Reflectance Ratio: A Photometric Invariant for Object Recognition

Shree K. Nayar and Ruud M. Bolle *

Abstract

Neighboring points on a smoothly curved surface have similar surface orientations and illumination conditions. Hence, their brightness values can be used to compute the ratio of their reflectance coefficients. Based on this observation, we develop an efficient algorithm that estimates a reflectance ratio for each region in an image with respect to its background. The region reflectance ratio represents a physical property of a region that is invariant to the illumination conditions. The ratio invariant is used to recognize objects from a single brightness image of a scene. We conclude with experimental results that demonstrate the power of using reflectance and geometric properties of objects, simultaneously.

1 Introduction

Object recognition has been an active area of machine vision research for the past two decades [1]. The traditional approach has been to recover geometric features from images and then use these features to hypothesize and verify the existence of three-dimensional objects in the image. Edges and vertices are examples of geometric features often used by recognition systems. In the past, little attention has been given to the use of other physical properties of objects for recognition. In addition to its geometry, an object may be characterized by intrinsic properties such as reflectance, roughness, and material type. Clearly, the representation of an object using all of these intrinsic properties is useful only if the recognition system is able to compute the properties from images.

In this paper, we present a method for computing the reflectance of regions in a scene, with respect to their backgrounds, from a single image. The result is a physical property of each scene region that is invariant to the intensity and direction of illumination. This photometric invariant, referred to as the *reflectance ratio*, provides valuable information for recognition tasks. The reflectance ratios (photometric features) of object regions and the spatial configuration (geometric features) of the regions are used to represent the object.

The problem of computing the reflectance of regions in a scene was first addressed by Land [2]. In general, image brightness is the product of surface reflectance and illumination. Land developed the retinex theory that suggests computational steps for recovering the reflectance (or lightness) of scene regions in the presence of varying illumination. Subsequently, several hardware implementations for the retinex theory were proposed [3], [4]. The main idea underlying Land's lightness computation is global consistency. The lightness value computed for any particular region must be consistent with those computed elsewhere in the image. However, realistic images include shadows, occlusions, and noise. Each one of these factors can cause a region boundary to go undetected or the computed lightness of a region to be erroneous. Such errors can greatly affect the lightness values computed for all other regions in the image. For this reason, Land's global method is not applicable to most real images.

In this paper, we develop an alternative scheme for computing the ratio of the reflectance of a region to that of its background. The image is first segmented into regions of constant (but unknown) reflectance. Next, a reflectance ratio is computed for each region and its background using only points that lie close to the region's boundary. In this approach, the reflectance ratio computed for any particular region is not affected by those computed for regions elsewhere in the image. Land's analysis [2] was restricted to planar (two-dimensional) scenes with patches of constant reflectance. In contrast, our derivation of the reflectance ratio is based on the analysis of regions that lie on curved surfaces. In the case of curved surfaces, image brightness variations result from both illumination variations as well as surface normal changes. For curved surfaces, our reflectance ratio invariant is valid when a region and its background have the same distribution (scattering) function but different reflectance coefficients (albedo).

Recently, Finlayson [5] proposed computing histograms using ratios in different color channels for object recognition. Histograms, however, are in general

*S. K. Nayar is with the Department of Computer Science, Columbia University, New York, NY 10027. R. M. Bolle is with the IBM T.J. Watson Research Center, Yorktown Heights, N.Y. 10598, U.S.A. This research was supported in part by DARPA Contract No. DACA 76-92-C-0007 and in part by the David and Lucile Packard Fellowship.

0-8186-3870-2/93 $3.00 © 1993 IEEE

sensitive to the scale and rotation of objects in the scene and hence are not effective for three-dimensional object recognition and pose estimation. Here, we use the reflectance ratio invariant to recognize objects from a single image. This approach is very effective in the case of man-made objects that have printed characters and pictures. Each object is assumed to have a set of regions, each with constant reflectance. The reflectance ratio and center of each region are used to represent objects using a hash table. Recognition and pose estimation algorithms are presented that use the reflectance ratios of scene regions to index the hash table. The result is a hypothesis for the existence of an object in the image. This hypothesis is verified using the reflectance ratios and locations of other regions in the scene. Recognition results are presented for realistic scenes with occlusion, shadows, and illumination variations. These results show the simultaneous use of reflectance and geometry to be a powerful approach to object recognition problems.

2 Reflectance Ratio Invariant

The reflectance of a surface depends on its roughness and material properties. In general, incident light is scattered by a surface in different directions. This distribution of reflected light can be described as a function of the angle of incidence, the angle of emittance, and the wavelength of incident light. Consider an infinitesimal surface patch with normal $\mathbf{n}$, illuminated with monochromatic light of wavelength λ from the direction s, and viewed from the direction $\mathbf{v}$. The reflectance of the surface patch can be expressed as: $r(\mathrm{s}, \mathbf{v}, \mathbf{n}, \lambda)$. Now consider an image of the surface patch. If the spectral distribution of the incident light is $e(\lambda)$ and the spectral response of the sensor is $s(\lambda)$, the image brightness value produced by the sensor is:

$$I = \int s(\lambda)\, e(\lambda)\, r(\mathrm{s}, \mathbf{v}, \mathbf{n}, \lambda)\, d\lambda \qquad (1)$$

If we assume the surface patch is illuminated by "white" light and the spectral response of the sensor is constant within the visible-light spectrum, then $s(\lambda) = s$ and $e(\lambda) = e$. We have:

$$I = s\ e\ \rho\ R(\mathrm{s}, \mathbf{v}, \mathbf{n}) \qquad (2)$$

where $\rho\, R(\mathrm{s}, \mathbf{v}, \mathbf{n})$ is the integral of $r(\mathrm{s}, \mathbf{v}, \mathbf{n}, \lambda)$ over the visible-light spectrum. We have decomposed the result into $R(.)$ which represents the dependence of surface reflectance on the geometry of illumination and sensing, and ρ which may be interpreted as the fraction of the incident light that is reflected in all directions by the surface. Incident light that is not reflected by the surface is absorbed or transmitted through the surface. Two surfaces with the same distribution function $R(.)$ can have different reflectance coefficients ρ.

As a result of the white-light assumption, the reflectance coefficient ρ is independent of wavelength. This enables us to represent the reflectance of the surface element with a single constant. The same can be achieved by using an alternative approach which does not require making assumptions about the spectral distribution of the incident light and the spectral response of the sensor. Consider a narrow-band filter with spectral response $f(\lambda)$, placed in front of the sensor. Image brightness is then:

$$I = \int f(\lambda)\, s(\lambda)\, e(\lambda)\, r(\mathrm{s}, \mathbf{v}, \mathbf{n}, \lambda)\, d\lambda \qquad (3)$$

Since the filter is narrow-band, it essentially passes a single wavelength λ' of reflected light. Its spectral response can therefore be expressed as:

$$f(\lambda) = \delta\,(\lambda' - \lambda) \qquad (4)$$

The image brightness measured with such a filter is:

$$I = s'\, e'\, r(\mathrm{s}, \mathbf{v}, \mathbf{n}, \lambda') \qquad (5)$$

where $s' = s(\lambda')$ and $e' = e(\lambda')$. Once again, the reflectance function can be decomposed into a geometrical function and a reflectance coefficient:

$$I = s'\, e'\, \rho'\, R'(\mathrm{s}, \mathbf{v}, \mathbf{n}) \qquad (6)$$

In this case, $R'(.)$ represents the distribution of reflected light for a particular wavelength of incident light. On the other hand, for white-light illumination, $R(.)$ represents the distribution computed as an average over the entire visible-light spectrum. However, the individual terms in both (2) and (6) represent similar effects. Since we have used the white-light illumination assumption in our experiments, we will use the following expression for image brightness in our discussion:

$$I = k\, \rho\, R(\mathrm{s}, \mathbf{v}, \mathbf{n}) \qquad (7)$$

The constant $k = s.e$ accounts for the brightness of the light source and the response of the sensor. The exact functional form of $R(\mathrm{s}, \mathbf{v}, \mathbf{n})$ is determined to a great extent by micro-structure of the surface; generally $R(\cdot)$ includes a diffuse component and a specular component [6]. Once again, the reflection coefficient ρ is the fraction of incident light that is reflected by the surface. It represents the reflective power of the surface and is sometimes referred to as surface albedo.

Consider two *neighboring* points on a surface (Figure 1). For a smooth continuous surface, the two points may be assumed to have the same surface normal vectors. Further, the two points have the same source and sensor directions. Hence, the brightness values, I_1 and I_2, of the two points may be written as:

$$I_1 = k\, \rho_1 R_1(\mathrm{s}, \mathbf{v}, \mathbf{n}) \qquad (8)$$

$$I_2 = k\, \rho_2 R_2(\mathrm{s}, \mathbf{v}, \mathbf{n}) \qquad (9)$$

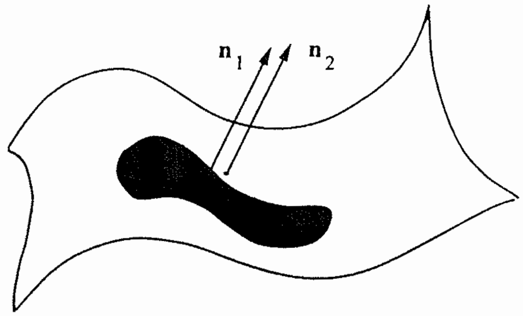


Figure 1: Neighboring points on a surface.

The main assumption made in computing the reflectance ratio is that the two points have the same scattering functions ($R_1 = R_2 = R$) but their reflectance coefficient ρ_1 and ρ_2 may differ. An example is that of two neighboring Lambertian points that have different albedo values because they lie in regions that have different shades or colors. Then, the image brightness values produced by the two points are:

$$\begin{aligned} I_1 &= k\,\rho_1 R(\mathrm{s}, \mathbf{v}, \mathbf{n}) \\ I_2 &= k\,\rho_2 R(\mathrm{s}, \mathbf{v}, \mathbf{n}) \end{aligned} \tag{10}$$

The ratio of the reflectance coefficients of the two points is:

$$p = I_1/I_2 = \rho_1/\rho_2 \tag{11}$$

Note that p is independent of the reflectance function, illumination direction and intensity, and the surface normal of the two points. It is a photometric invariant that is easy to compute and does not vary with the position and orientation of the surface with respect to the sensor and the source. Further, it represents an intrinsic surface property that can be used for object recognition.

We have assumed that the scene is illuminated by a single light source. Now consider the same scene illuminated by several light sources. The brightness of any point can be written as:

$$I = \rho\,[k_1 R(\mathbf{s}_1, \mathbf{v}, \mathbf{n}) + k_2 R(\mathbf{s}_2, \mathbf{v}, \mathbf{n}) ... + k_n R(\mathbf{s}_n, \mathbf{v}, \mathbf{n})] \tag{12}$$

where $\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_n$ are the directions of the n sources that are visible to the surface point under consideration and k_1, k_2, ..., k_n are proportional to the brightness of the n sources. Since the reflectance ratio is computed using neighboring points, it can be assumed that both points are illuminated by the same set of sources. Then, from (11) and (12) we see that the reflectance ratio p is unaffected by the presence of multiple light sources.

Note that by definition p is unbounded; if the second surface point is black, $I_2 = 0$, then $p = \infty$. From a computational perspective, this poses implementation problems. Hence, we use a different definition for p to make it a well-behaved function of the reflectance coefficients ρ_1 and ρ_2:

$$p = (I_1 - I_2)/(I_1 + I_2) = (\rho_1 - \rho_2)/(\rho_1 + \rho_2) \tag{13}$$

Now, we have $-1 \leq p \leq 1$. We will use this definition of the reflectance ratio in the following sections.

3 Reflectance Ratio of a Region

To this point, we have focused on two neighboring points. Now consider a surface region S that has constant reflectance coefficient ρ_1 and is surrounded by a background region with constant reflectance coefficient ρ_2. We are interested in computing the reflectance ratio $P(S)$ of the surface region S with respect to its background. The brightness of the entire region cannot be assumed constant for following two reasons. (a) The surface may be curved and hence the surface normal can vary substantially over the region. (b) While the illumination may be assumed to be locally constant, it may vary over the region. These factors can cause brightness variations, or shading, over the region and its background as well. However, the reflectance ratio can be accurately estimated using neighboring (or nearby) points that lie on either side of the boundary between the region and the background. The reflectance ratio for a region can then be determined as an average of the reflectance ratios computed along the boundary of the region. This computed ratio is also a photometric invariant; it is independent of the three-dimensional shape of the surface and the illumination conditions.

Details of the reflectance ratio algorithm are given in [8]. Due to space limitations, we will simply outline the main steps of the algorithm. The algorithm can be divided in two parts. First, a sequential labeling algorithm [7] is used to segment the image into connected regions. During sequential labeling, the reflectance ratio of neighboring pixels is used as a measure of the "connectivity" between the pixels. In the second stage, a reflectance ratio for each segmented region is computed as the mean of the ratios computed for all points on the boundary of the region. The algorithm is computationally efficient in that reflectance ratios of all scene regions are computed in just two raster scans of the image [8].

4 Object Recognition

In this section, we apply reflectance ratios to the problem of object recognition. The recognition methods presented here are effective for objects that have markings with different reflectance coefficients. Man-made objects with pictures and text printed on them are good examples of such objects.

Learning Object Models:

Since, our objective is to recover the three-dimensional pose of an object from a single brightness image, the object model must include reflectance ratios of the object as well as the three-dimensional coordinates of the centroids of each region. This is done using a range finder. We use the image sensor of the range finder to also obtain a brightness image of the object. As a result, the range and brightness images of the object are registered. The reflectance ratio algorithm is applied to the brightness image and the ratios ($\hat{P}_m$) and centroids ($\hat{\mathbf{x}}_m$) (in the image) of the object's regions are determined. Next, the range map is used to obtain the three-dimensional coordinates ($\hat{\mathbf{X}}_m$) of points of the object surface that correspond to the region centroids in the image. We assume that though the object surface may be curved, each constant reflectance region is small compared to the size of the object and hence can be assumed to be near-planar. Under this assumption, centroids of regions in the image correspond to centroids of the regions in the 3-D scene. Using the above approach, a ratio-centroid list $L_A = ((\hat{\mathbf{X}}_1, \hat{P}_1), (\hat{\mathbf{X}}_2, \hat{P}_2), \ldots, (\hat{\mathbf{X}}_m, \hat{P}_m), \ldots)$ is obtained for each object. Here, $\hat{\mathbf{X}}_m$, $m = 1, \ldots, M$ are the 3-D centroids of the regions and $\hat{P}_m$, $m = 1, \ldots, M$ are the reflectance ratios.

Next, a hash table [9] is initialized. All object models are stored in the same hash table. The indices in the hash table are invariants that can be computed from a single image of the scene. There are no useful geometric invariants that can be computed from the spatial arrangement of the region centroids. This is because object rotation in the scene changes the relative configuration of the region centroids in the image. Hence, we rely on the photometric invariance of reflectance ratios for indexing into the hash table. We select three regions, i, j, and k on the object and use their reflectance ratios to obtain an index $< \hat{P}_i, \hat{P}_j, \hat{P}_k >$. Indices are formed using only those region triplets (i, j, k) whose centroids in 3-D space lie within the radius of coherence D_A. This ensures that the number of indices generated is $O(N)$, with N the number of visible regions on the object, and not combinatorial in N.

Associated with each index in the hash table is an entry. In the entry are stored, the object identifier $\mathcal{M}_I$, and the 3-D coordinates of the centroids $(\hat{\mathbf{X}}_i, \hat{\mathbf{X}}_j, \hat{\mathbf{X}}_k)$ of the three regions used in the index. The entry also includes the ratio-centroid pairs $(\hat{\mathbf{X}}_m, \hat{P}_m)$, of other object regions that are used for object verification and pose estimation.

The above procedure is applied to all sets of three regions in the list L_A. Each object is typically represented by several indices and entries in the hash table. This process is repeated for all objects, $\mathcal{M}_I, I = 1, \ldots, \mathcal{O}$, of interest to the recognition system. The resulting hash table represents the complete object-model database which is ready for use by the recognition system.

Recognition and Pose Estimation

Though model acquisition requires the use of both a brightness and a range image of each object, recognition and pose estimation is accomplished using a *single* brightness image. The reflectance ratio algorithm is applied to the scene image to obtain the list $L_R = ((\mathbf{x}_1, P_1), (\mathbf{x}_2, P_2), \ldots)$. For recognition, a set of three regions is selected from the list L_R. Consider the three regions (i, j, k). This set is used only if the image centroids of the regions j and k lie within the radius of coherence D_R from the centroid of the region i. The ratios of the three regions are used to form the index $< P_i, P_j, P_k >$. If this index does not have an entry in the hash table, the next set of three regions is selected from L_R. If an entry does exist, we have a hypothesis for the object (say $\mathcal{M}_K$). The entry includes the 3-D centroids of the regions (i, j, k) and a set of centroid-ratio pairs for other regions on the object $\mathcal{M}_K$. Assuming the object hypothesis is correct, we have a correspondence between the image centroids $(\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k)$and the 3-D centroids $(\hat{\mathbf{X}}_i, \hat{\mathbf{X}}_j, \hat{\mathbf{X}}_k)$ in the entry. Under the weak-perspective assumption, the transformation $\mathbf{T}$ from the 3-D scene points to 2-D image points can be computed from the three corresponding 3-D centroids and image centroids using the alignment technique proposed by Huttenlocher and Ullman [10]. In general, however, there exist two solutions to the transformation [10]:

$$\mathbf{x} = \mathbf{T}_{K1}(\hat{\mathbf{X}}) \quad \text{and} \quad \mathbf{x} = \mathbf{T}_{K2}(\hat{\mathbf{X}}) \tag{14}$$

Weinshall [11] has shown that instead of computing these two transformations the inverse of the Grammian of the points $\hat{\mathbf{X}}_i$, $\hat{\mathbf{X}}_j$, and $\hat{\mathbf{X}}_k$ can be used to predict the image coordinates $\hat{\mathbf{x}}_o$ of a fourth 3-D point $\hat{\mathbf{X}}_o$ in the entry. Again, two solutions to $\hat{\mathbf{x}}_o$ exist but if the initial object hypothesis is correct, one of the two solutions is likely to be close to one of the centroids in the list L_R. Further, the relectance ratio $\hat{P}_o$ (in the entry) and P_o (in the list L_R) must be similar. The point $\hat{\mathbf{x}}_o$ is not gauranteed to be in the list L_R since it may not be visible to the sensor or it may be occluded by other objects in the scene. In any case, for the object to be verified, one or more projections of the 3-D regions in the entry must match in location and ratio with regions in the list L_R. If so, the object $\mathcal{M}_K$ has been recognized and its pose is given by either $\mathbf{T}_{K1}$ or $\mathbf{T}_{K2}$.

At this point, all regions used as indices and those that are verified are removed from the list L_R. A new set of three regions is selected from the list and used to form the next index. This process is repeated until either all objects in the hash table have been recognized or all regions in the list L_R have been explained.

5 Experiments

In this section, we present experimental results related to the invariance of reflectance ratios as well as the application of ratios to object recognition. Figure 2(a) shows a brightness image of an object with several constant reflectance regions. The image was obtained under ambient lighting conditions. Figure 2(b) shows the result obtained by applying the reflectance ratio algorithm. Ratio values between -1.0 and 1.0 are offset and scaled to lie between 0 and 255 image brightness levels. Note that all letters in the word "KRYLON" have similar ratio values. The white dot in the center of each region represents the centroid of the region.

The invariance of computed ratios to the illumination direction is illustrated in Figure 3. The direction of a single light source is varied (about the axis of the object) from -70 degrees (left of the object) to 20 degrees (right of the object) in increments of 10 degrees. As seen from the figure, the reflectance ratio for region "K" demonstrates remarkable invariance to illumination direction.

The recognition experiments were conducted on man-made objects with letters and pictures printed on them. The printed regions have reflectance coefficients that depend on the shade or color of the paint used to print them. Figure 4(a) shows the model acquisition results for a 3-D object. The range image was obtained using a light stripe range finder. The vertices of the triangle displayed are the centroids of three regions whose reflectance ratios were used as indices in the hash table. Other nearby regions that are included in the hash table entry for object verification and pose estimation and their centroids are indicated by black boxes. All regions used for model acquisition and recognition are assumed to be near-planar though the objects they lie on may be curved. Under this assumption, the centroid of an image region corresponds to the projection of the centroid of the region in 3-D space.

Recognition and pose estimation is done using a single brightness image of the scene (see Figure 4(b)). The scene consists of several 3-D objects in different orientations and positions. It includes occlusions, shadows, and non-uniform illumination. The reflectance ratio algorithm was applied to the scene image and a total of 18 constant reflectance regions were detected. The index triangle shown in the model image is found and verified in the scene image. The set of three regions in the scene image produce a hypothesis for the object. Other regions in the object model are used to verify this hypothesis using the alignment technique (14). The actual and projected centroids of the verification regions are indicated by black and white boxes, respectively. Some of the verification regions are not found in the scene image since they are occluded by other objects.

6 Conclusion

We have presented a photometric invariant, called region reflectance ratio, that is computed from a single brightness image. Reflectance ratios are used to recognize objects from images. This approach is in contrast to previous recognition methods that rely solely on geometric features for recognition and pose estimation. The recognition technique presented here is capable of automatically learning models of the objects of interest.

Acknowledgement

The authors would like to thank Ushir Shah for his assistance in implementing the model acquisition and recognition algorithms, and Daphna Weinshall for providing the alignment code.

References

[1] R. T. Chin and C. R. Dyer. Model-based recognition in robot vision. *ACM Computing Surveys*, 18(1), March 1986.

[2] E. H. Land. The retinex. *American Scientist*, 52(2):247–264, June 1964.

[3] E. H. Land and J. J. McCann. Lightness and retinex theory. *Journal of Optical Society of America*, 61(1):1–11, January 1971.

[4] B. K. P. Horn. Determining lightness from an image. *Computer Graphics and Image Processing*, 3(1):277–299, December 1974.

[5] G.D. Finlayson. *Colour Object Recognition.* M.S. thesis, Simon Fraser University, 1992.

[6] S. K. Nayar, K. Ikeuchi, and T. Kanade. Surface reflection: Physical and geometrical perspectives. *IEEE Trans. on Pattern Analysis and Machine Intell.*, 13(7):611–634, July 1991.

[7] B. K. P. Horn. *Robot Vision.* MIT Press, Cambridge, MA, 1986.

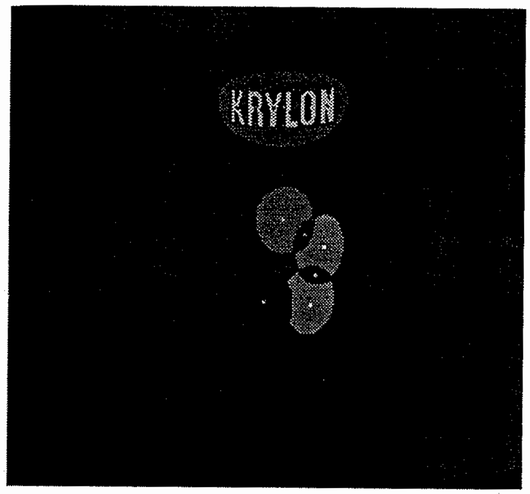
[8] S. K. Nayar and R. M. Bolle. Reflectance based object recognition. Technical Report CUCS-055-92, Department of Computer Science, Columbia University, September 1992.

[9] A.V. Aho, J.E. Hopcroft, and J.D. Ullman. *The Design and Analysis of Computer Algorithms.* Addison-Wesley Publishing Company, Reading, MA, 1974.

[10] D.P. Huttenlocher and S. Ullman. Recognizing solid objects by alignment with an image. *Int. Journal of Computer Vision*, 5(2):195–212, November 1990.

[11] D. Weinshall. Model-based invariants for 3d vision. Technical Report RC 17705, IBM Thomas J. Watson Research Center, December 1991.

(a) Brightness image.



(b) Ratios and centroids of object regions.

Figure 2: Reflectance ratios of regions computed from a single image.

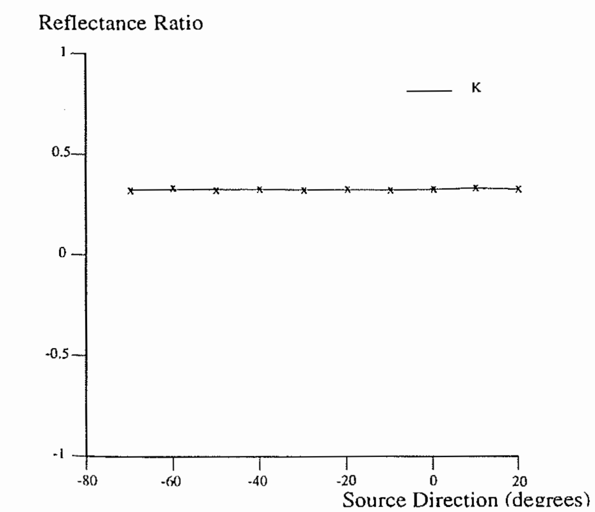
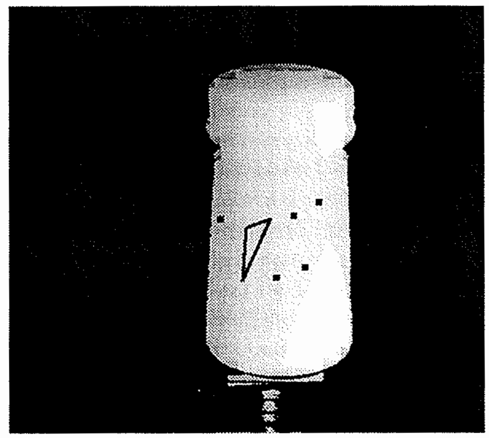


Figure 3: Invariance of reflectance ratios to the direction of illumination.

Brightness image



Range image

(a) Object model acquisition.

(b) Object recognition and pose estimation.

Figure 4: Model acquisition and object recognition results obtained for a three-dimensional recognition problem.