optimal pose is defined as the pose of a reference model (i.e. a model of the object under observation with an ideal dimension) where the distances of its control points to the projection lines are minimized. The error function to be minimized is the sum of the squared distances from the model control points to their corresponding projection lines. This solution is called the least-squares solution. The iteration terminates when the pose improvement of the feature points in an iterating cycle falls below a certain precision threshold.

A pose can be represented by a mathematical description of the reference model and a rigid transformation. The mathematical description of the reference model is specified in a pose chosen close to the expected pose of the object
 5 under observation. The rigid transformation consists of a rotation, which is represented by a 3x3 matrix R , and a translation, which is represented by a 3-vector T . The object in the resultant pose is obtained by multiplying the coordinates of the reference model by the rotation R and then
 10 adding them with the translation T .

The rotation is a non-linear transformation in terms of the rotation angles, $(\theta_x, \theta_y, \theta_z)$. To get a close-form solution, this transformation is linearized by restricting to small
 15 rotation angles as

$$R \approx \begin{bmatrix} 1 & \theta_z & -\theta_y \\ -\theta_z & 1 & \theta_x \\ \theta_y & -\theta_x & 1 \end{bmatrix} \quad (6)$$

The error involves in the linear approximation increases with
 20 the size of the rotation angle. Due to this small rotation assumption, the pose of the reference model has to be close to the optimal solution.

The approach to obtain the optimal pose is by finding the
 25 (R, T) which transforms the control points of the reference model from an initial pose to a final pose where the error function is minimized. Due to the rotational approximation,

the final pose may not exactly be the optimal pose. To obtain a closer solution to the optimal pose, another cycle of the above pose determination is executed where the final pose of the previous cycle becomes the initial pose of the current cycle. As the pose gets closer and closer to the optimal pose, the amount of rotation and the approximation error involved will also reduce accordingly. Therefore, the error in pose determination can be reduced from cycle to cycle until a satisfactory level of accuracy is reached.

10

The error involved in each cycle of pose determination can be estimated by the average travelling distance of each control point in a cycle. Practically, an average travelling distance of less than a tenth of the required accuracy is considered insignificant. Thus, once the average travelling distance reaches below this level, the iterating cycle is terminated.

15

The equations involved in a cycle of pose determination are presented below. Referring to Figure 3, the error term for each control point, $p_i(p_{ix}, p_{iy}, p_{iz})$, defined as the distance from the transformed control point, p_i' , to its correspondent projection line, is evaluated as

20

$$d_i = || p_i' - q_i || \sin \beta_i \quad (7)$$

25

where $q_i(q_{ix}, q_{iy}, q_{iz})$ is any point on the projection line (one example is the point (Q_x, Q_y, Q_z) , obtained by setting $t=0$ in equations (3), (4) and (5)); and β_i is the angle between the line $(p_i' - q_i)$ and the vector, $n_i(n_{ix}, n_{iy}, n_{iz})$, defined

as a unit directional vector pointing along the tangent of the projection line (this vector can be calculated by normalising the vector (P_x, P_y, P_z) whose parameters are obtained from equations (3), (4) and (5)). Written in vector notation, (7) becomes

$$d_i = || (Rp_i + T - q_i) \times n_i || \quad (8)$$

The error function for k control points is defined as

$$F = \sum_{i=1}^k d_i^2 \quad (9)$$

The linear matrix form of the expansion of equation (8) utilising the approximation (6) and setting $d_i=0$, is (the subscript i has been dropped for simplicity)

$$W_p y_{pose} = b_p \quad (10)$$

where

$$W_p = \begin{bmatrix} n_y p_y + n_z p_z & -n_y p_x & -n_z p_x & 0 & n_z & -n_y \\ -n_x p_y & n_x p_x + n_z p_z & -n_z p_y & -n_z & 0 & n_x \\ -n_x p_z & -n_y p_z & n_x p_x + n_y p_y & n_y & -n_x & 0 \end{bmatrix}$$

$$y_{pose} = [\theta_x \quad \theta_y \quad \theta_z \quad T_x \quad T_y \quad T_z]^T$$

$$b_p = \begin{bmatrix} n_z(q_y - p_y) - n_y(q_z - p_z) \\ -n_z(q_x - p_x) + n_x(q_z - p_z) \\ n_y(q_x - p_x) - n_x(q_y - p_y) \end{bmatrix}$$

The equations from the k control points are collected into the following matrix form,

$$W_{pose} = [W_{p_1}^T \ W_{p_2}^T \ \dots \ W_{p_k}^T]^T \quad (11)$$

$$b_{pose} = [b_{p_1}^T \ b_{p_2}^T \ \dots \ b_{p_k}^T]^T \quad (12)$$

5 where each sub-matrix denotes the W_{p_i} matrix generated in (10) by the related control point p_i . The error function (9) can be rewritten as

$$F = (W_{pose} \ y_{pose} - b_{pose})^T (W_{pose} \ y_{pose} - b_{pose}) \quad (13)$$

10

The pseudo-inverse formula for the least squares solution of error function (13) is

$$y_{pose} = (W_{pose}^T W_{pose})^{-1} W_{pose}^T b_{pose} \quad (14)$$

15

A reference to the pseudo-inverse formula can be found in:

D.H.Ballard and C.M.Brown, *Computer Vision*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey 07632, pp.485-6

20

Once the pose is found, the surface of object 14 is predicted using its mathematical surface equation. In process 32 of Figure 2, the feature points on this surface are detected on the camera images. Their projection lines are generated. The intersections between the surface and these projection lines give the coordinates of these feature points. In the following, a quadratic surface is assumed. Note that, the assumption of quadratic surface is only for illustrative

25

purpose and should not limit the current invention to any type of surface as long as it can be expressed mathematically. Close-form solutions to line-to-surface intersection for low-order surfaces can be generally obtained. However, for higher order and non-linear surfaces, iterative techniques are required to find the solution of intersection up to an approximation. The reference to these techniques can be found in:

10 W.H.Press, S.A.Teukolsky, W.T.Vetterling and B.P.Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd Edition, Cambridge University Press, Cambridge, 1992.

15 The following derives the solution to calculate the point of intersection between a line (which represents the projection line of a detected feature point) and a quadratic surface (which represents an object's surface) to illustrate the principle of the current method.

20

The quadratic equation for the curved surface is

$$Z = A + BX + CY + DX^2 + EXY + FY^2 \quad (15)$$

The line equations with a tangent vector, $n(P_x, P_y, P_z)$ and a point on the line, $q(Q_x, Q_y, Q_z)$ is expressed as

$$X = P_x t + Q_x \quad (16)$$

$$Y = P_y t + Q_y \quad (17)$$

$$Z = P_z t + Q_z \quad (18)$$

Substitute (16), (17), (18) into (15) for the point of intersection, we have

$$\begin{aligned}
 0 = & (A - Q_z + BQ_x + CQ_y + DQ_x^2 + EQ_xQ_y + FQ_y^2) + \\
 & (BP_x + CP_y - P_z + 2DP_xQ_x + EP_xQ_y + EQ_xP_y + 2FP_yQ_y)t + \\
 & (DP_x^2 + EP_xP_y + FP_y^2)t^2
 \end{aligned} \tag{19}$$

By taking

$$\begin{aligned}
 a &= DP_x^2 + EP_xP_y + FP_y^2 \\
 b &= BP_x + CP_y - P_z + 2DP_xQ_x + EP_xQ_y + EQ_xP_y + 2FP_yQ_y \\
 c &= A - Q_z + BQ_x + CQ_y + DQ_x^2 + EQ_xQ_y + FQ_y^2
 \end{aligned}$$

we can rewrite equation (19) as

$$at^2 + bt + c = 0 \tag{20}$$

The two solutions for t are

$$t_1 = \{-b + (b^2 - 4ac)^{1/2}\} / (2a) \tag{21}$$

$$t_2 = \{-b - (b^2 - 4ac)^{1/2}\} / (2a) \tag{22}$$

For each t_i , the point of intersection is calculated using equations (16), (17) and (18). Then, the point which falls in the expected range of the surface is selected as the feature point's coordinate.

In CRT geometry adjustment, the coordinates of feature points determined on the CRT surface are used to calculate geometrical properties such as widths, heights and angles.

Possible Utilization in Industry of The Invention

This invention allows the measurement of feature points on a curved surface for any number of camera images. The method of searching for the optimal pose by a rigid transformation
5 reduces the complexity of the computation to a linear function compared to the prior art method of quadratic complexity using geometry consistency. There is no need of an extra stage of pose parameters calculation as the search directly finds the pose parameters.

CLAIMS

1) A method for determining the positional and orientational deviations of an object in 3-D space to be inspected from its ideal locations, using one or more digital images captured by one or more 2-D sensors, and comprising the ordered steps of:

positioning an object within the view of the 2-D sensors, for the purpose of sensor calibration;

selecting six sets of 3-D world coordinates, then capturing the corresponding 2-D image coordinates using each 2-D sensor for each set of the chosen 3-D world coordinates;

performing the 2-D sensors calibration process, and formulating the mappings from 2-D image coordinates to 3-D world coordinates, by applying the Direct Linear Transformation (DLT) method;

generating reference data relating to at least three non-collinear control points of the reference model;

positioning the desired object to be inspected within the view of the 2-D sensors, for the purpose of determining the pose of the object;

capturing the 2-D digital images from one or more 2-D sensors, containing at least a total of three non-collinear control points of the object to be inspected;

formulating the projection lines for all available control points;

formulating an error function (9), equivalent to the summation of the squares of the distance from the transformed control points to the projection lines;

iteratively minimizing the error function (9) by successively updating the vectors V_{pose} (13) until the

deviations in the control points fall below a predefined threshold.

2) The method as in claim 1, wherein the object is a rigid body.

3) The method as in claim 1, wherein the 2-D sensor is a camera.

4) The method as in claim 1, wherein said step of pose determination involves minimizing the sum of the squared distances from the control points of the reference model to the 3-D lines.

5) A method for determining the locations of feature points on a surface of an object in 3-D space from one or more digital images captured by one or more 2-D sensors, the method comprising the steps of:

generating sensor calibration data relating the position and orientation of the 2-D sensor to the 3-D space;

generating reference data relating to at least three non-collinear control points of a reference model;

generating the as few as one digital image containing the at least three non-collinear control points of the object; locating each of the control points in the as few as one digital image;

computing at least three non-parallel 3-D lines as a function of the locations of the control points and the sensor calibration data, each of the lines passing through its respective control point on the object;

determining the pose of the object by minimizing the fitting errors of the control points of the reference

model to the 3-D lines;

determining the equation for the surface of the object;

locating feature points in the as few as one digital image;

computing 3-D lines as a function of the locations of the feature points and the sensor calibration data, each of the lines passing through its respective feature point on the object; and

determining the locations of the feature points by the intersections of their respective 3-D lines and the surface of the object.

6) The method as in claim 5, wherein the object is a rigid body.

7) The method as in claim 5, wherein the 2-D sensor is a camera.

8) The method as in claim 5, wherein said step of pose determination involves minimizing the sum of the squares of distances from the control points model to the 3-D lines.

9) The method as in claim 5, wherein said step of equation of the reference determination makes use of the pose information of the object.