Stereo in the Presence of Specular Reflection *

Dinkar N. Bhat
Department of Computer Science
Columbia University
New York, NY 10027
bhat@cs.columbia.edu

Shree. K. Nayar
Department of Computer Science
Columbia University
New York, NY 10027
nayar@cs.columbia.edu

Abstract

The problem of accurate depth estimation using stereo in the presence of specular reflection is addressed. Specular reflection, a fundamental and ubiquitous reflection mechanism, is viewpoint dependent and can cause large intensity differences at corresponding points, resulting in significant depth errors. We analyze the physics of specular reflection and the geometry of stereopsis which led us to a relationship between stereo vergence, surface roughness, and the likelihood of a correct match. Given a lower bound on surface roughness, an optimal binocular stereo configuration can be determined which maximizes precision in depth estimation despite specular reflection. However, surface roughness is difficult to estimate in unstructured environments. Therefore, trinocular configurations, independent of surface roughness, are determined such that at each scene point visible to all sensors, at least one stereo pair can compute produce depth. We have developed a simple algorithm to reconstruct depth from the multiple stereo pairs.

1 Introduction

Stereo is a direct and passive method of obtaining three-dimensional structure of the visual world which makes it attractive for applications like autonomous navigation and surveying. The robustness of a stereo system is characterized to a large extent by its ability to obtain accurate depth estimates of scenes comprising objects with different reflectance properties.

The stereo *correspondence problem* [1] is inherently underconstrained. Therefore, constraints have to be imposed by making assumptions regarding scene reflectance and structure. A common assumption is that intensities at corresponding points in the images are identical. Based on this supposition, various search based strategies have been developed which correlate image regions (*area-based*) [9], or image features (*feature-based*) [6]. However, this assumption is valid only when the surfaces in the scene are Lambertian. Corresponding point intensities are *not* identical in the presence of *specular reflection*, the specular intensity at any scene point being dependent on the viewing direction. This effect is more clearly manifest on smoother surfaces where *highlights* – bright regions due to specular reflection – shift on the surface even with slight changes in viewpoint. Thus, corresponding regions in stereo images can be poorly correlated, causing area-based schemes to compute incorrect depth. Similarly, when highlights are assumed to be real scene features and matched, feature-based schemes can fail. Figure 1 shows a stereo pair of a rendered cup, and depth obtained along two scanlines; one including a highlight and the other away from it. Depth was computed using a correlation based algorithm; hence erroneous at points where corresponding intensities are vastly different. Our paper deals with accurate depth estimation in the presence of specular reflection which is abundant in real scenes.

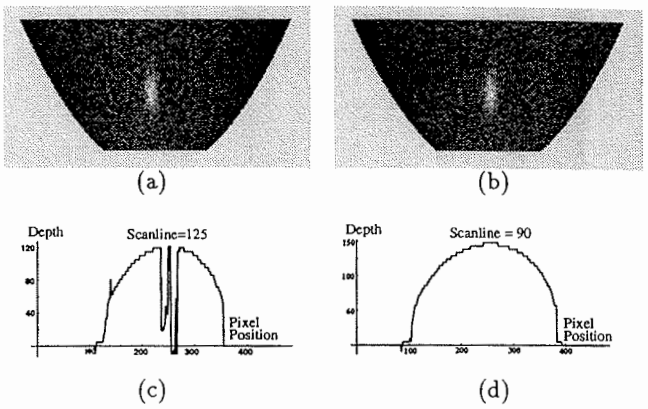


Figure 1: Rendered stereo pair and depth computed along two scanlines. (a) Left image; (b) Right image; (c) Depth along a scanline including the highlight; (d) Depth along a scanline away from the highlight. Large depth errors are observed in (c).

To overcome the problem of depth errors due to strong highlights, Brelstaff and Blake [4] suggested excising them from images before matching. Removal of highlights is difficult in images of real scenes and is an active area of research [7]. Ching *et al* [5] developed an empirical correlation based technique to detect and avoid specular reflection when the camera is active. On a different note, Blake [3] related the movement of a highlight to the Hessian of the surface which describes local surface geometry. The above techniques assume ideal specular reflection which is only an extreme case as surface roughness tends to zero.

Current stereo algorithms are therefore seriously deficient in dealing with specular reflection. In this paper, we address the problem of precise depth estimation in the presence of specular reflection from surfaces with macroscopic roughness. First, we seek an optimal *binocular stereo* configuration such that intensity differences at corresponding points is limited, while depth resolution is maximized. The optimal configuration is determined independent of surface normal and source direction, and its parameters are shown to be a function of surface roughness. Therefore, for a scene where the lower bound on roughness can be estimated – quite possible in structured environments – the two cameras can

*This work was supported in part by an NSF National Investigator Award and in part by ARPA contract DACA-76-92-C-007.

0-8186-7042-8/95 $4.00 © 1995 IEEE

be positioned so as to minimize mismatches without losing depth precision.

Next, we seek to avoid estimation of surface roughness since the measurement of surface roughness is often impractical. We determine *trinocular* configurations whose parameters are independent of surface roughness. The important characteristic of these configurations is that for each scene point in the common field of view of the sensors, *at least* one binocular pair provides the correct depth estimate. We have developed a practical correspondence algorithm to extract correct depth estimates of scene points from different pairs so as to yield an accurate depth map of the scene.

Our approach considers specular reflection from rough surfaces in the context of stereo. All previous methods have implicitly [5] or explicitly [3] assumed ideal specular reflection. We do not attempt to avoid or detect the immediate artifacts of specular reflection like strong highlights but rather perform accurate matching in their presence. Thus, preprocessing of images, like removal of highlights, is avoided. Our approach is not limited by any specific reflectance model or to any correspondence scheme. It is therefore easy to incorporate into existing stereo algorithms.

2 Reflection Mechanisms

Surfaces exhibit two forms of reflectance - *diffuse* and *specular*. Diffuse reflection occurs due to subsurface scattering of light. It is often assumed to be Lambertian, an assumption shown to be incorrect for surfaces with macroscopic roughness [10]. Nonetheless, the change in diffuse component with viewing direction is generally much less pronounced than the change in specular component.

2.1 Specular Reflection

Specular reflection occurs at the boundary between surface and medium. It comprises of two components - a spike and a lobe [8]. We do not deal with surfaces smooth in comparison to the wavelength of incident light as they are rare in real scenes. Hence, specular reflection refers to the lobe only. The specular lobe spreads in directions other than and including the specular direction, the width of the distribution depending on the roughness of the surface. This is described by the Torrance-Sparrow model [11] which is briefly outlined below.

A surface is viewed as a collection of planar microfacets, each behaving like a perfect mirror. A rough surface can be modelled using a probability distribution for the slopes of the microfacets. The slope distribution model uses a parameter σ which represents surface roughness. A smoother surface is characterized by a lower value for σ. Using this surface model, the specular intensity I_s at any point was shown as:

$$\begin{aligned} I_s &= \frac{K_s F\, G}{\hat{n}\cdot\hat{v}} exp(-\frac{1}{2\sigma^2}(cos^{-1}(\hat{h}\cdot\hat{n}))^2) \\ \hat{h} &= \frac{\hat{v}+\hat{s}}{||\hat{v}+\hat{s}||} \end{aligned} \quad (1)$$

where $\hat{v}$, $\hat{s}$ and $\hat{n}$ are unit vectors pointing along the viewing, source and normal directions, respectively; $\hat{h}$ is the bisector of $\hat{v}$ and $\hat{s}$, G is the Geometrical Attenuation Factor, and F is the Fresnel's coefficient. K_s accounts for the gain of the sensor measuring intensity, the source strength, normalization factors in the specular intensity expression, and the reflectivity of the surface. From (1), it can be deduced that: (a) when the surface is smooth, the distribution of I_s is concentrated in a small region around the specular direction, and (b) as the surface becomes rougher, the peak value of I_s decreases and the distribution of I_s widens.

2.2 Implications for Stereo

The total image intensity I_t for any point in the scene is given by the sum of diffuse and specular intensity components. Due to variation in each component with viewing direction, the total intensities of corresponding points in the stereo images are different. But, since the change in diffuse component is much smaller than the change in specular component, it follows that the overall intensity difference I_{diff} is approximately equal to the difference in specular intensities:

$$I_{diff} = |I_s^1 - I_s^2| \quad (2)$$

where, I_s^1 and I_s^2 are the specular intensities of the point in the two stereo images[1]. I_{diff} varies over the scene as the surface normal and roughness are generally not constant. The local variance of I_{diff} (in a window, for example) could be large if the viewing directions are chosen arbitrarily, resulting in wrong matches while computing stereo correspondence using linear correlation methods [2].

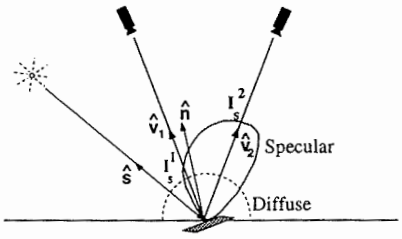


Figure 2: A binocular stereo configuration. Note that the specular intensity is different in the two sensors.

Figure 2 shows a binocular stereo configuration operating at a point with some surface roughness. The question then is: How far apart can the the viewing vectors be located at which I_{diff} exceeds a threshold? This upper limit is bound to be smaller for smoother surfaces since an equivalent change in viewing direction can cause a comparatively large change in I_s(Equation 1). We seek to ascertain this limit independent of surface normal and source direction since these are indeterminable except in highly structured environments.

3 Vergence

We discuss how specular intensity difference at scene points can be affected by camera parameters. When points are projected *orthographically*, as shown in Figure 3, corresponding rays are parallel to their respective optical axes. Thus, the angle between projected rays from all points in the scene can be simultaneously varied, by changing camera vergence β alone. θ_v is termed as *point vergence*. Point vergence is a controllable parameter, independent of surface normal, and affects specular intensity difference at scene points. The relation between point vergence and camera vergence for orthographic projection is simply, $\theta_v = \beta = \beta_1 + \beta_2$. In the case of *perspective* projection, viewing direction at each point in the scene, varies with respect to either viewpoint, i.e point vergence varies across the scene. To define a single

[1] We assume the scene is illuminated by a light source whose direction is fixed but unknown, the gain of the stereo cameras are identical while obtaining the images, and the response of each camera is linear with respect to scene radiance.

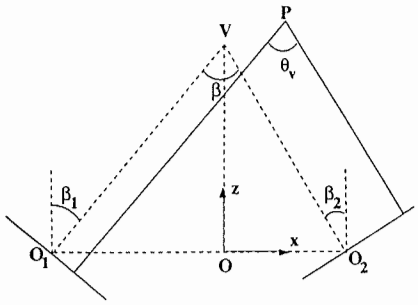


Figure 3: Point vergence and camera vergence under orthographic projection.

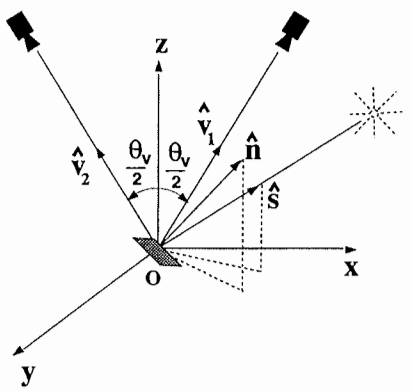


Figure 4: Coordinate system used for the stereo problem. Each scene point is mapped to the origin (O) of the coordinate system.

controllable parameter which affects specular intensity differences over the scene, point vergence can be averaged over a *workspace*[2]. This mean value is called the *field vergence*. If the workspace is defined explicitly in world coordinates, then a relation can be obtained between the field vergence and baseline [2]. Therefore, the baseline indirectly controls I_{diff} over the workspace.

Hereon, we will refer to both point vergence and field vergence as simply *vergence*. Vergence is related to *depth resolution*, an important design parameter. Depth accuracy, and hence resolution, are limited by spatial image quantization amongst other factors. The depth resolution attainable at any point is directly proportional to vergence [2], assuming quantization is the primary cause for matching errors. Achieving maximum depth resolution therefore conflicts with the requirement of minimizing intensity difference over the scene.

4 Binocular Stereo

Determining the maximum acceptable vergence in the presence of specular reflection can be formulated as a constrained optimization problem, as described in this section. We use a left-handed coordinate system (Figure 4) with every scene point mapped to the origin O, and its surface normal is described by a unit vector $\hat{n}$ pointing away from O. The aim is to attain maximum vergence in order to achieve best depth resolution. Hence, a suitable objective function f_{obj} is:

$$f_{obj} = \hat{v}_1 \cdot \hat{v}_2 \tag{3}$$

To limit specular intensity difference I_{diff} at every point in the scene, the following constraint ($c1$) is imposed:

$$I_{diff} < T \tag{4}$$

where T is a threshold. From a statistical perspective, restricting $Idiff$ amounts to limiting the variance of specular intensity difference in any local region [2]. The cameras are restricted to lie in the positive x-z plane, and tilt symmetrically about the z-axis. These constraints ($c2$) can be expressed as:

$$\begin{aligned} \hat{v}_1 \cdot \hat{j} &= \hat{v}_2 \cdot \hat{j} = 0 \\ \hat{v}_1 \cdot \hat{k} &= \hat{v}_2 \cdot \hat{k} > 0 \end{aligned} \tag{5}$$

[2]The workspace could be the entire stereo field of view.

where $\hat{i}$, $\hat{j}$ and $\hat{k}$ are unit vectors along the x,y and z axes, respectively. To avoid grazing incidence and viewing angles, constraints ($c3$) are imposed:

$$\hat{v}_1 \cdot \hat{n},\ \hat{v}_2 \cdot \hat{n},\ \hat{s} \cdot \hat{n} > 0 \tag{6}$$

The optimization problem[3] can now be stated as:

$$\begin{aligned} &Minimize:\ f_{obj} \\ &subject\ to\ constraints:\ (c1, c2, c3) \end{aligned} \tag{7}$$

The variables are $\hat{v}_1, \hat{v}_2, \hat{s}$ and $\hat{n}$. Solving the above problem, the optimal viewing directions $\hat{v}_1^{opt}$ and $\hat{v}_2^{opt}$ and hence the optimal vergence θ_v^{opt}, can be obtained independent of $\hat{s}$ and $\hat{n}$.

To demonstrate a particular solution, the expression for specular intensity given by (1) is used in constraint ($c1$). Dividing both sides by K_s, the constraint can be written as:

$$I_{diff}/K_s < T/K_s \tag{8}$$

It can be seen that T/K_s is an independent parameter. We call it the *relative threshold* which is related to image correspondence. Roughness σ is also unconstrained because surfaces in the scene are unknown. Thus, the *optimal vergence* ${\theta_v}^{opt}$ *is a function of surface roughness* σ *and relative threshold* T/K_s.

The optimization problem is solved numerically and the relationship obtained between ${\theta_v}^{opt}$, σ and T/K_s is shown in Figure 5. The salient features of this relationship are:

- The optimal vergence increases with roughness. The reason is that I_{diff} weakens with increasing roughness allowing larger vergence. The surface progressively behaves in a diffuse manner, and thus the effects of specular reflection on matching diminish.
- The optimal vergence also increases with relative threshold. This is perceivable because a larger threshold permits a larger variation in I_s.

[3]The dot product of the two viewing vectors represents the cosine of point vergence. Therefore, minimizing the value of the dot product amounts to maximizing vergence.

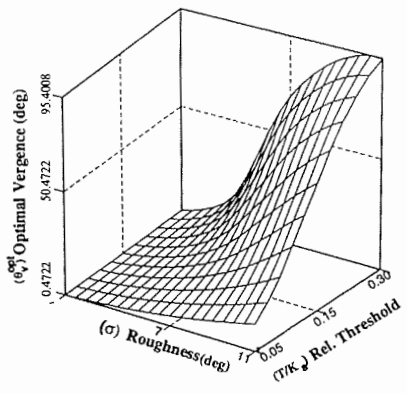


Figure 5: Graph illustrating the relationship between roughness, relative threshold and optimal vergence.

The monotonically increasing relationship of vergence with roughness implies that if the lowest roughness value in the scene is known, then the corresponding optimal vergence can be used for stereo. Arguing similarly, a conservative lower bound for the relative threshold is sufficient to sufficient to configure a system that produces low intensity difference for all scene points. Variations to the general problem can now be considered by modifying the constraints, however the approach to determine the optimal stereo configuration remains unchanged. For example, the normal vectors at all points could be constrained to lie in one plane, i.e all objects have translational symmetry and oriented appropriately.

Since we do not have a closed-form expression for optimal vergence in terms of relative threshold and roughness, we pursued a functional approximation. If $\tilde{\theta_v}^{opt}$ approximates ${\theta_v}^{opt}$, then

$$\tilde{\theta_v}^{opt} = \frac{a\,(T/K_s)^2\,\sigma^2}{(T/K_s)^2 + b\,\sigma^2} \tag{9}$$

where a and b are constants obtained numerically [2].

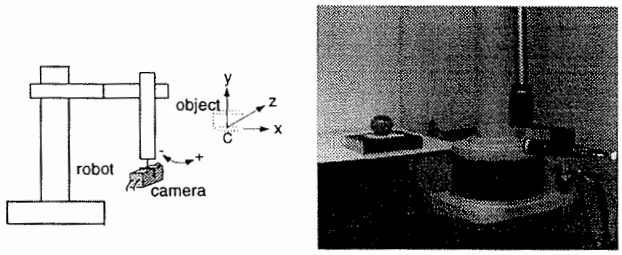
To make the relationship in Figure 5 usable, a correspondence operator is required which is sensitive to changes in the relative threshold and degrades gracefully. The *normalized correlation coefficient* (NCC) measures the degree of linear relationship between intensities in image windows. It is invariant to scaling of the intensities in the windows. With two matching windows W in the two images, containing N pixels and having intensities $I_1^{(i,j)}$ and $I_2^{(i,j)}$, $NCC = 1$ if $I_1^{(i,j)} = I_2^{(i,j)}, (i,j) \in W$, i.e if the corresponding surface is Lambertian[4]. However, due to specular reflection the intensities are not equal, hence NCC deviates from 1. The deviation is estimated by E, where $E \propto \frac{1}{N}\sum_{(i,j)\in W}(\frac{I_1^{(i,j)} - I_2^{(i,j)}}{K_s})^2$ assuming that all points in each window are identically scaled in specular intensity by K_s. Since we limit the intensity difference at all corresponding points by T, it follows that $E < (T/K_s)^2$; thus, NCC is sensitive to changing relative threshold. A closely related stereo operator, the sum of squared differences (SSD), can also be made sensitive to variations in relative threshold [2].

[4] We ignore noise, and geometrical distortion in the windows.

The exact value of the relative threshold when mismatches begin to occur (the breaking threshold) depends on diffuse texture of the surface which are diverse, making its estimation a hard problem. Note that this problem is inherent to stereo matching, and it is only natural that the threshold appears in our formulation. Adopting a conservative lower bound for the relative threshold results in small vergence which in turn implies poor depth resolution. We will show that the problem can be mitigated by using the trinocular stereo approach.

5 Experiments

We illustrate the effect of vergence on stereo matching using surfaces with different roughness. For these experiments, we use a 5 degree of freedom SCARA (Adept) robot (see Figure 6). The end-effector is equipped with a camera to obtain different viewing directions. We use two uniformly rough



(a) (b)

Figure 6: The experimental setup. (a) Diagram of a robot with a camera fixed to its end-effector. The coordinate system is also shown; (b) A photograph of the setting.

cylindrical objects wrapped with different surfaces (see Figure 7); a gift wrapper and a roughened xerox quality paper. Their surface roughness was measured [2], and the values obtained are $\sigma = 3.5°$ and $\sigma = 6.3°$, respectively. It can be also be seen from the the images in Figure 7.

In order to use approximately the same relative threshold, similar random patterns on the surfaces were marked. Images obtained at equal angles about the z-axis are matched along scanlines containing texture[5]. For each surface, depth obtained along a scanline at different vergence values is shown in Figure 7. It can be seen that for each surface large depth errors are computed at larger vergence: 8.0° and 11.0° respectively, although a higher vergence is acceptable for the rougher surface. For the smoother surface, the mismatches are confined to the highlight region over which the variation of I_{diff} is large.

6 Trinocular Stereo

While binocular stereo as described earlier is viable in structured environments where surface roughness can be estimated, it is not generally practical. Further, if the vergence corresponding to the lowest roughness estimate in the scene is used, then the depth resolution obtained for rougher surfaces is suboptimal. Thus, we seek an alternative scheme.

Figure 6 shows schematics of a trinocular system configured such that the intensity difference at a point, with varying surface roughness, is constrained to a threshold in *at least* one pair of views. Therefore, depth of the point can be

[5] We have imposed the scanline epipolarity constraint by ensuring that the robot moves in the x-z plane only, and by using imaging optics that approximates orthographic projection.

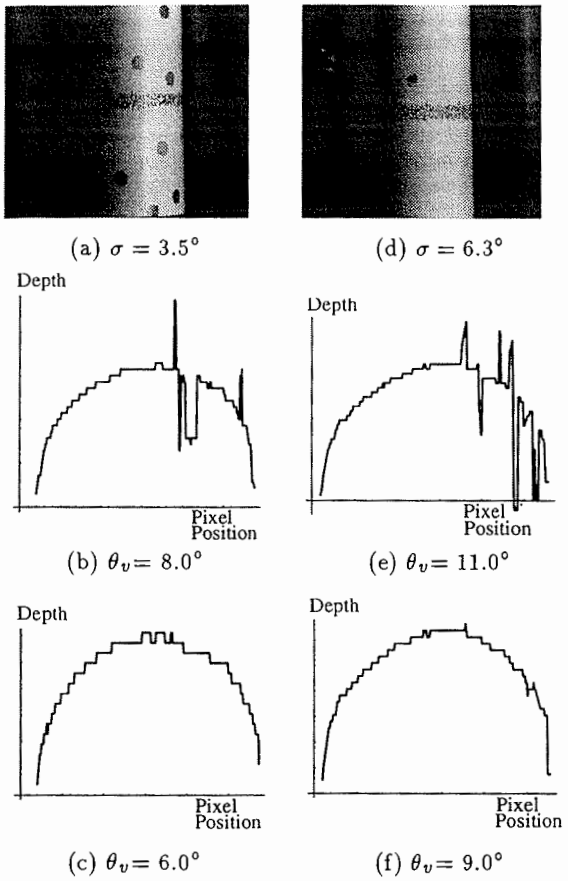


Figure 7: Effect of varying vergence on correspondence. (a-c) Image of the object with gift wrapper surface and depth obtained along a scanline using the vergence values shown. (d-f) Image of the object with rough xerox paper surface and depth obtained along a scanline using the vergence values shown. Notice that for both surfaces, depth is incorrectly recovered at larger vergence, although a relatively higher vergence is acceptable for the rougher surface.

accurately computed in at least one stereo pair, *regardless* of surface roughness. While the configuration need not be limited to three sensors [2], increasing the number of sensors makes stereo implementation cumbersome. We analyze a planar symmetric trinocular stereo system (see Figure 9) with α as a single configurable parameter. Therefore, the following geometrical constraints ($d1$) hold:

$$\begin{aligned} \hat{v}_1 \cdot \hat{v}_2 &= \hat{v}_2 \cdot \hat{v}_3 \\ \hat{v}_m \cdot \hat{k},\ \hat{v}_m \cdot \hat{n},\ \hat{s} \cdot \hat{n} &> 0 \\ \hat{v}_m \cdot \hat{j} &= 0,\ m = 1,2,3 \end{aligned} \tag{10}$$

For any scene point, I_{diff} must not be too large in at least one stereo pair. This constraint ($d2$) can be expressed as:

$$\exists (i,k)(|I_s^i - I_s^k| < T),\ k \neq i\,,\ 1 \leq i,k \leq 3 \tag{11}$$

Note that the two views which satisfy the above constraint can change from one scene point to the next. Therefore, if the constraint is satisfied for all scene points, then an algorithm can be designed that switches between different stereo pairs to construct a complete and accurate depth map.

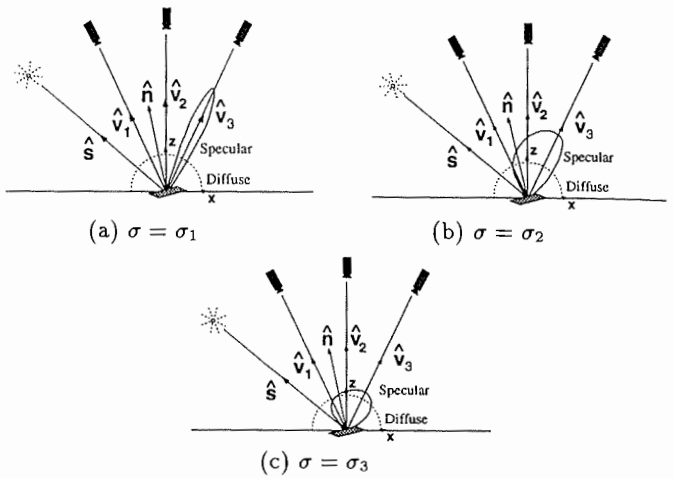


Figure 8: A trinocular configuration illustrated on three surfaces with $\sigma_1 < \sigma_2 < \sigma_3$. (a) $|I_s^1 - I_s^2| < T$, (b) $|I_s^2 - I_s^3| < T$, and (c) $|I_s^1 - I_s^2| < T$, $|I_s^2 - I_s^3| < T$ and $|I_s^1 - I_s^3| < T$.

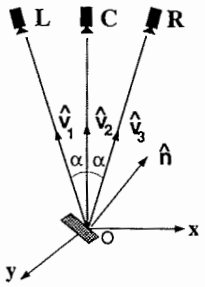
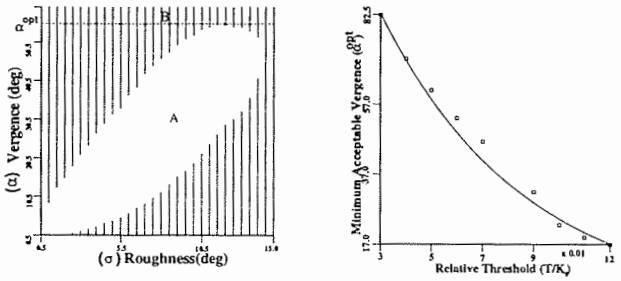


Figure 9: Layout of a trinocular configuration with the cameras labelled as L, C and R.

We analyze the following problem – Determine those values of the trinocular vergence parameter α which satisfy the constraints $d1$ and $d2$. Like in the case of binocular stereo, the relative threshold T/K_s and the roughness σ are free parameters. The problem is solved numerically, and Figure 10(a) illustrates the corresponding solution space (α vs σ) for a given value of T/K_s. The unshaded region marked A denotes unacceptable vergences while the shaded region represents acceptable vergence values. Notice that all $\alpha > \alpha^{opt}$ are acceptable values for *any roughness value*. In other words, α^{opt} denotes that vergence beyond which it is ensured that the intensity difference does not exceed the chosen value of threshold in at least one pair of views for any scene point. This is true for arbitrary surface roughness. α^{opt} is termed as the *minimum acceptable vergence*. Figure 10(b) illustrates the variation of α^{opt} with T/K_s. The monotonically decreasing relationship suggests that a conservative lower bound for T/K_s will give good depth resolution unlike binocular stereo.

Another advantage of trinocular stereo is that the depth resolution obtainable is greater than its binocular counterpart for smoother surfaces since $\alpha^o pt$ is much higher than the corresponding θ_v^{opt}. It can be observed that the binocular stereo solution is subsumed in Figure 10.



(a) (b)

Figure 10: The trinocular stereo solution space. (a) $\alpha - \sigma$ plot with relative threshold $(T/K_s) = 0.06$. Region B denotes the region $\alpha > \alpha^{opt}$ where all vergences are acceptable. (b) Variation of α^{opt} with T/K_s.

7 Reconstruction

This section describes an algorithm for matching three views obtained using a configuration with $\alpha > \alpha^{opt}$. The three designated stereo image pairs are, (L, R), (C, R) and (L, C). The essence of the algorithm lies in determining which of the three stereo pairs provides a "good" depth estimate for any point in the scene.

To evaluate goodness of a match, two confidence tests are used: (i) $C1$: Compare the NCC value obtained with a predefined threshold. Only if the normalized correlation value is higher, accept the match. At a wrong match, texture and shading between the windows being different, similarity is expected to be poor. (ii) $C2$: Let I_1 and I_2 denote two stereo images. If x_b is the current match in image I_2 for pixel x_a in I_1, then reverse the search and find the corresponding pixel for x_b by searching in I_1. This match must coincide with x_a if x_b and x_a are corresponding points.

Reconstruction Algorithm:

(1) Initialize the current stereo pair to (L, R). The reason for choosing this pair is that it yields maximum vergence thereby providing good depth resolution.

(2) Choose a pixel x_L in L with sufficient neighbouring texture, and perform the following steps:

(2.1) Find a matching pixel in R. Using confidence tests $C1$ and $C2$, evaluate the goodness of match. If the match is good, compute depth and go to step 2. If not, the current stereo pair (L, R) cannot be used for matching pixel x_L, and then perform the following steps:

(2.1.1) Set (L, C) as the current stereo pair, and find the corresponding pixel for x_L in C. Evaluate the confidence of matching using $C1$ and $C2$. If the match is good, then compute the corresponding pixel x_R in image R by transformation. Compute depth using x_L and x_R and go to step 2. If the match is not good, then the current stereo pair too has failed to establish correspondence, and hence perform the following step:

(2.1.1.1) Set (C, R) as the current stereo pair. If S is the search range, find that pixel x_C in C within the range $(x_L - S, x_L + S)$ which matches well with x_R in R, *and* together map onto x_L when transformed into the image coordinate system of L. The mapping under orthographic projection is given in [2]. Thus, we establish consistent correspondence for x_L in the three images. Compute depth using x_L and x_R and go to step 2. If no such consistent correspondence can be established, then depth cannot be computed at point x_L, hence go back to step 2 for processing the next pixel.

Note that the complexity of the algorithm is of the order of simple linear correlation computed twice over every pixel for which depth is being computed. The complexity increases in proportion to the number of switches between stereo pairs. If the surface geometry is known, then the switching sequence and the total complexity can actually be evaluated [2].

8 Experiments

We present trinocular stereo experiments with objects of different roughness. Here we *do not* estimate surface roughness as required in the case of binocular stereo. Figure 6 shows the photograph of the experimental stereo setup used. As with the experiments on binocular stereo (section 5), different vergence values are obtained by moving the camera in a circle about a center close to which objects are placed.

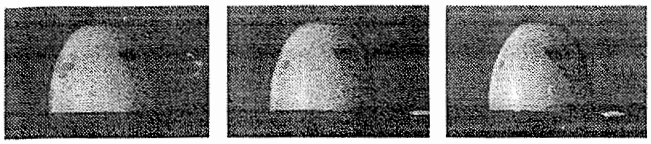
Figure 11 shows trinocular stereo images of an egg-shaped object. The object is relatively rough, as is perceivable from the spread out highlight region. Notice that the specular region shifts in the image space differently from the neighboring texture. The images were obtained using $\alpha = 7.5°$, i.e the binocular vergence with the left and right images is $15.0°$. This value was chosen to keep search ranges relatively small. A large value for α will necessitate a coarse to fine matching strategy which we have not implemented currently. We used a single distant light source in order to keep the experiments consistent with the theory. The performance of the reconstruction algorithm is first illustrated on one scanline. Figure 12 compares our algorithm with naive binocular stereo matching (using views L and R). It can be seen that our algorithm works well demonstrating robustness to specularities. A complete depth map is shown in Figure 13.

The second scene (Figure 14) contains two objects with different surfaces: a vase shaped object whose roughness varies over the surface, and a cylindrical object with unknown roughness. Again, $\alpha = 7.5°$ was used to capture the trinocular images. Figure 15 illustrates the depth map of the scene produced by the reconstruction algorithm. The experiments demonstrate that the algorithm works reasonably well in the case of objects with different reflectance characteristics, an essential requirement for a practical stereo algorithm.

9 Conclusion

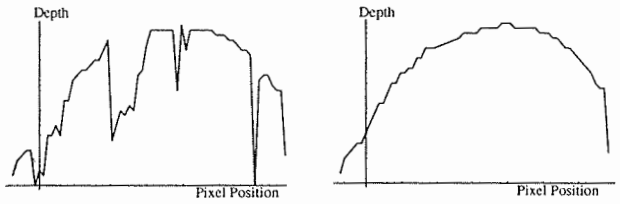
We conclude our paper by summarizing its main results and contributions :

- We have developed a physically based approach for reliable stereo in the presence of specular reflection.
- A scene independent binocular stereo solution was obtained by minimizing intensity differences at corresponding points while maximizing depth resolution. The solution was shown to be a function of surface roughness. Hence, this configuration is usable in structured environments where roughness can be assessed.



(a) Left (L) (b) Center (C) (c) Right (R)

Figure 11: Trinocular stereo images of an egg-shaped object, obtained using $\alpha = 7.5°$. The images are gamma corrected to enhance contrast for display.



(a) (b)

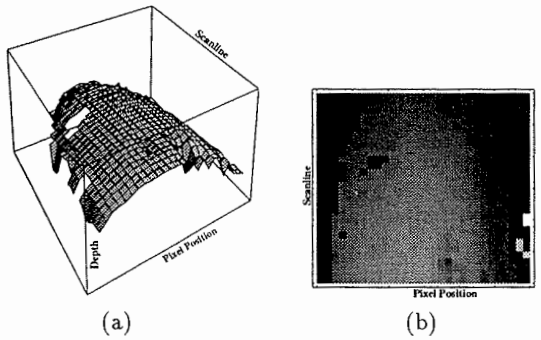
Figure 12: Depth computed for the egg shaped object along a scanline, (a) using views L and R; and (b) using the proposed reconstruction algorithm which uses all three views.

- Trinocular stereo configurations were derived to obviate the need for surface roughness measurement. These configurations can be used in scenes containing unknown objects with possibly varying reflectance properties.
- We have developed a simple algorithm for reconstructing accurate depth maps from three views of a scene that include specular reflections from surfaces of unknown roughness.

References

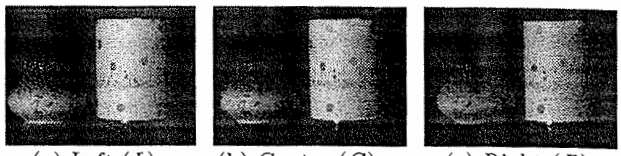
[1] S. T. Barnard and M. A. Fischler. Computational stereo. *ACM Computing Surveys*, 14(4):553–572, Dec. 1982.

[2] D. Bhat and S. K. Nayar. Stereo and specular reflection. Technical Report CUCS-030-94, Columbia Univ., Nov. 1994.



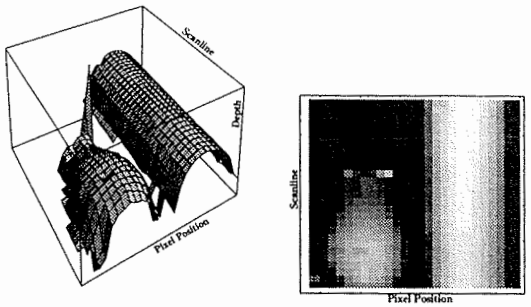
(a) (b)

Figure 13: The depth map of the egg-shaped object, computed using the proposed reconstruction algorithm. (a) A surface plot, and (b) A density plot.



(a) Left (L) (b) Center (C) (c) Right (R)

Figure 14: Trinocular stereo images of a scene with two objects of different roughness, obtained using $\alpha = 7.5°$. The center of the cylinder and the vase-shaped object exhibit significant specular reflection.



(a) (b)

Figure 15: The depth map of the two object scene, computed using the proposed algorithm. (a) A surface plot, and (b) a density plot.

[3] A. Blake. Specular stereo. *Proc. of the 9th Intl. Joint Conf. on Art. Intell. (IJCAI)*, pages 973–976, 1985.

[4] G. Brelstaff and A. Blake. Detecting specular reflections using Lambertian constraints. *Proc. of the IEEE Intl. Conf. on Comp. Vis. (ICCV)*, pages 297–302, 1988.

[5] W. S. Ching, P. S. Toh, K. Luk, and M. H. Er. Robust vergence with concurrent detection of occlusion and specular highlights. *Proc. of the IEEE Intl. Conf. on Comp. Vis. (ICCV).*, pages 384–394, 1993.

[6] U. R. Dhond and J. K. Aggarwal. A cost-benefit analysis of a third camera for stereo correspondence. *Intl. Jrnl of Comp. Vis. (IJCV)*, 6:39–58, 1991.

[7] S. K. Nayar, X. S. Fang, and T. Boult. Removal of specularities using color and polarization. *Proc. of the IEEE Conf. on Comp. Vis. and Patt. Recog. (CVPR)*, pages 583–590, 1993.

[8] S. K. Nayar, K. Ikeuchi, and T. Kanade. Surface reflection: Physical and geometrical perspectives. *IEEE Trans. on Patt. Anal. and Mach. Intell. (PAMI)*, 13(7):611–634, Jul. 1991.

[9] M. Okutomi and T. Kanade. A multiple-baseline stereo. *IEEE Trans. on Patt. Anal. and Mach. Intell. (PAMI)*, 15(4), Apr. 1993.

[10] M. Oren and S. K. Nayar. Seeing beyond Lambert's law. *Proc. of the Euro. Conf. on Comp. Vis. (ECCV)*, pages 269–280, 1994.

[11] K. E. Torrance and E. M. Sparrow. Theory for off-specular reflection from roughened surfaces. *Jrnl. of the Opt. Soc. of America*, 57:1105–1114, 1967.