Structured Highlight Inspection of Specular Surfaces

ARTHUR C. SANDERSON, SENIOR MEMBER, IEEE, LEE E. WEISS, MEMBER, IEEE,
AND SHREE K. NAYAR, STUDENT MEMBER, IEEE

***Abstract*—In this paper we describe an approach to illumination and imaging of specular surfaces which yields three-dimensional shape information. The structured highlight approach uses a scanned array of point sources and images of the resulting reflected highlights to compute local surface height and orientation. A prototype Structured Highlight INspection sYstem, called SHINY, has been implemented and demonstrates the determination of surface shape for several test objects including solder joints. The current SHINY system makes the distant source assumption and requires only one camera. A stereo structured highlight system using two cameras is proposed to determine surface element orientation for objects in a much larger field of view. Analysis and description of the algorithms are included here. The proposed structured highlight techniques are promising for many industrial tasks including inspection of machined parts and inspection of solder joints. Practical implementation will require additional attention to calibration methods, to coded scanning procedures, and to computation speed.**

***Index Terms*—Computer vision, inspection, robotics, shape measurement, solder joints, stereo vision.**

I. Introduction

THE interpretation of surface shape based on the image brightness pattern of one or more images depends intricately on the prior knowledge of surface properties, imaging geometry, and lighting conditions. Most approaches to this problem have assumed that surfaces are Lambertian, that is, incident light is scattered by the surface so that the perceived brightness is independent of the direction of view. Shape from shading (single light source) [1]–[3], photometric stereo (multiple light source) [4]–[7], as well as structured lighting techniques (patterned light source) [8] have generally used this assumption to recover three-dimensional shape from one or more images.

Many practical tasks in robot vision and inspection require interpretation of images of specular, or shiny, surfaces where the perceived brightness becomes a very strong function of viewing direction due to highlights or reflections from the source. In these cases, shape from shading techniques provide little or often erroneous information about surface shape because the underlying models assume independence of viewing direction. Inspection and handling of machined parts, inspection of the shape of solder joint surfaces, and inspecting surface smoothness of metallic or plastic sheets are examples of industrial tasks where surface specularity is a primary consideration.

Solder inspection has been an application of particular interest since it is a common industrial task which is also very laborious due to the size and nature of defects. Visual solder defects have been described and categorized [9], [10], and a variety of automated visual inspection techniques have been applied [10]–[13]. In general, these approaches have met with only partial success, and there is no generally accepted industrial standard for automated visual solder inspection. Approaches which apply conventional machine vision recognition algorithms have demonstrated high variability in gray level [10] and highlight [11] images. Approaches which incorporate shape features [10] or incorporate alternative imaging modalities such as infrared [13] or X-ray may be more consistent. The structured highlight inspection approach described here extracts surface shape features and is well-suited to the solder inspection application.

In this paper we describe an approach to illuminating and imaging specular surfaces which yields three-dimensional shape information. The "structured highlight" approach uses multiple point sources and images the resulting highlight patterns reflected from the surface. The basic principle is introduced in Fig. 1. Each light ray emitted by a point source is reflected off the specular surface such that the angle of incidence equals the angle of reflection. A fixed camera will image the reflected light ray only if it is positioned and oriented such that it is admitted by the camera's projective axis, either orthographic or perspective. Once the reflected light ray is observed as a highlight in the camera image, and the direction of the incident ray is known, the orientation of the surface element where the light was reflected may be found.

The structured highlight inspection approach uses an array of point sources to scan all possible positions and directions of incident light rays relative to a fixed camera, and therefore guarantees (within some bounds on angle of view) that each surface element of an object within the field may be scanned. The use of a parallel source, such as a laser, rather than a point source, becomes complex because all possible positions and directions must be scanned. A spherical array of fixed point sources guarantees that every point inside the sphere receives incident illumination from every possible angle. As shown in Fig. 1, the array of point sources is scanned to determine where the observed highlight originated, then the normal to the reflecting surface element at the origin can be determined

Manuscript received December 15, 1986; revised May 15, 1987. This work was supported in part by the Robotics Institute of Carnegie Mellon University and in part by the Westinghouse Electric Corporation.

The authors are with the Robotics Institute, Department of Electrical and Computer Engineering, Carnegie-Mellon University, Pittsburgh, PA 15213.

IEEE Log Number 8717223.

0162-8828/88/0100-0044$01.00 © 1988 IEEE

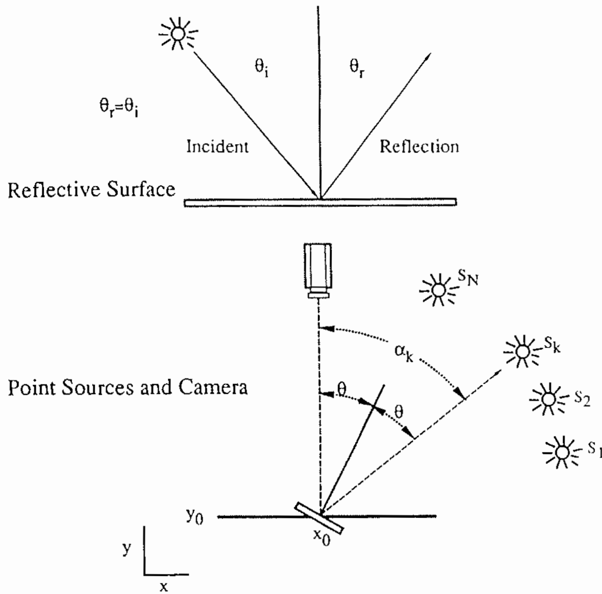


Fig. 1. Basic geometric relations for the structured highlight method. The angle of incidence of each ray is equal to the angle of reflection. An array of point sources guarantees that each surface element receives an incident ray from every direction. By identifying the source of an observed highlight, the surface orientation can be recovered.

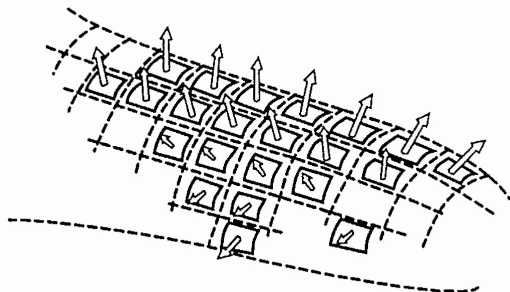


Fig. 2. Sketch of reconstructed surface based on the measurement of array of surface normals. Reflected highlights from each of the points are used to determine the local surface orientations. Fig. 12 shows a profile of the solder joint reconstructed from experimental measurements.

directly. For an object which is small relative to the radius of the point source array, all the surface normals may be obtained from a single scan of the array with a single camera. The objective is to inspect or reconstruct a surface from these elements in the manner shown schematically in Fig. 2.

In this paper, we describe an analysis of the structured highlight approach to inspection of specular surfaces. As background, we summarize previous work on extracting shape from specular surfaces and review the reflectance map representation of one specular surface model. Structured highlight inspection using the distant point source approximation is described as a discrete scanning of the reflectance map where the sampling distance and the sharpness of the specularity determine the accuracy of the estimation of the normals. An experimental prototype of the SHINY system for measurement of surface normals using the structured highlight technique has been used to demonstrate surface estimation for a calibrated sphere and for a solder joint. Structured highlight stereo is an extension of the current technique which does not require the distant point source assumption. Two fixed cameras acquire highlight patterns from the same source array, and a matching procedure is used to determine corresponding surface points between the two cameras. Corresponding points determine surface normal and surface element position.

II. Background

Previous approaches to determination of orientation of specular surfaces have relied on photometric methods which require the measurement of image brightness. Ikeuchi [14] proposed an approach which utilized graded illumination fields to illuminate the surface. By making the assumption of a distant source and known object position, the angular relation between an observed reflected brightness and the position in the source field with that brightness could be used to derive surface orientation. The approach requires measurement of reflected brightness with sufficient accuracy to isolate a particular position in the source field. The approach is analogous to illuminating with a set of point sources which are brightness encoded. Babu *et al.* [15] used a specular surface model to estimate orientation of an extended planar surface based on estimation of parameters in the shading model using contours of constant brightness in the image. The approach utilized a parallel source and perspective camera model so that observed brightness of reflected illumination depended on surface position relative to the camera axis.

Many approaches to surface shape recovery have utilized the reflectance map as a representational tool to describe the relation between local surface orientation and observed brightness. Consider a small flat surface patch at distance $-z(x, y)$ relative to a coordinate frame pointing away from the object. The reflectance map [2], [16] $R(p, q)$ describes a relationship between the observed image intensity R and the surface patch orientation in terms of p and q, the components of the local gradient of the surface:

$$p = dz/dx, \tag{1}$$

$$q = dz/dy. \tag{2}$$

Given a model of the reflectance characteristics of the surface, the direction of illumination of the surface, and the direction of the surface normal, we can calculate the observed brightness of the surface patch in the image. The resulting image consists of the array:

$$E(x, y) = R[p(x, y), q(x, y)], \tag{3}$$

where $E(x, y)$ is the image irradiance at (x, y) in the image plane. This equation depends on knowing the direction of illumination at every point. For parallel illumination (extended source) the direction of illumination is independent of surface patch position; however, for a point source, the direction of illumination may depend

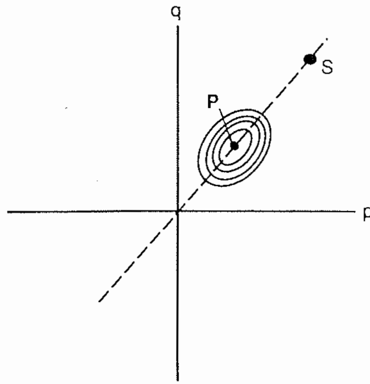


Fig. 3. Reflectance map for a single distant point source and nearly specular surface. Equal brightness contours define a narrow peak due to the specular reflection.

upon surface position as well. Similarly, in orthographic projection, the effective viewing direction is independent of position, while in perspective projection, the effective viewing direction depends explicitly on position.

The reflectance relation cannot, in general, be inverted to provide a unique solution for (p, q) given a single image $E(x, y)$. The equation can only be solved subject to additional constraints on surface geometry [1], or using multiple sources as in photometric stereo [4], [5]. Most commonly the Lambertian model is used:

$$E(x, y) = A(\boldsymbol{n} \cdot \boldsymbol{s}), \tag{4}$$

where $\boldsymbol{n}$ is the unit vector in the surface normal direction, $\boldsymbol{s}$ is the unit vector in the source direction at that point, and A is a constant related to illumination intensity and surface albedo.

An extension to the Lambertian surface model which incorporates a specular component of reflection [2] is described by the following equation:

$$E(x, y) = Ak(\boldsymbol{n} \cdot \boldsymbol{s}) + (A/2)(1 - k) \cdot \left[2(\boldsymbol{n} \cdot \boldsymbol{z})(\boldsymbol{n} \cdot \boldsymbol{s}) - (\boldsymbol{z} \cdot \boldsymbol{s})\right]^{n} \tag{5}$$

where $\boldsymbol{z}$ is the unit vector in the viewing direction, k represents the relative weight of Lambertian and specular components, and $\boldsymbol{n}$ is the sharpness of the specularity. The maximum of the specular component of this model occurs at:

$$(\boldsymbol{n} \cdot \boldsymbol{z}) = (\boldsymbol{n} \cdot \boldsymbol{s}) \tag{6}$$

or when $\boldsymbol{n}$ bisects the angle between $\boldsymbol{z}$ and $\boldsymbol{s}$. For a source at (Θ, ϕ), the surface normal corresponding to the specular peaks is at $(\Theta, \phi/2)$ as expected. In gradient space, for a source at (p_s, q_s) the corresponding surface normal is at:

$$p = p_s\left(\left[1 + p_s^2 + q_s^2\right]^{1/2} - 1\right)/(p_s^2 + q_s^2) \tag{7}$$

$$q = q_s\left(\left[1 + p_s^2 + q_s^2\right]^{1/2} - 1\right)/(p_s^2 + q_s^2). \tag{8}$$

The resulting reflectance map for a specular surface using the above model is shown in Fig. 3. The peak of the reflectance occurs at the indicated point P. The equal reflectance contours which surround the peak are indicated. The reflectance model is symmetric in the space of polar angles and would result in circular contours on the surface of a Gaussian sphere. In (p, q) space the symmetry is distorted and the resulting contours are nearly elliptical. The size of these ellipses is determined by the sharpness coefficient n in (5). The height of the reflectance peak is determined by the specularity coefficient k. The reflectance map shown in Fig. 3 provides the basis for the analysis of the structured highlight inspection technique in the next section.

III. Structured Highlight Inspection

The structured highlight approach uses an array of point sources to illuminate a specular surface. As shown in Fig. 1, the object is centered at the origin, and a single camera is used to view the reflected highlights. For purposes of discussion here, we will assume that the point sources are distant from the surface, that the surface is at a fixed reference height, and that the extent of the surface from the origin is much smaller than the distance to the source. Under this assumption, the angle of incidence of illumination is determined only by the position of the source and does not depend on the relative position of illumination on the surface. We will utilize this assumption throughout this section and generalize the approach in Section V.

Assume that a set of sources $\{S_k, k = 1, \cdots, N\}$ is positioned at polar angles (Θ_k, ϕ_k) at distance R with respect to the origin. Based on the discussion of the reflectance map in the previous section, each of the sources produces a local peak in the reflectance map as shown in Fig. 3. The resulting reflectance map for the whole array is shown in Fig. 4. The size and distribution of the elliptical contours depend on the surface specularity and the spacing of the discrete point sources. For a given specularity, the point sources are chosen to span the reflectance map, effectively sampling the space of orientations of the target surface. In practice, a threshold on image brightness is used to detect highlights which fall within a specified reflectance map contour. Scanning of the sources and detection of the corresponding highlights provides a sampling of the surface orientations of the object.

Each of the sources S_k sequentially illuminates the object, and a single camera image $I_k(x, y)$ is acquired for each. Each image is thresholded relative to I_0, and connected regions greater than some minimum size are stored. I_0 is chosen to correspond to a known reflectance map contour and therefore implicitly defines the tradeoff between orientation resolution and sampling density of the orientation space. The resulting set of image points:

$$\left\{h(x, y)_k\right\} = \left\{\text{highlight at } (x, y) \text{ due to } S_k\right\} \tag{9}$$

are the highlights due to $\{S_k\}$ where each highlight corresponds to a point on the object surface. More than one point $h(x, y)$ may be associated with the same source S_k. The mapping between the observed $h(x, y)$ to S_k is based on the geometric camera model and is used to compute

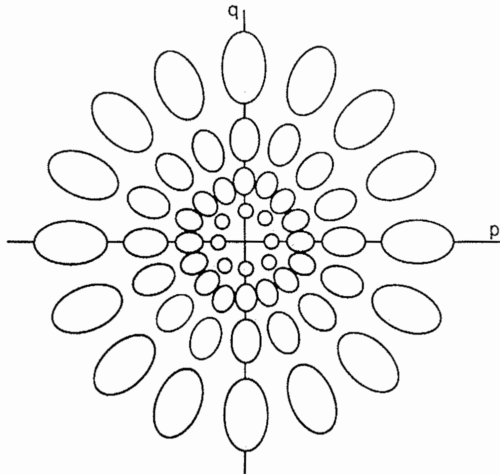


Fig. 4. Reflectance map for a spherical array of distant point sources. Scanning of the point sources corresponds to scanning of the reflectance map to determine local surface orientation.

the estimated surface normal at (x, y):

$$n(x, y) = v(\Theta_k/2, \phi_k) \quad (10)$$

where $v(\ )$ denotes the unit direction vector associated with the given polar angle. The resulting points $\{n(x, y)\}$ are nonuniform samples of the orientation vector over the surface of the object. These samples may then be used for comparison to a known model, as in the inspection case, or for reconstruction of the surface itself as shown schematically in Fig. 2. Examples of the surface normals derived for test objects are described in the next section.

Determination of the surface normally depends on the use of a camera model to relate source position, highlight position, and surface orientation and position. There are several possible camera models which could be used, and the choice of the camera model affects the accuracy of the surface orientation estimate subject to the distant source approximation. The perspective camera model in which all reflected rays pass through a focal point provides the most accurate measurements but requires more extensive calibration procedures. In the orthographic projection model, the focal point is assumed to be an infinite distance from the camera and all reflected rays are perpendicular to the image plane. We have also used a "fixed" camera model in which all rays are emitted from a single point on the reference plane and all surface normal estimates are computed with respect to that reference point.

Sensitivity of the surface orientation measurement to the distant source approximation has been investigated for these three camera models. We consider the case where the incidence, normal, and emittance vectors lie in one plane. The measured angle of the normal vector Θ can be written as a function of coordinates of the point source, the height Δh of the surface element from the reference plane, and the horizontal distance x of the surface element from the origin of the world frame. If Θ_o, Θ_p, Θ_f are the measured angles for the orthonormal, perspective, and the fixed model, respectively, and Θ_{act} is the actual orientation of the surface element, then

$$\Theta_o = \frac{1}{2}\left[\frac{\pi}{2} + \tan^{-1}\frac{b}{a - x}\right] \quad (11)$$

$$\Theta_p = \frac{1}{2}\left[\tan^{-1}\frac{f}{x} + \tan^{-1}\frac{b}{a - x}\right] \quad (12)$$

$$\Theta_f = \frac{1}{2}\left[\frac{\pi}{2} + \tan^{-1}\frac{b}{a}\right] \quad (13)$$

$$\Theta_{act} = \frac{1}{2}\left[\tan^{-1}\frac{f + \Delta h}{x} + \tan^{-1}\frac{b - \Delta h}{a - x}\right]. \quad (14)$$

The measurement errors can then be written as

$$\Theta_{eo} = \Theta_{act} - \Theta_o \quad (15)$$

$$\Theta_{ep} = \Theta_{act} - \Theta_p \quad (16)$$

$$\Theta_{ef} = \Theta_{act} - \Theta_f \quad (17)$$

for the orthonormal, perspective, and fixed model, respectively. Fig. 5 shows the sensitivity of the measurement error for each of the above models to the height Δh of the surface element from the reference plane and the distance x from the optical axis. The error in each case is the percentage error in measurement. The reference source distance $R = 1.7$ in. The value of x lies between -0.03 inches 0.03 inches. The error has been plotted for values of Δh ranging from 0 to 0.2 inches. For the case of surface mounted solder joints, Δh will not exceed approximately 0.05 inches. We observe that over this range the sensitivity plots for the perspective and the fixed model are similar. The orthonormal model is more sensitive to Δh and x and the absolute percentage error for any particular Δh and x is greater in this case than for the perspective and fixed model. For example, when $\Delta h = 0.05$ inches and $x = 0.2$ inches, the percentage errors are 1.25, 0.85, and 0.91 percent for the orthonormal, perspective, and fixed models, respectively. The above results are surprising since the fixed model was expected to be less accurate than the orthonormal model. The higher accuracy and relatively lower sensitivity of the fixed model is a consequence of the partial cancellation of errors in measurement due to non zero values of Δh with the errors in measurement due to the nonzero values of x. As expected, the perspective model is the most reliable among the three models. However, the fixed model is more accurate and less sensitive than the orthonormal model, and is therefore a good choice for practical implementation since it reduces the need for calibration as well as required computation. The measurement errors in the fixed model are at most 0.5 percent more than for the perspective model and therefore the fixed model can result in a maximum measurement error of 1.5 percent.

Other factors affecting measurement accuracy include point source positioning error and highlight measurement error. The sensitivity to point source positioning error has also been computed for the three camera models described above and these results are shown in Fig. 6. Again, the

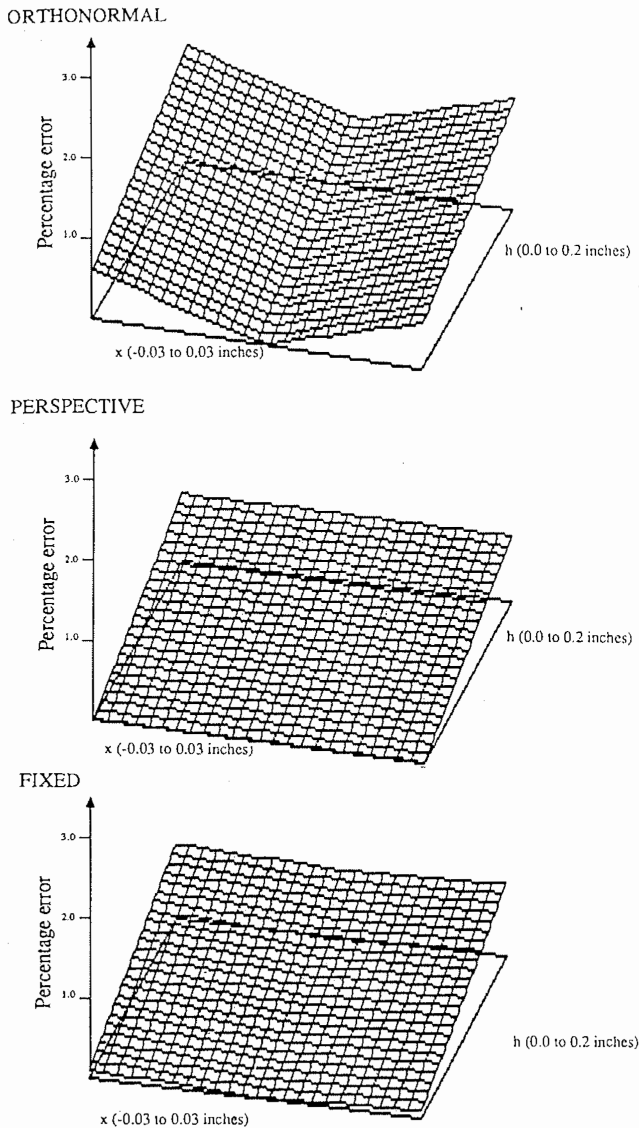


Fig. 5. Sensitivity maps to errors in surface height Δh and position of origin x.

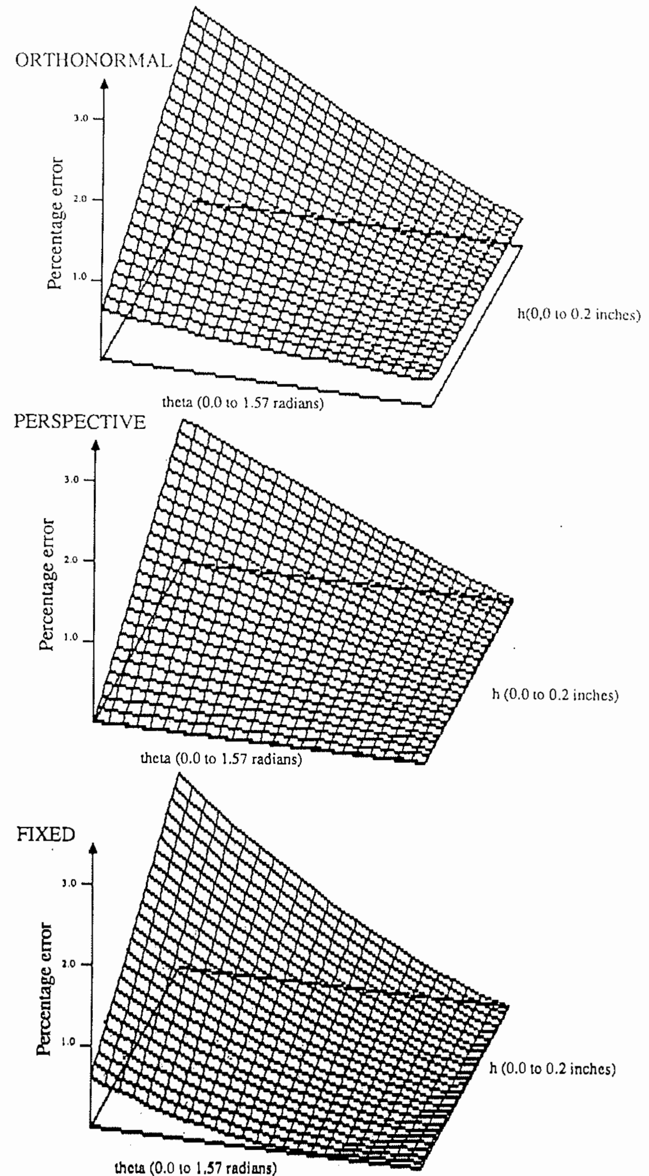


Fig. 6. Sensitivity maps to errors in point source orientation in Θ and surface height Δh.

perspective model is most accurate while the fixed model provides acceptable performance for small angular and height errors. Based on this analysis, we require point source angular errors of less than 0.1 radians. In our current system, this requires point sources of less than 0.080 inch diameter.

The highlight measurement error is directly related to the camera resolution, magnification, and highlight size. Based on the model sensitivity analysis above, we require an object size in the surface plane of less than $0.07h$ which in our current system requires object size of less than 0.12 in. The resulting surface resolution for a 512×512 camera at minimum magnification is $0.14h \times 10^{-3}$ inches. In the current system this results in 0.23×10^{-3} inch resolution based on camera resolution. This spatial resolution is further affected by the highlight detection scheme. The highlight is detected by thresholding and finding the center-of-gravity of the resulting connected region. The quality of the highlight image depends on camera and surface characteristics which are difficult to quantify. In our experimental work, repeatability of the highlight location is less than 10 pixels, and spatial resolution of less than 0.001 inches are achievable.

IV. Experiments

We have conducted a number of experiments on structured highlight inspection of small objects with shiny surfaces using the prototype SHINY system shown schematically in Fig. 7. The experimental apparatus consists of a semicircle of point sources fixed relative to a television camera and a scanning system which permits sequential addressing of individual point source elements. This apparatus provides a good distant source approximation to objects which are 0.120 inches or less. Individual images are digitized and processed using a vision system based on a SUN workstation. Highlights are extracted and tab-

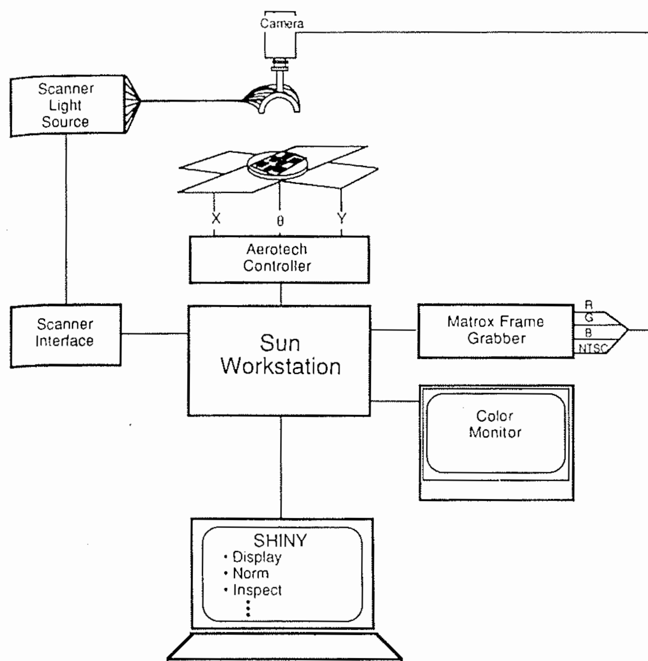


Fig. 7. Schematic diagram of the SHINY system.

ulated and surface normals are computed based on look-up tables derived from calibration experiments. The resulting hardware and software system called SHINY, or Structured Highlight INspection sYstem, is currently being used for analysis and feasibility studies in a number of industrial applications areas.

The point sources used in the system are based on a fiber optic array which was designed and fabricated at CMU. Each side of the array consists of ten 0.040 inch fiber optic light sources which act as the point sources. The smaller the fiber diameter, the closer is the approximation to a true point source. However, as the diameter decreases, the highlight intensity decreases making it more difficult to image the highlight. While the fiber size was chosen based upon experimental studies, smaller diameters are feasible. At their opposite end, the fibers are attached to an electromechanical scanning device which was also designed and fabricated at CMU. The scanner can be commanded to select a particular fiber and guide a light beam from a fiber optic cable into that fiber.

The height of the point source and imaging fixture above the reference surface is adjusted using a fine motion linear translational stage. The fixture is mechanically positioned with four "calibration posts." The post heights are set at 0.125 inches. Once in place, the translational stage is fixed and the posts are removed. The mechanical calibration posts are also used to attach a camera calibration plate to the fixture. The calibration plate has seven 0.008 inch holes drilled through its face, and mounted such that its upper surface lies on the reference plane. A calibration procedure to calculate camera focal length and image transformation parameter uses this plate. The location of the holes are precisely known relative to the calibration posts and thus to the world coordinate frame associated with the fixture. The calibration procedure solves the geometric equations for the six unknown parameters relating the image coordinate frame to the world coordinate frame. This coordinate transformation is then used as the basis for interpretation of highlights in terms of camera models described in the previous section.

The SHINY software environment is a command interpreter composed of basic imaging functions (e.g., frame grab, binarization, connectivity, etc.) [17] and the customized SHINY procedures. For calibration, the calibration plate is attached to the measurement fixture, and all seven light sources on the calibration plate are illuminated simultaneously. The image of the highlights corresponding to the sources is grabbed by the frame grabber, binarized, and the program "connectivity" finds the seven blobs and their *c*enter-*o*f-*g*ravity (COG) image locations. A one-to-one correspondence is established between the image blobs and the point source on the plate using the known configuration of the sources. The world coordinates of the sources and their respective image coordinates are used to solve for the transformation constants. Subsequently, the world and transformed coordinates of the remaining point sources are used to determine the best estimate of the focal point of the camera. In our implementation we have used a magnification and gradient fit.

The measurement procedure involves scanning the 21 point sources of the measurement fixture. For each point source illuminated, the image is grabbed by the frame buffer and binarized. Connectivity finds the COG's of the blobs in the image. When the size of a blob is relatively large, it is subdivided into smaller blobs. Each blob location is then transformed to its world coordinates using the previously computed transformations. The surface orientation at that point is then found using one of the three camera models. For display purposes, the surface orientations can be superimposed on the image as normal vectors. The displayed length of each normal is the projection of that unit vector onto the image plane. In addition, the vector is displayed in a color which identifies the light source that produced it.

The current implementation of SHINY has been used for experimental evaluation of the basic principles and demonstration of specific applications including solder inspection. The current system requires several seconds per point and is limited by the simple scan mechanism for point sources. While the current prototype system is too slow for most real inspection applications, refinements in the design of scanning mechanisms, vision hardware, and software would improve this performance, and adequate speed for many practical applications seems achievable.

In order to determine the measurement accuracy of the system we scanned a polished sphere of 1/16 inch diameter as shown schematically in Fig. 8. The surface orientation is measured at each point on the sphere, and this estimate is compared to the actual orientation. The sphere has a monotonically varying surface orientation in the x and y directions and no two points on the surface have the same orientation. Consequently, each point source produces a single blob in the image as seen in Fig. 9(a). For

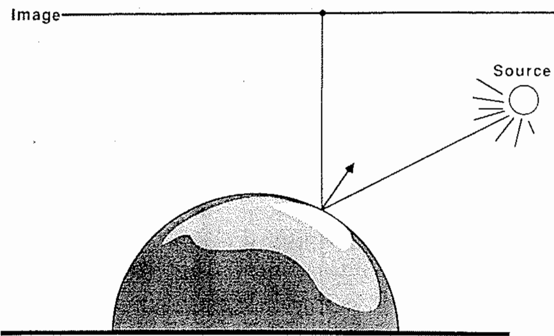
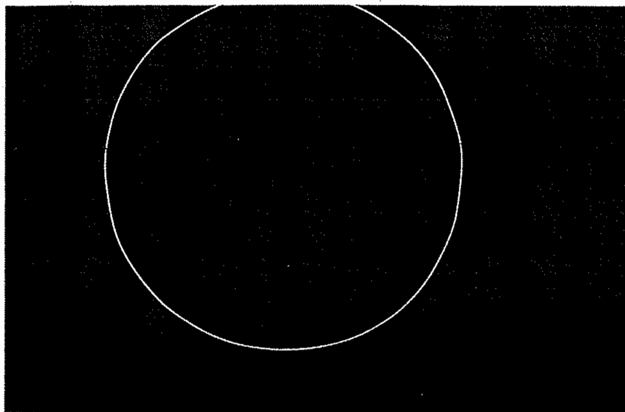
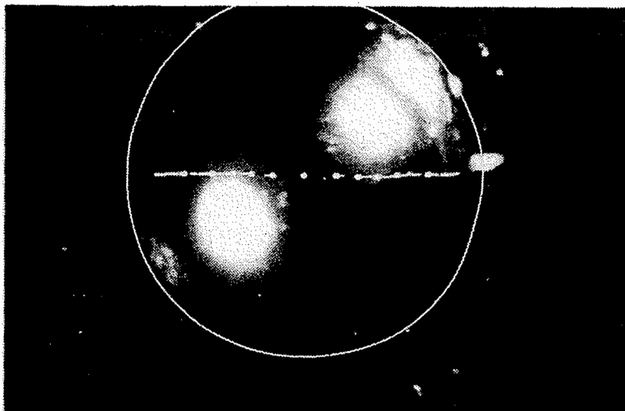


Fig. 8. Schematic view of the calibration hemisphere used to evaluate the structured highlight SHINY system. The hemisphere is 0.0625 inches in diameter and positioned at the origin of the coordinate frame.



(a)



(b)

Fig. 9. (a). The highlight and normal vector produced by a single source illuminating sphere. (b). Set of highlights and normal vectors for a single scan of the sphere. Vectors are superimposed in image of the sphere with general background illumination.

more complex surfaces, such as a solder joint, each point source may produce multiple highlight reflections. The percentage error in measurement for the example in Fig. 9(a) was 0.3 percent. A complete scan of the cross section of the sphere is shown in Fig. 9(b). For clarity every other highlight has been displayed. The overall accuracy of the system for a complete scan is demonstrated by finding the deviations of the points of intersections of the extended normal vectors from the physical center of the sphere. This

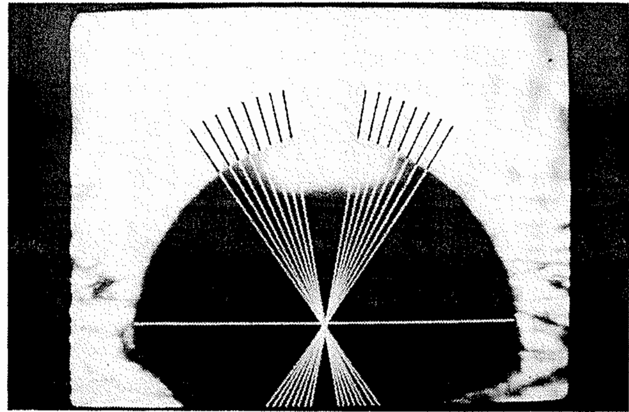


Fig. 10. Photo of normal vectors determined by the SHINY system. The vectors are extrapolated to the center and overlaid on a profile image of the object itself. The intersection of the normals demonstrates the accuracy of the system.

was done by taking a side view of the sphere and superimposing the measured normal vectors on the surface of Fig. 10. The figure shows that the vectors do indeed intersect at the center of the sphere.

Fig. 11 shows a cross-section diagram of a typical surface-mount solder joint. The wire lead is soldered to the surface of a printed circuit board. The shape of the solder surface provides strong indications of the quality of the solder connection itself. The solder joint itself is less than 0.050 inches long and is very tedious to inspect manually. Inspection and reconstruction of such a solder joint profile provides a useful basis for evaluation of solder joint quality.

The reconstruction of the object surface from a set of structured highlight normal vectors is difficult since the sampling of the surface points is nonuniform. Surface normals are measured at uniform orientation steps, not uniform spatial steps. The surface reconstruction therefore uses interpolation of intermediate points followed by smoothing over local regions of the surface. Interpolation is carried out by mapping measured normals into a one-dimensional array for each cross-section. Circular interpolation is applied to pairs of points which are sufficiently close in distance and angle. This results in a smooth transition between successive pairs. When the distance or angle difference betwen points is too large, a discontinuity in shape is assumed. At such points, the relative height differential may be unknown. Fig. 12 shows a reconstructed profile of a solder joint for a surface-mounted component based on this interpolation procedure. The total length of this lead is 0.050 inches and the resolution of surface normal measurement is less than 2° at approximately 0.001 inch spatial accuracy. The spatial resolution varies due to nonuniform spatial sampling. This profile requires approximately 1 minute to scan, estimate, interpolate, and reconstruct using the current system.

The reconstructed profile shown in Fig. 12 is a useful basis for classification of some types of solder defects. We have demonstrated capability to automatically extract features from these profiles and classify several solder de-

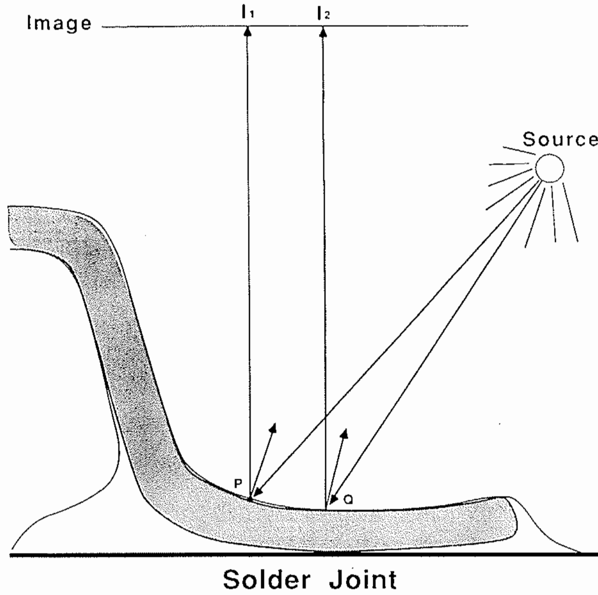


Fig. 11. Schematic diagram of a solder joint for a surface-mounted component on a circuit board. The shape of the solder surface provides information on the quality of the joint.

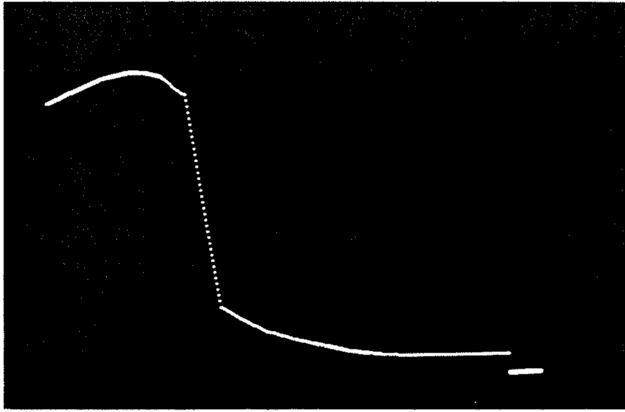


Fig. 12. Photo of reconstructed profile of solder joint. Measured surface normals are weighted and smoothed to produce this estimate.

fect types. This inspection and classification procedure will not be described in this paper.

V. Structured Highlight Stereo

In Section III, the distant source approximation was used to uniquely define the angle of incidence of illumination from a given source. In general, a given surface element may be in some larger working volume, at some point (x, y, z) relative to the origin where the distance $d(x, y, z)$ is not small relative to R, the distance to the source. Observation of the reflected highlight uniquely determines the (x, y) position of the surface element in the camera plane, but leaves a continuous set of (z, Θ, ϕ) combinations undetermined. In this section, we show that this ambiguity in surface orientation and surface height can be resolved using a second camera, and a stereo matching procedure to uniquely determine position and orientation of each surface element. We describe the structured highlight stereo approach first in the plane, then summarize its extension to three dimensions.

Fig. 13 shows the geometry of a point source illuminating a surface element at position (x, y) which is off the origin but in the plane of the source and camera axis. The camera axis is the z-axis and orthographic projection is assumed. In practice, the orthographic assumption may be replaced by a perspective model. Some point source in the array, S_k, illuminates the surface element and reflects a highlight into the camera image. The resulting highlight position determines the x position of the surface element. The height z of the element which satisfies this imaging condition depends on the surface orientation Θ. The analytical relationship between the height z and the orientation Θ may be derived by noting that all the elements which reflect light from S at (x, z) must lie on one of the family of parabolas:

$$(x - x_s)^2 = 4a(z - z_s + a), \tag{18}$$

where a is the parameter of the parabola family. The slope of the surface element at each point is therefore:

$$p = (x - x_s)/2a. \tag{19}$$

The resulting relationship between z and Θ may be expressed as:

$$\hat{z} = \cos\alpha + (\hat{x} - \sin\alpha)/\tan 2\Theta, \tag{20}$$

where

$$\hat{z} = z/R, \tag{21}$$

$$\hat{x} = R\cos\alpha. \tag{22}$$

Equivalently, we can express this relation in terms of the gradient space p as:

$$\hat{z} = \cos\alpha - [(\hat{x} - \sin\alpha)/2][p - 1/p]. \tag{23}$$

Examples of these relations z versus Θ are plotted in Fig. 14 for several different sources at angle α. The set of points on each curve are all consistent solutions to the reflection equations for the observed highlight. The ambiguity in the determination of p maps into a range of possible reflectance map interpretations for a given source.

The ambiguity among possible interpretations of the observed highlights may be removed by using a second camera in a stereo configuration as shown in Fig. 15. One camera C' is positioned at angle $+\beta$ with respect to the vertical, and a second camera C'' is positioned at angle $-\beta$ with respect to the vertical. Each of the sources S sequentially illuminates the object and two camera images, $I'(x, y)$ and $I''(x, y)$, are acquired for each. The thresholded images yield sets of highlight points $h'(x, y)$ and $h''(x, y)$ due to S. For the planar case, the observed points x' due to S' at α' and x'' due to S'' at α'' must satisfy the following equations:

$$z' = \cos\alpha' + [(x' - \sin\alpha')/\tan 2(\Theta - \beta)], \tag{24}$$

$$z'' = \cos\alpha'' + [(x'' - \sin\alpha'')/\tan 2(\Theta + \beta)], \tag{25}$$

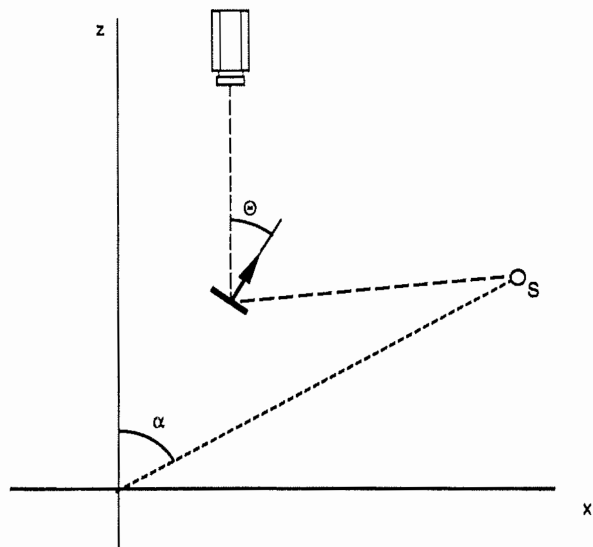


Fig. 13. Geometry of a point source illuminating a surface element at arbitrary position (x, z) in the plane. Determination of the local surface orientation θ is ambiguous from the observation of x. A second camera is required to resolve the ambiguity.

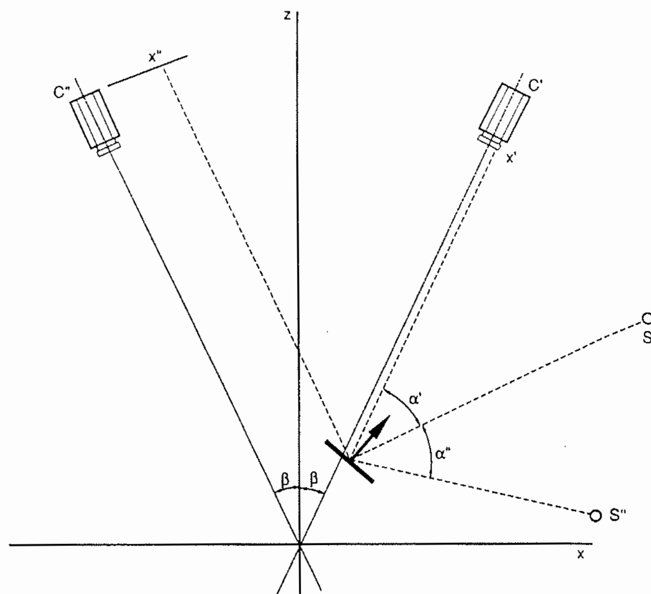


Fig. 15. Stereo structured highlight configuration. Two sets of highlight points are observed for each scan of the array. Use of two cameras resolves the ambiguity in (z, Θ).

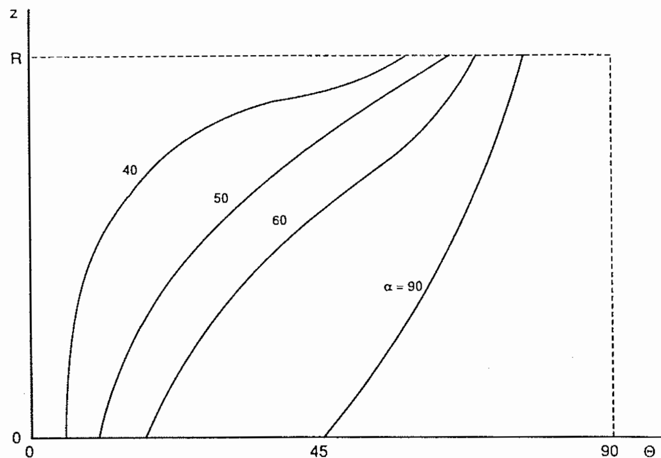


Fig. 14. Examples of plots of z versus θ for different observed source reflections. The set of points on each curve are all consistent solutions to the reflection equations for the observed highlight.

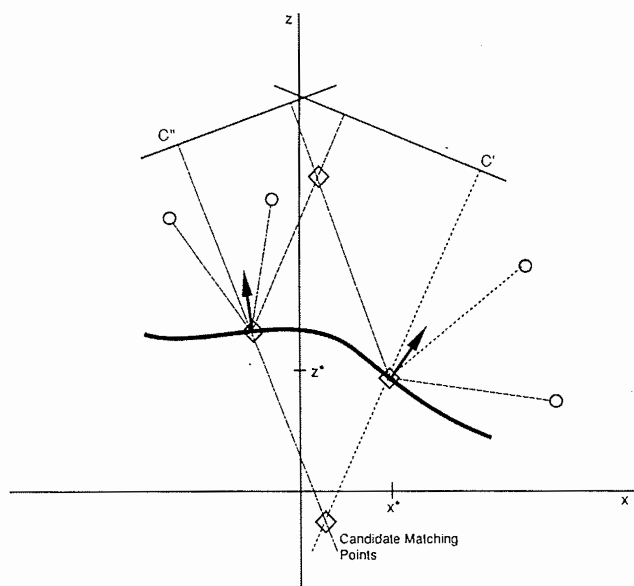


Fig. 16. There are many intersection points which are candidate matches between the two cameras. A matching procedure which compares consistent surface orientations is used to resolve which matches are retained.

where α' and α'', x' and x'', and z' and z'', are measured relative to the individual camera coordinate frames.

Solution of these equations to obtain a consistent (z, Θ) pair requires two steps: intersection finding and reflectance matching.

Intersection finding determines the candidate points for reflectance matching. Each observed highlight point x'' defines a ray in (x, z) space, and each intersection of two rays defines a candidate point where a reflecting surface element might be. As shown in Fig. 16, there are many more candidate points than actual matches. Solution to the ray equations gives the following intersection (x^*, z^*):

$$x^* = (x_0' + x_0'')/2 \cos \beta, \tag{26}$$

$$z^* = (x_0' - x_0'')/2 \sin \beta. \tag{27}$$

The intersection points may be transformed into camera coordinates:

$$x'^* = x', \tag{28}$$

$$z'^* = x^* \sin \beta + z^* \cos \beta, \tag{29}$$

$$x''^* = x'', \tag{30}$$

$$z''^* = -x^* \sin \beta + z^* \cos \beta. \tag{31}$$

Reflectance matching checks to see if there is a consistent interpretation of each set (x^*, z^*) in terms of the surface orientation at that point. To carry out this matching, we solve the equation:

$$z'(\Theta - \beta) - z'^* = z''(\Theta + \beta) - z''^* \tag{32}$$

for Θ^* using a Newton–Raphson gradient method. Then we recalculate the errors:

$$E'(\Theta^*) = z'(\Theta^* - \beta) - z'^*, \quad (33)$$

$$E''(\Theta^*) = z''(\Theta^* + \beta) - z''^*, \quad (34)$$

and check that E' and E'' are small according to some criterion. This solution and test are shown graphically in Fig. 17, where we have plotted $z'(\Theta - \beta) - z'^*$ and $z''(\Theta + \beta) - z''^*$ versus Θ and expect the intersection to be close to the $z = 0$ axis. If E' and E'' are sufficiently small, then the match is confirmed and the surface element at z^* is interpreted with orientation Θ^*. These calculations have been carried out for simulated cases such as that shown in Fig. 16, but not yet for experimental data.

As shown in Fig. 16, pairs of surface elements may generate intersections which may be confused with real surface elements. In general, the requirement in (32) is sufficient to guarantee that intersections which do not correspond to surface elements are rejected. However, there are singular point solutions to

$$z'(\Theta_1 - \beta) - z'^* = z''(\Theta_2 + \beta) - z''^* \quad (35)$$

such that $\Theta_1 = \Theta_2$ at an intersection z^* which would yield ambiguous interpretation of the point. Theoretically, such solutions will be infinitely rare since they require exact adjustment of orientations of two real surface elements which generate the highlights and then masquerade as a real element. However, in practical experiments, finite resolution of angular measurement will cause such misinterpretations to occur and will yield spurious data points. The frequency of occurrence will depend upon the resolution. These points occur in addition to the correct points and result in ambiguous assignment of a surface element along a ray. Similar ambiguity arises in other stereo techniques, such as random dot stereograms. This ambiguity arises less commonly in structured highlight stereo than in other stereo approaches due to the more exact matching criterion. However, similar techniques which impose connectivity and smoothness constraints may be used. A relaxation algorithm such as Marr and Poggio [18] introduced might be used to impose constraints about surface smoothness and converge to a smooth surface estimate.

Our approach to structured highlight stereo poses the same computational issues as other stereo methods. Given N point sources we observe $O(N)$ highlights and $O(N^2)$ intersections which are candidate surface points. The matching procedure is an $O(N^2)$ combinatorial search problem which tests the consistency of these candidate points subject to constraints. This problem has been addressed from several different perspectives leading to reduction in computational complexity. Reduction in complexity is achieved by selectively searching regions of points subject to connectivity and smoothness constraints. This underlying smooth surface model assumption may reduce computational complexity to $O(kN)$ for a surface tracking algorithm where k is a parameter of the local surface model. The choice of the implementation will depend on the application since smoothness and connectivity assumptions will depend on the nature of the specific class of objects of interest.

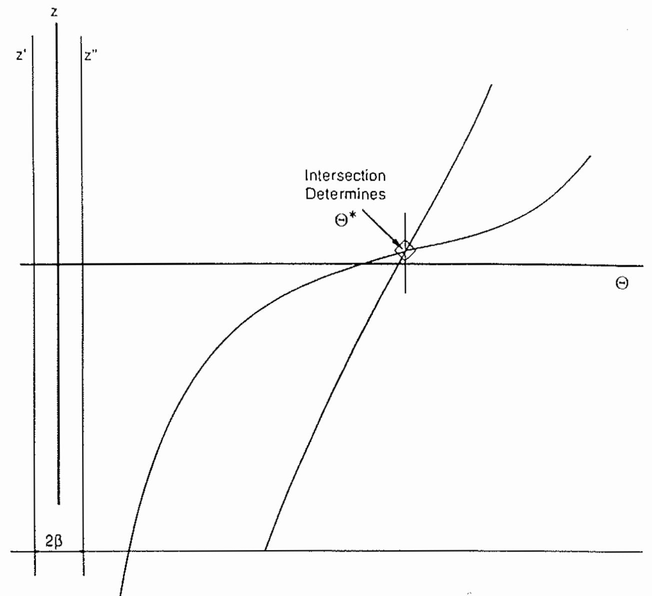


Fig. 17. Graphical representation of the correspondence matching test for stereo pairs. The individual $z - \Theta$ curves are shifted horizontally by 2β and shifted vertically by z'^*. The resulting intersection should coincide with the horizontal axis.

The extension of structured highlight stereo to three dimensions is straightforward, but requires some additional attention to the intersection and matching procedures. Given a source at position vector $\boldsymbol{p}_s$ and a surface element at position vector $\boldsymbol{p}$, define $\boldsymbol{r}$ to be the vector from source to surface element and a camera axis through S with z-axis z'. A visible highlight will result from a surface patch which is part of a family of paraboloids:

$$z' = (r'^2/4a) - a, \quad (36)$$

where a is the parameter of the parabola family, and

$$r' = (x'^2 + y'^2)^{1/2}. \quad (37)$$

The resulting relationship between element position (x', y', z') and surface gradient (p', q') is:

$$z' = x'p'/2 + y'q'/2 - r'/(2[p'^2 + q'^2]^{1/2}), \quad (38)$$

$$q'/p' = y'/x'. \quad (39)$$

Testing the *intersection* of rays in three dimensions is more difficult since exact intersections are unlikely to occur. Instead, we compute the distance:

$$D = (\boldsymbol{v}' - \boldsymbol{v}'')(\boldsymbol{u}' \times \boldsymbol{u}'')/u' \times u'', \quad (40)$$

between the unit direction vectors $\boldsymbol{u}'$ and $\boldsymbol{u}''$ in the direction of view, given the position vectors $\boldsymbol{v}'$ and $\boldsymbol{v}''$ of observed points in the image plane. Observed highlights with rays which are sufficiently close to intersection ($D < D_0$) are labeled as candidate matches. For candidate matches,

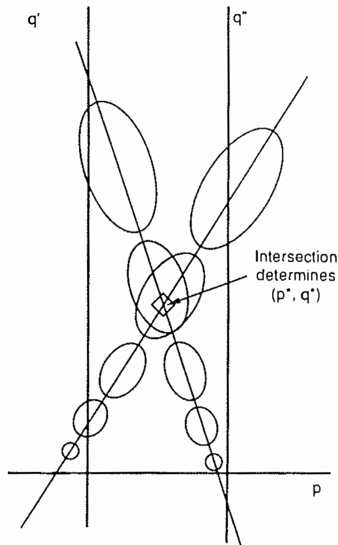


Fig. 18. Graphical interpretation of the full three-dimensional stereo matching procedure. Each camera generates a reflectance map with a locus of possible solutions. At the intersection point (p^*, q^*), the consistency of the z'^* values is checked to determine the validity of the match.

the distance in each camera coordinate frame may be found by:

$$z' = [\mathbf{r}'' - \mathbf{r}'] \cdot \mathbf{u}', \tag{41}$$

$$z'' = [\mathbf{r}' - \mathbf{r}''] \cdot \mathbf{u}''. \tag{42}$$

The *reflectance match* is tested from the candidate pair by:

$$\Delta z = [z'(p^*, q^*) - z'] - [z''(p^*, q^*) - z'']. \tag{43}$$

If Δz is close to zero, then the match is consistent with reflectance criteria. (p^*, q^*) are determined by simultaneous solution of:

$$(q - q') = (y'/x')(p - p'), \tag{44}$$

$$(q - q'') = (y''/x'')(p - p''). \tag{45}$$

The graphical interpretation of the three-dimensional matching procedure is shown in Fig. 18. Each camera image generates a reflectance map with a set of candidate matches along a line defined by (p', q'). The two reflectance maps are displaced by an amount due to β the relative angle of the two cameras. The intersection of the two reflectance map loci determines (p^*, q^*), and the consistency of the z'^* values with the z values predicted by ray intersection determines the quality of the match.

VI. Conclusions

In this paper we have described a structured highlight technique for estimation of orientation of specular surfaces. Experiments with the technique using the SHINY prototype system for a small object and distant source approximation demonstrate that accurate mapping of smooth surface contours may be achieved. Analysis of the general structured highlight approach shows that two cameras are required to determine surface orientation. In practical systems, with finite angular resolution, additional constraints may be required to resolve stereo matching ambiguities. A procedure for structured highlight stereo has been described and evaluated for simulated data.

The structured highlight approach is promising for many common industrial problems including inspection of surface shape of machined parts and inspection of solder joint surfaces. In practice, the accuracy and speed of the technique will strongly influence its range of application. In addition to the factors discussed above, the accuracy will be determined strongly by the calibration of the system. The relative position of coordinate frames attached to camera and source array must be known with high accuracy in order to determine surface orientation. For many inspection problems, the object itself must also be positioned with high relative accuracy in order to match measured contours with models. Implementation of the SHINY system described here required detailed attention to the calibration problem.

A second factor affecting practical implementation is speed. Many inspection tasks must be carried out at high speed, and any technique which would be comparable to human performance would require corresponding speed. The structured highlight approach is inherently slowed down by the need to scan across a source array. The process may be speeded up considerably by appropriate nonsequential coding of source displays. The low-level image problem itself is relatively simple and additional gains may be made through special hardware which implements the structured highlight algorithm directly. The use of two cameras for stereo structured highlight requires no further scanning time, but there is significantly more computation required by the intersection and matching procedures. Practical implementation of the stereo procedures might therefore require hardware acceleration of these procedures.

Acknowledgment

The authors are grateful to W. Satterfield and M. J. Dowling for assistance in preparation of the manuscript.

References

[1] B. K. P. Horn, "Determining shape from shading," in *The Psychology of Computer Vision*, P. H. Winston, Ed. New York: McGraw-Hill, 1975.

[2] ——, "Image intensity understanding," *Artificial Intell.*, vol. 8, no. 2, 1977.

[3] K. Ikeuchi, "Numerical shape from shading and occluding contours in a single view," Artificial Intelligence Lab., M.I.T., Cambridge, AI Memo 566, 1980.

[4] R. J. Woodham, "Photometric stereo: A reflectance map technique for determining surface orientation from image intensity," *Proc. SPIE*, vol. 155, pp. 136–143, 1978.

[5] ——, "Photometric method for determining surface orientation from multiple images," *Opt. Eng.*, vol. 19, no. 1, pp. 139–144, 1980.

[6] B. K. P. Horn and K. Ikeuchi, "The mechanical manipulation of randomly oriented parts," *Sci. Amer.*, vol. 251, no. 2, pp. 100–111, Aug. 1984.

[7] E. N. Coleman and R. Jain, "Obtaining 3-dimensional shape of textured and specular surface using four-source photometry," *Comput. Graphics Image Processing*, vol. 18, no. 4, pp. 309–328, Apr. 1982.

[8] G. Agin, "Representation and description of curved objects," Stanford Univ., AI Memo, Oct. 1972.

[9] J. G. Davy, "A comprehensive list of wave solder defects and their probable causes," *Brazing and Soldering*, no. 9, pp. 50–59, Autumn 1985.

[10] P. J. Besl, E. J. Delp, and R. Jain, "Automatic visual solder joint inspection," *IEEE J. Robotics Automation*, vol. RA-1, pp. 42–56, Mar. 1985.

[11] P. A. Merrill and M. D. Levine, "Visual inspection of solder joints by computer," McGill Univ., Rep. TR-85-1R, Jan. 1985.

[12] Y. Nakagawa, "Automatic visual inspection of solder joints on printed circuit boards," *SPIE Robot Vision*, vol. 336, pp. 121–127, 1982.

[13] R. Vanzetti, A. C. Traub, and J. S. Ele, "Hidden solder joint defects detected by laser infrared system," in *Proc. IPC 24th Annu. Meeting*, Apr. 1981, pp. 1–15.

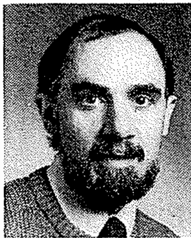
[14] K. Ikeuchi, "Determining surface orientations of specular surfaces by using the photometric stereo method," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-3, pp. 661–669, Nov. 1981.

[15] M. D. R. Babu, C-H. Lee, and A. Rosenfeld, "Determining plane orientation from specular reflectance," *Pattern Recognition*, vol. 18, no. 1, pp. 53–62, 1985.

[16] B. K. P. Horn and R. W. Sjoberg, "Calculating the reflectance map," *Appl. Opt.*, vol. 18, no. 1, pp. 1770–1779.

[17] R. Bracho, J. F. Schlag, and A. C. Sanderson, "POPEYE: A gray-level vision system for robotics applications," Robotics Inst., Carnegie-Mellon Univ., Rep. CMU-RI-TR-83-6.

[18] D. Marr and T. Poggio, "Cooperative computation of stereo disparity," *Science*, vol. 194, pp. 283–287, 1976.



Arthur C. Sanderson (S'66–M'68–SM'86) is Professor and Department Head of Electrical, Computer and Systems Engineering at Rensselaer Polytechnic Institute. He received the B.S. degree from Brown University, Providence, RI, in 1968, and the M.S. and Ph.D. degrees from Carnegie-Mellon University, Pittsburgh, PA, in 1970 and 1972, respectively.

He is Professor and Department Head of Electrical, Computer and Systems Engineering at Rensselaer Polytechnic Institute, Troy, NY. From 1968 to 1970 he was a Research Engineer at the Westinghouse Research Laboratories, and from 1972 to 1973 he was a Visiting Research Fellow at Delft University of Technology, Delft, The Netherlands. From 1973 to 1987, he was a faculty member at Carnegie-Mellon University in the Department of Electrical and Computer Engineering. At CMU, he participated in the founding and development of the Robotics Institute, and was Associate Director of the Institute from 1980 to 1987. From 1985 to 1987, he was on leave from CMU and held the position of Director of Information Sciences Research at Philips Laboratories, Briarcliff Manor, NY. He is the author of over 120 technical publications and proceedings. His current research interests include planning systems for robots and automation systems, sensor-based control, computer vision, and applications of knowledge-based systems.

Dr. Sanderson is Associate Editor of the *IEEE* JOURNAL OF ROBOTICS AND AUTOMATION. He was the Program Chairman for the 1987 IEEE International Conference on Robotics and Automation. He is a member of AAAI, SME, AAAS, and serves on the editorial boards of a number of other journals in the areas of robotics and computer vision.

Lee E. Weiss received (S'79–M'82) the B.S. degree in electrical engineering from the University of Pittsburgh, Pittsburgh, PA, in 1972, and the M.S. degree in bioengineering and the Ph.D. degree in electrical and computer engineering from Carnegie-Mellon University, Pittsburgh, PA, in 1974 and 1984, respectively.

From 1974 to 1976 he was a Research Assistant associated with numerous Pittsburgh hospitals. During this period he was involved with the development of several prosthetic and implantable devices, and with the signal processing aspects of computerized fetal monitoring systems. From 1976 to 1979 he was a bioengineer at ARCO (Atlantic Richfield) Medical Products Company where he was responsible for the design and manufacture of cardiac pacemaker heartleads. In 1979 he became the first graduate student to be supported by the Robotics Institute of Carnegie-Mellon, where he is currently a Research Scientist. He holds several patents in both the robotics and bioengineering fields. His research interests in robotics include dynamic sensory control, automated visual inspection, flexible assembly, and learning systems for automated workcells.

Shree K. Nayar (S'86) was born in Bangalore, India, on January 15, 1963. He received the B.S. degree in electrical engineering from the Birla Institute of Technology, Ranchi, India, in 1984, and the M.S. degree in electrical and computer engineering from the North Carolina State University, Raleigh, in 1986.

He is currently a Research Assistant at the Robotics Institute, Carnegie-Mellon University, Pittsburgh, PA, and is working toward the Ph.D. degree in electrical and computer engineering. His research interests include computer vision, robotics and artificial intelligence.