Sphereo: Determining Depth using Two Specular Spheres and a Single Camera

Shree K. Nayar

The Robotics Institute, **Carnegie** Mellon University. Pittsburgh, Pennsylvania 15213

ABSTRACT

This paper proposes a method for determining the depth of points in a three-dimensional scene. The concept is to we two spheres wuh highly specular surfaces W obtain two different perspectives of the scene. Both spheres are viewed by a single stationary camera, and each sphere reflects the world around it into the camera. Correspondence between points on the two spheres is established by matching features such as edges and image intensities, as in traditional stereopsis. Depth is recovered from each pair of corresponding points by triangulation. The we of a single fixed camera avoids the undesirable complexities that characterize the stereo calibration procedure. The measurable range of the system is greatly enhanced by the we of specular spheres and is not limited by the field of view of the camera. Experiments were conducted to determine the accuracy in depth measurement and the feasibility of practical implementation. The technique presented in this paper has been named "SPHEREO" as it wes two SPHeres, rather than two cameras, to emulate stEREOpsis.

1 INTRODUCTION

Three-dimensional object **recognition** is currently **an** active **area** of vision research. In many vision applications. the limitations **of** two-dimensional analysis have been realized. **For** example, in a typical bin picking operation, the position and shape of an object **must** be determined in three-dimensional space to enable the robot **to** securely *grasp* the object. Shape extraction is **an** essential part of a three-dimensional recognition system. Any ambiguity in the physical shape of **an** object generally renders the recognition problem more difficult. Hence, **the** advent of three-dimensional vision systems **has** created considerable interest in the development of **high** quality depth sensors.

Stereo is a popular technique for depth perception. It has generated much interest in the research community due **to** its strong resemblance **to** the mammalian approach to depth perception. In **stereopsis, images of** the **scene arc recorded** from two different perspectives. The two perspectives are obtained by **using two cameras** to observe the *scene*. **Features,** such **as edges, are extracted** from **both camera images and,** on **the basis** of the feature values. a point-—point **correspondence** is established between the **two imager** Range **or** depth **is recovered** from each pair of corresponding points by triangulation. The passive **nature** of stereopsis makes it **an** attractive depth perception method. It is suited to **most** applications, unlike "active" sensing methods such **as** radar, **laser** ranging. **and** structured light.

Stereo systems are **posed** with the acute problem **of** calibration. Corresponding points in the **two** images **are** **projections of** a single point in the three-dimensional scene. In order **to** triangulate **and** determine the three-dimensional **coordinates of** the scene point, the parameters of the ***two*** cameras must be **known. Therefore. for** a given configuration of the cameras, it is necessary **to** calibrate the inmnsic and extrinsic parameters **of** the *cameras.* Many researchers have studied the stereo calibration problem. One approach is to independently calibrate the ***two*** **cameras by** using a set of points at **known** locations in a common frame **of** reference. An alternative method **does** not rely ***on*** knowing the locations of the calibration points, but rather the correspondence between the points in the images. Gennery [3] proposed performing the calibration by a generalized least-squares adjustment. Errors **are** formulated by using the epipolar constraint. Minimizing the errors results in estimates **of** the *camera* parameters. Faugeras and Toscani [2] have suggested a recursive estimation of the camera parameters by using extended Kalman filtering.

The complexity **of** the calibration procedure has limited the applicability ***of*** stereo systems. Since it is computationally inefficient to **perform** the calibration on-tine, the relative positions **and** orientations **of the cameras** need **to** be rigidly fixed. In addition to **the** calibration problem, stereo systems are often limited **by** a **small** field **of** view. The depth of a point can be **measured** only if the point is seen by both cameras. Thmfore, the field **of** view of a stereo system is the intersection of the fields of view of the ***two*** cameras.

This paper describes "sphereo" as a new approach to stereo vision. **Two** spheres with highly reflective surfaces **are** placed in the view of a single camera. Reflections **of** the three-dimensional scene **are** recorded in the image of the spheres. Hence, a single **camera** image includes ***two*** different perspectives of the three-dimensional scene. **These two** perspectives **are** equivalent to ***images* obtained** from **two** different camera locations in stereopsis. The stereo calibration problem is avoided by using a single camera. However, **the** position of the ***two*** spheres must **be** **known** to **recover** depth **by** triangulation. **To** this end, a simple calibration procedure is presented in this paper that determines the **location** of the two spheres on-line. Each *camera* image contains information regarding the positions of the spheres and **the depth** of **points** in **the** scene **at** the *same* instant in time. The positions of **the** spheres are first determined **and** then **used** to compute the coordinates of points in **the** scene.

The field **of view** of a sphereo system is a great improvement **over that** of stereo systems. The use **of** specular sphms enables **the system to measure depth** outside **the camera's** field of view. The **surface** of a sphere completely **spans** the gradient space. Therefore, points **in** all **directions** are reflected by the spheres **into** the camera, **and** the **camera** is **used** only to observe the **surfaces** of **the** spheres. Such an imaging geometry, ***makes*** it possible to measure depth of **points** both inside and outside the camera's field **of** view.

Experiments were conducted to demonstrate the practical feasibility of the sphereo concept. Point sources of light were positioned at known locations in the scene, and a sphmo set-up was used to determine the three-dimensional coordinates of the sources. The measurement accuracy was estimated by comparing the experimentally determined source positions with the actual positions. The sphereo approach does not simplify the correspondence problem associated with stereo vision. This paper does not address the comspondence problem posed by complex scenes but focuses on the merits of a new approach to stereo imaging.

2. SPHEREO

2.1 Concept

The geometry of a sphm is completely defined by its radius. The sphere also possesses the property that no two points on its surface have the same surface normal. Figure (1) shows the reflection of light from a point source, off the surface of a specular sphere. In the case of specular reflection, the angle of incidence i equals the angle of reflection e. Let us assume an onhographic camera model; all light rays observed by the camera are parallel and are in the direction v. Under the above stated constraints, only a single point A on the surface of the sphere is capable of reflecting light from the point source into the camera. An alternative interpretation of this effect is as follows: if the position of the center of the sphere and its radius are known, then a bright point in the image can be projected out of the camera to intersect the surface of the sphere at the point A. The surface normal n at the surface point A is unique and is determined by the position of the sphae. Given the viewing direction v and the surface normal n, we can find the source direction s by using the specular reflectance model.

Sphereo uses two specular spheres of known radii and center positions, each reflecting the world in the direction of the camera. Figure (2) shows the two spheres illuminated by a point

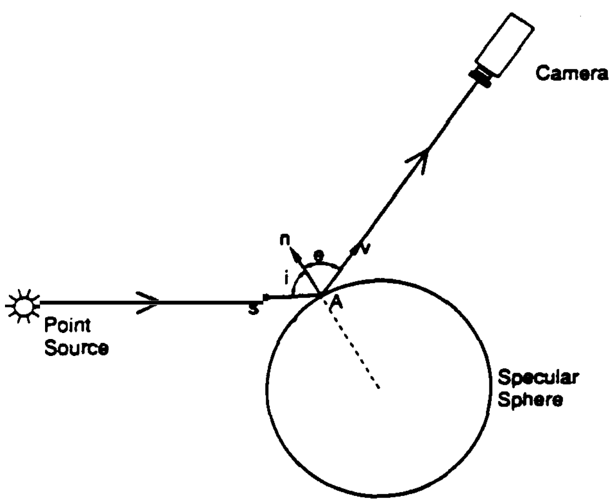

Figure 1: The specular sphere reflects light from the point source into the camera.

source. The resulting image has two discrete bright points, namely, I, and I_2. Lines perpendicular to the image plane are projected from I, and I_2 to intersect the spheres S, and S_2 at the points P and Q, respectively. The surface normal vectors n_1 and n_2 at points P and Q are computed by using the known radii and center locations of the spheres. Since the spheres are separated by a distance D, the surface normal vectors n_1 and n_2 differ in direction. Given n_1, n_2, and the viewing direction v, the source directions s_1 and s_2 are computed by using the specular reflectance model. The point source lies on the line L_1 passing through the point P in the direction s_1. The point source also lies on the line L_2 passing through the point Q in the direction s_2. Therefore, the point source location W is found at the point of intersection of the two lines L_1 and L_2. The point source has been used in the above discussion to explain the principle underlying the sphereo method. In practice, however, candidates for matching are not confined to bright image points and may also be characterized by features such as discontinuities in image intensity.

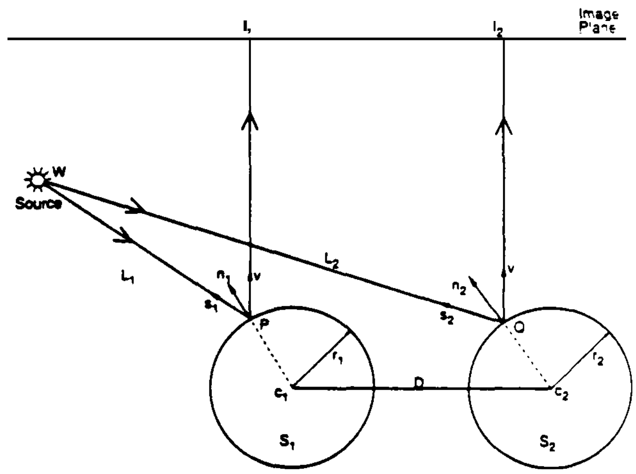

Figure 2: Sphereo: determining the position of a point using two specular spheres.

2.2 Finding the Spheres

Depth measurement using sphm is based on the knowledge of the radii and positions of the specular spheres. We will assume that the radii of the sphaes are known. The position of the spheres with respect to each other and the camera may be determined by a simple calibration procedure. Figure (3) shows the spheres S_1 and S_2 placed on the x-y plane of the world frame. The z coordinate of the center of each sphere is equal its radius r. However, the x and y coordinates of the center need to be determined. Four point sources PS_1, PS_2, PS_3, and PS_4 are symmetrically positioned about the optical axis O-O' of the camera. The point sources are coplanar and each source is at a distance q from the optical axis. Consider either of the two spheres S_1 and S_2. The distance d of the sphere center from the optical axis is small compared to the height h of the four sources from the x-y plane and the distance q of each source from the optical axis. Each point source produces a highlight in the image of the sphere. Let (X_i,Y_i) be the center of mass of the highlight

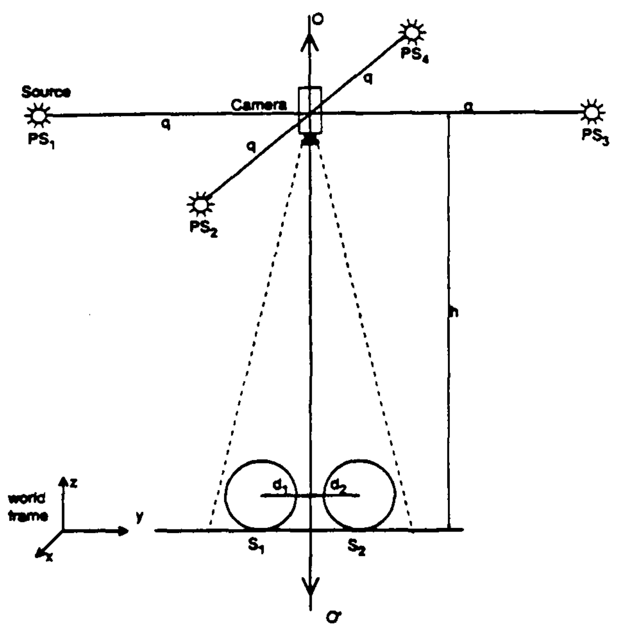

Figure 3: Calibration: point **sources** PS_1, PS_2, PS_3, and PS_4 arc used to find the spheres S_1 and S_2 in the camera image.

comsponding to point source PS_i. The centroid $O(X_c,Y_c)$ of the **four** highlights may be determined as:

$$X_c = \frac{1}{4}\sum_{i=1}^{4} X_i , \tag{1}$$

$$Y_c = \frac{1}{4}\sum_{i=1}^{4} Y_i .$$

Under the distant **source assumption, the** image point $O(X_c,Y_c)$ is the projection of the center of **the sphere** onto the *image*. In practice, it is **not** necessary **to use** exactly **four sources for the calibration.** Any number of **sources** may **be used as long as** they **are** coplanar and their **centroid** lies *on* the optical **axis.**

The next **step** is to **find** the world coordinates $C(x_c,y_c,z_c)$ of the sphere center **from** its image coordinates $O(X_c,Y_c)$. Transformations between **the** world **and the** image are determined by the inmnsic *and* extrinsic parameters of the camera. **During** the process of **image formation, a** point $P(x,y,z)$ in the world is **projected** onto the point $I(X,Y)$ in the image. The *camera* parameters **may** be **used to determine** $I(X,Y)$ **from** $P(x,y,z)$. **However,** it **is not** possible *to* **recover** a **world point** from an image point. **For** each **image** point $I(X,Y)$, **the camera** parameters *can* only determine the **equation of a line** in **the** world ***on*** which the point $P(x,y,z)$ lies. Therefore. the center of the sphere lies on the line:

$$x_c = a\,z_c + b , \tag{2}$$

$$y_c = c\,z_c + d ,$$

where the **transformation** parameters **a**, b, c, and d **are** determined by the camera parameters **and** the image coordinates X_c and Y_c. Since the sphere is placed on the *x-y* plane. $z_c = r$. Hence, the world coordinates (x_c,y_c,z_c) of the center are uniquely determined from equation (2). The **radius** of a sphere is **measured** in pixels *in* the image and inches in the world frame. The spheres S_1 and S_2 have radii R_1 and R_2 in the image and r_1 and r_2 in the worid. The centers of the spheres **will be** referred to **as** $O_1(X_{c1},Y_{c1})$ and $O_2(X_{c2},Y_{c2})$ in the image, and $C_1(x_{c1},y_{c1},z_{c1})$ **and** $C_2(x_{c2},y_{c2},z_{c2})$ in the world.

The simplicity **of** the calibration **procedure** described in this section *makes* it feasible **for** it **to be performed** on-line. If the calibration light **sources** *are* always active, each camera image obtained **describes** the **positions of** the spheres and the scene at the same instant in time.

2.3 Correswndence

Prior to computing **the** depth **of** scene points, the sphereo system is required *to* solve the well-known compspondence problem: the task **of determining** which point in the image of one sphae **corresponds to** a particular point in the image **of** the other sphere. **Features, such as edges,** *are* extracted **from** the two **circular** sections in **the** image **that** correspond *to* the projections **of the** spheres S_1 **and** S_2. A feature **value** at the image point $A(X,Y)$ may **be** expressed **as** $F\{A(X,Y)\}$. If the image points $I_1(X_1,Y_1)$ and $I_2(X_2,Y_2)$ constitute a pair **of** corresponding points. they must lie on different spheres and have matching feature **values:**

$$F\{I_1(X_1,Y_1)\} = F\{I_2(X_2,Y_2)\}, \tag{3}$$

where:

$$|I_1 - O_1| < R_1 ,$$

$$|I_2 - O_2| < R_2 .$$

2.4 Triangulation

Depth values are recovered from pain *of* corresponding image points **by** triangulation. **Consider** the image points I_1 and I_2 in **Figure (2). Since** both **points** *are* reflections of the same point in **the scene, they** satisfy **the** constraints given in equation (3) **and** thus **constitute a pair of** corresponding points. The point $I_1(X_1,Y_1)$ is the **projection of** the point $P(x_1,y_1,z_1)$. We know that $P(x_1,y_1,z_1)$ lies on **the** line:

$$x_1 = a\,z_1 + b , \tag{4}$$

$$y_1 = c\,z_1 + d ,$$

where a, b, c, **and** d **are determined by** the **camera parameters and** the **image** coordinates X_1 and Y_1. The point $P(x_1,y_1,z_1)$ **also** lies on the surface of sphere S_1. Therefore,

$$(x_1 - x_{c1})^2 + (y_1 - y_{c1})^2 + (z_1 - z_{c1})^2 = r_1^2 . \tag{5}$$

Equation (4) may **be** used to eliminate x_1 and y_1 in equation (5).

This results in a quadratic equation in z_1. As shown in Figure (3), the camera is positioned in the positive z direction and thus the point P lies on the upper hemisphere of S_1. Therefore, z_1 is the higher of the two roots of the quadratic equation. The x_1 and y_1 coordinates of P are then computed by using equation (4). At point P, the unit vector v_1 in the viewing direction is determined from equation (4) as:

$$v_1 = \frac{V_1}{|V_1|}, \tag{6}$$

where:

$$V_1 = (a, b, 1).$$

The unit surface normal vector n_1 at the point P on the sphere S_1 is computed as:

$$n_1 = \frac{P - c_1}{r_1}. \tag{7}$$

In order to find the location of the point W in Figure (2), we need to determine the direction of W as seen from point P. Let the unit vector in this direction be s_1. For specular reflections on the surface of the sphere, the angle of reflection equals the angle of incidence. This *specular constraint* may be used to relate the three vectors s_1, n_1, and v_1:

$$[\, n_1 \cdot v_1 \,]\, n_1 = \frac{v_1 + s_1}{2}. \tag{8}$$

The source direction s_1 is determined by rewriting equation (8) in the form:

$$s_1 = 2\,[\, n_1 \cdot v_1 \,]\, n_1 - v_1 . \tag{9}$$

On the same lines, source direction s_2 is computed from the image point I_2. A line is projected from I_2 to intersect sphere S_2 at the point $Q(x_2,y_2,z_2)$. The source direction s_2 is computed by using the specular constraint. The line L_1 in Figure (2) passes through point P in the direction s_1. The line L_2 passes through point Q in the direction s_2. The point W is found at the intersection of lines L_1 and L_2.

The accuracy of a sphereo system is related to the resolution of measured depth. As in the case of stereo, depth resolution is related to pixel resolution in the camera image. Figure (4) illustrates triangulation uncertainty in the two dimensions of the image plane. The pixels A and B are projections of the same scene point W and thus constitute a pair of matching image points. Uncertainity in the location of the point W is represented by the shaded region. Therefore, errors in triangulation result from image quantization; due to finite resolution in the image, the location of the point W can lie anywhere in the shaded region around the actual location. The area of the uncertainty region tends to increase with the distance of point W from the two spheres. The line joining the centers of the two spheres is called the sphereo baseline. The area and shape of the uncertainty region are also dependent on the baseline magnitude D and the baseline orientation θ. In three dimensions the uncertainty region is a volume bounded by a polyhedron.

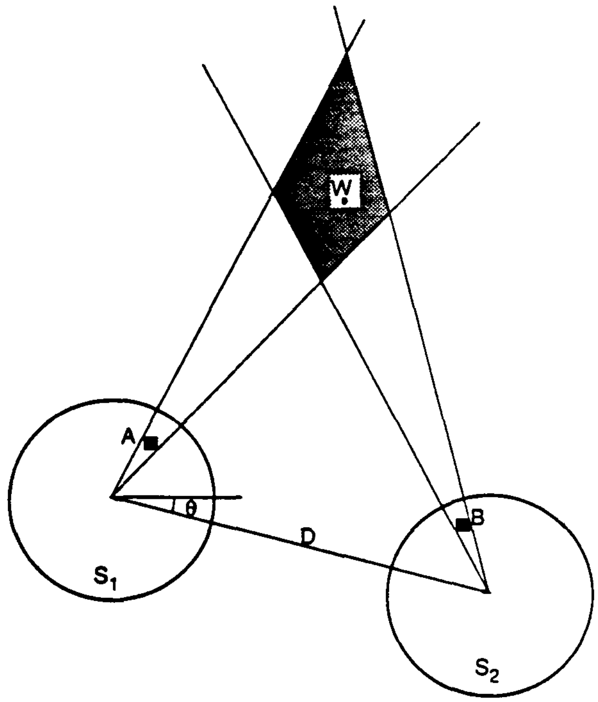

Figure 4: Sphereo triangulation uncertainty in two dimensions. Due to image quantization, the location of the point W can lie anywhere in the shaded region.

Factors such as finite pixel resolution and image noise, cause inaccuracies in the positions of extracted features. Consequently, triangulation in three dimensions may not result in the intersection of the projection lines L_1 and L_2 shown in Figure (2). The two lines may only pass each other in close proximity without actually intersecting. Rather than attempting to intersect the lines L_1 and L_2, a more practical approach is to determine the common normal L_3 between the lines L_1 and L_2, as shown in Figure (5). The unit vector s_3 in the direction of L_3 is computed as:

$$s_3 = \frac{S_3}{|S_3|}, \tag{10}$$

where:

$$S_3 = s_1 \times s_2 .$$

The line L_3 intersects lines L_1 and L_2 at the points $U(x,y,z)$ and $V(x,y,z)$, respectively. Let k, l, and m be the distances between the points U and P, V and Q, and U and V, respectively. We can express the coordinates of points U and V as:

$$U = P + k s_1 , \tag{11}$$

$$V = Q + l s_2 ,$$

$$V = U + m s_3 .$$

The parameters k, l, and m can be determined by eliminating U and V in the above set of equations:

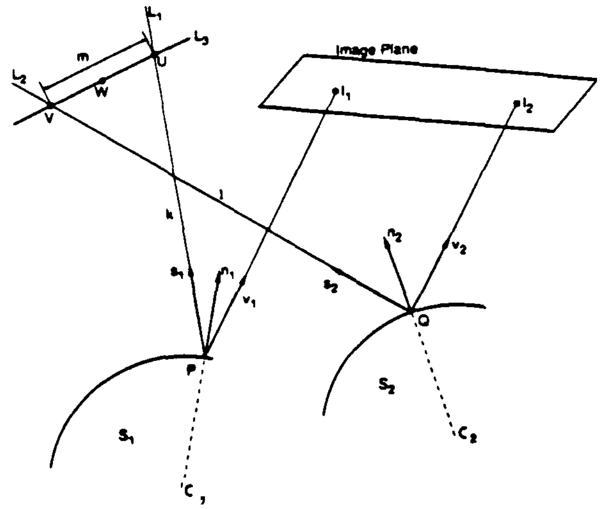

Figure 5: Triangulation in three dimensions. The mid-point W of the common normal L_3 is considered to be the best estimate of the point of intersection of lines L_1 and L_2.

$$k s_1 - l s_2 - m s_3 = Q - P. \tag{12}$$

Depth is computed only at points that have a high *measure of triangulation.* The triangulation measure is formulated as a function of the length m of the common normal and the distances k and l of the common normal from the spheres S_1 and S_2, respectively. In the current sphereo implementation, a successful match or intersection is found between corresponding image points when:

$$\frac{k+l}{2m} > T, \tag{13}$$

where T is an empirically determined threshold level. For a successful match, the point of intersection W is defined as the mid-point of the common normal:

$$W = \frac{U+V}{2}. \tag{14}$$

2.5 Search Space

Though corresponding points have matching feature values, triangulation of all pairs of matching image points is unnecessary. As in stereo, the imaging geometry may be used to impose physical constraints on the positions of corresponding points in the image. These constraints considerably reduce the search space for corresponding points. Consider a camera image of the two spheres, as shown in Figure (6). The projection of each sphere is a circular section in the image. Each point inside the circular section is the projection of a point on the sphere's surface. We shall denote the image sections corresponding to the spheres S_1 and S_2 by CS_1 and CS_2, respectively. If constraints are not used while finding corresponding points, the features computed at each point in CS_1 have to be compared with features at all points in CS_2. Therefore, the search space for the point in CS_2 that corresponds to a point in CS_1 is the entire two-dimensional section CS_2.

Consider the sphereo imaging geometry shown in Figure (2). The image point I_1 is the projection of the point W by sphere S_1. Given the point I_1, we know from the camera model and the specular constraint that W must lie on the line L_1. Therefore, the point I_2 that corresponds to I_1 must lie on the image projection of the line L_1 by the sphere S_2. This is the epipolar *constraint.* The image projection of line L_1 by sphere S_2 is called the *epipolar curve.* As shown in Figure (6), the search space for the point I_2 that corresponds to the point I_1 is reduced from the two-dimensional section CS_2 to a one-dimensional epipolar curve *AB.* If a feature match is determined between a point in CS_1 and a point on its epipolar curve in CS_2, then a high measure of triangulation is ensured.

Epipolar curves for each point in the section CS_1 can be pre-computed and stored in memory. Consider, for example, the epipolar curve corresponding to the image point I_1 in CS_1. As in the case of triangulation, the line L_1 is determined from I_1. A point U on L_1 may be expressed as:

$$U(k) = P + k s_1. \tag{15}$$

where k is the distance of U from P. The point $Q(k)$ on sphere S_2 that reflects $U(k)$ into the camera is determined by using the specular constraint and by assuming an orthographic camera projection. The point $Q(k)$ is then projected to the point $I_2(k)$ in the image plane, by using the camera parameters. The epipolar curve for I_1 is thus determined by computing $I_2(k)$ for all k in the interval $0 < k < k_{max}$, where k_{max} is greatest distance of a measured point from sphere S_1. The image coordinates of points on the epipolar curves are stored in memory. Matches for a point in CS_1 are obtained by comparing its feature value with those on its epipolar curve in CS_2.

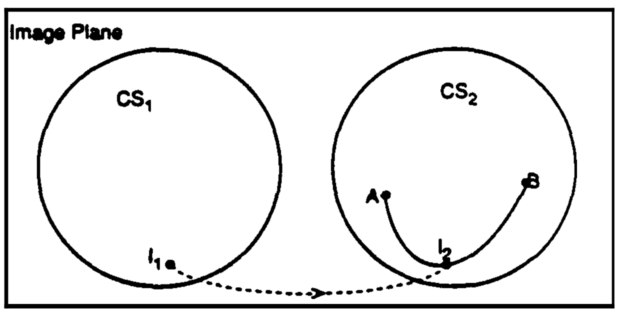

Figure 6 The epipolar constraint. The point I_2 in CS_2 that corresponds to the point I_1 in CS_1 lies on the epipolar curve *AB.*

2.6 Field of View

The field of view of a typical stereo system is shown in **Figure** (7). Depth can be **measured** only at those **points** that can be Seen in both the camera images. Therefore, the field of view of a s t e m system is the intersection **of** the fields of view of the **two** *cameras.* **A** large field of view *can* be obtained by minimizing the baseline ***D*** and keeping the viewing directions of the **two** cameras almost **equal.** Such an arrangement, however. **results** in lower depth resolution. **A high** depth **resolution** is obtained **by making** the viewing directions **of** the ***two* cameras** orthogonal to each other. **However,** this ***configuration*** drastically reduces the field of view **of** the **stereo** ***system***

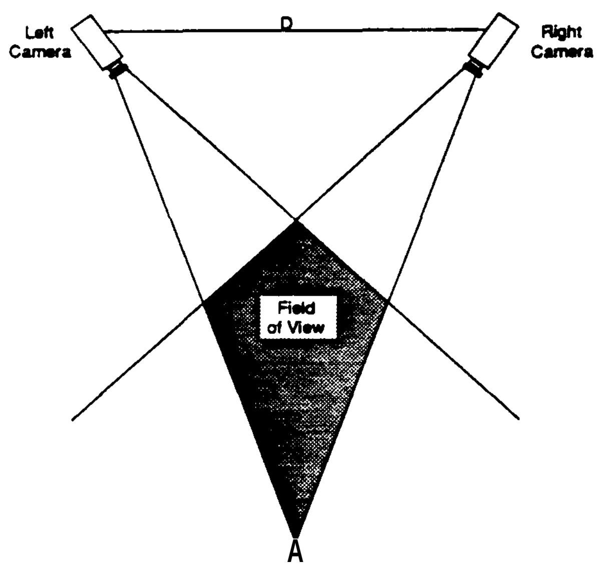

Figure **7:** Field of view of a stereo system.

The field of view of a sphereo **system** is a great improvement over that of stereo systems. **This** is primarily due *to* the use of specular spheres for **stereo imaging.** The surface of a **specular** sphere may be thought of **as** being constructed by an infinite number of **small** planar **mirrors.** Since no **two** points ***on*** a sphere have the ***same*** surface **normal.** each planar **mirror** faces in a unique direction. Also, the complete **set** of mirrors span the surface normal **space. Consider** a specular **sphere** placed in the view of a camera. Under the assumption of orthographic image projection, the viewing **direction** is constant **over** the entire field of view of the *camera.* Any non-occluded point in space would be reflected in the direction of the **camera** by a single point on the surface of the sphm. **Thmfore.** the field **of** view **of** a sphmo system ***consists*** of all points that *can* be refluxed onto the image plane by both spheres. Both spheres are placed in the focal plane of the *camera* **and** therefore **the** image projection **of points** in the scene is not affected **by** a **limited** depth of field of the camera. ***On*** the other hand, in stereo, objects must **be** placed clox to the focal planes of both *cameras* to avoid the blurring **of** image features.

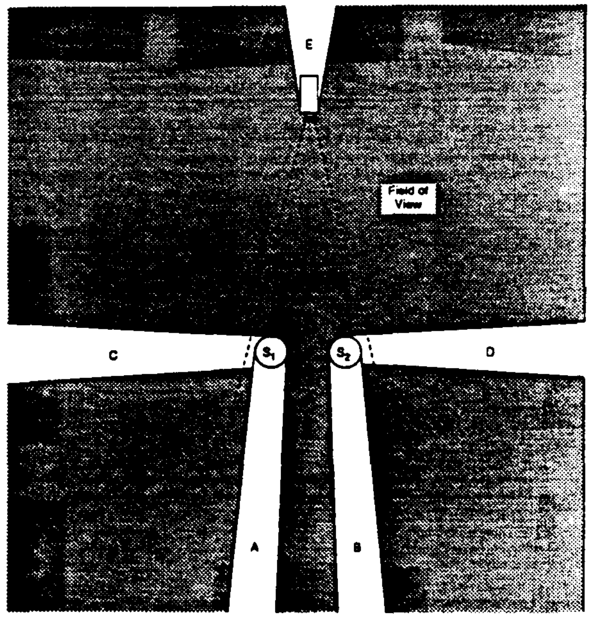

Figure 8: Field of view **of** the sphereo system.

In **Figure** (8), the **shaded** region denotes the sphereo field of view. The measurable range is not confined to the field of view of the camera Points in regions ***A*** and *B* **are** not reflected into the **camera** by spheres S_1 and S_2, respectively. Points in region *C* **are** occluded **by** S_1 and thus **are** not reflected into the camera by S_2. Similarly, points in region ***D*** **are** occluded by S_2 and **are nor** reflected into the camera by S_1. Region *E* consists of points that **arc** occluded by the camera and **arc** not visible to either sphere S, *or* S_2.

3. EXPERIMENTAL RESULTS

Experiments **were** conducted **to** demonstrate the practical feasibility of the sphmo concept. **Figure** (9) shows a photograph **of** the apparatus **used** for the current implementation. Two steel ball bearings, each **1 3/8** inches in diameter, **arc** used **as** specular spheres. The surface **of** each sphere is highly specular in reflection. A CCD Panasonic *camera* with a **510 x 492** pixel resolution is **used** *to* observe the **two** spheres. The physical **resolution** of **the camera and the** optical system is 0.011 inches **pa** image pixel. **Four** 1/4 **Watt,** ***5*** Volt bulbs **are** mounted on the *camera* lens and **arc used by** the calibration procedure to find the location **of** the **two** spheres in the image. Individual images **are** digitized and processed using a vision system based on a **SUN** work-station.

The sphereo concept was described **by** assuming an orthographic *camera* **projection** model. **More** accurate results are obtained **by** determining **the** extrinsic and intrinsic parameters of the *camera.* **The** current implementation of sphereo uses the single-camera calibration **technique** developed by Tsai[7]. **Tsai's** calibration algorithm is **structured to** determine the value of **each** parameter rather **than** just a transformation from image **to** scene **and** vice-versa. The parameters **are** computed in well-defined stages, **thus** making it easy ***to*** code during software

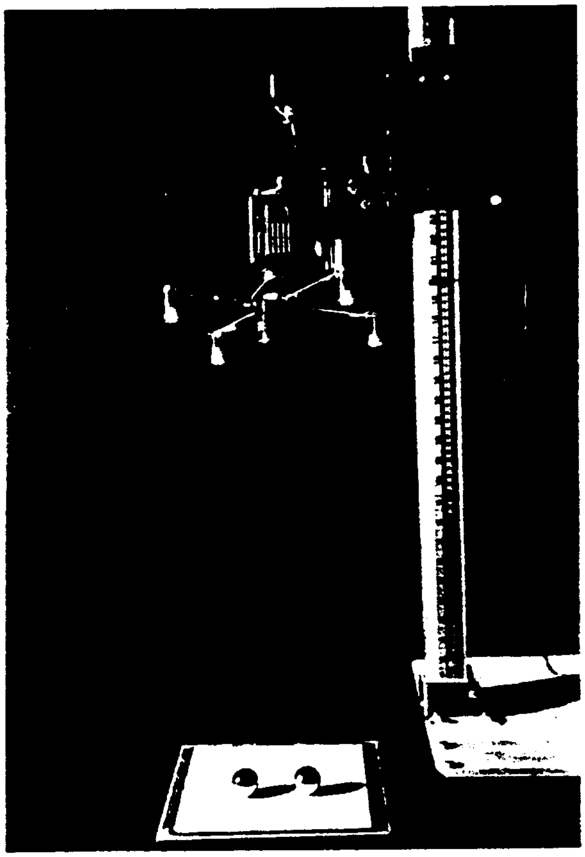

Figure 9: Photo of a prototype sphereo set-up. Four calibration lights arc mounted *on* the **camera** lens and **two** specular steel balls are placed in the camera's view.

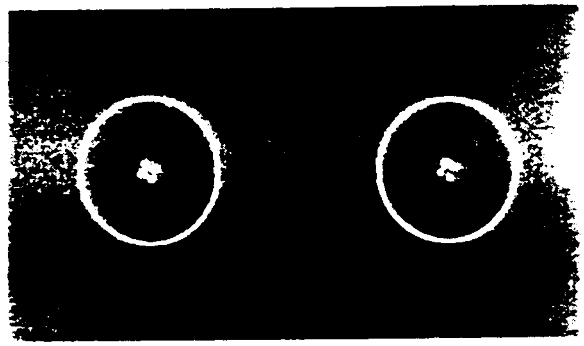

Figure 10: Photo of a *camera* image of the two spheres illuminated by the four calibration point sources. The centers of the spheres in the image **are** determined and circles **are drawn** around the centers **to** show the boundaries of the spheres in the image.

implementation. The computed parameters produced excellent msults for transformations **between** the image plane **and** the world frame.

The strength of a sphereo **system** lies in its ability to determine the location of the **two** specular spheres **by** using a simple calibration. The four light bulbs **mounted** on the *camera* lens produce four highlights *or* **bright points** on the **surface** of each sphere. **Under** the distant **source** assumption. the configuration. *or* relative positions, of the four highlights in the image of each sphere is the ***same* as** the configuration of the four light bulbs in the plane in which they lie. Also, the highlight configuration **on** each sphere is independent of the position of the sphere in the image. Once the four lights *an* rigidly fixed *on* the camera lens, the *configuration* of **the** four highlights is **known.** By using a template of four bright **points** in the expected **configuration.** we *can* determine the positions of the **two sets** of four highlights in the image. **As** explained in section 2, the centroid of each **set** of four highlights **determines** the centers of the spheres in **the** image.

Figure (10) shows **an** image of the **two** spheres resulting from the four-point illumination used for calibration. The centroids of the two **sets** of highlights **were** computed and **the** circles drawn around the centroids represent the boundaries of **spheres.** The **centroids** for the spheres were projected into the world **by** using the camera parameters, **and** the locations of the centers of the **two** sphms were **determined by the** system to **be** within a **distance** of **0.01** inches from their actual positions.

The aiangulation accuracy of the sphereo system **was** estimated using point sources. Low powered miniature lamps were placed at known locations in the world frame **and** the ***system* was** used to **measure** the coordinates of the lamps in three dimensions. Each lamp produces a single highlight in the image of each sphere. An **image** of the spheres is digitized **by** a frame-**grabber and** binarized **by** using a threshold. The highlights appear **as** bright blobs in the **binary** image, and the center-of-gravity ***(COG)* of** each blob is computed. The use of point **sources** greatly simplifies the correspondence problem **as** highlights *an* the only features **that need to** be extracted from the image. Corresponding **pairs** of highlight **COGs** that satisfy the epipolar constraint *an* triangulated **to obtain** position estimates in three dimensions. Figure(11) shows **an** incandescent light bulb **used to** generate a highlight *on* each of the **two** spheres. The pre-computed epipolar curve comsponding to the highlight **COG on** the left **sphere** is plotted on the right sphere. The epipolar curve **is a reflection** of the line that passes through the

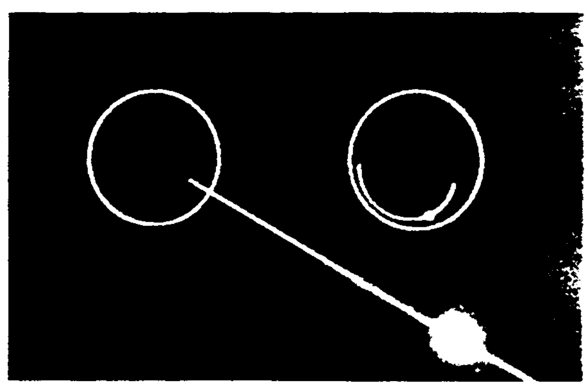

Figure 11: Highlights *on* the **two** spheres resulting from an incandescent *lamp.* The epipolar **curve** for the highlight point on the left sphere is plotted *on* the right sphere. **As** expected, the highlight on the right sphere lies *on* the epipolar curve.

highlight point on the left sphere and the location of the light bulb. *An* image projection of this line is displayed in Figure (11). As expected. the highlight on the right sphere lies *on* the epipolar curve corresponding to the highlight on the left sphere. Therefore. sphereo triangulation of the two highlights would result in an intersection at the location of the lamp.

The reliability of the matching process was tested by using a light display. The light display was constructed by mounting six light emitting diodes (LED) on a circuit board. Figure (12) shows the highlights *on* the *two* spheres resulting from the light display. The LEDs of the light display lie outside the camera's field of view and thus are not visible in the image shown in Figure (12). The circles represent the pre-determined boundaries of the two spheres. The COGs of the highlights resulting from the light display were determined and pain of highlight COGs were triangulated. World coordinates were computed for for COG pairs that produced high measures of aiangulation. The triangulation results for the image in Figure (12) *are* shown in Figure (13). The actual LED locations and the measured positions are both shown to illustrate the accuracy of the current implementation. The camera and the two specular spheres *are* also shown to give an idea of the relative positions and Sires of the spheres. the light display. and the camera. In Figure (13), the x, y. and z world coordinates of the LEDs determined by the system were found to be within 3.5% of the actual coordinates.

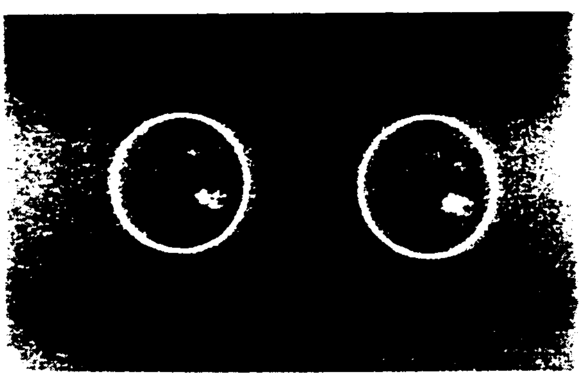

Figure 12: Sphereo image of a light display made of six LEDs.

The non-linear nature *of* the depth resolution grid, shown in Figure (4), makes it difficult to specify the accuracy of the sphereo system. *An* important *measure* of performance is the sensitivity of the system to the distance of the measured point from the *two* sphaes. Triangulation errors were computed for point source locations along a straight line starting from the origin of the world frame and moving in a particular direction. Figure (14) shows the percentage errors recorded for source locations on the x-axis of the world frame, and for the sphere locations shown in Figure (13) . By measurement error we refer to the Euclidean distance of the measured position from the actual position. As expected, the triangulation errors increase with the distance of the measured point from the sphms. The measured locations were found to be within 4% of the actual location. for points that are less than 3 feet from the world frame origin.

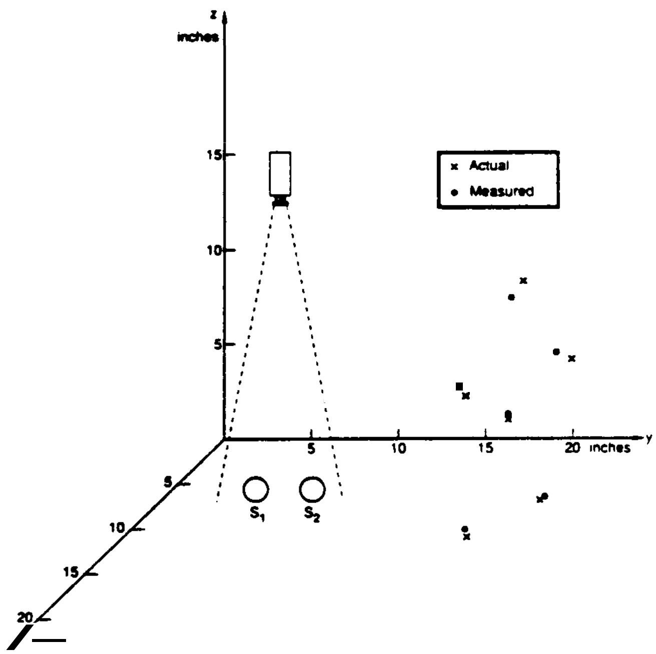

Figure 13: Positions of the sources on the light display were computed from the sphereo image shown in Figure (12). The measured and actual locations of the sources are plotted in the world coordinate frame.

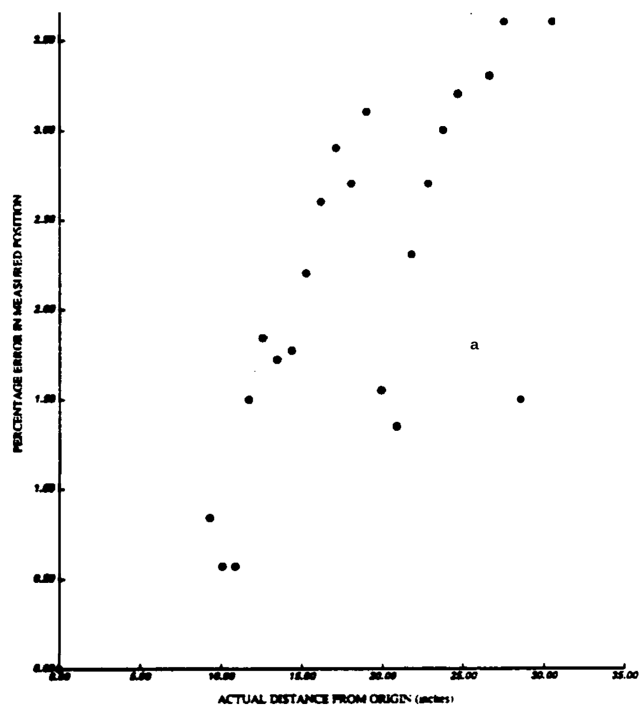

Figure 14: Percentage error in measured position of a scene point is plotted as a function of the actual distance of the scene point from the origin of the world reference frame.

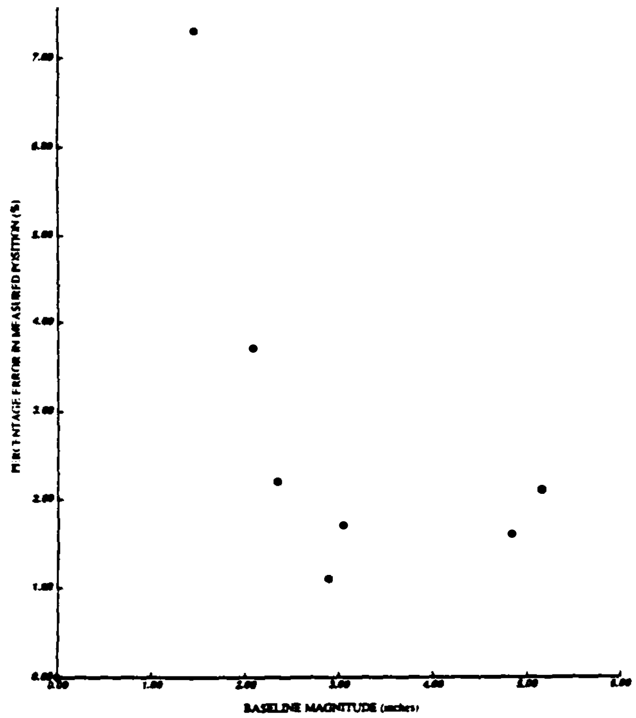

Figure 15: Percentage e m in **measured** position of a Scene point is plotted **as** a function of the magnitude ***D*** of the sphmo baseline.

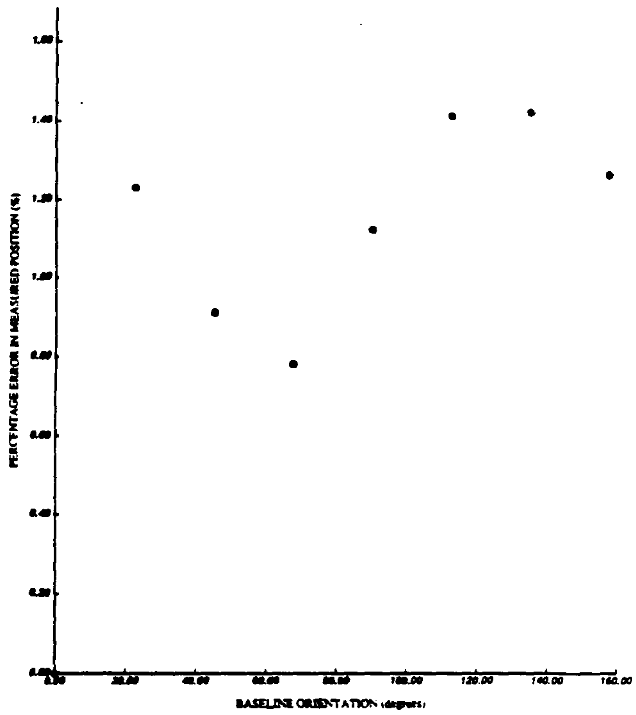

Figure 16: Percentage e m in measured position of a scene point is plotted **as** a function of the orientation θ of the spherco **bascline.**

The accuracy of a sphereo system is also dependent on the sphmo baseline D *and* its orientation θ **shown** in **Figure (4). As** the **two** spheres **are** brought closer **to** each other. the baseline decreases and the reflections of the world on the **two** spheres **become** less distinguishable. The view lines L_1 *and* L_2 **in Figure** *(5)* become almost parallel **to** each other and the measurements **arc** prone **to** greater errors. **This** effect is Seen in Figure **(15)** where the triangulation errors **arc** plotted **as** a function of the baseline magnitude for a constant point **source** position of x = 10.6 inches, **y = 123** inches. **and** z = **4.3** inches. **For** the **same** source location and baseline D = **2.34** inches, the measurement errors **are** also plotted for different baseline orientation values. **as shown** in **Figure (16). As** the baseline orientation is **varied** from 0 degrees ***to*** 180 **degrees,** the measurement **error** varies like a sine function between ***0.8%*** **and** 1.4%.

The experimental apparatus was **set** up using commercially available **and** inexpensive hardware **to** demonstrate **the** ease of implementation **of** sphereo systems. **From** the experiments, it was realized that a few changes in **the current** *system* would yield considerable **improvement** in **measurement** accuracy. Cameras with higher image pixel resolution, than **the one** used for the experiments, are now **commercially** available. **A** higher pixel resolution would improve **the** sensitivity ***of*** **a sphmo** *system* Spheres with greater **surface specularity and lesser shape** imperfections **than ball** bearings would produce more reliable **results. As** explained earlier on, point **sources** were **used** to Simplify the **correspondence problem.** However, **the** use of miniature lamps **as** point sources **comes** with the cost of imperfect highlights in **the** image. **The** shape of a highlight is dependent **on** the shape of the filament of the lamp that caused the highlight, the pixel **resolution** in the image, and noise levels in the **camera** output. The size of a highlight is related to **the** filament **size and** the sharpness of surface specularity of the spheres. In **the experiments discussed** above, no attention was given to the **shape of** highlights. and **each** highlight was simply represented **by** its centerof-gravity. **More** consistent highlight shapes **may be** obtained by using **miniature** arc lamps rather than incandescent lamps.

4 CONCLUSIONS

Sphereo has **been** presented **as** a new approach to stereo imaging. **A** prototype sphereo system **was** implemented **and** used **to measure** the position **of** point ***sources*** located in its field **of** view. The main **advantages** of sphereo, over conventional stereopsis, **are the** simplicity of the calibration procedure and the considerable **improvement** in field of view. The current implementation **uses two** spheres to ***obtain*** different perspectives of **the world.** Better **estimates of** depth may **be** obtained by placing more **than** two **spheres** in **the camera's** field of **view. The results** obtained from the experiments conducted **with** point sources are promising **and** provide motivation to apply the **sphereo method** to complex *scenes.*

ACKNOWLEDGEMENTS

The author would like to thank Arthur Sanderson, Katsushi Ikeuchi. and Takeo Kanade for their valuable comments and suggestions. The author is also grateful to Rachel Levine and Ken Mohnkern for their help in preparing the manuscript.

REFERENCES

1. S. T. Barnard and M. A. Fischler, "Computational Stem", *ACM Computing Surveys*, Vol. 14, No. 4, pp. 553-572, 1982.
2. O. D. Faugeras and G. Toscani, "The Calibration Problem for Stereo". *IEEE,* 1986.
3. D. B. Gennay, "Stereo-Camera Calibration". *Proceeding* Image *Understanding Workshop*, pp. 101-108, November, 1979.
4. D. Marr and T. Poggio. "A Computational Theory of Human Stereo Vision". *Proc. Roy. Soc. Lond.*, Vol. B 204, pp. 301-328, 1979.
5. L Matthies and S. Shafa, "Error modeling in stereo navigation," *IEEE* Journal *of Robotics and Automation*, Vol. RA-3, No. 3, pp. 239-248, 1987.
6. Y. Ohta and T. Kanade, "Stereo by Inaa- and Inter Scanline Search Using Dynamic Programming", *IEEE. Trans. Pattern Analysis and Machine Intelligence.* Vol. PAMI-7, No. 2. pp. 139-154, 1985.
7. R. Y. Tsai, "A Versatile Camera Calibration Technique for High Accuracy 3D Machine Vision Metrology using Off-the-Shelf TV Cameras and Lenses". *IBM Research Report RC 51342,* October. 1985.
8. Y. Yakimovsky, R. Cunningham, "A System for Extracting Three-Dimensional Measurements from a *Stereo* Pair of TV Cameras", *Computer Graphics and Image Processing.* 7, 195-210.